aws

Guida alla gestione

Amazon EMR

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

Amazon EMR: Guida alla gestione

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

I marchi e l'immagine commerciale di Amazon non possono essere utilizzati in relazione a prodotti o servizi che non siano di Amazon, in una qualsiasi modalità che possa causare confusione tra i clienti o in una qualsiasi modalità che denigri o discrediti Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà delle rispettive aziende, che possono o meno essere associate, collegate o sponsorizzate da Amazon.

Table of Contents

Che cos'è Amazon EMR? 1
Comprendere come creare e utilizzare i cluster Amazon EMR 1
Acquisire familiarità con cluster e nodi 2
Invio di lavori a un cluster 2
Elaborazione di dati 3
Comprensione del ciclo di vita del cluster 4
Vantaggi dell'utilizzo di Amazon EMR 5
Risparmio sui costi 6
AWS integrazione 6
Implementazione 7
Scalabilità e flessibilità 7
Affidabilità 8
Sicurezza 8
Monitoraggio 10
Interfacce di gestione 10
Architettura e livelli di servizio 11
Storage 12
Gestione delle risorse del cluster 12
Framework di elaborazione dati 13
Applicazioni e programmi 14
Prima di configurare Amazon EMR 15
Iscriviti per un Account AWS 15
Crea un utente con accesso amministrativo 15
Crea una coppia di EC2 chiavi Amazon per SSH 17
Passaggi successivi 17
Tutorial sulle nozioni di base 18
Configurazione del cluster Amazon EMR 18
Fase 1: configurare le risorse dati e avviare un cluster Amazon EMR 19
Preparare la archiviazione per Amazon EMR 19
Preparazione di un'applicazione con i dati di input per Amazon EMR 19
Avvio di un cluster Amazon EMR 22
Fase 2: invia il lavoro al tuo cluster Amazon EMR 25
Invia il lavoro e visualizza i risultati 25
Visualizzazione dei risultati 28

Fase 3: pulizia ... 32
Terminazione di un cluster ... 32
Eliminazione di risorse S3 ... 34
Passaggi successivi ... 35
Valutazione delle applicazioni di big data per Amazon EMR ... 35
Pianificazione dell'hardware, della rete e della sicurezza del cluster ... 35
Gestione di cluster ... 35
Uso di un'interfaccia diversa ... 35
Consultazione del blog tecnico su EMR ... 36
Gestione dei cluster Amazon EMR con la console ... 37
Funzionalità della console ... 37
Riepilogo delle differenze ... 38
Compatibilità dei cluster nella console ... 38
Creazione di cluster ... 38
Visualizzazione e ricerca di cluster ... 40
Visualizzazione o modifica dei dettagli del cluster ... 41
Differenze nell'utilizzo delle configurazioni di sicurezza ... 42
Amazon EMR Studio ... 44
Funzionalità principali ... 44
IDE di Amazon SageMaker Unified Studio ... 45
Cronologia delle funzionalità ... 45
Come funziona ... 46
Autenticazione e accesso utente ... 47
Controllo accessi ... 51
WorkSpace ... 52
Archiviazione di notebook ... 53
Caratteristiche, requisiti e limiti di EMR Studio ... 53
Considerazioni ... 53
Problemi noti ... 56
Limitazioni delle caratteristiche ... 57
Limiti del servizio ... 58
Best practice per VPC e sottorete per EMR Studio ... 58
Requisiti del cluster ... 59
Configurazione di EMR Studio ... 61
Autorizzazioni di amministratore per creare un EMR Studio ... 62
Configurazione di un EMR Studio ... 68

Monitora, aggiorna ed elimina le risorse di Amazon EMR Studio 140
Crittografia dei notebook dell'area di lavoro 147
Controllo del traffico di rete EMR Studio 150
Creazione di modelli di cluster 152
Accesso e autorizzazioni per i repository basati su Git 157
Ottimizzazione dei processi Spark 160
Utilizzo di EMR Studio 162
Scopri gli spazi di lavoro di EMR Studio 163
Collaborazione dei WorkSpace 171
Esecuzione di un Workspace con un ruolo di runtime 174
Esegui i notebook Amazon EMR Studio Workspace Workspace a livello di programmazione 179
Sfoglia i dati con SQL Explorer per EMR Studio 179
Collegamento di un calcolo a un WorkSpace 180
Collegamento di repository Git 188
Integrazione di Athena 192
CodeWhisperer integrazione 194
Debug di applicazioni e processi 196
Installazione di kernel e librerie 200
Comandi magic 201
Usare notebook multilingue con i kernel Spark 210
Notebook EMR 212
Notebook nella console 213
Informazioni sulla transizione 213
Che cosa occorre fare? 214
Vantaggi di Workspace 214
Autorizzazioni richieste 214
Considerazioni 216
Requisiti del cluster 216
Differenze nelle funzionalità in base alla versione del cluster 217
Limiti di notebook EMR collegati contemporaneamente 218
Versioni Jupyter Notebook e Python 219
Considerazioni relative alla sicurezza 219
Creazione di un notebook 220
Operazioni con Notebook EMR 223
Comprensione dello stato dei notebook 224

Utilizzo dell'editor di notebook ... 225
Modifica dei cluster ... 226
Eliminazione dei notebook e dei relativi file ... 227
Condivisione di file del notebook ... 228
Esempi di comandi programmatici per Notebooks EMR ... 229
Panoramica ... 229
Autorizzazioni ... 230
Limitazioni ... 231
Esempi ... 232
Esempi di comandi CLI per notebook ... 232
Script di esempio dell'SDK Boto3 ... 237
Script di esempio Ruby ... 240
Rappresentazione utente per Spark ... 242
Impostazione della rappresentazione utente Spark ... 243
Utilizzo del widget di monitoraggio dei processi Spark ... 244
Sicurezza ... 244
Installazione e utilizzo di kernel e librerie in EMR Studio ... 245
... 245
Installazione di kernel e librerie Python su un nodo primario del cluster ... 246
Considerazioni e limitazioni relative alle librerie con ambito notebook ... 248
Lavorare con le librerie con ambito notebook ... 249
Associazione di repository basati su Git con EMR Notebooks ... 250
Prerequisiti e considerazioni ... 251
Aggiunta di un repository basato su Git ad Amazon EMR ... 255
Aggiorna o elimina un repository basato su Git da un EMR Studio Workspace ... 255
Collega o scollega un repository basato su Git in EMR Studio ... 256
Crea un nuovo Notebook con un repository Git associato in EMR Studio ... 257
Utilizzare gli archivi Git in un notebook EMR Studio ... 258
Pianifica, configura e avvia i cluster Amazon EMR ... 260
Avvia rapidamente un cluster Amazon EMR ... 260
Configurazione della posizione e dello storage dei dati del cluster Amazon EMR ... 261
Scegli una AWS regione per il tuo cluster Amazon EMR ... 261
Lavorare con sistemi di storage e file system ... 263
Prepara i dati di input per l'elaborazione con Amazon EMR ... 267
Configurare una posizione per l'output del cluster Amazon EMR ... 287
Pianifica e configura i nodi primari nel tuo cluster Amazon EMR ... 293

Funzionalità che supportano l'elevata disponibilità in un cluster Amazon EMR e il loro funzionamento con applicazioni open source ... 295
Avvio di un cluster Amazon EMR con più nodi primari ... 304
Integrazione di Amazon EMR con EC2 i gruppi di collocamento ... 310
Considerazioni e best practice per un cluster con più nodi primari ... 318
Cluster EMR attivi AWS Outposts ... 321
Prerequisiti ... 321
Limitazioni ... 321
Considerazioni sulla connettività di rete ... 322
Creazione di un cluster Amazon EMR su AWS Outposts ... 323
Cluster EMR su Local Zones AWS ... 324
Tipi di istanze supportati ... 324
Creazione di un cluster Amazon EMR sulle Local Zones ... 325
Configura Docker per l'uso con i cluster Amazon EMR ... 326
Registri Docker ... 327
Configurazione dei registri Docker ... 327
Configurazione di YARN per accedere ad Amazon ECR su EMR 6.0.0 e versioni precedenti ... 329
Controlla la terminazione dei cluster Amazon EMR ... 331
Configurazione di un cluster Amazon EMR per continuare o terminare dopo l'esecuzione delle fasi ... 332
Utilizzo di una policy di terminazione automatica ... 335
Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali ... 341
Sostituzione di nodi non integri con Amazon EMR ... 347
Impostazioni predefinite per la sostituzione dei nodi e la protezione dalla terminazione ... 349
Configurazione della sostituzione di nodi non integri all'avvio di un cluster ... 349
Configurazione della sostituzione non corretta dei nodi in un cluster in esecuzione ... 351
Lavorare con AMIs ... 352
Panoramica ... 352
Utilizzo dell'AMI predefinita ... 353
Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR ... 531
Cambia la versione AL ... 544
Personalizzazione del volume root EBS ... 545
Configura le applicazioni all'avvio del cluster Amazon EMR ... 549

Creazione di operazioni di bootstrap 550
Configurazione dell'hardware e della rete del cluster Amazon EMR 555
Descrizione dei tipi di nodo 556
Configurazione dei tipi di EC2 istanze Amazon da utilizzare con Amazon EMR 559
Configurazione del logging e del debug dei cluster Amazon EMR 1539
File di log di default 1539
Archiviazione di file di log in Amazon S3 1540
Posizione dei log 1546
Etichetta e classifica le risorse del cluster Amazon EMR 1548
Restrizioni che si applicano all'etichettatura delle risorse in Amazon EMR 1549
Contrassegna le risorse Amazon EMR per la fatturazione 1550
Aggiungere tag a un cluster Amazon EMR 1550
Visualizza i tag su un cluster Amazon EMR 1552
Rimuovere i tag da un cluster Amazon EMR 1553
Integrazione di driver e applicazioni di terze parti su Amazon EMR 1554
Utilizzo degli strumenti di Business Intelligence con Amazon EMR 1555
Sicurezza 1556
Sicurezza della rete e dell'infrastruttura 1556
Aggiornamenti per l'AMI predefinita di Amazon Linux 1557
AWS Identity and Access Management con Amazon EMR 1558
Cluster single-tenant e multi-tenant 1559
Protezione dei dati 1560
Controllo dell'accesso ai dati 1560
Configurazioni di sicurezza 1561
Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI 1561
Specifica una configurazione di sicurezza 1593
Protezione dei dati 1594
Crittografa i dati inattivi e in transito con Amazon EMR 1595
IAM con Amazon EMR 1632
Destinatari 1632
Autenticazione con identità 1633
Gestione dell'accesso con policy 1637
Funzionamento di Amazon EMR con IAM 1639
Ruoli di runtime per le fasi di Amazon EMR 1647
Configurazione dei ruoli di servizio per Amazon EMR 1656
Esempi di policy basate su identità 1722

S3 Access Grants con Amazon EMR 1765
Panoramica 1765
Come funziona 1766
Considerazioni 1766
Avvia un cluster 1768
`Lake Formation` 1769
`fallbackToIAM` 1770
Autenticazione nei nodi cluster 1770
Usa una coppia di EC2 chiavi per le credenziali SSH per Amazon EMR 1771
Utilizzo dell'autenticazione Kerberos 1771
Utilizzo dell'autenticazione LDAP 1804
Integra Amazon EMR con il Centro identità 1815
Panoramica 1816
Funzionalità 1816
Nozioni di base 1816
Considerazioni 1825
Integrazione di Amazon EMR con Lake Formation 1826
Funzionamento di Amazon EMR con Lake Formation 1827
Prerequisiti 1827
Abilitazione di Amazon EMR con Lake Formation 1828
Hudi e Lake Formation 1832
Iceberg e Lake Formation 1835
Delta Lake e Lake Formation 1836
Considerazioni 1838
Utilizzo delle viste del catalogo dati di Glue in Amazon EMR 1840
Utilizzo di un ruolo di runtime del job Lake Formation con accesso completo alla tabella ... 1847
Integrazione di Amazon EMR con Apache Ranger 1855
Panoramica su Apache Ranger 1855
Considerazioni sull'utilizzo di Amazon EMR con Apache Ranger 1857
Configurazione di Amazon EMR per Apache Ranger 1859
Plugin Apache Ranger per scenari di integrazione Amazon EMR 1877
Risoluzione dei problemi di Apache Ranger 1896
Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR 1900
Utilizzo dei gruppi di sicurezza gestiti da Amazon EMR 1902
Utilizzo di gruppi di sicurezza aggiuntivi per un cluster Amazon EMR 1913
Specifica dei gruppi di sicurezza 1914

Gruppi di sicurezza per EMR Notebooks 1917
Blocco dell'accesso pubblico 1919
Convalida della conformità 1924
Resilienza 1925
Sicurezza dell'infrastruttura 1925
Connessione ad Amazon EMR utilizzando un endpoint VPC di interfaccia 1926
Gestione dei cluster Amazon EMR 1931
Connettiti a un cluster Amazon EMR 1931
Prima di connetterti 1932
Connect al nodo primario del cluster Amazon EMR tramite SSH 1934
Invia il lavoro a un cluster Amazon EMR 1959
Aggiunta di fasi con la console 1960
Aggiunta di fasi con la CLI 1963
Esecuzione di più fasi 1965
Visualizzazione dei passaggi dopo l'invio del lavoro a un cluster Amazon EMR 1966
Annulla i passaggi quando invii il lavoro a un cluster Amazon EMR 1967
Visualizza e monitora un cluster Amazon EMR mentre esegue il lavoro 1969
Visualizza lo stato e i dettagli del cluster Amazon EMR 1969
Debug avanzato in fasi con Amazon EMR 1975
Visualizza la cronologia delle applicazioni Amazon EMR 1977
Visualizza i file di log di Amazon EMR 1986
Visualizza le istanze di cluster in Amazon EC2 1990
CloudWatch eventi e metriche di Amazon EMR 1992
Visualizza i parametri delle applicazioni cluster utilizzando Ganglia con Amazon EMR 2079
CloudTrail registri 2080
Best practice per l'osservabilità dell'EMR 2084
Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione 2085
Considerazioni 2087
Dimensionamento gestito 2087
Dimensionamento automatico con una policy personalizzata 2130
Ridimensionare un cluster in esecuzione 2143
Timeout del provisioning 2151
Opzioni di scalabilità verso il basso per i cluster Amazon EMR 2156
Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa 2160
Termina dalla console 2161

Termina dalla CLI 2161
Termina dall'API 2162
Clonare un cluster 2163
Automatizza i cluster Amazon EMR ricorrenti con AWS Data Pipeline 2164
Tutorial Amazon EMR 2165
Amazon EMR attivo EC2 : monitoraggio avanzato con CloudWatch l'utilizzo di parametri e log personalizzati 2165
Panoramica 2165
Prerequisiti e contesto 2165
..... 2166
Monitora le applicazioni Apache Spark su Amazon EMR con Amazon CloudWatch 2176
Monitora lo stato delle applicazioni Amazon EMR con l'integrazione CloudWatch 2176
Risoluzione dei problemi relativi ai cluster Amazon EMR 2177
Strumenti per la risoluzione dei problemi 2177
Visualizzazione dei dettagli del cluster 2178
Visualizzazione dei dettagli degli errori 2178
Esecuzione di script e configurazione di processi 2179
Visualizzare file di log di 2179
Monitoraggio delle prestazioni del cluster 2180
Visualizzazione e riavvio di processi 2180
Visualizzazione dei processi in esecuzione 2181
Arresto e riavvio di processi 2182
Errori comuni 2185
Codici di errore 2186
Errori di risorse durante le operazioni del cluster Amazon EMR 2203
Errori di input e output del cluster durante le operazioni di Amazon EMR 2218
Errori di autorizzazione durante le operazioni del cluster Amazon EMR 2220
Errori del cluster Hive 2222
Errori VPC durante le operazioni del cluster Amazon EMR 2223
Errori di streaming del cluster Amazon EMR 2227
Amazon EMR: errori di cluster JAR personalizzati 2229
Errori di Amazon EMR AWS GovCloud (Stati Uniti occidentali) 2229
Ricerca di un cluster mancante 2230
Risoluzione dei problemi di un cluster con errori 2230
Fase 1: raccogliere dati sul problema con il cluster Amazon EMR 2231
Fase 2: Controllo dell'ambiente 2232

Fase 3: Esame dell'ultima modifica dello stato 2233
Fase 4: Esamina i file di log di Amazon EMR 2234
Fase 5: testare il cluster Amazon EMR passo dopo passo 2235
Risoluzione dei problemi che rallentano un cluster 2236
Fase 1: raccogliere dati sul problema con il cluster Amazon EMR 2237
Fase 2: Verificare l'ambiente del cluster EMR 2237
Fase 3: Esamina i file di log per il cluster Amazon EMR 2239
Fase 4: verifica dello stato del cluster e dell'istanza Amazon EMR 2241
Fase 5: Verifica della presenza di gruppi sospesi 2242
Fase 6: Rivedere le impostazioni di configurazione per il cluster Amazon EMR 2243
Fase 7: esaminare i dati di input per il cluster Amazon EMR 2246
Risolvi i problemi più comuni durante l'utilizzo di Amazon EMR con Lake Formation AWS 2246
Accesso al data lake non consentito 2246
Scadenza della sessione 2247
Nessuna autorizzazione per l'utente sulla tabella richiesta 2247
Interrogazione dei dati tra account condivisi con Lake Formation 2247
Inserimento, creazione e modifica di tabelle 2249
Scrivi applicazioni che avviano e gestiscono i cluster Amazon EMR 2250
End-to-end Esempio di codice sorgente Java di Amazon EMR 2250
Concetti comuni per le chiamate API Amazon EMR 2254
Endpoint per Amazon EMR 2255
Specifica dei parametri del cluster in Amazon EMR 2255
Zone di disponibilità in Amazon EMR 2256
Come utilizzare file e librerie aggiuntivi in cluster Amazon EMR 2256
SDKs Usalo per chiamare Amazon EMR APIs 2257
Utilizzo di AWS SDK per Java per creare un cluster Amazon EMR 2257
Gestione delle Service Quotas di Amazon EMR 2260
Cosa sono le Service Quotas di Amazon EMR 2260
Come gestire le Service Quotas di Amazon EMR 2261
Quando configurare gli eventi EMR in CloudWatch 2262
AWS Glossario 2265
Cronologia dei documenti 2266
..... mmcclxvii

Che cos'è Amazon EMR?

Amazon EMR, che in precedenza si chiamava Amazon Elastic MapReduce, è una piattaforma di cluster gestita che semplifica l'esecuzione di framework di big data, come Apache Hadoop e Apache Spark, per elaborare e analizzare grandi quantità di dati. AWS Utilizzando questi framework e i relativi progetti open source, è possibile elaborare i dati per scopi di analisi e carichi di lavoro di business intelligence. Amazon EMR consente inoltre di trasformare e spostare grandi quantità di dati all'interno e all'esterno di altri datastore e database AWS , come Amazon Simple Storage Service (Amazon S3) e Amazon DynamoDB.

Se si usa Amazon EMR per la prima volta, è consigliabile iniziare leggendo anche le seguenti sezioni:

- Amazon EMR: questo servizio fornisce i punti salienti di Amazon EMR, i dettagli del prodotto e informazioni sui prezzi.
- Tutorial: Nozioni di base su Amazon EMR: questo tutorial ti consente di iniziare a utilizzare rapidamente Amazon EMR.

In questa sezione

- Comprendere come creare e utilizzare i cluster Amazon EMR
- Vantaggi dell'utilizzo di Amazon EMR
- Architettura e livelli di servizio di Amazon EMR

Comprendere come creare e utilizzare i cluster Amazon EMR

Questo argomento fornisce una panoramica dei cluster Amazon EMR e include come inviare il lavoro a un cluster, come vengono elaborati i dati e le varie condizioni del cluster durante l'elaborazione.

In questo argomento

- Acquisire familiarità con cluster e nodi
- Invio di lavori a un cluster
- Elaborazione di dati
- Comprensione del ciclo di vita del cluster

Acquisire familiarità con cluster e nodi

Il componente centrale di Amazon EMR è il cluster. Un cluster è una raccolta di istanze Amazon Elastic Compute Cloud (Amazon EC2). Ogni istanza nel cluster è denominata nodo. Ogni nodo ha un ruolo all'interno del cluster, indicato come tipo di nodo. Amazon EMR installa anche diversi componenti software su ogni tipo di nodo, dando a ogni nodo un ruolo in un'applicazione distribuita come Apache Hadoop.

I tipi di nodo in Amazon EMR sono i seguenti:

- Nodo primario: un nodo che gestisce il cluster eseguendo componenti software per coordinare la distribuzione di dati e attività tra altri nodi per l'elaborazione. Il nodo primario tiene traccia dello stato delle attività e monitora lo stato del cluster. Ogni cluster dispone di un nodo primario ed è possibile creare un singolo nodo cluster con solo il nodo primario.
- Nodo principale: un nodo con componenti software che eseguono attività e archiviano dati nel File system distribuito Hadoop (HDFS) sul cluster. I cluster multi-nodo dispongono di almeno un nodo principale.
- Nodo attività: un nodo con componenti software che esegue solo le attività e non archivia i dati in HDFS. I nodi attività sono facoltativi.

Invio di lavori a un cluster

Quando si esegue un cluster su Amazon EMR, si hanno diverse opzioni circa il modo di specificare il lavoro che deve essere svolto.

- Fornisci l'intera definizione lavoro da svolgere in funzioni che specifichi come fasi al momento della creazione di un cluster. Questo viene generalmente fatto per i cluster che elaborano una determinata quantità di dati e poi terminano quando l'elaborazione è completa.
- Crea un cluster a lunga durata e utilizza la console Amazon EMR, l'API Amazon EMR o AWS CLI la procedura di invio, che può contenere uno o più lavori. Per ulteriori informazioni, consulta Invia il lavoro a un cluster Amazon EMR.
- Crea un cluster, connettiti al nodo primario e ad altri nodi in base alle esigenze attraverso SSH e utilizza le interfacce che le applicazioni installate forniscono per eseguire le attività e inviare le query, con script o in modo interattivo. Per ulteriori informazioni, consulta la Guida ai rilasci di Amazon EMR.

Elaborazione di dati

Quando avvii il cluster, devi scegliere i framework e le applicazioni da installare per le tue esigenze di elaborazione dati. Per elaborare i dati nel cluster Amazon EMR, è possibile inviare processi o query direttamente alle applicazioni installate, oppure è possibile eseguire fasi nel cluster.

Invio diretto di processi alle applicazioni

È possibile inviare processi e interagire direttamente con il software installato nel cluster Amazon EMR. A tale scopo, in genere, ci si connette al nodo primario attraverso una connessione sicura e si accede alle interfacce e agli strumenti disponibili per il software che è in esecuzione direttamente sul cluster. Per ulteriori informazioni, consulta Connettiti a un cluster Amazon EMR.

Esecuzione di fasi per elaborare i dati

È possibile inviare una o più fasi ordinate a un cluster Amazon EMR. Ogni fase è un'unità di lavoro che contiene le istruzioni per manipolare i dati per l'elaborazione da parte di software installato sul cluster.

Di seguito è riportato un processo di esempio che si articola in quattro fasi:

1. Invio di un set di dati di input per l'elaborazione.
2. Elaborazione dell'output della prima fase con un programma Pig.
3. Elaborazione di un secondo set di dati di input con un programma Hive.
4. Scrittura di un set di dati di output.

Generalmente, quando si elaborano dati in Amazon EMR, l'input è costituito da dati memorizzati come file nel file system sottostante scelto, ad esempio Amazon S3 o HDFS. Questi dati passano da una fase all'altra della sequenza di elaborazione. La fase finale scrive i dati di output in un percorso specificato, ad esempio un bucket Amazon S3.

Le fasi vengono eseguite nella sequenza seguente:

1. Una richiesta viene inviata per iniziare le fasi di elaborazione.
2. Lo stato di tutte le fasi è impostato su PENDING (IN SOSPESO).
3. Quando la prima fase nella sequenza inizia, il relativo stato diventa RUNNING (IN ESECUZIONE). Le altre fasi rimangono nello stato PENDING (IN SOSPESO).
4. Dopo il completamento della prima fase, il suo stato diventa COMPLETED (COMPLETATO).

5. La fase successiva nella sequenza inizia e il relativo stato diventa RUNNING (IN ESECUZIONE). Quando viene completata, il relativo stato diventa COMPLETED (COMPLETATO).
6. Questo schema si ripete per ogni fase fino al loro completamento e alla fine dell'elaborazione.

Il seguente diagramma rappresenta la sequenza delle fasi e il cambiamento di stato delle stesse durante la loro elaborazione.

Se una fase ha esito negativo durante l'elaborazione, lo stato diventa FAILED (NON RIUSCITO). È possibile stabilire il passaggio successivo per ogni fase. Per impostazione predefinita, le restanti fasi nella sequenza sono impostate sullo stato CANCELLED (ANNULLATO) e non eseguite in caso di insuccesso di una fase precedente. È inoltre possibile scegliere di ignorare l'errore e consentire alle fasi rimanenti di proseguire o di terminare il cluster immediatamente.

Il diagramma seguente rappresenta la sequenza delle fasi e il cambiamento di stato predefinito quando una fase ha esito negativo durante l'elaborazione.

Comprensione del ciclo di vita del cluster

Un cluster Amazon EMR andato a buon fine segue questa procedura:

1. Amazon EMR effettua innanzitutto il provisioning delle EC2 istanze nel cluster per ogni istanza in base alle tue specifiche. Per ulteriori informazioni, consulta Configurazione dell'hardware e della rete del cluster Amazon EMR. Per tutte le istanze, Amazon EMR usa l'AMI predefinita di Amazon EMR o un'AMI Amazon Linux personalizzata specificata dall'utente. Per ulteriori informazioni, consulta Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR. Durante questa fase, lo stato del cluster è `STARTING`.
2. Amazon EMR esegue le operazioni di bootstrap specificate per ogni istanza. È possibile utilizzare le operazioni di bootstrap per installare applicazioni personalizzate ed eseguire le personalizzazioni necessarie. Per ulteriori informazioni, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR. Durante questa fase, lo stato del cluster è `BOOTSTRAPPING`.
3. Amazon EMR installa le applicazioni native specificate al momento della creazione del cluster, ad esempio Hive, Spark, Hadoop e così via.
4. Una volta completate le operazioni di bootstrap e installate le applicazioni native, lo stato del cluster è `RUNNING`. A questo punto, puoi connetterti alle istanze del cluster e il cluster eseguirà

in sequenza tutte le fasi specificate durante la creazione del cluster. È possibile aggiungere ulteriori fasi, che vengono eseguite dopo il completamento di tutte le fasi precedenti. Per ulteriori informazioni, consulta Invia il lavoro a un cluster Amazon EMR.

5. Dopo l'esecuzione corretta delle fasi, il cluster entra nello stato `WAITING`. Se un cluster è configurato per terminare in automatico al completamento dell'ultima fase, passa allo stato `TERMINATING` e poi allo stato `TERMINATED`. Se il cluster è configurato per attendere, è necessario terminarlo manualmente quando non è più necessario. Quando viene terminato manualmente, il cluster passa allo stato `TERMINATING` e poi allo stato `TERMINATED`.

Un errore durante il ciclo di vita del cluster porta Amazon EMR a terminare il cluster e tutte le istanze correlate, a meno che non sia stata abilitata la protezione da cessazione. Se un cluster viene terminato a causa di un errore, i dati archiviati al suo interno vengono eliminati e lo stato del cluster viene impostato su `TERMINATED_WITH_ERRORS`. Se hai attivato la protezione dall'arresto, è possibile recuperare i dati dal cluster e quindi rimuovere la protezione e terminare il cluster. Per ulteriori informazioni, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

Il diagramma seguente rappresenta il ciclo di vita di un cluster e come ogni fase del ciclo di vita viene associata a un particolare stato del cluster.

Vantaggi dell'utilizzo di Amazon EMR

L'uso di Amazon EMR offre molti vantaggi. Questi includono la flessibilità offerta AWS e i risparmi sui costi disponibili rispetto alla creazione di risorse locali proprie. Questa sezione ne presenta una panoramica e fornisce link a ulteriori informazioni per approfondire l'argomento.

Argomenti

- Risparmio sui costi
- AWS integrazione
- Implementazione
- Scalabilità e flessibilità
- Affidabilità
- Sicurezza
- Monitoraggio

- Interfacce di gestione

Risparmio sui costi

I prezzi di Amazon EMR dipendono dal tipo e dal numero di EC2 istanze Amazon che distribuisci e dalla regione in cui avvii il cluster. I prezzi su richiesta offrono tariffe basse, ma è possibile ridurre ulteriormente i costi acquistando Istanze riservate o Istanze spot. In alcuni casi, le Istanze spot possono offrire risparmi significativi fino a un decimo dei prezzi su richiesta.

> **Note**
>
> Se utilizzi Amazon S3, Amazon Kinesis o DynamoDB con il cluster EMR, sono previsti costi aggiuntivi che vengono fatturati separatamente rispetto all'utilizzo di Amazon EMR.

> **Note**
>
> Quando si configura un cluster Amazon EMR in una sottorete privata, si consiglia di configurare anche gli endpoint VPC per Simple Storage Service (Amazon S3). Se il cluster EMR si trova in una sottorete privata senza endpoint VPC per Simple Storage Service (Amazon S3), verranno addebitati costi aggiuntivi del gateway NAT associati al traffico S3 perché il traffico tra il cluster EMR e S3 non rimarrà all'interno del VPC.

Per ulteriori informazioni su opzioni di prezzo e dettagli, consulta Prezzi di Amazon EMR.

AWS integrazione

Amazon EMR si integra con altri AWS servizi per fornire capacità e funzionalità relative al networking, allo storage, alla sicurezza e così via per il tuo cluster. Di seguito sono elencati diversi esempi di questa integrazione:

- Amazon EC2 per le istanze che comprendono i nodi del cluster
- Amazon Virtual Private Cloud (Amazon VPC) per configurare la rete virtuale in cui è possibile avviare le istanze
- Amazon S3 per archiviare i dati di input e output
- Amazon CloudWatch per monitorare le prestazioni dei cluster e configurare gli allarmi

- AWS Identity and Access Management (IAM) per configurare le autorizzazioni
- AWS CloudTrail per controllare le richieste fatte al servizio
- AWS Data Pipeline per pianificare e avviare i cluster
- AWS Lake Formation per scoprire, catalogare e proteggere i dati in un data lake Amazon S3

Implementazione

Il cluster EMR è composto da EC2 istanze che eseguono il lavoro che invii al cluster. Quando si avvia il cluster, Amazon EMR configura le istanze con le applicazioni scelte, ad esempio Apache Hadoop o Spark. Scegli la dimensione e il tipo di istanza più adatti alle esigenze di elaborazione del cluster: elaborazione in batch, query a bassa latenza, streaming di dati o archiviazione di grandi quantità di dati. Per ulteriori informazioni sui tipi di istanza disponibili per Amazon EMR, consulta Configurazione dell'hardware e della rete del cluster Amazon EMR.

Amazon EMR offre svariati modi per configurare software sul cluster. Ad esempio, è possibile installare una versione di Amazon EMR con un set scelto di applicazioni che possono includere framework versatili, come Hadoop, e applicazioni come Hive, Pig o Spark. È anche possibile installare una delle diverse distribuzioni di MapR. Amazon EMR usa Amazon Linux, che ti consente di installare il software sul tuo cluster manualmente sfruttando il gestore dei pacchetti yum o direttamente dalla fonte. Per ulteriori informazioni, consulta Configura le applicazioni all'avvio del cluster Amazon EMR.

Scalabilità e flessibilità

Amazon EMR fornisce flessibilità per ridurre o aumentare le dimensioni del cluster al variare delle esigenze di computing. È possibile ridimensionare il cluster per aggiungere istanze per i carichi di lavoro di picco e rimuovere le istanze per controllare i costi quando tali carichi di lavoro si riducono. Per ulteriori informazioni, consulta Ridimensiona manualmente un cluster Amazon EMR in esecuzione.

Amazon EMR fornisce anche la possibilità di eseguire più gruppi di istanze in modo da poter utilizzare le Istanze on demand in un gruppo per garantire la potenza di elaborazione insieme alle Istanze spot in un altro gruppo e completare i processi più velocemente e a costi inferiori. È anche possibile mescolare diversi tipi di istanza per sfruttare i prezzi migliori per un tipo di Istanza spot rispetto a un'altra. Per ulteriori informazioni, consulta Quando occorre utilizzare le istanze Spot?.

Inoltre, Amazon EMR offre la flessibilità di utilizzare diversi file system per i dati di input, output e intermedi. Ad esempio, puoi scegliere il File system distribuito Hadoop (HDFS) che viene eseguito

sui nodi primario e principali del cluster per elaborare i dati che devi archiviare oltre il ciclo di vita del cluster. È possibile scegliere il File system EMR (EMR File System, EMRFS) per utilizzare Amazon S3 come livello di dati per le applicazioni in esecuzione sul cluster in modo da poter separare il calcolo e l'archiviazione e mantenere i dati al di fuori del ciclo di vita del cluster. Come ulteriore vantaggio, EMRFS offre la possibilità di ridurre o aumentare le dimensioni per le esigenze di calcolo e archiviazione in modo indipendente. È possibile scalare le esigenze di calcolo ridimensionando il cluster e le esigenze di archiviazione con Amazon S3. Per ulteriori informazioni, consulta Utilizzo di sistemi di storage e file con Amazon EMR.

Affidabilità

Amazon EMR monitora i nodi del cluster e termina e sostituisce in automatico un'istanza in caso di esito negativo.

Amazon EMR fornisce opzioni di configurazione che controllano la modalità di terminazione del cluster (automatica o manuale). Se configuri la terminazione automatica del cluster, questa viene terminata una volta completate tutte le fasi. Si tratta di un cluster transitorio. Tuttavia, è possibile configurare il cluster in modo che continui a funzionare anche dopo il completamento dell'elaborazione, in modo da poter scegliere di terminarlo manualmente quando non è più necessario. In alternativa, è possibile creare un cluster, interagire direttamente con le applicazioni installate e quindi terminare manualmente il cluster quando non è più necessario. I cluster in questi esempi vengono definiti cluster di lunga durata.

Inoltre, è possibile configurare la protezione di terminazione per evitare che le istanze principali del cluster vengano terminate a causa di errori o problemi durante l'elaborazione. Quando la protezione di terminazione è abilitata, è possibile ripristinare i dati dalle istanze prima della terminazione. Le impostazioni predefinite di queste opzioni differiscono a seconda che si avvii il cluster utilizzando la console, la CLI o l'API. Per ulteriori informazioni, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

Sicurezza

Amazon EMR sfrutta altri AWS servizi, come IAM e Amazon VPC, e funzionalità come le coppie di EC2 chiavi Amazon, per aiutarti a proteggere cluster e dati.

IAM

Amazon EMR si integra con IAM per gestire le autorizzazioni. L'utente definisce le autorizzazioni utilizzando policy IAM da collegare a utenti o gruppi IAM. Le autorizzazioni definite nella policy

determinano le azioni che gli utenti o i membri del gruppo possono eseguire e le risorse a cui possono accedere. Per ulteriori informazioni, consulta Funzionamento di Amazon EMR con IAM.

Inoltre, Amazon EMR utilizza i ruoli IAM per il servizio Amazon EMR stesso e il profilo dell'istanza per le EC2 istanze. Questi ruoli concedono al servizio e alle istanze le autorizzazioni per accedere ad altri AWS servizi per tuo conto. Esiste un ruolo predefinito per il servizio Amazon EMR e un ruolo predefinito per il profilo dell' EC2istanza. I ruoli predefiniti utilizzano policy AWS gestite, che vengono create automaticamente la prima volta che si avvia un cluster EMR dalla console e si scelgono le autorizzazioni predefinite. È anche possibile creare i ruoli IAM predefiniti dalla AWS CLI. Se invece desideri gestire le autorizzazioni AWS, puoi scegliere ruoli personalizzati per il servizio e il profilo dell'istanza. Per ulteriori informazioni, consulta Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS.

Gruppi di sicurezza

Amazon EMR utilizza gruppi di sicurezza per controllare il traffico in entrata e in uscita verso le tue istanze. EC2 Quando avvii il cluster, Amazon EMR utilizza un gruppo di sicurezza per l'istanza principale e un gruppo di sicurezza condiviso dalle 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 istanze per regole più avanzate. Per ulteriori informazioni, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Crittografia

Amazon EMR supporta la crittografia lato server e lato client Amazon S3 facoltativa con EMRFS per proteggere i dati archiviati in Amazon S3. Con la crittografia lato server, Amazon S3 crittografa i dati dopo il caricamento.

Con la crittografia lato client, i processi di crittografia e decrittografia avvengono nel client EMRFS sul tuo cluster EMR. Puoi gestire la chiave principale per la crittografia lato client utilizzando il AWS Key Management Service (AWS KMS) o il tuo sistema di gestione delle chiavi.

Per ulteriori informazioni, consulta Configurazione della crittografia Amazon S3 con le proprietà EMRFS.

Amazon VPC

Amazon EMR supporta l'avvio di cluster in un cloud privato virtuale (Virtual Private Cloud, VPC) in Amazon VPC. Un VPC è una rete virtuale isolata AWS che offre la possibilità di controllare

aspetti avanzati della configurazione e dell'accesso alla rete. Per ulteriori informazioni, consulta Configurazione della rete in un VPC per Amazon EMR.

AWS CloudTrail

Amazon EMR si integra con CloudTrail la registrazione delle informazioni sulle richieste effettuate da o per conto del tuo account. AWS Con queste informazioni, puoi tenere traccia di chi e quando sta accedendo al cluster e dell'indirizzo IP da cui è stata effettuata la richiesta. Per ulteriori informazioni, consulta Registrazione delle chiamate AWS API EMR utilizzando AWS CloudTrail.

Coppie di EC2 chiavi Amazon

È possibile monitorare e interagire con il cluster creando una connessione sicura tra il computer remoto e il nodo primario. Per questa connessione dovrai utilizzare il protocollo di rete Secure Shell (SSH) oppure Kerberos per l'autenticazione. Se usi SSH, è necessaria una coppia di EC2 chiavi Amazon. Per ulteriori informazioni, consulta Usa una coppia di EC2 chiavi per le credenziali SSH per Amazon EMR.

Monitoraggio

Puoi utilizzare le interfacce di gestione e i file di log di Amazon EMR per risolvere problemi del cluster, come esiti negativi o errori. Amazon EMR consente di archiviare i file di log in Amazon S3 in modo da poter archiviare i log e risolvere eventuali problemi anche dopo la terminazione del cluster. Amazon EMR fornisce anche uno strumento opzionale per il debug nella console Amazon EMR per sfogliare i file di log in base a fasi, processi e attività. Per ulteriori informazioni, consulta Configurazione del logging e del debug dei cluster Amazon EMR.

Amazon EMR si integra con CloudWatch per tracciare i parametri delle prestazioni per il cluster e i lavori all'interno del cluster. Puoi configurare gli allarmi in base a diversi parametri, ad esempio se il cluster è inattivo o la percentuale di spazio di archiviazione utilizzata. Per ulteriori informazioni, consulta Monitoraggio dei parametri di Amazon EMR con CloudWatch.

Interfacce di gestione

Esistono vari modi per interagire con Amazon EMR:

- Console: un'interfaccia utente grafica che consente di avviare e gestire i cluster. Attraverso la console si compilano i moduli Web per specificare i dettagli dei cluster da avviare, visualizzare i dettagli dei cluster esistenti, eseguire il debug e terminare i cluster. L'uso della console

è il modo più semplice per iniziare a familiarizzare con Amazon EMR: infatti, non richiede competenze in termini di programmazione. La console è disponibile online a casa. https://console.aws.amazon.com/elasticmapreduce/

- AWS Command Line Interface (AWS CLI) — Un'applicazione client che esegui sul tuo computer locale per connetterti ad Amazon EMR e creare e gestire cluster. AWS CLI Contiene un set di comandi ricco di funzionalità specifici per Amazon EMR. Consente di scrivere script che automatizzano il processo di avvio e gestione dei cluster. Se preferisci lavorare da una riga di comando, usare la AWS CLI è l'opzione migliore. Per ulteriori informazioni, consulta Amazon EMR nella Guida di riferimento ai comandi della AWS CLI .
- Software Development Kit (SDK): SDKs fornisce funzioni che richiamano Amazon EMR per creare e gestire cluster. Permettono di scrivere applicazioni che automatizzano il processo di creazione e gestione dei cluster. Utilizzare gli SDK è l'opzione migliore per ampliare o personalizzare la funzionalità di Amazon EMR. Amazon EMR è attualmente disponibile nei seguenti formatiSDKs: Go, Java, .NET (C# e VB.NET), Node.js, PHP, Python e Ruby. Per ulteriori informazioni su questi argomenti SDKs, consulta Tools for AWS e librerie di esempio per codice e librerie di Amazon EMR.
- Web Service API: un'interfaccia di basso livello che è possibile utilizzare per chiamare il servizio Web direttamente, utilizzando JSON. Utilizzare l'API è l'opzione migliore per creare un SDK personalizzato che invochi Amazon EMR. Per ulteriori informazioni, consulta la Guida di riferimento alle API di Amazon EMR.

Architettura e livelli di servizio di Amazon EMR

L'architettura di servizio di Amazon EMR è composta da diversi livelli, ognuno dei quali fornisce determinate capacità e funzionalità al cluster. Questa sezione fornisce una panoramica dei livelli e i componenti di ciascuno di essi.

In questo argomento

- Storage
- Gestione delle risorse del cluster
- Framework di elaborazione dati
- Applicazioni e programmi

Storage

Il livello di archiviazione include diversi file system utilizzati con il cluster. Esistono diversi tipi di opzioni di archiviazione, come indicato di seguito.

File system distribuito Hadoop (HDFS)

File system distribuito Hadoop (HDFS) è un file di sistema distribuito e scalabile per Hadoop. HDFS distribuisce i dati che archivia tra le istanze del cluster, archiviando più copie dei dati su istanze diverse per garantire che non vadano persi in caso di guasto di una singola istanza. HDFS è uno spazio di archiviazione temporaneo che viene recuperato quando si termina un cluster. HDFS è utile per memorizzare nella cache i risultati intermedi durante l' MapReduce elaborazione o per carichi di lavoro con I/O casuali significativi.

Per ulteriori informazioni, consulta Opzioni e comportamento di storage delle istanze in Amazon EMR oppure la HDFS User Guide (Guida per l'utente di HDFS) sul sito Web di Apache Hadoop.

File system EMR (EMRFS)

Utilizzando il file system EMR (EMRFS), Amazon EMR estende Hadoop per aggiungere la possibilità di accedere direttamente ai dati memorizzati in Amazon S3 come se si trattasse di un file system del tipo di HDFS. È possibile utilizzare HDFS o Amazon S3 come file system nel cluster. Amazon S3 viene utilizzato principalmente per memorizzare i dati di input e di output, mentre i risultati intermedi sono memorizzati in HDFS.

File system locale

Il file system locale fa riferimento a un disco con connessione locale. Quando crei un cluster Hadoop, ogni nodo viene creato da un' EC2 istanza Amazon dotata di un blocco preconfigurato di storage su disco precollegato chiamato instance store. I dati sui volumi dell'instance store persistono solo durante il ciclo di vita dell'istanza Amazon corrispondente. EC2

Gestione delle risorse del cluster

Il livello di gestione delle risorse è responsabile della gestione delle risorse del cluster e della pianificazione dei processi per l'elaborazione dei dati.

Per impostazione predefinita, Amazon EMR usa YARN (Yet Another Resource Negotiator), che è un componente introdotto in Apache Hadoop 2.0 per gestire centralmente le risorse del cluster per

più framework di elaborazione dati. Tuttavia, esistono altri framework e applicazioni offerti in Amazon EMR che non utilizzano YARN come gestore di risorse. Ogni nodo di Amazon EMR può inoltre contare su un agente che gestisce i componenti YARN, preserva l'integrità del cluster e comunica con Amazon EMR.

Poiché le Istanze Spot vengono spesso utilizzate per eseguire nodi attività, Amazon EMR dispone delle funzionalità predefinite per la pianificazione dei processi YARN in modo che i processi in esecuzione non abbiano esito negativo quando i nodi attività in esecuzione sulle Istanze Spot vengono terminati. Amazon EMR esegue questa operazione consentendo ai processi master delle applicazioni di funzionare solo sui nodi principali. Il processo master dell'applicazione controlla i processi in esecuzione e deve rimanere attivo per tutta la durata del processo.

Amazon EMR rilascio 5.19.0 e successivi utilizzano la caratteristica integrata etichette nodo YARN per questo scopo. (Le versioni precedenti utilizzavano una patch di codice). Le proprietà nelle classificazioni di configurazione `yarn-site` e `capacity-scheduler` sono configurate per impostazione predefinita in modo che capacity-scheduler e fair-scheduler YARN sfruttino le etichette dei nodi. Amazon EMR etichetta in automatico i nodi principali con l'etichetta `CORE` e imposta le proprietà in modo che i master dell'applicazione siano pianificati solo sui nodi con l'etichetta CORE. La modifica manuale delle proprietà correlate nelle classificazioni di configurazione del sito di YARN e del pianificatore di capacità o direttamente nei file XML associati potrebbe interrompere o alterare questa funzionalità.

Framework di elaborazione dati

Il livello di framework di elaborazione dati è il motore utilizzato per elaborare e analizzare i dati. Sono disponibili molti framework che funzionano su YARN o che hanno una propria gestione delle risorse. Sono disponibili diversi framework per diversi tipi di esigenze di elaborazione, come batch, interattivi, in-memory, streaming e così via. Il framework scelto varia a seconda del caso d'uso. Questo ha un impatto sulle lingue e sulle interfacce disponibili dal livello dell'applicazione, che è il livello utilizzato per interagire con i dati che si desidera elaborare. I principali framework di elaborazione disponibili per Amazon EMR sono Hadoop e Spark. MapReduce

Hadoop MapReduce

Hadoop MapReduce è un modello di programmazione open source per il calcolo distribuito. Questo modello semplifica il processo di scrittura di applicazioni distribuite in parallelo gestendo tutta la logica, mentre le funzioni Map e Reduce vengono fornite dall'utente. La funzione Map mappa i dati in gruppi di coppie chiave-valore denominati risultati intermedi. La funzione Reduce combina i risultati

intermedi, applica algoritmi aggiuntivi e genera l'output finale. Sono disponibili diversi framework MapReduce, come Hive, che genera automaticamente i programmi Map e Reduce.

Per ulteriori informazioni, consulta How map and reduce operations are actually carried out (Come vengono eseguite le operazioni Map e Reduce) sul sito Web del wiki di Apache Hadoop.

Apache Spark

Spark è un framework di cluster e un modello di programmazione per l'elaborazione di carichi di lavoro di big data. Come Hadoop MapReduce, Spark è un sistema di elaborazione distribuito open source, ma utilizza grafici aciclici diretti per i piani di esecuzione e la memorizzazione nella cache in memoria per i set di dati. Quando si esegue Spark su Amazon EMR, è possibile utilizzare EMRFS per accedere direttamente ai dati in Amazon S3. Spark supporta più moduli di query interattivi, come SparkSQL.

Per ulteriori informazioni, consulta Apache Spark su cluster Amazon EMR nella Guida ai rilasci di Amazon EMR.

Applicazioni e programmi

Amazon EMR supporta numerose applicazioni come Hive, Pig e la libreria Spark Streaming per fornire funzionalità quali l'utilizzo di linguaggi di livello superiore per creare carichi di lavoro di elaborazione, l'utilizzo di algoritmi di machine learning, la creazione di applicazioni di elaborazione del flusso di lavoro e la creazione di data warehouse. Inoltre, Amazon EMR supporta anche i progetti open source che sfruttano le proprie funzionalità di gestione dei cluster anziché utilizzare YARN.

È possibile interagire con le applicazioni eseguite in Amazon EMR attraverso varie librerie e lingue. Ad esempio, puoi usare Java, Hive o Pig con MapReduce o Spark Streaming, Spark SQL MLlib e GraphX con Spark.

Per ulteriori informazioni, consulta la Guida ai rilasci di Amazon EMR.

Prima di configurare Amazon EMR

Completa le attività preliminari descritte in questa sezione prima di avviare un cluster Amazon EMR per la prima volta. Queste includono la configurazione del tuo AWS account, se necessario, e l'adozione di misure per configurare comunicazioni sicure.

Iscriviti per un Account AWS

Se non ne hai uno Account AWS, completa i seguenti passaggi per crearne uno.

Per iscriverti a un Account AWS

1. Apri la https://portal.aws.amazon.com/billing/registrazione.
2. Segui le istruzioni online.

   Nel corso della procedura di registrazione riceverai una telefonata o un messaggio di testo e ti verrà chiesto di inserire un codice di verifica attraverso la tastiera del telefono.

   Quando ti iscrivi a un Account AWS, Utente root dell'account AWSviene creato un. L'utente root dispone dell'accesso a tutte le risorse e tutti i Servizi AWS nell'account. Come best practice di sicurezza, assegna l'accesso amministrativo a un utente e utilizza solo l'utente root per eseguire attività che richiedono l'accesso di un utente root.

AWS ti invia un'email di conferma dopo il completamento della procedura di registrazione. In qualsiasi momento, puoi visualizzare l'attività corrente del tuo account e gestirlo accedendo a https://aws.amazon.com/e scegliendo Il mio account.

Crea un utente con accesso amministrativo

Dopo esserti registrato Account AWS, proteggi Utente root dell'account AWS AWS IAM Identity Center, abilita e crea un utente amministrativo in modo da non utilizzare l'utente root per le attività quotidiane.

Proteggi i tuoi Utente root dell'account AWS

1. Accedi AWS Management Consolecome proprietario dell'account scegliendo Utente root e inserendo il tuo indirizzo Account AWS email. Nella pagina successiva, inserisci la password.

Per informazioni sull'accesso utilizzando un utente root, consulta la pagina Signing in as the root user della Guida per l'utente di Accedi ad AWS .

2. Abilita l'autenticazione a più fattori (MFA) per l'utente root.

   Per istruzioni, consulta Abilitare un dispositivo MFA virtuale per l'utente Account AWS root (console) nella Guida per l'utente IAM.

Crea un utente con accesso amministrativo

1. Abilita Centro identità IAM.

   Per istruzioni, consulta Abilitazione di AWS IAM Identity Center nella Guida per l'utente di AWS IAM Identity Center .

2. In IAM Identity Center, assegna l'accesso amministrativo a un utente.

   Per un tutorial sull'utilizzo di IAM Identity Center directory come fonte di identità, consulta Configurare l'accesso utente con l'impostazione predefinita IAM Identity Center directory nella Guida per l'AWS IAM Identity Center utente.

Accesso come utente amministratore

- Per accedere con l'utente IAM Identity Center, utilizza l'URL di accesso che è stato inviato al tuo indirizzo e-mail quando hai creato l'utente IAM Identity Center.

  Per informazioni sull'accesso utilizzando un utente IAM Identity Center, consulta AWS Accedere al portale di accesso nella Guida per l'Accedi ad AWS utente.

Assegna l'accesso a ulteriori utenti

1. In IAM Identity Center, crea un set di autorizzazioni conforme alla best practice dell'applicazione di autorizzazioni con il privilegio minimo.

   Segui le istruzioni riportate nella pagina Creazione di un set di autorizzazioni nella Guida per l'utente di AWS IAM Identity Center .

2. Assegna al gruppo prima gli utenti e poi l'accesso con autenticazione unica (Single Sign-On).

   Per istruzioni, consulta Aggiungere gruppi nella Guida per l'utente di AWS IAM Identity Center .

Crea una coppia di EC2 chiavi Amazon per SSH

> **Note**
>
> Con Amazon EMR versione 5.10.0 o versioni successive, puoi configurare Kerberos per autenticare utenti e connessioni SSH su un cluster. Per ulteriori informazioni, consulta Utilizzo di Kerberos per l'autenticazione con Amazon EMR.

Per autenticarsi e connettersi ai nodi di un cluster tramite un canale sicuro utilizzando il protocollo Secure Shell (SSH), crea una EC2 coppia di chiavi Amazon Elastic Compute Cloud (Amazon) prima di avviare il cluster. Puoi anche creare un cluster senza una coppia di chiavi. Questa operazione viene di solito effettuata con cluster transitori che si avviano, eseguono fasi e quindi terminano in automatico.

Se...	Allora...
Disponi già di una coppia di EC2 chiavi Amazon che desideri utilizzare oppure non devi autenticarti nel tuo cluster.	Salta questo passaggio.
Occorre creare una coppia di chiavi.	Vedi Creazione della tua coppia di key pair con Amazon EC2.

Passaggi successivi

- Per ulteriori informazioni sulla creazione di un cluster di esempio, consulta Tutorial: Nozioni di base su Amazon EMR.
- Per ulteriori informazioni su come configurare un cluster personalizzato e controllare l'accesso a quest'ultimo, consulta Pianifica, configura e avvia i cluster Amazon EMR e Sicurezza in Amazon EMR.

Tutorial: Nozioni di base su Amazon EMR

Segui un flusso di lavoro per configurare rapidamente un cluster Amazon EMR ed eseguire un'applicazione Spark.

Configurazione del cluster Amazon EMR

Con Amazon EMR è possibile configurare un cluster per elaborare e analizzare i dati con framework di big data in pochi minuti. Questo tutorial mostra come avviare un cluster di esempio utilizzando Spark e come eseguire un semplice PySpark script archiviato in un bucket Amazon S3. Copre le attività essenziali di Amazon EMR in tre categorie principali del flusso di lavoro: pianificazione e configurazione, gestione e pulizia.

Nel corso del tutorial sono disponibili collegamenti ad argomenti più dettagliati e idee per ulteriori fasi nella sezione Passaggi successivi. In caso di domande o problemi, contatta il team Amazon EMR sul nostro Forum di discussione.

Prerequisiti

- Prima di avviare un cluster Amazon EMR, assicurati di completare le attività descritte in Prima di configurare Amazon EMR.

Costo

- Il cluster di esempio che crei viene eseguito in un ambiente reale. Il cluster accumula costi minimi. Per evitare costi aggiuntivi, assicurati di completare le processi di pulizia nell'ultima fase di questo tutorial. I costi maturano alla tariffa al secondo in base ai prezzi di Amazon EMR. I costi variano anche in base alla Regione. Per ulteriori informazioni, consulta Prezzi di Amazon EMR.
- Potrebbero esserci degli costi minimi per i file di piccole dimensioni che archivi su Amazon S3. Alcuni o tutti i costi per Amazon S3 potrebbero non essere addebitati se rientri nei limiti di utilizzo del AWS piano gratuito. Per ulteriori informazioni, consulta Prezzi di Amazon S3 e Piano gratuito di AWS.

Fase 1: configurare le risorse dati e avviare un cluster Amazon EMR

Preparare la archiviazione per Amazon EMR

Quando utilizzi Amazon EMR, puoi scegliere tra una varietà di file system per archiviare dati di input, dati di output e file di log. In questo tutorial, utilizzi EMRFS per archiviare i dati in un bucket S3. EMRFS è un'implementazione del file system Hadoop che permette di leggere e scrivere file standard su Amazon S3. Per ulteriori informazioni, consulta Utilizzo di sistemi di storage e file con Amazon EMR.

Per creare un bucket per questo tutorial, segui le istruzioni in Come creare un bucket S3? nella Guida per l'utente della console Amazon Simple Storage Service. Crea il bucket nella stessa AWS regione in cui prevedi di lanciare il tuo cluster Amazon EMR. Ad esempio, US West (Oregon) us-west-2 (Stati Uniti occidentali (Oregon) us-west-2).

I bucket e le cartelle utilizzati con Amazon EMR presentano le seguenti limitazioni:

- I nomi devono includere lettere minuscole, numeri, punti (.) e trattini (-).
- I nomi non devono terminare con numeri.
- Il nome del bucket deve essere univoco per tutti gli account AWS .
- Una cartella di output deve essere vuota.

Preparazione di un'applicazione con i dati di input per Amazon EMR

Il modo più comune per preparare un'applicazione per Amazon EMR è caricare l'applicazione e i relativi dati di input su Amazon S3. Poi, quando invii il lavoro al cluster, specificare le posizioni Amazon S3 per i vostri script e dati.

In questo passaggio, carichi uno PySpark script di esempio nel tuo bucket Amazon S3. Ti abbiamo fornito uno PySpark script da utilizzare. Lo script elabora i dati di ispezione degli stabilimenti alimentari e restituisce un file dei risultati nel bucket S3. Il file di risultati elenca i primi dieci stabilimenti alimentari con il maggior numero di violazioni di tipo "Red (Rosso)".

Inoltre, carichi dati di input di esempio su Amazon S3 per l'elaborazione dello PySpark script. I dati di input sono una versione modificata dei risultati delle ispezioni condotte dal 2006 al 2020 dal Dipartimento di sanità pubblica della Contea di King, Washington. Per ulteriori informazioni, consulta

Dati pubblici della Contea di King: dati di ispezione sugli stabilimenti alimentari. Non scaricate i dati dei ristoranti di questo tutorial direttamente dal sito web di King County, perché si tratta di un file molto grande. Di seguito forniamo un download con un file con meno record, per facilitare il completamento del tutorial. Di seguito sono riportare righe di esempio del set di dati:

```
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
```

Per preparare lo PySpark script di esempio per EMR

1. Copia il codice di esempio fornito di seguito in un nuovo file nell'editor che hai scelto.

```
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. Salva il file con nome `health_violations.py`.
3. Carica `health_violations.py` su Amazon S3 nel bucket che hai creato per questo tutorial. Per istruzioni, consulta Caricamento di un oggetto in un bucket nella Guida alle operazioni di base di Amazon Simple Storage Service.

Preparazione dei dati di input di esempio per EMR

1. Scarica il file zip food_establishment_data.zip.
2. Decomprimere e salvare `food_establishment_data.zip` come `food_establishment_data.csv` sulla tua macchina.
3. Carica il file CSV nel bucket S3 creato per questo tutorial. Per istruzioni, consulta Caricamento di un oggetto in un bucket nella Guida alle operazioni di base di Amazon Simple Storage Service.

Per ulteriori informazioni sulla configurazione di dati per EMR, consulta Prepara i dati di input per l'elaborazione con Amazon EMR.

Avvio di un cluster Amazon EMR

Dopo aver preparato una posizione di archiviazione e l'applicazione, è possibile avviare un cluster Amazon EMR di esempio. In questa fase, avvii un cluster Apache Spark utilizzando l'ultima versione di rilascio di Amazon EMR.

Console

Per avviare un cluster con Spark installato con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Nella pagina Crea Cluster, prendi nota dei valori predefiniti per Versione, Tipo di istanza, Numero di istanze e Autorizzazioni. Questi campi si compilano in automatico con i valori che funzionano per i cluster per uso generico.
4. Nel campo Nome cluster, inserisci un nome di cluster univoco per aiutarti a identificare il cluster, ad esempio. `My first cluster` Il nome del cluster non può contenere i caratteri <, >, $, | o `(backtick).
5. In Applicazioni, seleziona l'opzione Spark per installare Spark nel cluster.

   > **Note**
   >
   > Scegli le applicazioni che desideri nel cluster Amazon EMR prima di avviarlo. Non è possibile aggiungere o rimuovere applicazioni da un cluster dopo l'avvio.

6. In Registri del cluster, seleziona la casella di spunta Pubblica log specifici del cluster su Amazon S3. Sostituisci il valore di Posizione di Amazon S3 con il bucket Amazon S3 che hai creato seguito da **`/logs`**. Ad esempio, **`s3://amzn-s3-demo-bucket/logs`**. Aggiungendo **`/logs`** si crea una nuova cartella chiamata "logs" nel tuo bucket, nella quale Amazon EMR copierà i file di log del cluster.
7. In Configurazione e autorizzazioni di sicurezza, scegli la tua EC2 key pair. Nella stessa sezione, seleziona il menu a discesa Service role for Amazon EMR e scegli EMR_. DefaultRole Quindi, seleziona il menu a discesa del profilo IAM role for instance e scegli EMR_ _. EC2 DefaultRole
8. Scegli Crea cluster per avviare il cluster e apri la pagina dei dettagli del cluster.

9. Cerca l'indicazione Stato accanto al nome del cluster. Lo stato cambia da Avvio in corso a In esecuzione a In attesa mentre Amazon EMR effettua il provisioning del cluster. Potrebbe essere necessario scegliere l'icona di aggiornamento a destra o aggiornare il browser per visualizzare gli aggiornamenti di stato.

Quando lo stato del cluster passa a In attesa, il cluster è attivo, in esecuzione e pronto per accettare lavoro. Per ulteriori informazioni sulla lettura del riepilogo di un cluster, consulta Visualizza lo stato e i dettagli del cluster Amazon EMR. Per ulteriori informazioni sullo stato del cluster, consulta Comprensione del ciclo di vita del cluster.

CLI

Per avviare un cluster con Spark installato con il AWS CLI

1. Crea ruoli predefiniti IAM che puoi utilizzare per creare il cluster utilizzando il comando seguente.

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

Per ulteriori informazioni su come `create-default-roles`, consulta la Guida di riferimento ai comandi della AWS CLI.

2. Crea un cluster Spark con il comando seguente. Inserisci un nome per il tuo cluster con l'`--name`opzione e specifica il nome della tua EC2 key pair con l'`--ec2-attributes`opzione.

```
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
```

Prendi nota degli altri valori richiesti per `--instance-type`, `--instance-count` e `--use-default-roles`. Questi valori sono stati scelti per i cluster generici. Per ulteriori informazioni su come `create-cluster`, consulta la Guida di riferimento ai comandi della AWS CLI.

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

L'output restituito dovrebbe essere simile al seguente. L'output mostra il `ClusterId` e `ClusterArn` del nuovo cluster. Nota il tuo `ClusterId`. Utilizza il `ClusterId` per verificare lo stato del cluster e inviare il lavoro.

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

3. Verifica lo stato del cluster con il comando seguente.

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

Verrà visualizzato un output come il seguente con l'oggetto `Status` per il nuovo cluster.

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

La valore `State` cambia da `STARTING` a `RUNNING` a `WAITING` mentre Amazon EMR effettua il provisioning del cluster.

Lo stato del cluster passa a **`WAITING`** quando il cluster è attivo, in esecuzione e pronto per accettare lavoro. Per ulteriori informazioni sullo stato del cluster, consulta Comprensione del ciclo di vita del cluster.

Fase 2: invia il lavoro al tuo cluster Amazon EMR

Invia il lavoro e visualizza i risultati

Dopo aver avviato un cluster, è possibile inviare il lavoro al cluster in esecuzione per elaborare e analizzare i dati. Invia lavoro a un cluster Amazon EMR come una fase. Una fase è un'unità di lavoro costituita da uno o più operazioni. Ad esempio, potresti inviare una fase per calcolare valori o per trasferire ed elaborare dati. Puoi inviare fasi quando crei un cluster o a un cluster in esecuzione. In questa parte del tutorial, invii `health_violations.py` come fase al cluster in esecuzione. Per ulteriori informazioni sulle fasi, consulta Invia il lavoro a un cluster Amazon EMR.

Console

Per inviare un'applicazione Spark, procedi passo con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi seleziona il cluster a cui desideri inviare il lavoro. Lo stato del cluster deve essere In attesa.
3. Scegli la scheda Steps (Fasi), quindi scegli Add step (Aggiungi fase).
4. Configura la fase in base alle seguenti linee guida:
   - Per Type (Tipo), scegli Spark application (Applicazione Spark). Dovresti visualizzare altri campi per Deploy mode (Modalità di implementazione), Application location (Percorso dell'applicazione) e Spark-submit options (Opzioni Spark-submit).
   - In Name (Nome), inserisci un nuovo nome. Se in un cluster sono presenti più fasi, l'assegnazione di un nome a ciascuna fase aiuta a tenerne traccia.
   - Per Deploy mode (Modalità di implementazione), lascia il valore predefinito Cluster mode (Modalità cluster). Per ulteriori informazioni sulle modalità di implementazione di Spark, consulta la sezione Cluster mode overview (Panoramica della modalità cluster) nella documentazione di Apache Spark.

- Per Posizione dell'applicazione, inserisci la posizione dello `health_violations.py` script in Amazon S3, ad esempio. *`s3://amzn-s3-demo-bucket/health_violations.py`*
- Lascia vuoto il campo Spark-submit options (Opzioni Spark-submit). Per ulteriori informazioni sulle opzioni di `spark-submit`, consulta la sezione Launching applications with spark-submit (Avvio di applicazioni con spark-submit).
- Nel campo Argomenti, inserisci i seguenti argomenti e valori:

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

  Sostituisci *`s3://amzn-s3-demo-bucket/food_establishment_data.csv`* con l'URI del bucket S3 i dati di input in cui hai preparato. Preparazione di un'applicazione con i dati di input per Amazon EMR

  Sostituiscilo *`amzn-s3-demo-bucket`* con il nome del bucket che hai creato per questo tutorial e sostituiscilo *`myOutputFolder`* con un nome per la cartella di output del cluster.
- Per Action on failure (Operazione in caso di errore), accetta l'opzione predefinita, ossia Continue (Continua). In questo modo, se la fase riscontra un errore, il cluster continua a funzionare.

5. Seleziona Aggiungi per inviare la fase. La fase viene visualizzata nella console con lo stato In attesa.
6. Monitora lo stato della fase. Dovrebbe passare da Pending (In attesa) a Running (In esecuzione) a Completed (Completata). Per aggiornare lo stato nella console, scegli l'icona di aggiornamento a destra di Filter (Filtro). Lo script richiede circa un minuto per l'esecuzione. Quando lo stato diventa Completed (Completata), significa che la fase è stata completata correttamente.

CLI

Per inviare un'applicazione Spark, procedi passo con il AWS CLI

1. Accertati di disporre del `ClusterId` del cluster che hai avviato in Avvio di un cluster Amazon EMR. In alternativa, puoi recuperare l'ID del cluster con il comando seguente.

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

2. Invia `health_violations.py` come fase con il comando `add-steps` e il `ClusterId`.

   - Puoi specificare un nome per la tua fase *"My Spark Application"* sostituendo. Nell'`Args`array, *s3://amzn-s3-demo-bucket/health_violations.py* sostituiscilo con la posizione dell'`health_violations.py`applicazione.
   - Sostituisci *s3://amzn-s3-demo-bucket/food_establishment_data.csv* con la posizione S3 del tuo `food_establishment_data.csv` set di dati.
   - Sostituiscilo *s3://amzn-s3-demo-bucket/MyOutputFolder* con il percorso S3 del bucket designato e un nome per la cartella di output del cluster.
   - `ActionOnFailure=CONTINUE` indica che il cluster continua a funzionare qualora la fase avesse esito negativo.

```
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>]
```

   Per ulteriori informazioni sull'invio di fasi mediante la CLI, consulta la Guida di riferimento ai comandi della AWS CLI.

   Dopo aver inviato la fase, dovresti visualizzare l'output simile al seguente con un elenco di `StepIds`. Dal momento che hai inviato una sola fase, vedrai un solo ID nell'elenco. Copia il tuo ID della fase. Utilizzare l'ID della fase per verificare lo stato della fase.

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

3. Esegui una query sullo stato della fase con il comando `describe-step`.

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

L'output restituito dovrebbe essere simile al seguente con le informazioni sul fase.

```
{
    "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"
        }
    }
}
```

Lo `State` della fase cambia da `PENDING` a `RUNNING` a `COMPLETED` durante la sua esecuzione. La fase richiede circa un minuto per essere eseguita, quindi potrebbe essere necessario controllarne lo stato più volte.

Saprai che la fase è terminata correttamente quando il relativo `State` passa a **`COMPLETED`**.

Per ulteriori informazioni sul ciclo di vita delle fasi, consulta Esecuzione di fasi per elaborare i dati.

Visualizzazione dei risultati

Se l'esecuzione della fase è stata riuscita, puoi visualizzarne i risultati nella cartella output di Amazon S3.

Visualizzazione dei risultati di **`health_violations.py`**

1. Apri la console Amazon S3 all'indirizzo. https://console.aws.amazon.com/s3/
2. Scegli il Bucket name (Nome bucket) e la cartella di output che hai specificato all'invio della fase. Ad esempio, *`amzn-s3-demo-bucket`* e poi*`myOutputFolder`*.
3. Verifica che nella cartella di output siano presenti i seguenti elementi:
   - Un oggetto di piccole dimensioni denominato `_SUCCESS`.
   - Un file CSV che inizia con il prefisso `part-` che contiene i risultati.
4. Scegli l'oggetto con i risultati, quindi scegli Download per salvare i risultati nel file system locale.
5. Apri i risultati nell'editor che preferisci. Il file di output elenca i primi dieci stabilimenti alimentari con il maggior numero di violazioni di tipo "red (rosso)". Il file di output mostra anche il numero totale di violazioni di tipo "red (rosso)" per ogni stabilimento.

   Di seguito è riportato un esempio di risultati `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
```

Per ulteriori informazioni sull'output dei cluster Amazon EMR, consulta Configurare una posizione per l'output del cluster Amazon EMR.

(Facoltativo) Eseguire la connessione al cluster Amazon EMR in esecuzione

Quando utilizzi Amazon EMR, potresti voler connetterti a un cluster in esecuzione per leggere i file di log, eseguire il debug del cluster o utilizzare strumenti CLI come la shell Spark. Amazon EMR consente di connetterti a un cluster utilizzando il protocollo Secure Shell (SSH). Questa sezione illustra come configurare SSH, connettersi al cluster e visualizzare i file di log per Spark. Per ulteriori informazioni sulla connessione a un cluster, consulta Autenticazione nei nodi cluster Amazon EMR.

Autorizza le connessioni SSH al cluster

Prima di connetterti al cluster, devi modificare i gruppi di sicurezza del cluster per autorizzare le connessioni SSH in entrata. I gruppi EC2 di sicurezza di Amazon agiscono come firewall virtuali per controllare il traffico in entrata e in uscita verso il cluster. Quando hai creato il cluster per questo tutorial, Amazon EMR ha creato i seguenti gruppi di sicurezza per tuo conto:

ElasticMapReduce-padrone

Il gruppo di sicurezza gestito da Amazon EMR predefinito associato al nodo primario. In un cluster Amazon EMR, il nodo principale è un' EC2istanza Amazon che gestisce il cluster.

ElasticMapReduce-schiavo

Il gruppo di sicurezza predefinito associato ai nodi principali e attività.

Console

Per consentire l'accesso SSH a fonti attendibili per il gruppo di sicurezza principale con la console

Per modificare i gruppi di sicurezza, devi disporre dell'autorizzazione per gestire i gruppi di sicurezza per il VPC in cui si trova il cluster. Per ulteriori informazioni, consulta Modifica delle autorizzazioni per un utente e la politica di esempio che consente la gestione dei gruppi di EC2 sicurezza nella Guida per l'utente IAM.

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli il cluster che desideri aggiornare. Si apre la pagina dei dettagli del cluster. In questa pagina dovrebbe essere preselezionata la scheda Properties (Proprietà) .
3. In Rete nella scheda Proprietà, seleziona la freccia accanto ai gruppi EC2 di sicurezza (firewall) per espandere questa sezione. In Primary node (Nodo primario), seleziona il collegamento al gruppo di sicurezza. Dopo avere completato i seguenti passaggi, puoi facoltativamente tornare a questo passaggio, scegliere Core and task nodes (Nodi core e attività) e ripetere i passaggi seguenti per consentire al client SSH di accedere ai nodi core e attività.
4. Verrà aperta la EC2 console. Scegli la scheda Regole in entrata e quindi Modifica le regole in entrata.

5. Verifica la presenza di una regola in entrata che consenta l'accesso pubblico con le seguenti impostazioni. Se esiste, scegli Elimina per rimuoverla.

   - Tipo

     SSH
   - Porta

     22
   - Origine

     Personalizzata 0.0.0.0/0

   > **Warning**
   >
   > Prima di dicembre 2020, il gruppo di sicurezza ElasticMapReduce -master disponeva di una regola preconfigurata per consentire il traffico in entrata sulla porta 22 da tutte le fonti. Questa regola è stata creata per semplificare le connessioni SSH iniziali al nodo principale. Ti consigliamo di rimuovere questa regola in entrata e limitare il traffico a origini affidabili.

6. Scorri fino alla fine dell'elenco di regole e seleziona Aggiungi regola.
7. Per Tipo, seleziona SSH. Selezionando SSH si inserisce in automatico TCP per Protocollo e 22 per Intervallo porta.
8. Per origine, seleziona Il mio IP per aggiungere in automatico il tuo indirizzo IP come indirizzo di origine. Puoi anche aggiungere un intervallo di indirizzi IP affidabili del client Custom (Personalizzato), o creare regole aggiuntive per altri client. In molti ambienti di rete, gli indirizzi IP vengono allocati in modo dinamico, perciò, nel futuro, potrebbe essere necessario aggiornare gli indirizzi IP per client affidabili.
9. Scegli Save (Salva).
10. Facoltativamente, scegli Nodi principali e di attività dall'elenco e ripeti le fasi descritte in precedenza per consentire l'accesso SSH dei client ai nodi principali e attività.

Connect al cluster utilizzando il AWS CLI

A prescindere dal sistema operativo, è possibile creare una connessione SSH al cluster utilizzando la AWS CLI.

Per connettersi al cluster e visualizzare i file di registro utilizzando il AWS CLI

1. Utilizza il comando seguente per aprire una connessione SSH al cluster. Sostituisci *<mykeypair.key>* con il percorso completo e il nome del file della coppia di chiavi. Ad esempio, `C:\Users\<username>\.ssh\mykeypair.pem`.

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

2. Navigare a `/mnt/var/log/spark` per accedere ai registri Spark sul nodo principale del cluster. Quindi visualizza i file in quella posizione. Per un elenco di file di log aggiuntivi sul nodo principale, consulta Visualizzazione di file di log sul nodo primario.

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

Usa Amazon SageMaker AI Unified Studio per gestire il tuo cluster Amazon EMR

Amazon EMR on EC2 è anche un tipo di elaborazione supportato per Amazon SageMaker AI Unified Studio. Per informazioni su come utilizzare e gestire EMR sulle EC2 EC2 risorse in Unified Studio, consulta Managing Amazon EMR. Amazon SageMaker AI

Fase 3: pulizia delle risorse Amazon EMR

Terminazione di un cluster

Ora che hai inviato il lavoro al cluster e visualizzato i risultati della tua PySpark richiesta, puoi terminare il cluster. L'interruzione di un cluster interrompe tutti i costi Amazon EMR e le istanze Amazon associati al cluster. EC2

Amazon EMR mantiene gratuitamente i metadati relativi al cluster per due mesi dopo la terminazione del cluster. I metadati archiviati ti aiutano clonare il cluster per un nuovo lavoro o rivedere la configurazione del cluster a scopo di riferimento. I metadati non includono i dati che il cluster scrive in S3 o che sono archiviati in HDFS sul cluster.

> **Note**
>
> La console di Amazon EMR non consente di eliminare un cluster dalla visualizzazione elenco in seguito alla terminazione del cluster. Un cluster terminato scompare dalla console quando Amazon EMR ne cancella i metadati.

Console

Per terminare il cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. Seleziona Cluster, quindi scegli il cluster da terminare.
3. Nel menu a discesa Operazioni, scegli Termina cluster.
4. Nella finestra di dialogo, scegli Termina. A seconda della configurazione del cluster, la terminazione potrebbe richiedere da 5 a 10 minuti. Per ulteriori informazioni sui cluster Amazon EMR, consulta la sezione Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

CLI

Per terminare il cluster con AWS CLI

1. Avvia il processo di terminazione del cluster con il comando seguente. Sostituiscilo *<myClusterId>* con l'ID del cluster di esempio. Il comando non restituisce output.

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

2. Per verificare che il processo di terminazione del cluster è in corso di esecuzione, controlla lo stato del cluster con il comando seguente.

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

Di seguito è riportato un esempio di output in formato JSON. Lo `Status` del cluster dovrebbe passare da **`TERMINATING`** a **`TERMINATED`**. La terminazione potrebbe richiedere da 5 a 10 minuti a seconda della configurazione del cluster. Per ulteriori informazioni sulla terminazione

di un cluster Amazon EMR, consulta Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

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

Eliminazione di risorse S3

Per evitare costi aggiuntivi, devi eliminare il bucket Amazon S3. L'eliminazione del bucket rimuove tutte le risorse di Amazon S3 per questo tutorial. Il bucket deve contenere:

- La PySpark sceneggiatura
- Il set di dati di input
- La cartella dei risultati di output
- La cartella dei file di log

Potrebbe essere necessario eseguire ulteriori passaggi per eliminare i file archiviati se PySpark lo script o l'output sono stati salvati in una posizione diversa.

> **Note**
>
> Il cluster deve essere terminato prima di eliminare il bucket. In caso contrario, potrebbe non essere consentito svuotare il bucket.

Per eliminare il bucket, segui le istruzioni in Come eliminare un bucket S3? nella Guida per l'utente di Amazon Simple Storage Service.

Passaggi successivi

Ora hai avviato il tuo primo cluster Amazon EMR dall'inizio alla fine. Hai inoltre completato processi EMR essenziali come la preparazione e l'invio di applicazioni di Big Data, la visualizzazione dei risultati e la terminazione di un cluster.

Utilizza gli argomenti seguenti per ulteriori informazioni su come personalizzare il flusso di lavoro Amazon EMR.

Valutazione delle applicazioni di big data per Amazon EMR

Scopri e confronta le applicazioni di big data che si possono installare in un cluster nella Guida ai rilasci di Amazon EMR. La Guida ai rilasci descrive in dettaglio ogni versione di EMR e include suggerimenti per l'utilizzo di framework come Spark e Hadoop su Amazon EMR.

Pianificazione dell'hardware, della rete e della sicurezza del cluster

In questo tutorial hai creato un cluster EMR semplice senza configurare opzioni avanzate. Le opzioni avanzate consentono di specificare i tipi di EC2 istanze Amazon, la rete di cluster e la sicurezza dei cluster. Per ulteriori informazioni sulla pianificazione e sull'avvio di un cluster che soddisfi i tuoi requisiti, consulta Pianifica, configura e avvia i cluster Amazon EMR e Sicurezza in Amazon EMR.

Gestione di cluster

Immergiti nel lavoro con i cluster in esecuzione Gestione dei cluster Amazon EMR. Per gestire un cluster, è possibile connettersi al cluster, eseguire il debug delle fasi e tenere traccia delle processi e l'integrità del cluster. Inoltre, puoi regolare le risorse del cluster in risposta alle richieste dei carichi di lavoro con Dimensionamento gestito da EMR.

Uso di un'interfaccia diversa

Oltre alla console Amazon EMR, puoi gestire Amazon EMR utilizzando l'API del AWS Command Line Interface servizio Web o una delle tante supportate. AWS SDKs Per ulteriori informazioni, consulta Interfacce di gestione.

Inoltre, puoi interagire in molti modi con le applicazioni installate nei cluster Amazon EMR. Alcune applicazioni come Apache Hadoop pubblicano interfacce Web visualizzabili. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Consultazione del blog tecnico su EMR

Per esempi dettagliati e discussioni tecniche approfondite sulle nuove caratteristiche di Amazon EMR, consulta il Blog sui Big Data AWS.

Gestione dei cluster Amazon EMR con la console

La console offre un'interfaccia aggiornata che offre un modo intuitivo per gestire l'ambiente Amazon EMR e offre un comodo accesso alla documentazione, alle informazioni sui prodotti e ad altre risorse.

Funzionalità della console

La console Amazon EMR è disponibile al seguente URL:

- URL della console: https://console.aws.amazon.com /emr

La tabella seguente elenca lo stato dei principali componenti della console Amazon EMR.

Componente della console Amazon EMR	Console	
EMR Studio	✓	
Creazione e gestione dei cluster	✓	
Blocco dell'accesso pubblico	✓	
Monitora CloudWatch gli eventi Amazon	✓	
Configurazioni di sicurezza	✓	
Cluster virtuali (Amazon EMR su EKS)	✓	
Visualizza e gestisci le sottoreti Amazon Virtual Private Cloud [1]	✓	
Notebook [2]	✓	

[1] Nella console, puoi visualizzare e gestire le sottoreti Amazon VPC nella sezione Rete quando crei un cluster.

2 Notebooks EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Riepilogo delle differenze

Questa sezione descrive le funzionalità dell'esperienza della console Amazon EMR. Queste funzionalità rientrano nelle seguenti categorie:

- Compatibilità dei cluster nella console
- Creazione di cluster
- Visualizzazione o modifica dei dettagli del cluster
- Visualizzazione e ricerca di cluster
- Differenze nell'utilizzo delle configurazioni di sicurezza

Compatibilità dei cluster nella console

In alcuni casi, un cluster creato potrebbe non essere compatibile con la console. L'elenco seguente descrive i requisiti di compatibilità per la console Amazon EMR.

- La console supporta i cluster creati nelle versioni 5.20.1 e successive di Amazon EMR.
- Puoi clonare i cluster che utilizzano il ridimensionamento automatico nella console, ma puoi creare nuovi cluster solo se desideri ridimensionarli manualmente o utilizzare la scalabilità gestita.

Per creare e lavorare con i cluster della versione 5.20.1 e precedenti, puoi utilizzare () o l'SDK. AWS Command Line Interface AWS CLI AWS

Creazione di cluster

Funzionalità	Console	
Terminologia: tipi di nodi del cluster Amazon EMR	Primario (primary), core (core), attività (task)	

Funzionalità	Console	
Versioni supportate da Amazon EMR [1]	Amazon EMR rilascio 5.20.1 e successivi	
Avvio rapido di un cluster	Utilizzate il pulsante Crea cluster nel pannello Riepilogo . Il nome del cluster non può contenere i caratteri <, >, $, | o `(backtick).	
Configurazione di un timeout di provisioning Spot	Definire un periodo di timeout per il provisioning delle istanze per ogni parco del cluster.	
Ruoli di servizio e ruolo del profilo di EC2 istanza Amazon	La console non crea ruoli predefiniti; devi creare ruoli con la console IAM o selezionare un ruolo IAM già creato	
Visibilità del cluster	Dall'interno della console Amazon EMR non è possibile rendere visibile un cluster a tutti gli utenti; è la policy IAM a determina re l'accesso al cluster	
Reti: configurazione di sottoreti private	Gli endpoint Amazon S3 e i gateway NAT devono essere configura ti dalle rispettive console Amazon S3 e Amazon VPC	

Funzionalità	Console	
Visualizzazione coerente del file system EMR (EMRFS CV)	Con il rilascio di Amazon S3 Strong read-after-write Consistency il 1° dicembre 2020, non è necessario utilizzare EMRFS CV con i cluster EMR	
Debug	È possibile eseguire il debug dei processi utilizzando l'interfaccia utente dell'applicazione nella pagina dei dettagli del cluster	

[1] Non è possibile creare o modificare cluster utilizzando versioni precedenti ad Amazon EMR 5.20.1 nella console, ma tutti i cluster esistenti creati utilizzando versioni precedenti alla 5.20.1 continueranno a funzionare. Per creare e modificare cluster con versioni di Amazon EMR precedenti alla 5.20.1, utilizza l'API o la CLI. Puoi visualizzare tutti i cluster utilizzando la console, ma le console create prima della 5.20.1 potrebbero non essere compatibili con le nuove funzionalità.

Visualizzazione e ricerca di cluster

La tabella seguente illustra come utilizzare la console Amazon EMR per visualizzare, visualizzare e cercare cluster.

> **Note**
>
> L'applicazione di un filtro dati all'elenco dei cluster interroga l'intero database. Tuttavia, quando si inserisce una stringa di testo nella casella di ricerca, la ricerca si applica solo ai risultati che l'elenco ha caricato su lato client.

Funzionalità	Console	

Funzionalità	Console	
Visualizzazione dei dettagli del cluster	È possibile selezionare l'ID del cluster per visualizzare dettagli completi del cluster come opzioni di configurazione UIs, applicazione persistente e registri.	
Ricerca di cluster	Utilizza un unico campo di ricerca per inserire query di ricerca di testo e per creare e applicare filtri di dati come "Status = Any active status" (Stato = Qualsiasi stato attivo).	
Individuazione di cluster falliti	Per cercare i cluster con errori, applica il filtro Status (Stato) = Terminated with errors (Terminato con errori).	

Visualizzazione o modifica dei dettagli del cluster

Funzionalità	Console	
Visualizzazione delle istanze presenti nei gruppi di istanze e nei parchi istanze, oltre alle opzioni di dimensionamento, provisioning, ridimensionamento e terminazione	Visualizza le opzioni e i dettagli delle istanze nella scheda Istanze. Visualizza le opzioni di terminazione nella scheda Proprietà.	

Funzionalità	Console	
Visualizzazione di app UIs, registri e configurazioni (Interfaccia utente Apache Spark, servizio di cronologia Spark, interfaccia utente Apache Tez, server della cronologia YARN)	Visualizza le configurazioni del cluster nella scheda Configurazioni. Avvia un'interfaccia utente di applicazione attiva e persistente per visualizzare i log di un'applicazione dalla scheda Applicazioni.	
Esportazione di un cluster nella CLI	Opzione disponibile nei menu dei dettagli del cluster e di visualizzazione dell'elenco delle operazioni come "View command for cloning cluster" (Visualizza comando per la clonazione del cluster)	

Differenze nell'utilizzo delle configurazioni di sicurezza

Funzionalità	Console	
Configurazioni di sicurezza della clonazione	✓	
Governance federata con Trino e Apache Ranger	✓	
Utilizzo di un ruolo di runtime per inviare il lavoro a un cluster[1]	✓	
	Punti di accesso Amazon S3	

Funzionalità	Console	
Autorizzazione per l'accesso ai dati del file system EMR (EMRFS)		
AWS Lake Formation controlli di accesso	Ruoli di runtime	

[1] Per trasmettere un ruolo durante l'invio delle fasi, il cluster deve utilizzare una configurazione di sicurezza collegata a una policy di autorizzazione IAM in modo che l'utente IAM possa trasmettere solo i ruoli approvati e i processi possano accedere alle risorse Amazon EMR. Per ulteriori informazioni, consulta Ruoli di runtime per le fasi di Amazon EMR.

Amazon EMR Studio

Amazon EMR Studio è un ambiente di sviluppo integrato (IDE) basato sul Web per notebook Jupyter completamente gestiti che vengono eseguiti su cluster Amazon EMR. Puoi configurare un EMR Studio per il tuo team per sviluppare, visualizzare ed eseguire il debug di applicazioni scritte in R, Python, Scala e. PySpark EMR Studio è integrato con AWS Identity and Access Management (IAM) e IAM Identity Center in modo che gli utenti possano accedere utilizzando le proprie credenziali aziendali.

È possibile creare un EMR Studio in modo totalmente gratuito. Quando si utilizza EMR Studio, si applicano i costi applicabili per l'archiviazione Amazon S3 e per i cluster Amazon EMR. Per i dettagli e le caratteristiche principali del prodotto, consulta la pagina del servizio per Amazon EMR Studio.

Funzionalità principali di EMR Studio

Amazon EMR Studio offre le seguenti funzionalità:

- Autentica gli utenti con AWS Identity and Access Management (IAM), AWS IAM Identity Center con o senza una propagazione affidabile dell'identità e il tuo provider di identità aziendale.
- Accedi e avvia i cluster Amazon EMR on-demand per eseguire i processi di notebook Jupyter.
- Connetti ai cluster Amazon EMR su EKS per inviare il lavoro durante l'esecuzione del processo.
- Esplora e salva notebook di esempio. Per ulteriori informazioni sui taccuini di esempio, vedere l'archivio di esempi di notebook EMR Studio. GitHub
- Analizza i dati usando Python, Spark Scala PySpark, Spark R o SparkSQL e installa kernel e librerie personalizzati.
- Collabora in tempo reale con altri utenti nello stesso Workspace. Per ulteriori informazioni, consulta Configurazione della collaborazione Workspace in EMR Studio.
- Utilizza SQL Explorer di EMR Studio per sfogliare il catalogo dati, eseguire query SQL e scaricare i risultati prima di lavorare con i dati in un notebook.
- Esegui notebook parametrizzati nell'ambito dei flussi di lavoro pianificati con uno strumento di orchestrazione come Apache Airflow o Amazon Managed Workflows per Apache Airflow. Per ulteriori informazioni, consulta Orchestrare i lavori di analisi sui notebook EMR utilizzando MWAA nel blog Big Data. AWS
- GitHub Collega BitBucket repository di codici come e.

- Monitora ed esegui il debug dei processi utilizzando Spark History Server, l'interfaccia utente Tez o il server della cronologia YARN.

EMR Studio è idoneo all'HIPAA ed è certificato secondo HITRUST CSF e SOC 2. Per ulteriori informazioni sulla conformità HIPAA per i servizi, vedere. AWS https://aws.amazon.com/compliance/hipaa-compliance/ Per ulteriori informazioni sulla conformità HITRUST CSF per AWS i servizi, vedere. https://aws.amazon.com/compliance/hitrust/

Anche FedRamp EMR Studio è conforme. Per ulteriori informazioni sui programmi di conformità a cui è conforme Amazon EMR, consulta Convalida della conformità per Amazon EMR. Per ulteriori informazioni sui programmi di conformità aggiuntivi per i AWS servizi, consulta AWS Services in Scope by Compliance Program.

Ambiente di sviluppo integrato Amazon SageMaker Unified Studio

Amazon SageMaker Unified Studio fornisce un ambiente di sviluppo integrato (IDE) per i notebook Jupyter che viene eseguito su Amazon EMR su cluster o utilizzando connessioni di elaborazione EMR Serverless EC2 . Combinando la potenza di Amazon EMR con le funzionalità del end-to-end flusso di lavoro di Amazon SageMaker Unified Studio, i team possono semplificare la preparazione dei dati, lo sviluppo di pipeline e la sperimentazione del machine learning in un unico ambiente. Amazon EMR SageMaker rivoluziona l'elaborazione dei big data supportando framework open source come Apache Spark, Trino e Apache Flink. Elimina le complessità di gestione dell'infrastruttura scalando al contempo i carichi di lavoro di analisi senza problemi. Per ulteriori informazioni, consulta Amazon EMR.

Cronologia delle funzionalità di Amazon EMR Studio

Questa tabella elenca gli aggiornamenti della funzionalità di dimensionamento gestito da Amazon EMR.

Data di rilascio	Funzionalità
5 gennaio 2024	È stato aggiunto il supporto per EMR Studio in AWS GovCloud (Stati Uniti orientali) e AWS GovCloud (Stati Uniti occidentali).
26 novembre 2023	È stato aggiunto il supporto per la propagazione delle identità attendibili per EMR Studio con autenticazione IAM Identity Center.

Data di rilascio	Funzionalità
26 ottobre 2023	Aggiunta la possibilità di creare un'applicazione EMR Serverless con funzionalità interattive.
28 febbraio 2023	È stato aggiunto il supporto chiave AWS KMS gestito dal cliente per l'archiviazione dei registri delle applicazioni per le applicazioni EMR Serverless.
23 febbraio 2023	Aggiunta creazione di ruoli IAM con un solo clic per l'invio di processi EMR serverless. Aggiunta ricerca ECR per quando si seleziona un'immagine personalizzata per le applicazioni EMR serverless.
27 gennaio 2023	I notebook di esecuzione headless possono tracciare l'avanzam ento dell'esecuzione di ogni cella con `%execute_notebook` magic.
23 gennaio 2023	Le applicazioni persistenti sono state ottimizzate per consentire tempi di avvio più rapidi.

Come funziona Amazon EMR Studio

Amazon EMR Studio è una risorsa Amazon EMR creata per un team di utenti. Ogni Studio è un ambiente di sviluppo integrato autonomo basato sul Web per notebook Jupyter che vengono eseguiti su cluster Amazon EMR. Gli utenti registrano ad Studio utilizzando le proprie credenziali aziendali.

Ogni EMR Studio creato utilizza le seguenti risorse: AWS

- Amazon Virtual Private Cloud (VPC) con sottoreti: gli utenti eseguono kernel e applicazioni Studio su Amazon EMR e Amazon EMR su cluster EKS nel VPC specificato. EMR Studio può connettersi a qualsiasi cluster nelle sottoreti specificate al momento della creazione dello Studio.
- Ruoli IAM e policy di autorizzazione: per gestire le autorizzazioni utente, è possibile creare policy di autorizzazione IAM collegate all'identità IAM di un utente o a un ruolo utente. EMR Studio utilizza anche un ruolo di servizio IAM e gruppi di sicurezza per interagire con altri servizi. AWS Per ulteriori informazioni, consultare Controllo accessi e Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio.

- Gruppi di sicurezza: EMR Studio utilizza gruppi di sicurezza per stabilire un canale di rete sicuro tra Studio e un cluster EMR.
- Una posizione di backup di Amazon S3: EMR Studio salva il lavoro del notebook in una posizione Amazon S3.

Nelle fasi seguenti viene descritto come creare e amministrare un EMR Studio:

1. Crea uno Studio nel tuo Account AWS con l'autenticazione IAM o IAM Identity Center. Per istruzioni, consultare Configurazione di un EMR Studio.
2. Assegnazione di utenti e gruppi a Studio Utilizza le policy di autorizzazione per impostare autorizzazioni granulari per ogni utente. Per ulteriori informazioni, consulta l'argomento Assegnazione e gestione degli utenti di EMR Studio.
3. Inizia a monitorare le azioni di EMR Studio con AWS CloudTrail gli eventi. Per ulteriori informazioni, consulta Monitoraggio delle operazioni di Amazon EMR Studio.
4. Fornisci più opzioni di cluster agli utenti di Studio con modelli cluster e Amazon EMR su endpoint gestiti EKS.

Autenticazione e accesso utente

Amazon EMR Studio supporta due modalità di autenticazione: la modalità di autenticazione IAM e la modalità di autenticazione IAM Identity Center. La modalità IAM utilizza AWS Identity and Access Management (IAM), mentre la modalità IAM Identity Center utilizza AWS IAM Identity Center. Quando crei un EMR Studio, scegli la modalità di autenticazione per tutti gli utenti di tale Studio.

Modalità di autenticazione IAM

Con la modalità di autenticazione IAM, è possibile utilizzare l'autenticazione IAM o la federazione IAM.

L'autenticazione IAM consente di gestire le identità IAM come utenti, gruppi e ruoli in IAM. Concedi agli utenti l'accesso a uno Studio con policy di autorizzazione IAM e controllo dell'accesso basato su attributi (ABAC).

La federazione IAM ti consente di stabilire un rapporto di fiducia tra un provider di identità (IdP) di terze parti e AWS di gestire le identità degli utenti tramite il tuo IdP.

Modalità di autenticazione IAM Identity Center

La modalità di autenticazione IAM Identity Center consente agli utenti federati di accedere a un EMR Studio. È possibile utilizzare IAM Identity Center per l'autenticazione di utenti e gruppi della directory IAM Identity Center, della directory aziendale esistente o di un gestore dell'identità digitale (IdP) esterno come Azure Active Directory (AD). Quindi gestisci gli utenti con il provider di identità (IdP).

EMR Studio supporta l'uso dei seguenti gestori dell'identità digitale per IAM Identity Center:

- AWS Managed Microsoft AD e Active Directory autogestita: per ulteriori informazioni, vedi Connect to your Microsoft AD directory.
- Provider basati su SAML: per un elenco completo, consulta Provider di identità supportati.
- La directory IAM Identity Center: per ulteriori informazioni, consulta Gestire le identità in IAM Identity Center e Trusted Identity Propagation tra le applicazioni nella Guida per l'AWS IAM Identity Center utente.

In che modo l'autenticazione influisce sull'accesso e sull'assegnazione dell'utente

La modalità di autenticazione scelta per Amazon EMR Studio influisce sul modo in cui gli utenti accedono a uno Studio, sul modo in cui assegni un utente a uno Studio e sul modo in cui l'utente autorizza gli utenti (concede loro autorizzazioni) per eseguire operazioni come la creazione di nuovi cluster Amazon EMR.

Nella tabella seguente vengono riepilogati i metodi di accesso per EMR Studio in base alla modalità di autenticazione.

Opzioni di accesso a EMR Studio secondo la modalità di autenticazione

Modalità di autenticazione	Metodo di accesso	Descrizione
• IAM (autenticazione e federazione) • IAM Identity Center	URL di EMR Studio	Gli utenti accedono a uno Studio utilizzando l'URL di accesso allo Studio. Ad esempio, `https://xxxxxxxxxxxxxxxxxxxxxx.emrstudio-prod.us-east-1.amazonaws.com` . Gli utenti immettono le credenziali IAM quando si utilizza l'autenticazione IAM. Quando utilizzi la federazione IAM o IAM Identity Center, EMR

Modalità di autenticazione	Metodo di accesso	Descrizione
		Studio reindirizza gli utenti all'URL di accesso del gestore dell'identità digitale per inserire le credenziali. Nel contesto della federazione delle identità, questa opzione di accesso è chiamata l'accesso avviato dal Service Provider (SP).
• IAM (federazione) • IAM Identity Center	Portale del provider di identità (IdP)	Gli utenti accedono al portale del provider di identità, ad esempio il portale di Azure e avviano la console Amazon EMR. Dopo aver avviato la console Amazon EMR, gli utenti selezionano e aprono uno Studio dal Elenco degli Studio. È inoltre possibile configurare EMR Studio come applicazione SAML in modo che gli utenti possano accedere a uno Studio specifico dal portale del provider di identità. Per istruzioni, consulta Per configurare un EMR Studio come applicazione SAML nel portale IdP. Nel contesto della federazione delle identità, questa opzione di accesso è chiamata accesso avviato dal provider di identità (IdP).
• IAM (autenticazione)	AWS Management Console	Gli utenti accedono AWS Management Console utilizzando le credenziali IAM e aprono uno Studio dall'elenco Studios nella console Amazon EMR.

La tabella seguente illustra l'assegnazione e l'autorizzazione dell'utente per EMR Studio in base alla modalità di autenticazione.

Assegnazione e autorizzazione utente di EMR Studio secondo la modalità di autenticazione

Modalità di autenticazione	Assegnazione dell'utente	Autorizzazione dell'utente
IAM (autenticazione e federazione)	Consenti l'operazione `CreateStudioPresignedUrl` in una policy di autorizzazione IAM collegata a un'identità IAM (utente, gruppo o ruolo). Per gli utenti federati, consenti l'operazione `CreateStudioPresignedUrl` in un IAM nel policy di autorizzazione configurata per il ruolo IAM utilizzato per la federazione. Utilizzare il controllo di accesso basato su attributi (ABAC) per specificare lo Studio o gli studi a cui l'utente può accedere. Per istruzioni, consultare Assegnare un utente o un gruppo a un EMR Studio.	Definire le policy di autorizzazione IAM che consentono determinate azioni di EMR Studio. Per gli utenti nativi, collega la policy di autorizzazione IAM a un'identità IAM (utente, gruppo o ruolo). Per gli utenti federati, consenti le operazioni di Studio nel policy di autorizzazione configurata per il ruolo IAM utilizzato per la federazione. Per ulteriori informazioni, consulta Configurazione delle autorizzazioni utente di EMR Studio per Amazon o EC2 Amazon EKS.
Centro identità IAM	Per gli Studio creati con `IdCUserAssignment` impostati su `REQUIRED`, mappa gli utenti allo Studio con una policy di sessione specificata. Per ulteriori informazioni, consulta Assegnare un utente o un gruppo a un EMR Studio. Per gli Studio creati con `IdCUserAssignment` impostati	Facoltativo: definire le policy di sessione IAM che consentono alcune operazioni di EMR Studio. Mappare una policy di sessione a un utente quando si assegna l'utente a un Studio. Per ulteriori informazioni, consulta Autorizzazioni utente per la modalità di autenticazione IAM Identity Center.

Modalità di autenticazione	Assegnazione dell'utente	Autorizzazione dell'utente
	su `OPTIONAL`, qualsiasi utente o gruppo di Identity Center può accedere allo Studio.	

Controllo accessi

In Amazon EMR Studio, configuri l'autorizzazione utente (autorizzazioni) con policy basate su identità AWS Identity and Access Management (IAM). In queste policy basate su identità, specificare quali operazioni e risorse sono consentite, nonché le condizioni in base alle quali le operazioni sono consentite.

Autorizzazioni utente per la modalità di autenticazione IAM

Per impostare le autorizzazioni utente quando si utilizza l'autenticazione IAM per EMR Studio, è possibile consentire operazioni come `elasticmapreduce:RunJobFlow` in una policy di autorizzazione IAM. Puoi creare una o più policy di autorizzazione da utilizzare. Ad esempio, è possibile creare una policy di base che non consente a un utente di creare nuovi cluster Amazon EMR e un'altra policy che consente la creazione di cluster. Per un elenco di tutte le operazioni di Studio, consulta AWS Identity and Access Management autorizzazioni per gli utenti di EMR Studio.

Autorizzazioni utente per la modalità di autenticazione IAM Identity Center

Quando utilizzi l'autenticazione IAM Identity Center, crei un singolo ruolo utente di EMR Studio. Il ruolo utente è un ruolo IAM dedicato che uno Studio assume quando un utente accede.

Aallegare le policy di sessione IAM al ruolo utente di EMR Studio. Un policy di sessione è un tipo speciale di policy di autorizzazione IAM che limita ciò che un utente federato può fare durante una sessione di accesso a Studio. Le policy di sessione consentono di impostare autorizzazioni specifiche per un utente o un gruppo senza la necessità di creare più ruoli utente per EMR Studio.

Durante l'assegnazione di utenti e gruppi allo Studio, potrai mappare una policy di sessione a tale utente o gruppo per applicare le autorizzazione granulari. Potrai inoltre aggiornare la policy di sessione utente o un gruppo in qualsiasi momento. Amazon EMR archivia ogni mappatura delle policy di sessione che hai creato.

Per ulteriori informazioni sulle policy di sessione, consulta Policy e autorizzazioni nella Guida per l'utente di AWS Identity and Access Management .

WorkSpace

I WorkSpace sono gli elementi costitutivi principali di Amazon EMR Studio. Per organizzare i notebook, gli utenti creano uno o più istanze WorkSpace in uno Studio. Per ulteriori informazioni, consulta Scopri gli spazi di lavoro di EMR Studio.

Analogamente alle aree di lavoro in JupyterLab, un'area di lavoro preserva lo stato del lavoro sui notebook. Tuttavia, l'interfaccia utente di Workspace estende l'JupyterLabinterfaccia open source con strumenti aggiuntivi che consentono di creare e collegare cluster EMR, eseguire processi, esplorare notebook di esempio e collegare repository Git.

L'elenco seguente include le caratteristiche principali di EMR Studio WorkSpaces:

- La visibilità del WorkSpace è basata sullo Studio. Le istanze WorkSpace creati in uno Studio non sono visibili in altri Studio.
- Per impostazione predefinita, un WorkSpace è condiviso e può essere visualizzato da tutti gli utenti di Studio. Tuttavia, solo un utente alla volta può aprire e lavorare in un WorkSpace. Lavorare contemporaneamente con altri utenti è possibile con Configurazione della collaborazione Workspace in EMR Studio
- È possibile collaborare contemporaneamente con altri utenti in un WorkSpace quando si abilita la collaborazione di WorkSpace. Per ulteriori informazioni, consulta Configurazione della collaborazione Workspace in EMR Studio.
- I notebook in un WorkSpace condividono lo stesso cluster EMR per eseguire i comandi. Puoi collegare un Workspace a un cluster Amazon EMR in esecuzione su EC2 Amazon o a un cluster virtuale ed endpoint gestito Amazon EMR su EKS.
- I WorkSpace possono passare a un'altra zona di disponibilità associata alle sottoreti di uno Studio. È possibile arrestare e riavviare un WorkSpace per richiedere il processo di failover. Quando si riavvia un WorkSpace, EMR Studio avvia il WorkSpace in una zona di disponibilità diversa nel VPC di Studio quando Studio è configurato con l'accesso a più zone di disponibilità. Se Studio dispone di una sola zona di disponibilità, EMR Studio tenta di avviare il workspace in una sottorete diversa. Per ulteriori informazioni, consulta Risolvi i problemi di connettività di WorkSpace.
- Un workspace può connettersi a cluster in una qualsiasi delle sottoreti associate a uno Studio.

Per ulteriori informazioni sulla creazione e configurazione di WorkSpace EMR Studio, consulta Scopri gli spazi di lavoro di EMR Studio.

Archiviazione di notebook in Amazon EMR Studio

Quando utilizzi un'istanza WorkSpace, EMR Studio salva automaticamente le celle nei file notebook ad una cadenza regolare al posizione Amazon S3 associato al tuo Studio. Questo processo di backup mantiene il lavoro tra le sessioni in modo da poter tornare in un secondo momento senza commettere modifiche a un repository Git. Per ulteriori informazioni, consulta Salva i contenuti di Workspace in EMR Studio.

Quando elimini un file notebook da un WorkSpace, EMR Studio elimina automaticamente la versione di backup da Amazon S3. Tuttavia, se elimini un'istanza WorkSpace senza prima eliminare i file notebook, i file notebook rimangono in Amazon S3 e continuano ad addebitare costi di archiviazione. Per ulteriori informazioni, consulta Eliminare un workspace e i file del taccuino in EMR Studio.

Caratteristiche, requisiti e limiti di EMR Studio

Questo argomento include gli elementi da considerare quando si lavora con Amazon EMR Studio, tra cui considerazioni su regioni e strumenti, requisiti dei cluster e limitazioni tecniche.

Considerazioni

Quando lavori con EMR Studio, tieni in considerazione i seguenti aspetti:

- EMR Studio è disponibile nelle seguenti versioni: Regioni AWS
  - Stati Uniti orientali (Ohio) (us-east-2)
  - Stati Uniti orientali (Virginia settentrionale) (us-east-1)
  - Stati Uniti occidentali (California settentrionale) (us-west-1)
  - Stati Uniti occidentali (Oregon) (us-west-2)
  - Africa (Città del Capo) (af-south-1)
  - Asia Pacifico (Hong Kong) ap-east-1
  - Asia Pacifico (Giacarta) (ap-southeast-3) *
  - Asia Pacifico (Melbourne) (ap-southeast-4) *
  - Asia Pacifico (Mumbai) (ap-south-1)
  - Asia Pacifico (Osaka) (ap-northeast-3) *

- Asia Pacifico (Seoul) (ap-northeast-2)
- Asia Pacifico (Singapore) (ap-southeast-1)
- Asia Pacifico (Sydney) (ap-southeast-2)
- Asia Pacifico (Tokyo) (ap-northeast-1)
- Canada (Centrale) (ca-central-1)
- Europa (Francoforte) (eu-central-1)
- Europa (Irlanda) (eu-west-1)
- Europa (Londra) (eu-west-2)
- Europa (Milano) (eu-south-1)
- Europe (Parigi) (eu-west-3)
- Europa (Spagna) (eu-south-2)
- Europa (Stoccolma) (eu-north-1)
- Europa (Zurigo) (eu-central-2) *
- Israele (Tel Aviv) (il-central-1) *
- Medio Oriente (Emirati Arabi Uniti) (me-central-1) *
- Sud America (San Paolo) (sa-east-1)
- AWS GovCloud (Stati Uniti orientali) (-1) gov-us-east
- AWS GovCloud (Stati Uniti occidentali) (gov-us-west-1)

  * L'interfaccia utente live di Spark non è supportata in queste regioni.

- Per consentire agli utenti di effettuare il provisioning di nuovi cluster EMR in esecuzione su Amazon EC2 for a Workspace, puoi associare EMR Studio a un set di modelli di cluster. Gli amministratori possono definire modelli di cluster con Service Catalog e scegliere se un utente o un gruppo può accedere ai modelli o a nessuno dei modelli all'interno di uno Studio.
- Quando definisci le autorizzazioni di accesso ai file di notebook archiviati in Amazon S3 o leggi segreti AWS Secrets Manager da, usa il ruolo del servizio Amazon EMR. Le politicy di sessione non sono supportate con queste autorizzazioni.
- È possibile creare più EMR Studios per controllare l'accesso ai cluster EMR in diversi. VPCs
- Utilizza il AWS CLI per configurare Amazon EMR su cluster EKS. È quindi possibile utilizzare l'interfaccia Studio per collegare cluster ai Workspace con un endpoint gestito per eseguire processi notebook.

- Quando utilizzi la propagazione delle identità attendibili con Amazon EMR, ci sono altre considerazioni che si applicano anche a EMR Studio. Per ulteriori informazioni, consulta Considerazioni e limitazioni per Amazon EMR con l'integrazione del Centro identità.
- EMR Studio non supporta i seguenti comandi magic Python:
  - `%alias`
  - `%alias_magic`
  - `%automagic`
  - `%macro`
  - `%%js`
  - `%%javascript`
  - Modifica di `proxy_user` mediante `%configure`
  - Modifica di `KERNEL_USERNAME` mediante `%env` o `%set_env`
- Amazon EMR sui cluster EKS non supporta i comandi SparkMagic per EMR Studio.
- Per scrivere istruzioni Scala a più righe nelle celle del notebook, assicurarsi che tutte le righe tranne l'ultima finiscano con un punto. Nell'esempio seguente viene utilizzata la sintassi corretta per le istruzioni Scala a più righe.

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

- Per aumentare la sicurezza delle applicazioni off-console che potresti utilizzare con Amazon EMR, i domini di hosting delle applicazioni sono registrati nella Public Suffix List (PSL). Alcuni esempi di questi domini di hosting includono: `emrstudio-prod.us-east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Per maggiore sicurezza, se hai bisogno di impostare cookie sensibili nel nome di dominio predefinito, consigliamo di utilizzare i cookie con un prefisso `__Host-`. Questa pratica ti aiuterà a difendere il tuo dominio dai tentativi CSRF (cross-site request forgery). Per ulteriori informazioni, consulta la pagina Set-Cookie in Mozilla Developer Network.
- Gli endpoint Amazon EMR Studio Workspaces e Persistent UI utilizzano moduli crittografici convalidati FIPS 140 per facilitare l'adozione del servizio per encryption-in-transit carichi di lavoro regolamentati. Per ulteriori informazioni sugli endpoint dell'interfaccia utente persistente, consulta Visualizzazione delle interfacce utente delle applicazioni persistenti in Amazon EMR. Per ulteriori informazioni sui notebook, consulta la panoramica dei notebook Amazon EMR.

Problemi noti

- Un EMR Studio che utilizza IAM Identity Center con la propagazione delle identità attendibili abilitata può associarsi solo a cluster EMR che utilizzano la propagazione delle identità attendibili.
- Assicurati di disattivare gli strumenti di gestione proxy come FoxyProxy o SwitchyOmega nel browser prima di creare uno Studio. I proxy attivi possono causare errori quando scegli Create Studio (Crea Studio) e tradursi in un messaggio di errore Network Failure (Errore di rete).
- I kernel che vengono eseguiti su cluster Amazon EMR su EKS possono non avviarsi a causa di problemi di timeout. Se si verifica un errore o un problema durante l'avvio del kernel, è necessario chiudere il file notebook, arrestare il kernel e in seguito riaprire il file notebook.
- L'operazione Restart kernel (Riavvia kernel) non funziona come previsto quando si usa un cluster Amazon EMR su EKS. Dopo aver selezionato Restart kernel (Riavvia kernel), aggiorna il Workspace affinché il riavvio abbia effetto.
- Se un Workspace non è collegato a un cluster, viene visualizzato un messaggio di errore quando un utente dello Studio apre un file notebook e tenta di selezionare un kernel. Puoi ignorare questo messaggio di errore scegliendo Ok, ma è necessario collegare il Workspace a un cluster e selezionare un kernel prima di poter eseguire il codice del notebook.
- Quando utilizzi Amazon EMR 6.2.0 con una configurazione di sicurezza per impostare la protezione del cluster, l'interfaccia del Workspace appare vuota e non funziona come previsto. Se desideri configurare la crittografia dei dati o l'autorizzazione Amazon S3 per EMRFS per un cluster, consigliamo di utilizzare un'altra versione di Amazon EMR supportata. EMR Studio funziona con Amazon EMR versione 5.32.0 (Amazon EMR serie 5.x) o 6.2.0 (Amazon EMR serie 6.x) e versioni successive.
- Quando Esegui il debug di Amazon EMR in esecuzione su Amazon Jobs EC2 , i collegamenti all'interfaccia utente Spark sul cluster potrebbero non funzionare o non essere visualizzati. Per rigenerare i collegamenti, crea una nuova cella del notebook ed esegui il comando `%%info`.
- Jupyter Enterprise Gateway non elimina i kernel inattivi sul nodo primario di un cluster nelle seguenti versioni Amazon EMR: 5.32.0, 5.33.0, 6.2.0 e 6.3.0. I kernel inattivi consumano risorse di elaborazione e possono causa l'interruzione dei cluster a esecuzione prolungata. È possibile configurare l'eliminazione del kernel inattivo per Jupyter Enterprise Gateway utilizzando il seguente script di esempio. Puoi Connect al nodo primario del cluster Amazon EMR tramite SSH oppure inviare lo script come fase. Per ulteriori informazioni, consulta Esecuzione di comandi e script su un cluster 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
```

- Quando utilizzi una policy di terminazione automatica con Amazon EMR versioni 5.32.0, 5.33.0, 6.2.0 o 6.3.0, Amazon EMR contrassegna un cluster come inattivo e potrebbe terminarlo in automatico anche se disponi di un kernel Python3 attivo. Questo perché l'esecuzione di un kernel Python3 non invia un processo Spark sul cluster. Per utilizzare la terminazione automatica con un kernel Python3, consigliamo di utilizzare Amazon EMR versione 6.4.0 o successive. Per ulteriori informazioni sulla terminazione automatica, consulta Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR.
- Quando usi `%%display` per visualizzare uno Spark DataFrame in una tabella, le tabelle molto larghe potrebbero essere troncate. È possibile fare clic con il pulsante destro del mouse sull'output e selezionare Creare nuova vista per l'output per ottenere una schermata scorrevole dell'output.
- L'avvio di un kernel basato su Spark PySpark, come Spark o SparkR, avvia una sessione Spark e l'esecuzione di una cella in un notebook mette in coda i lavori Spark in quella sessione. Quando interrompi una cella in esecuzione, il processo Spark continua a essere eseguito. Per interrompere il processo Spark, è necessario utilizzare l'interfaccia utente Spark sul cluster. Per istruzioni sulla modalità di connessione all'interfaccia utente di Spark, consulta Debug di applicazioni e processi con EMR Studio.
- L'utilizzo di Amazon EMR Studio Workspaces come utente root in un Account AWS causa un errore. `403: Forbidden` Questo perché la configurazione di Jupyter Enterprise Gateway in Amazon EMR non consente l'accesso all'utente root. Ti consigliamo di non utilizzare l'utente root per le tue attività quotidiane. Per altre opzioni di autenticazione, consulta AWS Identity and Access Management Amazon EMR.

Limitazioni delle caratteristiche

Amazon EMR Studio non supporta le seguenti caratteristiche di Amazon EMR:

- Collegamento ed esecuzione di processi su cluster EMR con una configurazione di sicurezza che specifica l'autenticazione Kerberos
- Cluster con più nodi primari

- Cluster che utilizzano EC2 istanze Amazon basate su AWS Graviton2 per versioni di Amazon EMR 6.x precedenti alla 6.9.0 e versioni 5.x precedenti alla 5.36.1

Le seguenti funzionalità non sono supportate da uno Studio che utilizza la propagazione delle identità attendibili:

- Creazione di cluster EMR senza un modello.
- Utilizzo di applicazioni EMR serverless.
- Avvio di cluster Amazon EMR su EKS.
- Utilizzo di un ruolo di runtime.
- Abilitazione della collaborazione con SQL Explorer o Workspace.

Limiti del servizio per EMR Studio

La tabella seguente mostra i limiti del servizio per EMR Studio.

Elemento	Limite
EMR Studio	AWS Massimo 100 per account
Sottoreti	Un massimo di 5 associate a ciascun EMR Studio
Gruppi IAM Identity Center	Un massimo di 5 assegnati a ciascun EMR Studio
Utenti IAM Identity Center	Un massimo di 100 assegnati a ciascun EMR Studio

Best practice per VPC e sottorete per EMR Studio

Utilizza le seguenti best practice per configurare un Amazon Virtual Private Cloud (Amazon VPC) con sottoreti per EMR Studio:

- È possibile specificare un massimo di cinque sottoreti in un VPC da associare allo Studio. Si consiglia di fornire più sottoreti in diverse zone di disponibilità per supportare la disponibilità di Workspace e consentire agli utenti di Studio l'accesso ai cluster in diverse zone di disponibilità. Per ulteriori informazioni sull'utilizzo delle sottoreti e delle zone di disponibilità, consulta VPCs e sottoreti nella Guida per l'utente. VPCs Amazon Virtual Private Cloud

- Le sottoreti specificate dovrebbero essere in grado di comunicare tra loro.
- Per consentire agli utenti di collegare un Workspace a repository Git ospitati pubblicamente, è necessario specificare solo sottoreti private che hanno accesso a Internet tramite Network Address Translation (NAT). Per ulteriori informazioni sulla configurazione di una sottorete privata per Amazon EMR, consulta Sottoreti private.
- Quando si utilizza Amazon EMR su EKS con EMR Studio, deve esserci almeno una sottorete in comune tra Studio e il cluster Amazon EKS utilizzato per registrare il cluster virtuale. In caso contrario, l'endpoint gestito non verrà visualizzato come opzione nei Workspace Studio. Puoi creare un cluster Amazon EKS e associarlo a una sotterete appartenente allo Studio o creare uno Studio e specificare le sottoreti del tuo cluster EKS.
- Se intendi utilizzare Amazon EMR su EKS con EMR Studio, scegli lo stesso VPC dei nodi worker del cluster Amazon EKS.

Requisiti del cluster Amazon EMR

Cluster Amazon EMR in esecuzione su Amazon EC2

Tutti i cluster Amazon EMR in esecuzione su Amazon EC2 che crei per un EMR Studio Workspace devono soddisfare i seguenti requisiti. I cluster creati utilizzando l'interfaccia di EMR Studio soddisfano in automatico questi requisiti.

- Il cluster deve utilizzare Amazon EMR versione 5.32.0 (Amazon EMR serie 5.x) o 6.2.0 (Amazon EMR serie 6.x) o versioni successive. È possibile creare un cluster utilizzando la console Amazon EMR o SDK e quindi collegarlo a un EMR Studio Workspace. AWS Command Line Interface Gli utenti dello Studio possono anche creare e allegare un cluster durante la creazione o l'utilizzo di un Workspace Amazon EMR. Per ulteriori informazioni, consulta Collegamento di un calcolo a un WorkSpace EMR Studio.
- Il cluster deve essere all'interno di Amazon Virtual Private Cloud. La piattaforma EC2 -Classic non è supportata.
- Sul cluster devono essere installati Spark, Livy e Jupyter Enterprise Gateway. Se prevedi di utilizzare il cluster per SQL Explorer, devi installare sia Presto che Spark.
- Per utilizzare SQL Explorer, il cluster deve utilizzare Amazon EMR versione 5.34.0 o successive oppure versione 6.4.0 e disporre di Presto installato. Se desideri specificare AWS Glue Data Catalog come metastore Hive per Presto, devi configurarlo sul cluster. Per ulteriori informazioni, consulta Utilizzo di Presto con AWS Glue Data Catalog.

- Il cluster deve trovarsi in una sottorete privata con Network Address Translation (NAT) per utilizzare repository Git in hosting pubblico con EMR Studio.

Quando si lavora con EMR Studio, si consigliano le seguenti configurazioni del cluster.

- Imposta la modalità di implementazione per le sessioni Spark nella modalità cluster. La modalità cluster posiziona i processi principali dell'applicazione sui nodi principali e non sul nodo primario di un cluster. Così facendo, si alleggerisce il nodo primario delle potenziali pressioni in termini di memoria. Per ulteriori informazioni, consulta Panoramica della modalità cluster nella documentazione di Apache Spark.
- Modificare il timeout Livy dall'impostazione predefinita di un'ora a sei ore come nella seguente configurazione di esempio.

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

- Creare flotte di istanze diverse con un massimo di 30 istanze e selezionare più tipi di istanza nel parco istanze Spot. Ad esempio, è possibile specificare i seguenti tipi di istanza ottimizzati per la memoria per i carichi di lavoro Spark: r5.2x, r5.4x, r5.8x, r5.12x, r5.16x, r4.2x, r4.4x, r4.8x, r4.12, ecc. Per ulteriori informazioni, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR.
- Utilizza la strategia di allocazione ottimizzata in termini di capacità per le istanze Spot per aiutare Amazon EMR a selezionare le istanze in modo efficace sulla base di informazioni sulla capacità in tempo reale fornite da Amazon. EC2 Per ulteriori informazioni, consulta Strategia di allocazione per parchi istanze.
- Abilita il dimensionamento gestito sul cluster. Impostare il parametro dei nodi principali massimi sulla capacità minima persistente che si prevede di utilizzare e configurare la scalabilità su un parco istanze di processi ben diversificato in esecuzione su istanze Spot per risparmiare sui costi. Per ulteriori informazioni, consulta Utilizzo del dimensionamento gestito in Amazon EMR.

Ti invitiamo a mantenere abilitato Amazon EMR Block Public Access e questo per limitare il traffico SSH in entrata a fonti attendibili. L'accesso in ingresso a un cluster consente agli utenti di eseguire

notebook sul cluster. Per ulteriori informazioni, consultare Utilizzo del blocco dell'accesso pubblico di Amazon EMR e Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Cluster Amazon EMR su EKS

Oltre ai cluster EMR in esecuzione su Amazon EC2, puoi configurare e gestire Amazon EMR su cluster EKS per EMR Studio utilizzando. AWS CLI Imposta Amazon EMR sui cluster EKS utilizzando le seguenti linee guida:

- Crea un endpoint HTTPS gestito per Amazon EMR su cluster EKS. Gli utenti collegano un Workspace a un endpoint gestito. Il cluster Amazon Elastic Kubernetes Service (EKS) utilizzato per registrare il cluster virtuale deve disporre di una sottorete privata per supportare gli endpoint gestiti.
- Utilizza un cluster Amazon EKS con almeno una sottorete privata e una Network Address Translation (NAT) quando desideri utilizzare i repository Git in hosting pubblico.
- Evita di usare Arm Amazon Linux ottimizzato per Amazon EKS AMIs, che non sono supportati per Amazon EMR sugli endpoint gestiti EKS.
- Evita AWS Fargate di utilizzare solo cluster Amazon EKS, che non sono supportati.

Configurazione di Amazon EMR Studio

Questa sezione è dedicata agli amministratori di EMR Studio. Illustra come configurare un EMR Studio per il tuo team e fornisce istruzioni su processi come l'assegnazione di utenti e gruppi, la configurazione di modelli di cluster e l'ottimizzazione di Apache Spark per EMR Studio.

Argomenti

- Autorizzazioni di amministratore per creare e gestire un EMR Studio
- Configurazione di un EMR Studio
- Monitora, aggiorna ed elimina le risorse di Amazon EMR Studio
- Crittografia di notebook e file dell'area di lavoro EMR Studio
- Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio
- Crea AWS CloudFormation modelli per Amazon EMR Studio
- Definizione di accesso e autorizzazioni per i repository basati su Git
- Ottimizzazione dei processi Spark in EMR Studio

Autorizzazioni di amministratore per creare e gestire un EMR Studio

Le autorizzazioni IAM descritte in questa pagina consentono di creare e gestire un EMR Studio. Per informazioni dettagliate su ciascuna autorizzazione necessaria, consulta Autorizzazioni necessarie per gestire un EMR Studio.

Autorizzazioni necessarie per gestire un EMR Studio

Nella tabella seguente sono elencate le operazioni relative alla creazione e alla gestione di un EMR Studio. Nella tabella vengono inoltre visualizzate le autorizzazioni necessarie per ciascuna operazione.

> **Note**
>
> Le operazioni di `SessionMapping` di IAM Identity Center e Studio ti occorrono soltanto quando utilizzi la modalità di autenticazione IAM Identity Center.

Autorizzazioni per creare e gestire un EMR Studio

Operazione	Autorizzazioni
Creazione di uno Studio	`"elasticmapreduce:CreateStudio", "sso:CreateApplication", "sso:PutApplicationAuthenticationMethod", "sso:PutApplicationGrant", "sso:PutApplicationAccessScope", "sso:PutApplicationAssignmentConfiguration", "iam:PassRole"`
Descrizione di uno Studio	`"elasticmapreduce:DescribeStudio", "sso:GetManagedApplicationInstance"`
Elencazione degli Studio	`"elasticmapreduce:ListStudios"`
Eliminazione di uno Studio	`"elasticmapreduce:DeleteStudio",`

Operazione	Autorizzazioni
	"sso:DeleteApplication", "sso:DeleteApplicationAuthentica tionMethod", "sso:DeleteApplicationAccessScope", "sso:DeleteApplicationGrant"
Additional permissions required when you use IAM Identity Center mode	
Assegnazione di utenti o gruppi a uno Studio	"elasticmapreduce:CreateStudioSessio nMapping", "sso:GetProfile", "sso:ListDirectoryAssociations", "sso:ListProfiles", "sso:AssociateProfile", "sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:ListInstances", "sso:CreateApplicationAssignment", "sso:DescribeInstance", "organizations:DescribeOrga nization", "organizations:ListDelegatedAdmini strators", "sso:CreateInstance", "sso:DescribeRegisteredRegions", "sso:GetSharedSsoConfiguration", "iam:ListPolicies"
Recupero dei dettagli delle assegnazioni Studio per un utente o un gruppo	"sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:DescribeApplication", "elasticmapreduce:GetStudioSessio nMapping"

Operazione	Autorizzazioni
Elencazione di tutti gli utenti e i gruppi assegnati a uno Studio	`"elasticmapreduce:ListStudioSessionMappings"`
Aggiornamento della policy di sessione associata a un utente o a un gruppo assegnato a uno Studio	`"sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:DescribeApplication", "sso:DescribeInstance", "elasticmapreduce:UpdateStudioSessionMapping"`
Rimozione di un utente o un gruppo da uno Studio	`"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"`

Creare una policy con le autorizzazioni di amministrazione per EMR Studio

1. Segui le istruzioni in Creazione di policy IAM per creare una policy utilizzando uno dei seguenti esempi. Le autorizzazioni di cui hai bisogno dipendono dalla tua modalità di autenticazione per EMR Studio.

   Inserisci valori personalizzati per questi elementi:

   - Sostituisci *<your-resource-ARN>* per specificare l'Amazon Resource Name (ARN) dell'oggetto o degli oggetti coperti dall'istruzione per i tuoi casi d'uso.

- Sostituiscilo *<region>* con il codice del Regione AWS luogo in cui intendi creare lo Studio.
- Sostituisci *<aws-account_id>* con l'ID dell' AWS account per lo Studio.
- Sostituisci *<EMRStudio-Service-Role>* e *<EMRStudio-User-Role>* con i nomi del tuo ruolo di servizio EMR Studio e del ruolo utente di EMR Studio.

Example Policy di esempio: autorizzazioni di amministrazione quando si utilizza la modalità di autenticazione 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 Policy di esempio: autorizzazioni di amministrazione quando si utilizza la modalità di autenticazione IAM Identity Center

> **Note**
>
> Le directory Identity Center e Identity Center APIs non supportano la specificazione di un ARN nell'elemento risorsa di una dichiarazione di policy IAM. Per consentire l'accesso a IAM Identity Center e IAM Identity Center Directory, le seguenti autorizzazioni specificano tutte le risorse, "Resource":"*", per le operazioni di IAM Identity Center. Per ulteriori informazioni, consulta la sezione Operazioni, risorse e chiavi di condizione per la directory IAM Identity Center.

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. Collega la policy all'identità IAM (utente, ruolo o gruppo). Per ricevere istruzioni, consulta Aggiunta e rimozione di autorizzazioni per identità IAM.

Configurazione di un EMR Studio

Completa i seguenti passaggi per configurare un EMR Studio.

Prima di iniziare

> **Note**
>
> Se intendi utilizzare EMR Studio con Amazon EMR su EKS, ti consigliamo di configurare Amazon EMR su EKS per EMR Studio prima di configurare uno Studio.

Prima di configurare un EMR Studio, assicurati di disporre dei seguenti elementi:

- Un Account AWS. Per istruzioni, consultare Prima di configurare Amazon EMR.

- Autorizzazioni per creare e gestire un EMR Studio. Per ulteriori informazioni, consulta the section called “Autorizzazioni di amministratore per creare un EMR Studio”.
- Un bucket Amazon S3 in cui EMR Studio può eseguire il backup dei Workspace e dei file notebook nel tuo Studio. Per le istruzioni, consulta Creazione di un bucket nella Amazon Simple Storage Service Guida per l'utente (S3).
- Se desideri collegarti a un cluster Amazon EMR su o EC2 Amazon EMR su EKS o utilizzare repository Git, hai bisogno di un Amazon Virtual Private Cloud (VPC) per Studio e un massimo di cinque sottoreti. Non è necessario un VPC per utilizzare EMR Studio con EMR Serverless. Per suggerimenti su come configurare la rete, consulta Best practice per VPC e sottorete per EMR Studio.

Configurazione di un EMR Studio

1. Scelta di una modalità di autenticazione per Amazon EMR Studio
2. Creare le seguenti risorse EMR Studio.
   - Creazione di un ruolo di servizio EMR Studio
   - Configurazione delle autorizzazioni utente di EMR Studio per Amazon o EC2 Amazon EKS
   - (Facoltativo) Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio.
3. Creazione di un EMR Studio
4. Assegnare un utente o un gruppo a un EMR Studio

Una volta completate queste fasi di impostazione, puoi Utilizzare un Amazon EMR Studio.

Scelta di una modalità di autenticazione per Amazon EMR Studio

EMR Studio supporta due modalità di autenticazione: la modalità di autenticazione IAM e la modalità di autenticazione IAM Identity Center. La modalità IAM utilizza AWS Identity and Access Management (IAM), mentre la modalità IAM Identity Center utilizza. AWS IAM Identity Center Quando crei un EMR Studio, scegli la modalità di autenticazione per tutti gli utenti di tale Studio. Per ulteriori informazioni sulle differenti modalità di autenticazione, consulta Autenticazione e accesso utente.

Utilizzare la tabella seguente per scegliere una modalità di autenticazione per EMR Studio.

Se hai...	Consigliamo…
Già familiarità con o hai precedentemente configurato l'autenticazione o la federazione IAM	Modalità di autenticazione IAM, che offre i vantaggi seguenti: • Fornisce una configurazione rapida per EMR Studio se si già gestiscono identità come utenti e gruppi in IAM. • Funziona con i provider di identità compatibili con with OpenID Connect (OIDC) o Security Assertion Markup Language 2.0 (SAML 2.0). • Supporta l'utilizzo di più provider di identità con lo stesso Account AWS. • Disponibile in un'ampia gamma di Regioni AWS. • Conforme a SOC 2.
Nuovo utente del AWS nostro Amazon EMR	Modalità di autenticazione IAM Identity Center, che offre le seguenti caratteristiche: • Supporta una facile assegnazione di utenti e gruppi alle AWS risorse. • Funziona con Microsoft Active Directory e provider di identità SAML 2.0. • Facilita la configurazione della federazione di più account in modo da non dover configura re separatamente la federazione per ciascun Account AWS membro dell'organizzazione.

Configurare una modalità di autenticazione IAM per Amazon EMR Studio

Con la modalità di autenticazione IAM, è possibile utilizzare l'autenticazione IAM o la federazione IAM. L'autenticazione IAM consente di gestire le identità IAM come utenti, gruppi e ruoli in IAM. Concedi agli utenti l'accesso a uno Studio con policy di autorizzazione IAM e controllo dell'accesso basato su attributi (ABAC). La federazione IAM ti consente di stabilire un rapporto di fiducia tra un provider di identità (IdP) di terze parti e AWS di gestire le identità degli utenti tramite il tuo IdP.

> **Note**
>
> Se utilizzi già IAM per controllare l'accesso alle AWS risorse o se hai già configurato il tuo provider di identità (IdP) per IAM, consulta Autorizzazioni utente per la modalità di autenticazione IAM per impostare le autorizzazioni utente quando utilizzi la modalità di autenticazione IAM per EMR Studio.

Uso della federazione IAM per Amazon EMR Studio

Per utilizzare la federazione IAM per EMR Studio, crei una relazione di fiducia tra il tuo Account AWS e il tuo provider di identità (IdP) e consenti agli utenti federati di accedere a. AWS Management Console Le fasi che intraprendi per creare questa relazione di fiducia differiscono a seconda dello standard federativo del tuo IdP.

In generale, si completano le seguenti attività per configurare la federazione con un IdP esterno. Per istruzioni complete, consulta Abilitare gli utenti federati SAML 2.0 per accedere alla AWS Management Console e Abilitare il gestore identità personalizzato per accedere alla AWS Management Console nella Guida per l'utente di AWS Identity and Access Management .

1. Raccogli informazioni dal tuo gestore dell'identità digitale. Questo di solito significa generare un documento di metadati per convalidare le richieste di autenticazione SAML dal tuo gestore dell'identità digitale.
2. Crea un'entità IAM del provider di identità per archiviare informazioni sul tuo gestore dell'identità digitale. Per ricevere istruzioni, consulta Creazione di provider di identità IAM.
3. Creare uno o più ruoli IAM per il gestore dell'identità digitale. EMR Studio assegna un ruolo a un utente federato quando l'utente effettua l'accesso. Il ruolo consente all'IdP di richiedere credenziali di sicurezza provvisorie per l'accesso a AWS. Per ricevere istruzioni, consulta Creazione di un ruolo per un provider di identità di terza parte (federazione). Le politiche di autorizzazione assegnate al ruolo determinano cosa possono fare gli utenti federati in AWS e in EMR Studio. Per ulteriori informazioni, consulta Autorizzazioni utente per la modalità di autenticazione IAM.
4. (Per i provider SAML) Completa il trust SAML configurando il tuo IdP con le informazioni AWS e i ruoli che desideri vengano assunti dagli utenti federati. Questo processo di configurazione crea fiducia tra il tuo AWS IdP e. Per ulteriori informazioni, consulta Configurazione del provider di identità SAML 2.0 con una relazione di attendibilità e aggiungendo attestazioni.

Per configurare un EMR Studio come applicazione SAML nel portale del provider di identità

È possibile configurare un particolare EMR Studio come applicazione SAML utilizzando un deep link allo Studio. In questo modo, gli utenti accedono al tuo portale del provider di identità e lanciano uno Studio specifico invece di navigare attraverso la console Amazon EMR.

- Utilizza il seguente formato per configurare un deep link a EMR Studio come URL di destinazione dopo la verifica dell'asserzione SAML.

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

Configurazione della modalità di autenticazione IAM Identity Center per Amazon EMR Studio

AWS IAM Identity Center Per prepararsi a EMR Studio, è necessario configurare l'origine dell'identità ed effettuare il provisioning di utenti e gruppi. L'allocazione è il processo che rende disponibili le informazioni su utenti e gruppi per l'utilizzo da parte di IAM Identity Center e delle applicazioni che utilizzano IAM Identity Center. Per ulteriori informazioni, consulta Provisioning di utenti e gruppi.

EMR Studio supporta l'uso dei seguenti gestori dell'identità digitale per IAM Identity Center:

- AWS Managed Microsoft AD e Active Directory autogestita: per ulteriori informazioni, vedi Connect to your Microsoft AD directory.
- Provider basati su SAML: per un elenco completo, consulta Provider di identità supportati.
- La directory di IAM Identity Center: per ulteriori informazioni, consulta la sezione Gestione delle identità in IAM Identity Center.

Configurazione di IAM Identity Center per EMR Studio

1. Per configurare IAM Identity Center per EMR Studio, sono necessari i seguenti elementi:

   - Un account di gestione nell' AWS organizzazione se si utilizzano più account nell'organizzazione.

   > **Note**
   >
   > Per abilitare IAM Identity Center e allocare utenti e gruppi, dovresti utilizzare esclusivamente l'account di gestione. Dopo aver configurato IAM Identity Center,

puoi utilizzare un account membro per creare un EMR Studio e assegnare utenti e gruppi. Per ulteriori informazioni sulla AWS terminologia, vedere AWS Organizations Terminologia e concetti.

- Se hai abilitato IAM Identity Center prima del 25 novembre 2019, potresti dover abilitare le applicazioni che utilizzano IAM Identity Center per gli account della tua AWS organizzazione. Per ulteriori informazioni, consulta Abilitare le applicazioni integrate con IAM Identity Center negli AWS account.
- Accertati di soddisfare i prerequisiti elencati nella pagina Prerequisiti di IAM Identity Center.

2. Segui le istruzioni in Abilita IAM Identity Center per abilitare IAM Identity Center nel Regione AWS luogo in cui desideri creare EMR Studio.
3. Collega IAM Identity Center al tuo gestore dell'identità digitale (IdP) e alloca gli utenti e i gruppi che desideri assegnare allo Studio.

Se utilizzi...	Esegui questa operazione…
Una directory Microsoft AD	1. Segui le istruzioni in Connect to your Microsoft AD directory per connettere la tua AWS Managed Microsoft AD directory o Active Directory autogestita utilizzando AWS Directory Service. 2. Per allocare utenti e gruppi per IAM Identity Center, puoi sincronizzare i dati di identità dall'AD di origine a IAM Identity Center. Puoi sincronizzare le identità della tua AD di origine in molti modi. Un possibile modo è quello di assegnare utenti o gruppi AD a un account AWS nell' organizzazione. Per ricevere istruzioni, consulta Accesso single sign-on. La sincronizzazione può richiedere fino a due ore. Dopo avere completato questa fase, verranno visualizzati gli utenti e i gruppi sincronizzati nell'archivio identità.

Se utilizzi...	Esegui questa operazione…
	Note Gli utenti e i gruppi non vengono visualizzati nel tuo Identity Store finché non sincronizzi le informazioni su utenti e gruppi o utilizzi just-in-time (JIT) il provisioning degli utenti. Per ulteriori informazioni, consulta Provisioning quando gli utenti provengono da Active Directory. 3. (Facoltativo) Dopo aver sincronizzato gli utenti e i gruppi AD, puoi rimuovere il loro accesso al tuo AWS account configurato nel passaggio precedente. Per ricevere istruzioni, consulta Rimozione dell'accesso degli utenti.
Un provider di identità esterno	Segui le istruzioni in Connessione al provider di identità esterno.
La directory di IAM Identity Center	Quando crei utenti e gruppi in IAM Identity Center, l'allocazione è automatica. Per ulteriori informazioni, consulta la sezione Gestione delle identità in IAM Identity Center.

Ora puoi assegnare utenti e gruppi dall'archivio identità a un EMR Studio. Per istruzioni, consultare Assegnare un utente o un gruppo a un EMR Studio.

Creazione di un ruolo di servizio EMR Studio

Informazioni sul ruolo di servizio EMR Studio

Ogni EMR Studio utilizza un ruolo IAM con autorizzazioni che consentono allo Studio di interagire con altri servizi. AWS Questo ruolo di servizio deve includere autorizzazioni che consentano a EMR Studio di stabilire un canale di rete sicuro tra Workspace e cluster, di archiviare i file Amazon S3 Control del notebook e di accedervi AWS Secrets Manager mentre collega un Workspace a un repository Git.

Utilizza il ruolo di servizio Studio (anziché le policy di sessione) per definire tutte le autorizzazioni di accesso ad Amazon S3 per l'archiviazione dei file notebook e per definire autorizzazioni di accesso AWS Secrets Manager .

Come creare un ruolo di servizio per EMR Studio su Amazon o EC2 Amazon EKS

1. Segui le istruzioni in Creazione di un ruolo per delegare le autorizzazioni a un AWS servizio per creare il ruolo di servizio con la seguente politica di attendibilità.

   > **Important**
   >
   > La policy di affidabilità riportata di seguito include le chiavi di condizione globale `aws:SourceArn` e `aws:SourceAccount` per limitare le autorizzazioni che concedi a EMR Studio a determinate risorse del tuo account. In questo modo puoi proteggerti dal problema del "confused deputy".

   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. Rimuovere le autorizzazioni di ruolo predefinite. Quindi, includere le autorizzazioni dalla seguente policy di autorizzazione IAM di esempio. In alternativa, è possibile creare una policy personalizzata che utilizzi Autorizzazioni del ruolo di servizio EMR Studio.

> **Important**
>
> - Affinché il controllo degli accessi EC2 basato su tag di Amazon funzioni con EMR Studio, devi impostare l'accesso per `ModifyNetworkInterfaceAttribute` l'API come mostrato nella seguente politica.
> - Affinché EMR Studio funzioni con il ruolo di servizio, le istruzioni seguenti devono rimanere invariate: `AllowAddingEMRTagsDuringDefaultSecurityGroupCreation` e `AllowAddingTagsDuringEC2ENICreation`.
> - Per utilizzare la policy di esempio, è necessario taggare le seguenti risorse con la chiave **`"for-use-with-amazon-emr-managed-policies"`** e il valore **`"true"`**.
>   - Amazon Virtual Private Cloud (VPC) designato per EMR Studio.
>   - Ogni sottorete che si desidera utilizzare con Studio.
>   - Qualsiasi gruppo di sicurezza EMR Studio personalizzato. È necessario applicare tag a tutti i gruppi di sicurezza creati durante il periodo di anteprima di EMR Studio se si desidera continuare a utilizzarli.
>   - Segreti AWS Secrets Manager mantenuti dagli utenti di Studio per collegare i repository Git a un Workspace.
>
> Puoi applicare i tag alle risorse utilizzando la scheda Tag nella schermata delle risorse pertinente nella sezione AWS Management Console.

Se applicabile, modifica `*` in `"Resource":"*"` nella policy seguente per specificare il nome della risorsa Amazon (ARN) delle risorse che l'istruzione copre per i tuoi casi d'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. Offri al ruolo di servizio l'accesso in lettura e scrittura alla tua posizione Amazon S3 per EMR Studio. Utilizza il seguente set minimo di autorizzazioni. Per ulteriori informazioni, consulta l'esempio Amazon S3: consente l'accesso in lettura e scrittura agli oggetti di un bucket S3, in modo programmatico e nella console.

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

```
"s3:DeleteObject"
```

Se si crittografa il bucket Amazon S3, devi includere le seguenti autorizzazioni per AWS Key Management Service.

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

4. Se desideri controllare l'accesso ai segreti Git a livello di utente, aggiungi autorizzazioni basate su tag a `secretsmanager:GetSecretValue` nella policy del ruolo utente di EMR Studio e rimuovi le autorizzazioni alla policy `secretsmanager:GetSecretValue` dalla policey del ruolo di servizio di EMR Studio. Per ulteriori informazioni sull'impostazione di autorizzazioni utente granulari, consulta Creazione di policy di autorizzazione per gli utenti di EMR Studio.

Ruolo di servizio minimo per EMR Serverless

Se desideri eseguire carichi di lavoro interattivi con EMR Serverless tramite notebook di EMR Studio, utilizza la stessa policy di attendibilità utilizzata per configurare EMR Studio nella sezione precedente, Come creare un ruolo di servizio per EMR Studio su Amazon o EC2 Amazon EKS.

Per la tua policy IAM, la policy minima valida prevede le seguenti autorizzazioni. Aggiorna *bucket-name* con il nome del bucket che intendi utilizzare quando configuri EMR Studio e WorkSpace. EMR Studio utilizza il bucket per il backup delle istanze WorkSpaces e dei file notebook nel Studio.

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"
      ]
    }
  ]
}
```

Se intendi utilizzare un bucket Amazon S3 crittografato, aggiungi le seguenti autorizzazioni al policy:

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

Autorizzazioni del ruolo di servizio EMR Studio

Nella tabella seguente sono elencate le operazioni eseguite da EMR Studio utilizzando il ruolo di servizio, insieme alle operazioni IAM necessarie per ogni operazione.

Operazione	Azioni

Stabilisci un canale di rete protetto tra un Workspace e un cluster EMR ed esegui le azioni di eliminazione necessarie.	`"ec2:CreateNetworkInterface",`
	`"ec2:CreateNetworkInterfacePermission",`
	`"ec2:DeleteNetworkInterface",`
	`"ec2:DeleteNetworkInterfacePermission",`
	`"ec2:DescribeNetworkInterfaces",`
	`"ec2:ModifyNetworkInterfaceAttribute",`
	`"ec2:AuthorizeSecurityGroupEgress",`
	`"ec2:AuthorizeSecurityGroupIngress",`

Operazione	Azioni
	"ec2:CreateSecurityGroup", "ec2:DescribeSecurityGroups", "ec2:RevokeSecurityGroupEgress", "ec2:DescribeTags", "ec2:DescribeInstances", "ec2:DescribeSubnets", "ec2:DescribeVpcs", "elasticmapreduce:ListInstances", "elasticmapreduce:DescribeCluster", "elasticmapreduce:ListSteps"
Usa le credenziali Git archiviate in AWS Secrets Manager per collegare i repository Git a un Workspace.	"secretsmanager:GetSecretValue"
Applica i AWS tag all'interfaccia di rete e ai gruppi di sicurezza predefiniti creati da EMR Studio durante la configurazione del canale di rete sicuro. Per ulteriori informazioni, consulta Assegnazione di tag alle risorse AWS.	"ec2:CreateTags"
Accedi o carica i file notebook e i metadati in Amazon S3.	"s3:PutObject", "s3:GetObject", "s3:GetEncryptionConfiguration", "s3:ListBucket", "s3:DeleteObject" Se utilizzi un bucket Amazon S3 crittografato, devi includere le seguenti autorizzazioni. "kms:Decrypt", "kms:GenerateDataKey", "kms:ReEncryptFrom", "kms:ReEncryptTo", "kms:DescribeKey"

Operazione	Azioni
Abilita e configura la collaborazione di Workspace.	`"iam:GetUser", "iam:GetRole", "iam:ListUsers", "iam:ListRoles", "sso:GetManagedApplicationInstance", "sso-directory:SearchUsers", "sso:DescribeApplication", "sso:DescribeInstance"`
Crittografa i notebook e i file dell'area di lavoro EMR Studio utilizzando chiavi gestite dal cliente (CMK) con AWS Key Management Service	`"kms:Decrypt", "kms:GenerateDataKey", "kms:ReEncryptFrom", "kms:ReEncryptTo", "kms:DescribeKey"`

Configurazione delle autorizzazioni utente di EMR Studio per Amazon o EC2 Amazon EKS

È necessario configurare le policy di autorizzazione utente per Amazon EMR Studio in modo da poter impostare autorizzazioni granulari per utenti e gruppi. Per informazioni sul funzionamento delle autorizzazioni utente in EMR Studio, consulta Controllo accessi in Come funziona Amazon EMR Studio.

> **Note**
>
> Le autorizzazioni descritte in questa sezione non applicano il controllo dell'accesso ai dati. Per gestire l'accesso ai set di dati di input, è necessario configurare le autorizzazioni per i cluster utilizzati da Studio. Per ulteriori informazioni, consulta Sicurezza in Amazon EMR.

Creazione di un ruolo utente di EMR Studio per la modalità di autenticazione IAM Identity Center

Quando utilizzi la modalità di autenticazione IAM Identity Center per EMR Studio, devi creare un ruolo utente.

Creazione di un ruolo utente per EMR Studio

1. Segui le istruzioni in Creazione di un ruolo per delegare le autorizzazioni a un AWS servizio nella Guida per l'AWS Identity and Access Management utente per creare un ruolo utente.

   Quando crei il ruolo, utilizza la seguente policy di relazione di attendibilità.

   JSON

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

2. Rimuovere le autorizzazioni e le policy di ruolo predefinite.
3. Prima di assegnare utenti e gruppi a uno Studio, collega le policy di sessione di EMR Studio al ruolo utente. Per ricevere istruzioni su come creare policy di sessione, consulta Creazione di policy di autorizzazione per gli utenti di EMR Studio.

Creazione di policy di autorizzazione per gli utenti di EMR Studio

Fai riferimento alle sezioni seguenti per creare policy di autorizzazione per EMR Studio.

Argomenti

- Creazione delle policy di autorizzazione
- Imposta la proprietà per la collaborazione di Workspace
- Creazione di una politcy di segreti Git a livello di utente
- Collegamento della policy di autorizzazione all'identità IAM

> **Note**
>
> Per impostare le autorizzazioni di accesso di Amazon S3 per la memorizzazione dei file notebook e le autorizzazioni di accesso di AWS Secrets Manager per leggere i segreti durante il collegamento dei workspace ai repository Git, utilizza il ruolo di servizio EMR Studio.

Creazione delle policy di autorizzazione

Crea una o più policy di autorizzazione IAM che specificano le operazioni che possono essere compiute da un utente in Studio. Ad esempio, utilizzando le policy di esempio in questa pagina, puoi creare tre policy separate per utenti di Studio base, intermedi e avanzati.

Per una suddivisione di ciascuna operazione di Studio che un utente può eseguire e per il numero minimo di azioni IAM necessarie per eseguire ciascuna operazione, consulta AWS Identity and Access Management autorizzazioni per gli utenti di EMR Studio. Per i passaggi per creare le policy, consulta Creazione di policy IAM nella Guida per l'utente di IAM.

La policy di autorizzazione deve includere le istruzioni seguenti.

```
{
            "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"
}
```

Imposta la proprietà per la collaborazione di Workspace

La collaborazione di Workspace consente a più utenti di lavorare contemporaneamente nello stesso Workspace e può essere configurata con il pannello Collaboration (Collaborazione) nell'interfaccia utente di Workspace. Per visualizzare e utilizzare il pannello Collaboration (Collaborazione), un

utente deve disporre delle seguenti autorizzazioni. Tutti gli utenti con queste autorizzazioni possono visualizzare e utilizzare il pannello Collaboration (Collaborazione).

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

Per limitare l'accesso al pannello Collaboration (Collaborazione), puoi utilizzare il controllo degli accessi basato su tag. Quando un utente crea un Workspace, EMR Studio applica un tag di default con una chiave `creatorUserId` il cui valore è l'ID dell'utente che crea il Workspace.

> **Note**
>
> EMR Studio aggiunge il tag `creatorUserId` ai workspace creati dopo il 16 novembre 2021. Per limitare chi può configurare la collaborazione per i workspace creati prima di questa data, si consiglia di aggiungere manualmente il tag `creatorUserId` al workspace e di utilizzare il controllo degli accessi basato su tag nelle policy di autorizzazione utente.

La seguente istruzione di esempio consente a un utente di configurare la collaborazione per tutti i Workspace con la chiave di tag `creatorUserId` il cui valore corrisponde all'ID dell'utente (indicato dalla variabile `aws:userId` della policy). In altre parole, l'istruzione consente a un utente di configurare la collaborazione per Workspace creati. Per ulteriori informazioni sulle variabili della policy, consulta Elementi delle policy IAM: variabili e tag nella Guida per l'utente di IAM.

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

```
        }
    }
```

Creazione di una politcy di segreti Git a livello di utente

Argomenti

- Per utilizzare le autorizzazioni a livello di utente
- Per passare dalle autorizzazioni a livello di servizio alle autorizzazioni a livello di utente
- Per utilizzare le autorizzazioni a livello di servizio

Per utilizzare le autorizzazioni a livello di utente

EMR Studio aggiunge automaticamente il tag `for-use-with-amazon-emr-managed-user-policies` quando crea segreti Git. Se desideri controllare l'accesso ai segreti Git a livello di utente, aggiungi le autorizzazioni basate su tag alla policy del ruolo utente di EMR Studio con `secretsmanager:GetSecretValue` come mostrato nella sezione Per passare dalle autorizzazioni a livello di servizio alle autorizzazioni a livello di utente seguente.

Se disponi di autorizzazioni esistenti per `secretsmanager:GetSecretValue` nella policy del ruolo di servizio di EMR Studio, devi rimuovere tali autorizzazioni.

Per passare dalle autorizzazioni a livello di servizio alle autorizzazioni a livello di utente

> **Note**
>
> Il tag `for-use-with-amazon-emr-managed-user-policies` assicura che le autorizzazioni del Passaggio 1 di seguito concedano al creatore del workspace l'accesso al segreto Git. Tuttavia, se hai collegato i repository Git prima del 1° settembre 2023, ai segreti Git corrispondenti verrà negato l'accesso perché non hanno il tag `for-use-with-amazon-emr-managed-user-policies` applicato. Per applicare le autorizzazioni a livello utente, devi ricreare i vecchi segreti JupyterLab e collegare nuovamente i repository Git appropriati. Per ulteriori informazioni sulle variabili della policy, consulta Elementi delle policy IAM: variabili e tag nella Guida per l'utente di IAM.

1. Aggiungi le seguenti autorizzazioni alla policy del ruolo utente di EMR Studio. Viene utilizzata la chiave `for-use-with-amazon-emr-managed-user-policies` con il valore `"${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. Se presente, rimuovi la seguente autorizzazione dalla policy del ruolo di servizio di EMR Studio. Poiché la polcy del ruolo di servizio si applica a tutti i segreti definiti da ciascun utente, è necessario eseguire questa operazione una sola volta.

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

Per utilizzare le autorizzazioni a livello di servizio

A partire dal 1° settembre 2023, EMR Studio aggiunge automaticamente il tag `for-use-with-amazon-emr-managed-user-policies` per il controllo degli accessi a livello di utente. Poiché si tratta di una funzionalità aggiuntiva, è possibile continuare a utilizzare l'accesso a livello di servizio disponibile tramite l'autorizzazione `GetSecretValue` nel ruolo di servizio di EMR Studio.

Per i segreti creati prima del 1° settembre 2023, EMR Studio non ha aggiunto il tag `for-use-with-amazon-emr-managed-user-policies`. Per continuare a utilizzare le autorizzazioni a livello

di servizio, è sufficiente mantenere il ruolo di servizio di EMR Studio e le autorizzazioni del ruolo utente esistenti. Tuttavia, per limitare chi può accedere a un singolo segreto, si consiglia di seguire la procedura riportata in Per utilizzare le autorizzazioni a livello di utente per aggiungere manualmente il tag `for-use-with-amazon-emr-managed-user-policies` ai segreti e di utilizzare il controllo degli accessi basato su tag nelle policy di autorizzazione utente.

Per ulteriori informazioni sulle variabili della policy, consulta Elementi delle policy IAM: variabili e tag nella Guida per l'utente di IAM.

Collegamento della policy di autorizzazione all'identità IAM

La tabella seguente riepiloga l'identità IAM a cui si collega una policy di autorizzazione, a seconda della modalità di autenticazione di EMR Studio. Per ricevere istruzioni su come allegare una policy, consulta Aggiunta e rimozione di autorizzazioni per identità IAM.

Se utilizzi...	Collega la policy a…
Autenticazione IAM	Le identità IAM (utenti, gruppi di utenti o ruoli). Ad esempio, prova a collegare una policy di autorizzazione a un utente nel Account AWS.
Federazione IAM con un provider di identità esterno (IdP)	Il ruolo o i ruoli IAM creati per il tuo IdP esterno. Per esempio, una federazione IAM per SAML 2.0. EMR Studio utilizza le autorizzazioni collegate ai ruoli IAM per l'utente o gli utenti con accesso federato a uno Studio.
IAM Identity Center	Il tuo ruolo utente di Amazon EMR Studio.

Esempi di policy utente

La seguente policy utente di base consente la maggior parte delle operazioni di EMR Studio, ma non consente a un utente di creare nuovi cluster Amazon EMR.

Policy di base

> **Important**
>
> La policy di esempio non include l'autorizzazione `CreateStudioPresignedUrl`, che è necessario concedere a un utente quando si utilizza la modalità di autenticazione IAM. Per ulteriori informazioni, consulta Assegnare un utente o un gruppo a un EMR Studio.

La policy di esempio include elementi `Condition` per applicare il controllo degli accessi basato su tag (TBAC) in modo da poter utilizzare la policy con il ruolo di servizio di esempio per EMR Studio. Per ulteriori informazioni, consulta Creazione di un ruolo di servizio 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 seguente policy utente intermedia consente la maggior parte delle operazioni di EMR Studio e consente a un utente di creare nuovi cluster Amazon EMR utilizzando un modello di cluster.

Policy intermedia

> **Important**
>
> La policy di esempio non include l'autorizzazione `CreateStudioPresignedUrl`, che è necessario concedere a un utente quando si utilizza la modalità di autenticazione IAM. Per ulteriori informazioni, consulta Assegnare un utente o un gruppo a un EMR Studio.

La policy di esempio include elementi `Condition` per applicare il controllo degli accessi basato su tag (TBAC) in modo da poter utilizzare la policy con il ruolo di servizio di esempio per EMR Studio. Per ulteriori informazioni, consulta Creazione di un ruolo di servizio 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 seguente policy utente avanzata consente tutte le operazioni di EMR Studio e consente a un utente di creare nuovi cluster Amazon EMR utilizzando un modello di cluster o fornendo una configurazione cluster.

Policy avanzata

> **Important**
>
> La policy di esempio non include l'autorizzazione `CreateStudioPresignedUrl`, che è necessario concedere a un utente quando si utilizza la modalità di autenticazione IAM. Per ulteriori informazioni, consulta Assegnare un utente o un gruppo a un EMR Studio.

La policy di esempio include elementi `Condition` per applicare il controllo degli accessi basato su tag (TBAC) in modo da poter utilizzare la policy con il ruolo di servizio di esempio per EMR Studio. Per ulteriori informazioni, consulta Creazione di un ruolo di servizio 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 seguente policy utente contiene le autorizzazioni utente minime necessarie per utilizzare un'applicazione interattiva EMR Serverless con EMR Studio WorkSpaces.

Policy EMR Serverless interattiva

In questo esempio di policy che prevede le autorizzazioni utente per le applicazioni interattive EMR Serverless con EMR Studio, sostituisci i segnaposto *serverless-runtime-role* per *emr-studio-service-role* e con il ruolo di servizio EMR Studio e il ruolo di runtime EMR Serverless corretti.

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 autorizzazioni per gli utenti di EMR Studio

La tabella seguente include ogni operazione di Amazon EMR Studio che un utente potrebbe eseguire ed elenca le azioni IAM minime necessarie per eseguire tale operazione. Consenti queste operazioni nelle policy di autorizzazione IAM (quando utilizzi l'autenticazione IAM) o nelle policy di sessione del ruolo utente (quando utilizzi l'autenticazione IAM Identity Center) per EMR Studio.

La tabella riporta inoltre le operazioni consentite in ciascuna delle policy di autorizzazione di esempio per EMR Studio. Per ulteriori informazioni su policy di autorizzazione di esempio, consulta Creazione di policy di autorizzazione per gli utenti di EMR Studio.

Azione	Base	Intermedia	Avanzata	Azioni associate
Creazione ed eliminazione di Workspace	Sì	Sì	Sì	`"elasticmapreduce:CreateEditor", "elasticmapreduce:DescribeEditor", "elasticmapreduce:ListEditors", "elasticmapreduce:DeleteEditor"`
Visualizza il pannello Collaborazione, abilita la collaborazione di WorkSpace e aggiungi collaboratori. Per ulteriori informazioni, consulta Imposta la proprietà per la collaborazione di Workspace.	Sì	Sì	Sì	`"elasticmapreduce:UpdateEditor", "elasticmapreduce:PutWorkspaceAccess", "elasticmapreduce:DeleteWorkspaceAccess", "elasticmapreduce:ListWorkspaceAccessIdentities"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Visualizza un elenco di bucket di Amazon S3 Control archiviazione nello stesso account di Studio quando crei un nuovo cluster EMR e accedi ai log dei contenitori quando usi un'interfaccia utente Web per il debug delle applicazioni	Sì	Sì	Sì	`"s3:ListAllMyBuckets", "s3:ListBucket", "s3:GetBucketLocation", "s3:GetObject"`
Accesso ai Workspace	Sì	Sì	Sì	`"elasticmapreduce: DescribeEditor", "elasticmapreduce:ListEditors", "elasticmapreduce:StartEditor", "elasticmapreduce:StopEditor", "elasticmapreduce:OpenEditorInConsole"`
Collegamento o scollegamento di cluster Amazon EMR esistenti associati al Workspace	Sì	Sì	Sì	`"elasticmapreduce: AttachEditor", "elasticmapreduce:DetachEditor", "elasticmapreduce:ListClusters", "elasticmapreduce: DescribeCluster", "elasticmapreduce: ListInstanceGroups", "elasticmapreduce:ListBootstrapActions"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Collegamento o scollegamento di cluster Amazon EMR su EKS	Sì	Sì	Sì	`"elasticmapreduce:AttachEditor", "elasticmapreduce:DetachEditor", "emr-containers:ListVirtualClusters", "emr-containers:DescribeVirtualCluster", "emr-containers:ListManagedEndpoints", "emr-containers:DescribeManagedEndpoint", "emr-containers:GetManagedEndpointSessionCredentials"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Collegamento o scollegamento delle applicazioni EMR Serverless associate al Workspace	No	Sì	Sì	`"elasticmapreduce:AttachEditor", "elasticmapreduce:DetachEditor", "emr-serverless:GetApplication", "emr-serverless:StartApplication", "emr-serverless:ListApplications", "emr-serverless:GetDashboardForJobRun", "emr-serverless:AccessInteractiveEndpoints", "iam:PassRole"` L'autorizzazione `PassRole` è necessaria per passare il ruolo di runtime del processo EMR Serverless. Per ulteriori informazioni, consulta la sezione Ruoli di runtime del processo nella Guida per l'utente di Amazon EMR serverless.

Azione	Base	Intermedia	Avanzata	Azioni associate
Esegui il debug di Amazon EMR EC2 sui job con interfacce utente persistenti delle applicazioni	Sì	Sì	Sì	`"elasticmapreduce:CreatePersistentAppUI", "elasticmapreduce:DescribePersistentAppUI", "elasticmapreduce:GetPersistentAppUIPresignedURL", "elasticmapreduce:ListClusters", "elasticmapreduce:ListSteps", "elasticmapreduce:DescribeCluster", "s3:ListBucket", "s3:GetObject"`
Esegui il debug di Amazon EMR EC2 sui job con interfacce utente di applicazioni on-cluster	Sì	Sì	Sì	`"elasticmapreduce:GetOnClusterAppUIPresignedURL"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Esecuzione del debug delle esecuzioni di processo Amazon EMR su EKS utilizzando Spark History Server	Sì	Sì	Sì	`"elasticmapreduce:CreatePersistentAppUI", "elasticmapreduce:DescribePersistentAppUI", "elasticmapreduce:GetPersistentAppUIPresignedURL", "emr-containers:ListVirtualClusters", "emr-containers:DescribeVirtualCluster", "emr-containers:ListJobRuns", "emr-containers:DescribeJobRun", "s3:ListBucket", "s3:GetObject"`
Creazione ed eliminazione di repository Git	Sì	Sì	Sì	`"elasticmapreduce:CreateRepository", "elasticmapreduce:DeleteRepository", "elasticmapreduce:ListRepositories", "elasticmapreduce:DescribeRepository", "secretsmanager:CreateSecret", "secretsmanager:ListSecrets", "secretsmanager:TagResource"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Collegamento e scollegamento di repository Git	Sì	Sì	Sì	`"elasticmapreduce:LinkRepository", "elasticmapreduce:UnlinkRepository", "elasticmapreduce:ListRepositories", "elasticmapreduce:DescribeRepository"`
Creazione di nuovi cluster da modelli di cluster predefiniti	No	Sì	Sì	`"servicecatalog:SearchProducts", "servicecatalog:DescribeProduct", "servicecatalog:DescribeProductView", "servicecatalog:DescribeProvisioningParameters", "servicecatalog:ProvisionProduct", "servicecatalog:UpdateProvisionedProduct", "servicecatalog:ListProvisioningArtifacts", "servicecatalog:DescribeRecord", "servicecatalog:ListLaunchPaths", "cloudformation:DescribeStackResources", "elasticmapreduce:ListClusters", "elasticmapreduce:DescribeCluster"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Fornisci una configurazione del cluster per creare nuovi cluster.	No	No	Sì	`"elasticmapreduce:RunJobFlow", "iam:PassRole", "elasticmapreduce:ListClusters", "elasticmapreduce:DescribeCluster"`
Assegna un utente a uno Studio quando utilizzi la modalità di autenticazione IAM.	No	No	No	`"elasticmapreduce:CreateStudioPresignedUrl"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Descrivi oggetti di rete.	Sì	Sì	Sì	JSON

```
{
  "Version":
 "2012-10-17",
  "Statement": [
    {
      "Sid": "Describe
Network",
      "Effect":
 "Allow",
      "Action": [
        "ec2:Desc
ribeVpcs",
        "ec2:Desc
ribeSubnets",
        "ec2:Desc
ribeSecurityGroups"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Azione	Base	Intermedia	Avanzata	Azioni associate
Elenca i ruoli IAM.	Sì	Sì	Sì	JSON `{ "Version": "2012-10-17", "Statement": [ { "Sid": "ListIAMRoles", "Effect": "Allow", "Action": [ "iam:ListRoles" ], "Resource": [ "*" ] } ] }`
Connettiti a EMR Studio da Amazon SageMaker AI Studio e usa l'interfaccia visiva Data Wrangler.	No	No	Sì	`"sagemaker:CreatePresignedDomainUrl", "sagemaker:DescribeDomain", "sagemaker:ListDomains", "sagemaker:ListUserProfiles"`
Usa Amazon CodeWhisperer nel tuo EMR Studio.	No	No	Sì	`"codewhisperer:GenerateRecommendations"`

Azione	Base	Intermedia	Avanzata	Azioni associate
Accedi all'editor SQL di Amazon Athena dal tuo EMR Studio. Questo elenco potrebbe non includere tutte le autorizzazioni necessarie per utilizzare tutte le funzionalità di Athena. Per la maggior parte dell' up-to-dateelenco, consulta la politica di accesso completo di Athena.	No	No	Sì	`"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",`

Azione	Base	Intermedia	Avanzata	Azioni associate
				"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"

Creazione di un EMR Studio

Puoi creare un EMR Studio per il tuo team utilizzando la console Amazon EMR o la AWS CLI. La creazione di un'istanza Studio fa parte della configurazione di Amazon EMR Studio.

Prerequisiti

Prima di creare uno Studio, assicurati di aver completato i processi precedenti in Configurazione di un EMR Studio.

Per creare uno Studio utilizzando AWS CLI, è necessario che sia installata la versione più recente. Per ulteriori informazioni, consulta Installare o aggiornare la versione più recente della AWS CLI.

> **Important**
>
> Disattivate gli strumenti di gestione dei proxy come FoxyProxy o SwitchyOmega presenti nel browser prima di creare uno Studio. I proxy attivi possono generare un messaggio di errore Errore di rete quando scegli Crea Studio.

Amazon EMR ti offre una semplice esperienza di console per creare uno Studio, in modo da poter iniziare rapidamente con le impostazioni predefinite, eseguire carichi di lavoro interattivi o lavori in batch con le impostazioni predefinite. La creazione di EMR Studio crea anche un'applicazione EMR Serverless pronta per i lavori interattivi.

Se desideri il pieno controllo sulle impostazioni di Studio, puoi scegliere Personalizzato, che ti consente di configurare tutte le impostazioni aggiuntive.

Interactive workloads

Per creare un EMR Studio per carichi di lavoro interattivi

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. In EMR Studio, nella barra di navigazione a sinistra, scegli Nozioni di base. È inoltre possibile creare un nuovo Studio dalla pagina Studio.
3. Amazon EMR fornisce impostazioni predefinite per te se stai creando un EMR Studio per carichi di lavoro interattivi, ma puoi modificare queste impostazioni. Le impostazioni configurabili includono il nome di EMR Studio, la posizione S3 del Workspace, il ruolo di servizio da utilizzare, gli spazi di lavoro che si desidera utilizzare, il nome dell'applicazione EMR Serverless e il ruolo di runtime associato.

4. Scegli Create Studio e avvia Workspace per terminare e accedere alla pagina Studios. Il nuovo Studio è visibile nell'elenco con dettagli quali Nome Studio, Data di creazione e URL di accesso allo Studio. Il tuo spazio di lavoro si apre in una nuova scheda del browser.

Batch jobs

Per creare un EMR Studio per carichi di lavoro interattivi

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. In EMR Studio, nella barra di navigazione a sinistra, scegli Nozioni di base. È inoltre possibile creare un nuovo Studio dalla pagina Studio.
3. Amazon EMR fornisce impostazioni predefinite per te se stai creando un EMR Studio per lavori in batch, ma puoi modificare queste impostazioni. Le impostazioni configurabili includono il nome di EMR Studio, il nome dell'applicazione EMR Serverless e il ruolo di runtime associato.
4. Scegli Create Studio e avvia Workspace per terminare e accedere alla pagina Studios. Il nuovo Studio è visibile nell'elenco con dettagli quali Nome Studio, Data di creazione e URL di accesso allo Studio. EMR Studio si apre in una nuova scheda del browser.

Custom settings

Per creare un EMR Studio con impostazioni personalizzate

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. In EMR Studio, nella barra di navigazione a sinistra, scegli Nozioni di base. È inoltre possibile creare un nuovo Studio dalla pagina Studio.
3. Seleziona Crea uno Studio per aprire la pagina Crea uno Studio.
4. Inserisci il nome di uno Studio.
5. Scegli di creare un nuovo bucket S3 o utilizzare una posizione esistente.
6. Scegli l'area di lavoro da aggiungere allo Studio. Puoi aggiungere fino a 3 aree di lavoro.
7. In Autenticazione, scegli una modalità di autenticazione per Studio e fornisci informazioni in base alla seguente tabella. Per ulteriori informazioni sull'autenticazione per EMR Studio, consulta Scelta di una modalità di autenticazione per Amazon EMR Studio.

Se utilizzi...	Esegui questa operazione…
Autenticazione o federazione IAM	Il metodo di autenticazione predefinito è AWS Identity and Access Management (IAM). Nella parte inferiore dello schermo, puoi anche aggiungere tag per consentire a utenti specifici di accedere allo Studio, come descritto in Assegnare un utente o un gruppo a un EMR Studio. Se desideri che gli utenti federati accedano utilizzando l'URL di Studio e le credenziali per il tuo provider di identità (IdP), seleziona il tuo IdP dall'elenco a discesa e inserisci l'URL di accesso del provider di identità (IdP) e il nome del parametro. RelayState Per un elenco di RelayState nomi URLs e autenticazioni IdP, consulta. RelayState Parametri e autenticazione del provider di identità URLs

Se utilizzi...	Esegui questa operazione…
Autenticazione IAM Identity Center	Seleziona il Ruolo di servizio e il Ruolo utente di EMR Studio. Per ulteriori informazioni, consultare Creazione di un ruolo di servizio EMR Studio e Creazione di un ruolo utente di EMR Studio per la modalità di autenticazione IAM Identity Center. Quando utilizzi l'autenticazione IAM Identity Center (precedentemente AWS Single Sign On) per lo Studio, puoi scegliere di semplificare l'esperienza di accesso per gli utenti con l'opzione Enable trusted Identity Propagation. Con la propagazione affidabile delle identità, gli utenti possono accedere con le proprie credenziali Identity Center e far sì che le proprie identità vengano propagate ai servizi downstream quando utilizzano Studio. AWS Nella sezione Accesso alle applicazioni, puoi anche specificare se tutti gli utenti e i gruppi del tuo Identity Center devono avere accesso allo Studio o se solo gli utenti e i gruppi assegnati da te possono accedere allo Studio. Per ulteriori informazioniIntegra Amazon EMR con AWS IAM Identity Center, consulta la sezione Trusted Identity Propagation tra le applicazioni nella Guida per l'utente di IAM Identity Center.AWS

8. Per VPC, scegli un Amazon Virtual Private Cloud (VPC) per lo Studio dall'elenco a discesa.
9. In Subnets (Sottoreti) seleziona un massimo di cinque sottoreti in VPC da associare allo Studio. È possibile aggiungere altre sottoreti dopo aver creato lo Studio.

10. Per Security groups (Gruppi di sicurezza), scegli i gruppi di sicurezza di default o i gruppi di sicurezza personalizzati. Per ulteriori informazioni, consulta Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio.

Se scegli...	Esegui questa operazione…
I gruppi di sicurezza EMR Studio predefiniti	Per abilitare il collegamento ai repository basati su Git per Studio, scegli Enable e clusters/endpoints Git repository. In caso contrario, scegli Abilita cluster/endpoint.
Gruppi di sicurezza personalizzati per lo Studio	• In Cluster/endpoint security group (Gruppo di sicurezza cluster/endpoint), seleziona il gruppo di sicurezza del motore configurato dall'elenco a discesa. Lo Studio utilizza questo gruppo di sicurezza per consentire l'accesso in ingresso dai WorkSpace collegati. • In WorkSpace security group (Gruppo di sicurezza dell'istanza WorkSpace), seleziona il gruppo di sicurezza dell'istanza WorkSpace configurata dall'elenco a discesa. Il tuo Studio utilizza questo gruppo di sicurezza con WorkSpace per fornire l'accesso in uscita ai cluster Amazon EMR collegati e ai repository Git ospitati pubblicamente.

11. Aggiungi tag al tuo Studio e ad altre risorse. Per ulteriori informazioni sui tag, consulta Tag clusters.
12. Scegli Create Studio e avvia Workspace per terminare e accedere alla pagina Studios. Il nuovo Studio è visibile nell'elenco con dettagli quali Nome Studio, Data di creazione e URL di accesso allo Studio.

Dopo aver creato uno Studio, segui le istruzioni riportate in Assegnare un utente o un gruppo a un EMR Studio.

CLI

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

Example – Creazione di un EMR Studio che utilizza IAM per l'autenticazione

Il seguente AWS CLI comando di esempio crea un EMR Studio con modalità di autenticazione IAM. Quando utilizzi l'autenticazione o la federazione IAM per lo Studio, non specifichi un `--user-role`.

Per consentire agli utenti federati di accedere utilizzando l'URL di Studio e le credenziali per il tuo gestore dell'identità digitale (IdP), specifica `--idp-auth-url` e `--idp-relay-state-parameter-name`. Per un elenco di RelayState nomi URLs e autenticazioni IdP, consulta. RelayState Parametri e autenticazione del provider di identità 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 – Creazione di un EMR Studio che utilizza Identity Center per l'autenticazione

Il comando di AWS CLI esempio seguente crea uno studio EMR che utilizza la modalità di autenticazione IAM Identity Center. Quando utilizzi l'autenticazione IAM Identity Center, devi specificare un `--user-role`.

Per ulteriori informazioni sulla modalità di autenticazione IAM Identity Center, consulta la sezione Configurazione della modalità di autenticazione IAM Identity Center per 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 – Output della CLI per **`aws emr create-studio`**

Di seguito è riportato un esempio dell'output visualizzato dopo la creazione di uno Studio.

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

Per ulteriori informazioni sul comando `create-studio`, consulta la Guida di riferimento ai comandi della AWS CLI .

RelayState Parametri e autenticazione del provider di identità URLs

Quando utilizzi la federazione IAM e desideri che gli utenti accedano utilizzando l'URL di Studio e le credenziali per il tuo provider di identità (IdP), puoi specificare l'URL di accesso RelayStatee il nome del parametro del tuo provider di identità (IdP) quando lo fai. Creazione di un EMR Studio

La tabella seguente mostra l'URL di autenticazione standard e il nome del RelayState parametro per alcuni provider di identità più diffusi.

Provider di identità	Parametro	Autenticazione URL
Auth0	`RelayState`	`https://<sub_domain> .auth0.com/samlp/<app_id>`

Provider di identità	Parametro	Autenticazione URL
Account 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>`*

Assegnazione e gestione degli utenti di EMR Studio

Dopo aver creato un EMR Studio, è possibile assegnare utenti e gruppi ad esso. Il metodo utilizzato per assegnare, aggiornare e rimuovere utenti dipende dalla modalità di autenticazione di Studio.

- Quando utilizzi la modalità di autenticazione IAM, devi configurare l'assegnazione utente di EMR Studio e le autorizzazioni in IAM o con IAM e il provider di identità.
- Con la modalità di autenticazione IAM Identity Center, utilizzi la console di gestione Amazon EMR o AWS CLI per gestire gli utenti.

Per ulteriori informazioni sull'autenticazione per Amazon EMR Studio, consulta Scelta di una modalità di autenticazione per Amazon EMR Studio.

Assegnare un utente o un gruppo a un EMR Studio

IAM

Quando si utilizza Configurare una modalità di autenticazione IAM per Amazon EMR Studio è necessario consentire l'operazione `CreateStudioPresignedUrl` nella policy di autorizzazione IAM di un utente e limitare l'utente a un particolare Studio. È possibile includere `CreateStudioPresignedUrl` nell'Autorizzazioni utente per la modalità di autenticazione IAM o usare una policy separata.

Per limitare un utente a uno Studio (o a un set di Studio), è possibile utilizzare il controllo di accesso basato sugli attributi (ABAC) o specificare un nome della risorsa Amazon (ARN) di uno Studio nell'elemento `Resource` della policy di autorizzazione dell'utente.

Example Assegnazione di un utente a uno Studio con un ARN di Studio

La seguente policy di esempio fornisce a un utente l'accesso a un particolare EMR Studio consentendo l'operazione `CreateStudioPresignedUrl` e specificando il nome della risorsa Amazon (ARN) di Studio nel 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 Assegnazione di un utente a uno Studio con ABAC per l'autenticazione IAM

Esistono diversi modi per configurare il controllo di accesso basato sugli attributi (ABAC) per uno Studio. Ad esempio, è possibile allegare uno o più tag a un EMR Studio e quindi creare una policy IAM che limita l'operazione `CreateStudioPresignedUrl` per uno Studio particolare o un insieme di Studio con tali tag.

Puoi aggiungere tag durante o dopo la creazione di Studio. Per aggiungere tag a uno Studio esistente, puoi utilizzare il comando AWS CLI `emr add-tags`. L'esempio seguente aggiunge un tag con la coppia chiave-valore `Team = Data Analytics` a un EMR Studio.

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

La seguente policy di autorizzazione di esempio consente l'operazione `CreateStudioPresignedUrl` per EMR Studio con la coppia chiave-valore di tag `Team = DataAnalytics`. Per ulteriori informazioni sull'uso dei tag per controllare l'accesso, consulta Controllo dell'accesso a e per utenti e ruoli IAM mediante i tag o Controllo dell'accesso alle risorse AWS mediante i tag.

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 Assegna un utente a uno Studio utilizzando la chiave aws: SourceIdentity global condition

Quando usi la federazione IAM, puoi utilizzare la chiave di condizione globale `aws:SourceIdentity` in una policy di autorizzazione per consentire agli utenti di accedere a Studio quando assumono il ruolo IAM per la federazione.

È necessario innanzitutto configurare il proprio provider di identità (IdP) per restituire una stringa di identificazione, ad esempio un indirizzo e-mail o un nome utente, quando un utente autentica e assume il ruolo IAM per la federazione. IAM imposta la chiave di condizione globale `aws:SourceIdentity` alla stringa di identificazione restituita dal tuo IdP.

Per maggiori informazioni, consulta il post sul blog How to related IAM role activity to corporate identity nel AWS Security Blog e la voce aws: SourceIdentity entry in the global condition keys reference.

La seguente politica di esempio consente l'`CreateStudioPresignedUrl`azione e offre agli utenti un accesso `aws:SourceIdentity` che corrisponde all'*<example-source-identity>*accesso a EMR Studio specificato da. *<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

Quando si assegna un utente o un gruppo a un EMR Studio, specificare una policy di sessione che definisca autorizzazioni granulari, ad esempio la possibilità di creare un nuovo cluster EMR, per tale utente o gruppo. Amazon EMR archivia le mappature delle policy di sessione. È possibile aggiornare le policy di sessione a un utente o a un gruppo dopo l'assegnazione.

> **Note**
>
> Le autorizzazioni finali per un utente o un gruppo sono un'intersezione tra le autorizzazioni definite nel ruolo utente di EMR Studio e le autorizzazioni definite nella policy di sessione per tale utente o gruppo. Se un utente appartiene a più gruppi assegnati allo Studio, EMR Studio utilizza un'unione di autorizzazioni per quel determinato utente.

Assegnazione di utenti o gruppi a un EMR Studio con la console Amazon EMR

1. Passa alla nuova console Amazon EMR e seleziona Passa alla vecchia console dalla barra di navigazione laterale. Per ulteriori informazioni su cosa aspettarti quando passi alla vecchia console, consulta Utilizzo della vecchia console.
2. Scegli EMR Studio dalla barra di navigazione a sinistra.
3. Scegli il nome del tuo Studio dal menu Studios (Studio) oppure seleziona lo Studio e scegli View details (Visualizza dettagli) per aprire la pagina dei dettagli dello Studio.
4. Seleziona Add Users (Aggiungi utenti) per visualizzare la tabella di ricerca Users (Utenti) e Groups (Gruppi).
5. Seleziona la scheda Users (Utenti) oppure la scheda Groups (Gruppi) e inserisci un termine di ricerca nella barra di ricerca per trovare un utente o un gruppo.
6. Seleziona uno o più utenti o gruppi dall'elenco dei risultati di ricerca. Puoi passare dalla scheda Utenti alla scheda Gruppi e viceversa.
7. Dopo aver selezionato gli utenti e i gruppi da aggiungere allo Studio, scegli Add (Aggiungi). Dovresti visualizzare gli utenti e i gruppi nell'elenco Utenti dello Studio. Potrebbero essere necessari alcuni secondi prima che l'elenco venga aggiornato.

8. Segui le istruzioni in Aggiornamento delle autorizzazioni per un utente o un gruppo assegnato a uno Studio per perfezionare le autorizzazioni Studio per un utente o un gruppo.

Assegnazione di un utente o un gruppo a EMR Studio tramite la AWS CLI

Inserisci valori personalizzati per i seguenti argomenti `create-studio-session-mapping`. Per ulteriori informazioni sul comando `create-studio-session-mapping`, consulta la Guida di riferimento ai comandi della AWS CLI .

- **`--studio-id`**: l'ID dello Studio a cui si desidera assegnare l'utente o il gruppo. Per ricevere istruzioni su come recuperare l'ID dello Studio, consulta Visualizzazione dei dettagli dello Studio.
- **`--identity-name`**: il nome dell'utente o del gruppo dall'archivio identità. Per ulteriori informazioni, consulta UserNamela sezione dedicata agli utenti e DisplayNameai gruppi nell'Identity Store API Reference.
- **`--identity-type`**: utilizza `USER` o `GROUP` per specificare il tipo di identità.
- **`--session-policy-arn`**: il nome della risorsa Amazon (ARN) per la policy di sessione che desideri associare all'utente o al gruppo. Ad esempio, **`arn:aws:iam::`*`<aws-account-id>`*`:policy/`*`EMRStudio_Advanced_User_Policy`**. Per ulteriori informazioni, consulta Creazione di policy di autorizzazione per gli utenti di 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**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per la leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

Utilizza il comando `get-studio-session-mapping` per verificare la nuova assegnazione. Sostituiscilo *<example-identity-name>* con il nome IAM Identity Center dell'utente o del gruppo che hai aggiornato.

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

Aggiornamento delle autorizzazioni per un utente o un gruppo assegnato a uno Studio

IAM

Per aggiornare le autorizzazioni utente o di gruppo quando si utilizza la modalità di autenticazione IAM, utilizzare IAM per modificare le policy di autorizzazione IAM collegate alle identità IAM (utenti, gruppi o ruoli).

Per ulteriori informazioni, consulta Autorizzazioni utente per la modalità di autenticazione IAM.

IAM Identity Center

Aggiornamento delle autorizzazioni di EMR Studio per un utente o un gruppo mediante la console

1. Passa alla nuova console Amazon EMR e seleziona Passa alla vecchia console dalla barra di navigazione laterale. Per ulteriori informazioni su cosa aspettarti quando passi alla vecchia console, consulta Utilizzo della vecchia console.
2. Scegli EMR Studio dalla barra di navigazione a sinistra.
3. Scegli il nome del tuo Studio dal menu Studios (Studio) oppure seleziona lo Studio e scegli View details (Visualizza dettagli) per aprire la pagina dei dettagli dello Studio.
4. Nell'elenco Studio users (Utenti dello Studio) nella pagina dei dettagli dello Studio, cerca l'utente o il gruppo che desideri aggiornare. Puoi eseguire la ricerca per nome o tipo di identità.
5. Seleziona l'utente o il gruppo che desideri aggiornare e scegli Assign policy (Assegna policy) per aprire la finestra di dialogo Session policy (Policy di sessione).
6. Seleziona una policy da applicare all'utente o al gruppo scelto nella fase 5 e scegli Apply policy (Applica policy). L'elenco Studio users (Utenti dello Studio) dovrebbe mostrare il nome della policy nella colonna Session policy (Policy di sessione) per l'utente o il gruppo aggiornato.

Per aggiornare le autorizzazioni di EMR Studio per un utente o un gruppo utilizzando il AWS CLI

Inserisci valori personalizzati per i seguenti argomenti `update-studio-session-mappings`. Per ulteriori informazioni sul comando `update-studio-session-mappings`, consulta la Guida di riferimento ai comandi della 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> \
```

Utilizza il comando `get-studio-session-mapping` per verificare la nuova assegnazione della policy di sessione. Sostituiscilo *<example-identity-name>* con il nome IAM Identity Center dell'utente o del gruppo che hai aggiornato.

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

Rimozione di un utente o un gruppo da uno Studio

IAM

Per rimuovere un utente o un gruppo da EMR Studio quando si utilizza la modalità di autenticazione IAM, è necessario revocare l'accesso dell'utente a Studio riconfigurando la policy di autorizzazione IAM dell'utente.

Nella seguente policy di esempio, supponiamo che tu disponga di un EMR Studio con la coppia chiave-valore di tag `Team = Quality Assurance`. Secondo la policy, l'utente può accedere a Studio taggati con la chiave `Team` il cui valore è uguale a `Data Analytics` o `Quality Assurance`. Per rimuovere l'utente dallo Studio taggato con `Team = Quality Assurance`, rimuovi `Quality Assurance` dall'elenco dei valori dei tag.

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

Rimozione di un utente o un gruppo da un EMR Studio mediante la console

1. Passa alla nuova console Amazon EMR e seleziona Passa alla vecchia console dalla barra di navigazione laterale. Per ulteriori informazioni su cosa aspettarti quando passi alla vecchia console, consulta Utilizzo della vecchia console.
2. Scegli EMR Studio dalla barra di navigazione a sinistra.
3. Scegli il nome del tuo Studio dal menu Studios (Studio) oppure seleziona lo Studio e scegli View details (Visualizza dettagli) per aprire la pagina dei dettagli dello Studio.
4. Nell'elenco Studio users (Utenti dello Studio) nella pagina dei dettagli dello Studio, individua l'utente o il gruppo che desideri rimuovere dallo Studio. Puoi eseguire la ricerca per nome o tipo di identità.
5. Seleziona l'utente o il gruppo che desideri eliminare, quindi scegli Delete (Elimina) e conferma. L'utente o il gruppo eliminato scompare dall'elenco Utenti dello Studio.

Rimozione di un utente o di un gruppo da un EMR Studio utilizzando AWS CLI

Inserisci valori personalizzati per i seguenti argomenti `delete-studio-session-mapping`. Per ulteriori informazioni sul comando `delete-studio-session-mapping`, consulta la Guida di riferimento ai comandi della 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> \
```

Monitora, aggiorna ed elimina le risorse di Amazon EMR Studio

Questa sezione include istruzioni per monitorare, aggiornare o eliminare una risorsa EMR Studio. Per informazioni su come assegnare utenti o aggiornare le autorizzazioni utente, consulta Assegnazione e gestione degli utenti di EMR Studio.

Visualizzazione dei dettagli dello Studio

Console

Visualizzazione dei dettagli di un EMR Studio con la nuova console

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. In EMR Studio, nella barra di navigazione a sinistra, scegli Studio.
3. Seleziona lo Studio dall'elenco Studio per aprire la pagina dei suoi dettagli. La pagina dei dettagli dello Studio include le informazioni di Configurazione dello Studio, quali Descrizione, VPC e Sottoreti dello Studio.

CLI

Per recuperare i dettagli di un EMR Studio by Studio ID utilizzando il AWS CLI

Utilizzare il seguente `describe-studio` AWS CLI comando per recuperare informazioni dettagliate su un particolare EMR Studio. Per ulteriori informazioni, consulta la Guida di riferimento ai comandi della AWS CLI .

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

Recupero di un elenco degli EMR Studio utilizzando AWS CLI

Utilizza il seguente comando `list-studios` AWS CLI . Per ulteriori informazioni, consulta la Guida di riferimento ai comandi della AWS CLI .

```
aws emr list-studios
```

Di seguito è riportato un valore restituito di esempio per il comando `list-studios` in formato JSON.

```
{
    "Studios": [
        {
            "AuthMode": "IAM",
            "VpcId": "vpc-b21XXXXX",
            "Name": "example-studio-name",
            "Url": "https://es-7HWP74SNGDXXXXXXXXXXXXXXX.emrstudio-prod.us-
east-1.amazonaws.com",
            "CreationTime": 1605672582.781,
            "StudioId": "es-7HWP74SNGDXXXXXXXXXXXXXXX",
            "Description": "example studio description"
        }
    ]
}
```

Monitoraggio delle operazioni di Amazon EMR Studio

Visualizzazione dell'attività di EMR Studio e dell'API

EMR Studio è integrato con AWS CloudTrail, un servizio che fornisce un registro delle azioni intraprese da un utente, da un ruolo IAM o da un altro AWS servizio in EMR Studio. CloudTrail acquisisce le chiamate API per EMR Studio come eventi. È possibile visualizzare gli eventi utilizzando la CloudTrail console all'indirizzo. https://console.aws.amazon.com/cloudtrail/

Gli eventi EMR Studio forniscono informazioni come quale utente Studio o IAM effettua una richiesta e il tipo di richiesta.

> **Note**
>
> Le operazioni on-cluster come l'esecuzione di processi notebook non emetto AWS CloudTrail.

Puoi anche creare un percorso per la distribuzione continua degli CloudTrail eventi EMR Studio a un bucket Amazon S3. Per ulteriori informazioni, consulta la Guida per l'utente di AWS CloudTrail.

CloudTrail Evento di esempio: un utente chiama l'API DescribeStudio

Di seguito è riportato un AWS CloudTrail evento di esempio che viene creato quando un utente chiama l'DescribeStudioAPI. `admin` CloudTrail registra il nome utente come`admin`.

> **Note**
>
> Per proteggere i dettagli di Studio, l'evento API EMR Studio per DescribeStudio esclude un valore per. `responseElements`

```
{
   "eventVersion":"1.08",
   "userIdentity":{
      "type":"IAMUser",
      "principalId":"AIDXXXXXXXXXXXXXXXXX",
      "arn":"arn:aws:iam::653XXXXXXXX:user/admin",
      "accountId":"653XXXXXXXX",
      "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-9O5XXXXXXXXXXXXXXXXXXXX"
   },
   "responseElements":null,
   "requestID":"0fxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx",
```

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

Visualizzazione dell'attività degli utenti e dei processi Spark

Per visualizzare l'attività dei processi Spark da parte degli utenti di Amazon EMR Studio, è possibile configurare la rappresentazione utente in un cluster. Con la rappresentazione utente, ogni processo Spark inviato da un Workspace è associato all'utente Studio che ha eseguito il codice.

Quando la rappresentazione utente è abilitata, Amazon EMR crea una directory utente HDFS nel nodo primario del cluster per ogni utente che esegue il codice nel Workspace. Ad esempio, se l'utente `studio-user-1@example.com` esegue il codice, puoi collegarti al nodo primario per riscontrare che `hadoop fs -ls /user` ha una directory per `studio-user-1@example.com`.

Per impostare la rappresentazione utente di Spark, imposta le seguenti proprietà nelle classificazioni di configurazione:

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

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

Per visualizzare le pagine del server della cronologia, consulta Debug di applicazioni e processi con EMR Studio. È inoltre possibile connettersi al nodo primario del cluster utilizzando SSH per visualizzare le interfacce Web dell'applicazione. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Aggiornamento di un Amazon EMR Studio

Dopo aver creato un EMR Studio, puoi aggiornare i seguenti attributi utilizzando la AWS CLI:

- Nome
- Descrizione
- Percorso S3 predefinito
- Sottoreti

Per aggiornare un EMR Studio utilizzando AWS CLI

Usa il `update-studio` AWS CLI comando per aggiornare un EMR Studio. Per ulteriori informazioni, consulta la sezione relativa alle informazioni di riferimento ai comandi della AWS CLI .

> **Note**
>
> Puoi associare uno Studio a un massimo di 5 sottoreti. Queste sottoreti devono appartenere allo stesso VPC dello Studio. L'elenco di IDs sottoreti inviato al `update-studio` comando può includere una nuova sottorete IDs, ma deve includere anche tutta la sottorete IDs già associata a Studio. Non è possibile rimuovere le sottoreti da uno Studio.

```
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> \
```

Per verificare le modifiche, utilizzate il `describe-studio` AWS CLI comando e specificate il vostro ID Studio. Per ulteriori informazioni, consulta la Guida di riferimento ai comandi della AWS CLI .

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

Eliminazione di un Amazon EMR Studio e di Workspace

Quando elimini uno Studio, EMR Studio elimina tutte le assegnazioni di utenti e gruppi IAM Identity Center che sono associate allo Studio.

> **Note**
>
> Quando elimini uno Studio, Amazon EMR non elimina i Workspace a esso associati. È necessario eliminare separatamente i Workspace nello Studio.

Eliminazione dei Workspace

Console

Poiché ogni Workspace EMR Studio è un'istanza di notebook EMR, è possibile utilizzare la console di gestione Amazon EMR per eliminare i Workspace. Puoi eliminare i Workspace utilizzando la console Amazon EMR prima o dopo l'eliminazione di Studio

Eliminazione di un Workspace mediante la console Amazon EMR

1. Passa alla nuova console Amazon EMR e seleziona Passa alla vecchia console dalla barra di navigazione laterale. Per ulteriori informazioni su cosa aspettarti quando passi alla vecchia console, consulta Utilizzo della vecchia console.
2. Seleziona Notebook.
3. Seleziona i Workspace che intendi eliminare.
4. Seleziona Elimina e quindi nuovamente Elimina per confermare.
5. Segui le istruzioni per l'Eliminazione di oggetti nella Guida per l'utente della console Amazon Simple Storage Service per rimuovere i file notebook associati al Workspace eliminato da Amazon S3.

EMR Studio UI

From the Workspace UI

Eliminazione di un Workspace e dei file di backup associati da EMR Studio

1. Accedi al tuo EMR Studio con l'URL di accesso allo Studio e seleziona Workspace dal riquadro di navigazione a sinistra.

2. Individua il Workspace nell'elenco, quindi seleziona la casella di spunta accanto al relativo nome. È possibile selezionare più Workspace da eliminare contemporaneamente.
3. Dall'elenco Workspace, seleziona Elimina in alto a destra per confermare che desideri eliminare i Workspace selezionati. Seleziona Elimina per confermare.
4. Se desideri rimuovere i file notebook associati al WorkSpace eliminato da Amazon S3, segui le istruzioni per l'Eliminazione di oggetti nella Guida per l'utente della console Amazon Simple Storage Service. Se non hai creato lo Studio, contatta l'amministratore dello Studio per determinare la posizione del backup di Amazon S3 per il Workspace eliminato.

From the Workspaces list

Eliminazione di un WorkSpace e dei file di backup associati dall'elenco dei WorkSpace

1. Vai all'elenco dei Workspace nella console.
2. Seleziona il WorkSpace che desideri eliminare dall'elenco, quindi scegli Azioni.
3. Scegli Elimina.
4. Se desideri rimuovere i file notebook associati al WorkSpace eliminato da Amazon S3, segui le istruzioni per l'Eliminazione di oggetti nella Guida per l'utente della console Amazon Simple Storage Service. Se non hai creato lo Studio, contatta l'amministratore dello Studio per determinare la posizione del backup di Amazon S3 per il Workspace eliminato.

Eliminazione di un EMR Studio

Console

Eliminazione di un EMR Studio con la nuova console

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. In EMR Studio, nella barra di navigazione a sinistra, scegli Studios (Studio).
3. Seleziona lo Studio dall'elenco degli Studios (Studio) tramite l'interruttore a sinistra del nome dello Studio. Scegliere Delete (Elimina).

Old console

Eliminazione di un EMR Studio con la vecchia console

1. Apri la console Amazon EMR da casahttps://console.aws.amazon.com/elasticmapreduce/.
2. Seleziona EMR Studio dalla barra di navigazione a sinistra.
3. Seleziona lo Studio dall'elenco Studio e scegli Elimina.

CLI

Per eliminare un EMR Studio con AWS CLI

Usa il `delete-studio` AWS CLI comando per eliminare un EMR Studio. Per ulteriori informazioni, consulta la sezione relativa alle informazioni di riferimento ai comandi di AWS CLI .

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

Crittografia di notebook e file dell'area di lavoro EMR Studio

In EMR Studio, è possibile creare e configurare diverse aree di lavoro per organizzare ed eseguire notebook. Questi spazi di lavoro archiviano notebook e file correlati nel bucket Amazon S3 specificato. Per impostazione predefinita, questi file sono crittografati con chiavi gestite da Amazon S3 (SSE-S3) con crittografia lato server come livello base di crittografia. Puoi anche scegliere di utilizzare chiavi KMS gestite dal cliente (SSE-KMS) per crittografare i tuoi file. Puoi farlo utilizzando la console di gestione Amazon EMR o tramite l' AWS SDK AWS CLI and durante la creazione di un EMR Studio.

La crittografia dello storage dell'area di lavoro di EMR Studio è disponibile in tutte le regioni in cui è disponibile EMR Studio.

Prerequisiti

Prima di poter crittografare gli appunti e i file dell'area di lavoro di EMR Studio, è necessario AWS Key Management Service creare una chiave CMK (Symmetric Customer Manager Key) nella stessa Account AWS regione di EMR Studio.

La politica delle risorse dell'utente AWS KMS deve disporre delle autorizzazioni di accesso necessarie per il ruolo di servizio di EMR Studio. Di seguito è riportato un esempio di policy IAM che concede autorizzazioni di accesso minime per la crittografia dello storage EMR Studio Workspace:

```
{
    "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"
        }
    }
}
```

Il ruolo del servizio EMR Studio deve inoltre disporre delle autorizzazioni di accesso per utilizzare la chiave. AWS KMS Di seguito è riportato un esempio di policy IAM che concede le autorizzazioni di accesso minime per la crittografia dello storage EMR Studio Workspace:

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"
      ]
    }
  ]
}
```

Crea un nuovo EMR Studio

Segui questi passaggi per creare un nuovo EMR Studio che utilizzi la crittografia dello storage del workspace.

1. Apri la console di Amazon EMR all'indirizzo https://console.aws.amazon.com/elasticmapreduce/.
2. Scegli Studios, quindi scegli Create Studio.
3. Per la posizione S3 per lo storage, inserisci o scegli un percorso Amazon S3. Questa è la posizione Amazon S3 in cui Amazon EMR archivia i notebook e i file dell'area di lavoro.
4. Per il ruolo Service, inserisci o scegli un ruolo IAM. Questo è il ruolo IAM che Amazon EMR assume.
5. Scegli Encrypt Workspace files con la tua chiave. AWS KMS
6. Inserisci o scegli una AWS KMS chiave da utilizzare per crittografare i notebook e i file dell'area di lavoro in Amazon S3.
7. Scegli Create Studio o Create Studio e Launch Workspaces.
8. Scegli Encrypt Workspace file con la tua chiave. AWS KMS
9. Inserisci o scegli uno AWS KMS da utilizzare per crittografare i notebook e i file dell'area di lavoro in Amazon S3.
10. Seleziona Salva modifiche.

I passaggi seguenti mostrano come aggiornare un EMR Studio e configurare la crittografia dello storage del workspace.

1. Apri la console di Amazon EMR all'indirizzo https://console.aws.amazon.com/elasticmapreduce/.
2. Scegli un EMR Studio esistente dall'elenco, quindi scegli Modifica.
3. Scegli Encrypt Workspace file con la tua chiave. AWS KMS
4. Inserisci o scegli uno AWS KMS da utilizzare per crittografare i notebook e i file dell'area di lavoro in Amazon S3.

5. Seleziona Salva modifiche.

Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio

Informazioni sui gruppi di sicurezza EMR Studio

Amazon EMR Studio utilizza due gruppi di sicurezza per controllare il traffico di rete tra Workspaces in Studio e un cluster Amazon EMR collegato in esecuzione su Amazon: EC2

- Un gruppo di sicurezza del motore che utilizza la porta 18888 per comunicare con un cluster Amazon EMR collegato in esecuzione su Amazon. EC2
- Un gruppo di sicurezza del Workspace associato ai Workspace in uno Studio. Questo gruppo di sicurezza include una regola HTTPS in uscita per consentire al Workspace di instradare il traffico a Internet e deve consentire il traffico in uscita verso Internet sulla porta 443 per abilitare il collegamento di repository Git a un Workspace.

EMR Studio utilizza questi gruppi di sicurezza oltre a tutti i gruppi di sicurezza associati a un cluster EMR collegato a un Workspace.

È necessario creare questi gruppi di sicurezza quando si utilizza il AWS CLI per creare uno Studio.

> **Note**
>
> È possibile personalizzare i gruppi di sicurezza per EMR Studio con regole personalizzate per l'ambiente, ma è necessario includere le regole riportate in questa pagina. Il gruppo di sicurezza del Workspace non può consentire traffico in entrata e il gruppo di sicurezza del motore deve consentire il traffico in entrata dal gruppo di sicurezza del Workspace.

Uso dei gruppi di sicurezza EMR Studio predefiniti

Quando utilizzi la console Amazon EMR, puoi scegliere i seguenti gruppi di sicurezza predefiniti. I gruppi di sicurezza predefiniti vengono creati da EMR Studio per tuo conto e includono le regole minime in entrata e in uscita richieste per i Workspace in un EMR Studio.

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

Prerequisiti

Per creare i gruppi di sicurezza per EMR Studio, è necessario un cloud privato virtuale (VPC) Amazon per lo Studio. Il VPC viene scelto quando crei i gruppi di sicurezza. Dovrebbe corrispondere al VPC specificato durante la creazione dello Studio. Se intendi utilizzare Amazon EMR su EKS con EMR Studio, scegli il VPC per i nodi worker del cluster Amazon EKS.

Istruzioni

Segui le istruzioni in Creazione di un gruppo di sicurezza nella Amazon EC2 User Guide for Linux Instances per creare un gruppo di sicurezza del motore e un gruppo di sicurezza Workspace nel tuo VPC. I gruppi di sicurezza devono includere le regole riassunte nelle tabelle seguenti.

Quando crei gruppi di sicurezza per EMR Studio, tieni presente la dicitura IDs per entrambi. Quando crei uno Studio, puoi specificare ciascun gruppo di sicurezza in base all'ID.

Gruppo di sicurezza del motore

EMR Studio utilizza la porta 18888 per comunicare con un cluster collegato.

Regole in entrata

Tipo	Protocollo	Porta	Destinazione	Descrizione
TCP	TCP	18888	Il gruppo di sicurezza del Workspace EMR Studio.	Consenti il traffico da qualsiasi risorsa nel gruppo di sicurezza del Workspace per EMR Studio.

Gruppo di sicurezza del Workspace

Questo gruppo di sicurezza è associato ai Workspace in un EMR Studio.

Regole in uscita

Tipo	Protocollo	Porta	Destinazione	Descrizione
TCP	TCP	18888	Il gruppo di sicurezza del	Consenti il traffico verso qualsiasi

Tipo	Protocollo	Porta	Destinazione	Descrizione
			motore EMR Studio.	risorsa nel gruppo di sicurezza del motore per EMR Studio.
HTTPS	TCP	443	0.0.0.0/0	Consenti il traffico verso Internet per collegare repository Git in hosting pubblico ai Workspace.

Crea AWS CloudFormation modelli per Amazon EMR Studio

Informazioni sui modelli di cluster EMR Studio

Puoi creare AWS CloudFormation modelli per aiutare gli utenti di EMR Studio a lanciare nuovi cluster Amazon EMR in un workspace. CloudFormation i modelli sono file di testo formattati in JSON o YAML. In un modello, descrivi una pila di AWS risorse e spieghi CloudFormation come fornirti tali risorse. Per EMR Studio, puoi creare uno o più modelli che descrivono un cluster Amazon EMR.

Organizzi i tuoi modelli in AWS Service Catalog. AWS Service Catalog consente di creare e gestire servizi IT comunemente distribuiti denominati products on AWS. Raccogli i tuoi modelli come prodotti in un portfolio che condividi con gli utenti di EMR Studio. Dopo aver creato modelli di cluster, gli utenti di Studio possono avviare un nuovo cluster per un Workspace con uno dei modelli. Gli utenti devono disporre delle autorizzazioni per creare nuovi cluster dai modelli. Puoi impostare le autorizzazioni utente nel tuo Policy di autorizzazione di EMR Studio.

Per ulteriori informazioni sui CloudFormation modelli, consulta Templates nella Guida per l'AWS CloudFormation utente. Per ulteriori informazioni su AWS Service Catalog, consulta What is AWS Service Catalog.

Il video seguente illustra come configurare modelli di cluster in AWS Service Catalog per EMR Studio. Per ulteriori informazioni, consulta il post del blog Build a self-service environment for each line of business using Amazon EMR and Service Catalog (Creazione di un ambiente self-service per ogni linea di business utilizzando Amazon EMR e Service Catalog).

Parametri di modello opzionali

Puoi includere opzioni aggiuntive nella sezione `Parameters` del modello. Parametri consente agli utenti di Studio di immettere o selezionare valori personalizzati per un cluster. Ad esempio, prova a aggiungere un parametro che consente agli utenti di selezionare una release Amazon EMR specifica. Per ulteriori informazioni, consulta Parametri nella Guida per l'utente di AWS CloudFormation .

La sezione `Parameters` di esempio seguente definisce parametri di input aggiuntivi come il `ClusterName`, versione `EmrRelease` e `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"
```

Quando si aggiungono parametri, gli utenti di Studio visualizzano ulteriori opzioni di modulo dopo aver selezionato un modello di cluster. L'immagine seguente mostra opzioni di modulo aggiuntive per la EmrReleaseversione, ClusterName, e InstanceType.

Prerequisiti

Prima di creare un modello di cluster, assicurati di disporre delle autorizzazioni IAM per accedere alla vista console amministratore di Service Catalog. Sono inoltre necessarie le autorizzazioni IAM richieste per eseguire i processi amministrativi di Service Catalog. Per ulteriori informazioni, consulta Concessione di autorizzazioni ad amministratori Service Catalog.

Creazione di modelli di cluster EMR

Creazione di modelli di cluster EMR con Service Catalog

1. Crea uno o più CloudFormation modelli. Il luogo in cui si archiviano i modelli dipende da te. Poiché i modelli sono file di testo formattati, puoi caricarli su Amazon S3 o conservarli nel file system locale. Per ulteriori informazioni sui CloudFormation modelli, consulta Templates nella Guida AWS CloudFormation per l'utente.

   Utilizza le regole seguenti per assegnare un nome ai modelli o controllare i nomi in base al pattern `[a-zA-Z0-9][a-zA-Z0-9._-]*`.

   - I nomi dei modelli devono iniziare con una lettera o un numero.
   - I nomi dei modelli possono includere solo di lettere, numeri, punti (.), trattini bassi (_) e trattini alti (-).

   Ogni modello del cluster che si crea deve includere le seguenti opzioni:

   Parametri di input

   - ClusterName — Un nome per il cluster per aiutare gli utenti a identificarlo dopo il provisioning.

   Output

   - `ClusterId`: l'ID del cluster EMR di cui è stato appena effettuato il provisioning.

   Di seguito è riportato un AWS CloudFormation modello di esempio in formato YAML per un cluster con due nodi. Il modello di esempio include le opzioni del modello richieste e definisce i parametri di input aggiuntivi per `EmrRelease` e `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 un portfolio per i tuoi modelli di cluster nello stesso AWS account di Studio.

   a. Apri la AWS Service Catalog console all'indirizzo https://console.aws.amazon.com/servicecatalog/.

   b. Dal menu di navigazione a sinistra, scegli Portfolio.

   c. Nella pagina Crea un portfolio, immetti le informazioni richieste.

   d. Scegli Crea. AWS Service Catalog crea il portfolio e visualizza i dettagli del portfolio.

3. Utilizza la procedura seguente per aggiungere i modelli di cluster come prodotti AWS Service Catalog .

   a. Passare alla pagina Prodotti in Amministrazione nella console di gestione AWS Service Catalog

   b. Scegliere Upload new product (Carica un nuovo prodotto).

   c. Inserisci un Product name (Nome del prodotto) e Owner (Proprietario).

   d. Specificare il file modello in Version details (Dettagli della versione).

   e. Scegliere Review (Revisione) per rivedere le impostazioni del prodotto, quindi scegliere Create product (Crea prodotto).

4. Completa la procedura seguente per aggiungere i prodotti al portfolio.

   a. Passare alla pagina Products (Prodotti) della console di gestione AWS Service Catalog

   b. Scegli il tuo prodotto, scegli Actions (Operazioni), quindi scegli Add product to portfolio (Aggiungi prodotto al portfolio).

   c. Scegli un portfolio, quindi scegli Add product to portfolio (Aggiungi prodotto al portfolio).

5. Creare un vincolo di avvio per i prodotti. Un vincolo di avvio è un ruolo IAM che specifica le autorizzazioni utente per l'avvio di un prodotto. Puoi personalizzare i tuoi vincoli di lancio, ma devi consentire le autorizzazioni di utilizzo, CloudFormation Amazon EMR e. AWS Service Catalog Per ulteriori informazioni e istruzioni, consulta Vincoli di avvio di Service Catalog.

6. Applica il vincolo di avvio a ciascun prodotto del tuo portfolio. È necessario applicare il vincolo di avvio a ciascun prodotto singolarmente.

   a. Seleziona il tuo portfolio dal Portfolio della console di gestione AWS Service Catalog .

   b. Scegliere la scheda Vincoli, quindi Crea vincolo.

   c. Scegliere il tuo prodotto e scegliere Avvia in Tipo di vincolo. Scegli Continua.

d. Seleziona il ruolo di vincolo di avvio nella sezione Vincolo di avvio, quindi scegli Crea.

7. Concedere l'accesso al portfolio.

a. Seleziona il tuo portafoglio dal Portfolios (Portafogli) della console di gestione AWS Service Catalog

b. Espandi la scheda Gruppi, ruoli e utenti e scegli Aggiungi gruppi, ruoli, utenti.

c. Cerca il tuo ruolo EMR Studio IAM nella scheda Ruoli, seleziona il tuo ruolo e scegli Aggiungi accesso.

Se utilizzi...	Concedi l'accesso a...
Autenticazione IAM	I tuoi utenti nativi
Federazione IAM	Il tuo ruolo IAM per la federazione
Federazione IAM Identity Center	Il tuo ruolo utente di EMR Studio

Definizione di accesso e autorizzazioni per i repository basati su Git

EMR Studio supporta i seguenti servizi basati su Git:

- AWS CodeCommit
- GitHub
- Bitbucket
- GitLab

Per consentire agli utenti di EMR Studio di associare un repository Git a un Workspace, imposta i seguenti requisiti di accesso e autorizzazione. È inoltre possibile configurare i repository basati su Git ospitati in una rete privata seguendo le istruzioni riportate in Configurazione di un repository Git ospitato privatamente per EMR Studio.

Accesso a Internet del cluster

Sia i cluster Amazon EMR in esecuzione su Amazon che i cluster EC2 Amazon EMR su EKS collegati a Studio Workspaces devono trovarsi in una sottorete privata che utilizza un gateway NAT (Network Address Translation) oppure devono essere in grado di accedere a Internet tramite

un gateway privato virtuale. Per ulteriori informazioni, consulta Opzioni Amazon VPC all'avvio di un cluster.

I gruppi di sicurezza utilizzati con EMR Studio devono includere anche una regola in uscita che consenta ai Workspace di instradare il traffico a Internet da un cluster EMR collegato. Per ulteriori informazioni, consulta Definizione di gruppi di sicurezza per controllare il traffico di rete EMR Studio.

> **Important**
>
> Se l'interfaccia di rete si trova in una sottorete pubblica, non sarà in grado di comunicare con Internet tramite un Gateway Internet (IGW).

Autorizzazioni per AWS Secrets Manager

Per consentire agli utenti di EMR Studio di accedere ai repository Git con segreti archiviati in AWS Secrets Manager, aggiungi una policy di autorizzazione al ruolo di servizio per EMR Studio che consente l'operazione `secretsmanager:GetSecretValue`.

Per ulteriori informazioni su come collegare repository basati su Git ai Workspace, consulta Collegamento di repository basati su Git a un WorkSpace EMR Studio.

Configurazione di un repository Git ospitato privatamente per EMR Studio

Utilizza le seguenti istruzioni per configurare repository ospitati privatamente per Amazon EMR Studio. Fornire un file di configurazione con informazioni sui server DNS e Git. EMR Studio utilizza queste informazioni per configurare Workspace in grado di instradare il traffico ai repository autogestiti.

> **Note**
>
> Se si configura `DnsServerIpV4`, EMR Studio utilizza il tuo server DNS per risolvere entrambi i `GitServerDnsName` e il tuo endpoint Amazon EMR, ad esempio `elasticmapreduce.us-east-1.amazonaws.com`. Per configurare un endpoint per Amazon EMR, connettiti al tuo endpoint tramite il VPC che stai utilizzando con Studio. Ciò garantisce che l'endpoint Amazon EMR viene risolto in un IP privato. Per ulteriori

informazioni, consulta Connessione ad Amazon EMR utilizzando un endpoint VPC di interfaccia.

Prerequisiti

Prima di configurare un repository Git ospitato in livello privato per EMR Studio, è necessario un archivio Amazon S3 in cui EMR Studio possa eseguire il backup di Workspace e file notebook nello Studio. Utilizza lo stesso bucket S3 specificato durante la creazione di uno Studio.

Configurazione di uno o più repository Git ospitati privatamente per EMR Studio

1. Crea un file di configurazione utilizzando il seguente modello. Includi i seguenti valori per ogni server Git che desideri specificare nella configurazione:

   - **`DnsServerIpV4`**- L' IPv4 indirizzo del tuo server DNS. Se si forniscono valori per `DnsServerIpV4` e `GitServerIpV4List`, il valore per `DnsServerIpV4` ha la precedenza e EMR Studio utilizza `DnsServerIpV4` per risolvere il `GitServerDnsName`.

     > **Note**
     >
     > Per utilizzare repository Git ospitati privatamente, il server DNS deve consentire l'accesso in ingresso da EMR Studio. Si consiglia di proteggere il server DNS da altri accessi non autorizzati.

   - **`GitServerDnsName`**: il nome DNS del server Git. Ad esempio `"git.example.com"`.
   - **`GitServerIpV4List`**- Un elenco di IPv4 indirizzi che appartengono ai tuoi server 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. Salva il file di configurazione come `configuration.json`.
3. Carica il file di configurazione nella posizione di storage predefinita di Amazon S3 in una cartella chiamata. `life-cycle-configuration` Ad esempio, se la posizione S3 predefinita è `s3://amzn-s3-demo-bucket/workspace`, il file di configurazione sarà in `s3://amzn-s3-demo-bucket/workspace/`**`life-cycle-configuration/configuration.json`**.

   > **Important**
   >
   > Si consiglia di limitare l'accesso alla tua cartella `life-cycle-configuration` agli amministratori dello Studio e al ruolo di servizio EMR Studio, nonché di proteggere `configuration.json` contro l'accesso non autorizzato. Per ricevere istruzioni, consulta Controllo dell'accesso a un bucket con policy utente o Best practice di sicurezza per Amazon S3.

   Per istruzioni sul caricamento, consulta Creazione di una cartella e Caricamento degli oggetti nella Guida per l'utente di Amazon Simple Storage. Per applicare la configurazione a un Workspace esistente, chiudi e riavvia il Workspace dopo aver caricato il file di configurazione su Amazon S3.

Ottimizzazione dei processi Spark in EMR Studio

Quando si esegue un processo Spark utilizzando EMR Studio, è possibile eseguire alcune fase per garantire l'ottimizzazione delle risorse del cluster Amazon EMR.

Prolunga la tua sessione Livy

Se utilizzi Apache Livy insieme a Spark sul cluster Amazon EMR, ti consigliamo di aumentare il timeout della sessione Livy effettuando una delle seguenti operazioni:

- Quando crei un cluster Amazon EMR, imposta questa classificazione di configurazione nel campo Enter Configuration (Immettere la configurazione).

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

- Per un cluster EMR già in esecuzione, connettiti al cluster utilizzando `ssh` e imposta la classificazione di configurazione `livy-conf` in `/etc/livy/conf/livy.conf`.

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

  Potrebbe essere necessario riavviare Livy dopo aver modificato la configurazione.
- Se non vuoi che la tua sessione di Livy venga scaduta, imposta la proprietà `livy.server.session.timeout-check` a `false` in `/etc/livy/conf/livy.conf`.

Esecuzione di Spark in modalità cluster

In modalità cluster, il driver Spark viene eseguito su un nodo principale anziché sul nodo primario, il che migliora l'utilizzo delle risorse sul nodo primario.

Per eseguire l'applicazione Spark in modalità cluster anziché nella modalità client predefinita, scegli la modalità Cluster quando imposti Modalità di implementazione durante la configurazione della

fase Spark nel tuo nuovo cluster Amazon EMR. Per ulteriori informazioni, consulta Panoramica della modalità cluster nella documentazione di Apache Spark.

Aumento della memoria del driver Spark

Per aumentare la memoria del driver Spark, configura la sessione Spark utilizzando il comando magic `%%configure` nel notebook EMR, come nell'esempio seguente.

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

Utilizzare un Amazon EMR Studio

Questa sezione contiene argomenti che ti aiutano a configurare e interagire con Amazon EMR Studio.

Il video seguente illustra informazioni pratiche come creare un nuovo istanza WorkSpace e come avviare un nuovo cluster Amazon EMR utilizzando un modello di cluster. Il video mostra anche un notebook di esempio.

Questa sezione include i seguenti argomenti per informazioni su come lavorare in un EMR Studio:

- Scopri gli spazi di lavoro di EMR Studio
- Configurazione della collaborazione Workspace in EMR Studio
- Esecuzione di un WorkSpace EMR Studio con un ruolo di runtime
- Esegui i notebook Amazon EMR Studio Workspace Workspace a livello di programmazione
- Sfoglia i dati con SQL Explorer per EMR Studio
- Collegamento di un calcolo a un WorkSpace EMR Studio
- Collegamento di repository basati su Git a un WorkSpace EMR Studio
- Utilizzo dell'editor Amazon Athena SQL in EMR Studio
- CodeWhisperer Integrazione di Amazon con EMR Studio Workspaces
- Debug di applicazioni e processi con EMR Studio
- Installazione di kernel e librerie in un'istanza WorkSpace di EMR Studio
- Migliora i kernel con magic i comandi in EMR Studio
- Usare notebook multilingue con i kernel Spark

Scopri gli spazi di lavoro di EMR Studio

Quando utilizzi un EMR Studio, puoi creare e configurare diversi Workspace per organizzare ed eseguire notebook. Questa sezione descrive la creazione e l'utilizzo delle istanze WorkSpace. Per una panoramica concettuale, consulta WorkSpace nella pagina Come funziona Amazon EMR Studio.

Questa sezione comprende i seguenti argomenti per informazioni su come utilizzare i WorkSpace EMR Studio:

- Creazione di un WorkSpace EMR Studio
- Avvio di un'area di lavoro in EMR Studio
- Comprendi l'interfaccia utente di Workspace in EMR Studio
- Esplora esempi di notebook in uno spazio di lavoro EMR Studio
- Salva i contenuti di Workspace in EMR Studio
- Eliminare un workspace e i file del taccuino in EMR Studio
- Informazioni sullo stato del WorkSpace
- Risolvi i problemi di connettività di WorkSpace

Creazione di un WorkSpace EMR Studio

È possibile creare WorkSpace EMR Studio per eseguire il codice del notebook utilizzando l'interfaccia EMR Studio.

Creazione di un WorkSpace in un EMR Studio

1. Accedi al tuo EMR Studio.
2. Scegli Crea un WorkSpace.
3. Immetti un WorkSpace name (Nome WorkSpace) e una Description (Descrizione). L'assegnazione di un nome a un'istanza WorkSpace consente di identificarlo facilmente nella pagina WorkSpaces (istanze WorkSpace).
4. Se desideri lavorare con altri utenti di Studio in questo WorkSpace in tempo reale, abilita la collaborazione di WorkSpace. Puoi configurare i collaboratori dopo aver avviato il WorkSpace.
5. Se desideri collegare un cluster a un WorkSpace, espandi la sezione Configurazione avanzata. Puoi collegare un cluster in un secondo momento, se preferisci. Per ulteriori informazioni, consulta Collegamento di un calcolo a un WorkSpace EMR Studio.

> **Note**
>
> Per eseguire il provisioning di un nuovo cluster, devi accedere alle autorizzazioni di accesso da parte dell'amministratore.

Scegli una delle opzioni del cluster per l'istanza WorkSpace e collega il cluster. Per ulteriori informazioni sul provisioning di un cluster dopo la creazione di un WorkSpace, consulta Creazione e collegamento di un nuovo cluster EMR a un Workspace EMR Studio.

6. Scegli Crea un WorkSpace in basso a destra nella pagina.

Dopo aver creato l'istanza WorkSpace, EMR Studio aprirà la pagina WorkSpaces (istanze WorkSpace). Verrà visualizzato un banner verde di esito positivo nella parte superiore della pagina e sarà possibile trovare l'istanza WorkSpace appena creato nell'elenco.

Per impostazione predefinita, un WorkSpace è condiviso e può essere visualizzato da tutti gli utenti di Studio. Tuttavia, solo un utente alla volta può aprire e lavorare in un WorkSpace. Lavorare contemporaneamente con altri utenti è possibile con Configurazione della collaborazione Workspace in EMR Studio

Avvio di un'area di lavoro in EMR Studio

Per iniziare a lavorare con i file notebook, avvia un WorkSpace per accedere all'editor del notebook. La pagina WorkSpace di uno Studio elenca tutti i WorkSpace ai quali hai accesso con dettagli quali Name (Nome), Status (Stato), Creation time (Ora di creazione) e Last Modified (Ultima modifica).

> **Note**
>
> Se avevi notebook EMR nella vecchia console Amazon EMR, puoi trovarli nella console come EMR Studio Workspaces. Gli utenti dei notebook EMR necessitano di autorizzazioni aggiuntive per i ruoli IAM per accedere o creare WorkSpace. Se di recente hai creato un notebook nella vecchia console, potrebbe essere necessario aggiornare l'elenco delle aree di lavoro per visualizzarlo nella console. Per ulteriori informazioni sulla transizione, consulta I notebook Amazon EMR sono disponibili come aree di lavoro Amazon EMR Studio nella console. e Gestione dei cluster Amazon EMR con la console

Avvio di un'istanza WorkSpace per la modifica e l'esecuzione di notebook

1. Sulla pagina WorkSpaces (istanze WorkSpace) dello Studio, trova l'istanza WorkSpace. Puoi filtrare l'elenco in base alla parola chiave o al valore della colonna.
2. Scegli il nome dell'istanza WorkSpace per aprirla in una nuova scheda del browser. Potrebbero essere necessari alcuni minuti prima che il WorkSpace venga aperto se è Idle (Inattivo). In alternativa, seleziona la riga per il WorkSpace, quindi seleziona Avvia WorkSpace. Puoi scegliere tra le seguenti opzioni di avvio:

   - Avvio rapido: avvia rapidamente il tuo WorkSpace con le opzioni predefinite. Scegli Avvio rapido se desideri collegare i cluster al Workspace in. JupyterLab
   - Avvia con opzioni: avvia il tuo WorkSpace con opzioni personalizzate. Puoi scegliere di avviarlo in Jupyter oppure JupyterLab collegare il tuo Workspace a un cluster EMR e selezionare i tuoi gruppi di sicurezza.

> **Note**
>
> Solo un utente alla volta può aprire e lavorare in un WorkSpace. Se si seleziona un'istanza WorkSpace già in uso, EMR Studio visualizza una notifica quando si tenta di aprirlo. La colonna User (Utente) sulla pagina WorkSpaces (istanze WorkSpace) mostra l'utente che sta attualmente utilizzando l'istanza WorkSpace.

Comprendi l'interfaccia utente di Workspace in EMR Studio

L'interfaccia utente di EMR Studio Workspace si basa sull'JupyterLabinterfaccia con schede contrassegnate da icone nella barra laterale sinistra. Spostando il mouse su un'icona, è possibile visualizzare un tooltip che mostra il nome della scheda. Seleziona le schede dalla barra laterale sinistra per accedere ai seguenti pannelli.

- Browser di file: visualizza i file e le directory nell'istanza WorkSpace, nonché i file e le directory dei repository Git collegati.
- Kernel e terminali in esecuzione: elenca tutti i kernel e i terminali in esecuzione nell'istanza WorkSpace. Per ulteriori informazioni, consulta la sezione Gestione dei kernel e dei terminali nella documentazione ufficiale. JupyterLab

- Git: fornisce un'interfaccia utente grafica per l'esecuzione di comandi nei repository Git collegati all'istanza WorkSpace. Questo pannello è un' JupyterLab estensione chiamata jupyterlab-git. Per ulteriori informazioni, consulta jupyterlab-git.
- Cluster EMR: consente di collegare o scollegare un cluster dal WorkSpace per eseguire il codice del notebook. Il pannello di configurazione del cluster EMR fornisce inoltre opzioni di configurazione avanzate che consentono di creare e collegare un nuovo cluster nell'istanza WorkSpace. Per ulteriori informazioni, consulta Creazione e collegamento di un nuovo cluster EMR a un Workspace EMR Studio.
- Repository Git Amazon EMR: consente di collegare il WorkSpace a un massimo di tre repository Git. Per dettagli e istruzioni, consulta Collegamento di repository basati su Git a un WorkSpace EMR Studio.
- Notebook di esempio: fornisce un elenco di notebook di esempio che è possibile salvare nell'istanza WorkSpace. Puoi accedere agli esempi anche scegliendo Notebook Examples (Notebook di esempio) sulla pagina Launcher (Utilità di avvio) dell'istanza WorkSpace.
- Comandi: offre un modo basato sulla tastiera per cercare ed eseguire comandi. JupyterLab Per ulteriori informazioni, consulta la pagina della palette Comandi nella documentazione. JupyterLab
- Strumenti notebook: consente di selezionare e impostare opzioni quali il tipo di cella a scorrimento e i metadati. L'opzione Notebook Tools (Strumenti notebook) viene visualizzata nella barra laterale sinistra dopo aver aperto un file notebook.
- Open Tabs (Schede aperte): elenca i documenti e le attività aperti nell'area di lavoro principale in modo da poter passare a una scheda aperta. Per ulteriori informazioni, consultate la pagina Schede e modalità documento singolo nella JupyterLab documentazione.
- Collaboration (Collaborazione): consente di abilitare o disabilitare la collaborazione di WorkSpace e di gestire i collaboratori. Per visualizzare il pannello Collaboration (Collaborazione), devi disporre delle autorizzazioni necessarie. Per ulteriori informazioni, consulta Imposta la proprietà per la collaborazione di WorkSpace.

Esplora esempi di notebook in uno spazio di lavoro EMR Studio

Ogni WorkSpace EMR Studio include una serie di notebook di esempio che è possibile utilizzare per esplorare le caratteristiche di EMR Studio. Per modificare o eseguire un notebook di esempio, è possibile salvarlo nell'istanza WorkSpace.

Salvataggio di un notebook di esempio nell'istanza WorkSpace

1. Nella barra laterale sinistra, scegli l'opzione Notebook Examples (Notebook di esempio) per aprire il riquadro Notebook Examples (Notebook di esempio). Puoi accedere agli esempi anche scegliendo Notebook Examples (Notebook di esempio) sulla pagina Launcher (Utilità di avvio) dell'istanza WorkSpace.
2. Scegli un notebook di esempio per visualizzarlo in anteprima nell'area di lavoro principale. L'esempio è di sola lettura.
3. Per salvare il notebook di esempio nell'istanza WorkSpace, seleziona Save to WorkSpace (Salva nell'istanza WorkSpace). EMR Studio salva l'esempio nella directory principale. Dopo aver salvato un notebook di esempio nell'istanza WorkSpace, è possibile rinominarlo, modificarlo ed eseguirlo.

Per ulteriori informazioni sugli esempi di notebook, vedere l'archivio degli esempi GitHub di notebook EMR Studio.

Salva i contenuti di Workspace in EMR Studio

Quando lavori nell'editor del notebook di un'istanza WorkSpace, EMR Studio salva il contenuto delle celle del notebook e l'output nel percorso Amazon S3 associato al Studio. Questo processo di backup conserva il lavoro tra una sessione e l'altra.

Puoi anche salvare un notebook premendo CTRL+S nella scheda aperta del notebook o utilizzando una delle opzioni di salvataggio in File.

Un altro modo per eseguire il backup dei file notebook nell'istanza WorkSpace consiste nell'associare l'istanza WorkSpace a un repository basato su Git e sincronizzare le modifiche con il repository remoto. In questo modo, è inoltre possibile salvare e condividere i notebook con i membri del team che utilizzano un'istanza WorkSpace o Studio diverso. Per istruzioni, consultare Collegamento di repository basati su Git a un WorkSpace EMR Studio.

Eliminare un workspace e i file del taccuino in EMR Studio

Quando elimini un file notebook da un'istanza WorkSpace di EMR Studio, elimini il file dal menu File browser (Browser di file) ed EMR Studio rimuove la sua copia di backup in Amazon S3. Non è necessario adottare ulteriori misure per evitare addebiti di archiviazione quando si elimina un file da un'istanza WorkSpace.

Quando elimini un intero WorkSpace, i file e le cartelle del notebook rimarranno nella posizione di archiviazione di Amazon S3. I file continuano ad aumentare i costi di archiviazione. Per evitare costi di archiviazione, rimuovi tutti i file e le cartelle di backup associati al WorkSpace eliminato da Amazon S3.

Eliminazione di un file notebook da un WorkSpace EMR Studio

1. Seleziona il riquadro File browser (Browser di file) dalla barra laterale sinistra nell'istanza WorkSpace.
2. Seleziona il file o la cartella da eliminare. Fai clic con il pulsante destro del mouse sulla selezione, quindi scegli Delete (Elimina). Il file scompare dall'elenco. EMR Studio rimuove il file o la cartella da Amazon S3.

From the Workspace UI

Eliminazione di un WorkSpace e dei file di backup associati da EMR Studio

1. Accedi al tuo EMR Studio con l'URL di accesso allo Studio e seleziona Workspace dal riquadro di navigazione a sinistra.
2. Individua il Workspace nell'elenco, quindi seleziona la casella di spunta accanto al relativo nome. È possibile selezionare più Workspace da eliminare contemporaneamente.
3. Dall'elenco Workspace, seleziona Elimina in alto a destra per confermare che desideri eliminare i Workspace selezionati. Seleziona Elimina per confermare.
4. Se desideri rimuovere i file notebook associati al WorkSpace eliminato da Amazon S3, segui le istruzioni per l'Eliminazione di oggetti nella Guida per l'utente della console Amazon Simple Storage Service. Se non hai creato lo Studio, contatta l'amministratore dello Studio per determinare la posizione del backup di Amazon S3 per il Workspace eliminato.

From the Workspaces list

Eliminazione di un WorkSpace e dei file di backup associati dall'elenco dei WorkSpace

1. Vai all'elenco dei Workspace nella console.
2. Seleziona il WorkSpace che desideri eliminare dall'elenco, quindi scegli Azioni.
3. Scegli Elimina.

4. Se desideri rimuovere i file notebook associati al WorkSpace eliminato da Amazon S3, segui le istruzioni per l'Eliminazione di oggetti nella Guida per l'utente della console Amazon Simple Storage Service. Se non hai creato lo Studio, contatta l'amministratore dello Studio per determinare la posizione del backup di Amazon S3 per il Workspace eliminato.

Informazioni sullo stato del WorkSpace

Una volta creato, l'istanza WorkSpace di EMR Studio viene visualizzata come una riga nell'elenco WorkSpaces (isteanze WorkSpace) nel tuo Studio con il nome, lo stato, l'ora di creazione e il timestamp dell'ultima modifica. La tabella seguente descrive i stati della istanza WorkSpace.

Stato	Descrizione
Avvio di	Il WorkSpace è in fase di preparazione, ma non è ancora pronto per l'uso. Non è possibile aprire un WorkSpace quando il relativo stato è Starting (Avvio in corso).
Pronto	È possibile aprire il WorkSpace per utilizzare l'editor del notebook, ma è necessario collegare il WorkSpace a un cluster EMR prima di poter eseguire il codice del notebook.
Collegamento	Il WorkSpace è collegato a un cluster.
Collegato	Il WorkSpace è collegato a un cluster EMR ed è pronto per la scrittura e l'esecuzione del codice del notebook. Se lo stato di un WorkSpace non è Attached (Collegato), è necessario collegarlo a un cluster prima di poter eseguire il codice del notebook.
Idle	Il WorkSpace viene arrestato. Per riattivare un WorkSpace inattivo, occorre selezionarlo dall'elenco WorkSpaces (WorkSpace). Quando si seleziona il WorkSpace, il relativo stato

Stato	Descrizione
	cambia da Idle (Inattivo) a Starting (Avvio in corso) a Ready (Pronto).
Stopping (In arresto)	Il WorkSpace è in fase di arresto e verrà impostato su Inattivo. Quando si arresta un WorkSpace, vengono terminati tutti i kernel dei notebook corrispondenti. EMR Studio arresta i notebook che sono rimasti inattivi per molto tempo.
Deleting (Eliminazione in corso)	Quando si elimina un WorkSpace, EMR Studio lo contrassegna per l'eliminazione e avvia il processo di eliminazione. Al termine del processo di eliminazione, il WorkSpace scompare dall'elenco. Quando elimini un WorkSpace, i file e le cartelle del relativo notebook rimarranno nella posizione di archiviazione di Amazon S3.

Risolvi i problemi di connettività di WorkSpace

Per risolvere i problemi di connettività di WorkSpace, è possibile arrestare e riavviare un WorkSpace. Quando si riavvia un workspace, EMR Studio avvia il WorkSpace in una zona di disponibilità diversa o in una subnet diversa associata a Studio.

Per arrestare e riavviare un EMR Studio WorkSpace

1. Chiudere WorkSpace nel browser.
2. Vai all'elenco dei WorkSpace nella console.
3. Selezionare WorkSpace dall'elenco e scegliere Operazioni.
4. Scegliere Stop e attendere che lo stato del WorkSpace passi da Stopping (In fase di arresto) a Idle (Inattivo).
5. Scegliere di nuovo Actions (Operazioni), quindi scegliere Start (Avvio) per riavviare il WorkSpace.

6. Attendere che lo stato del WorkSpace cambi da Starting (In avvio) a Ready (Pronto), quindi scegliere il nome del WorkSpace per riaprirlo in una nuova scheda del browser.

Configurazione della collaborazione Workspace in EMR Studio

La collaborazione di WorkSpace consente di scrivere ed eseguire il codice del notebook contemporaneamente con altri membri del team. Quando lavori nello stesso file del notebook, vedrai le modifiche man mano che i collaboratori le apportano. Puoi abilitare la collaborazione durante la creazione di un WorkSpace o attivare e disattivare la collaborazione in un WorkSpace esistente.

> **Note**
>
> La collaborazione con i workspace di EMR Studio non è supportata con le applicazioni interattive EMR serverless o se è abilitata la propagazione delle identità attendibili.

Prerequisiti

Prima di configurare la collaborazione per un WorkSpace, assicurati di completare le seguenti attività:

- Assicurati che l'amministratore di EMR Studio ti abbia fornito le autorizzazioni necessarie. Ad esempio, la seguente istruzione di esempio consente a un utente di configurare la collaborazione per tutti i WorkSpace con la chiave di tag `creatorUserId` il cui valore corrisponde all'ID dell'utente (indicato dalla variabile `aws:userId` della policy).

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

```
}
```

- Assicurati che il ruolo di servizio associato a EMR Studio disponga delle autorizzazioni necessarie per abilitare e configurare la collaborazione di WorkSpace, come nella seguente istruzione di esempio.

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

  Per ulteriori informazioni, consulta Creazione di un ruolo di servizio EMR Studio.

Abilitazione della collaborazione di WorkSpace e aggiunta di collaboratori

1. Nel tuo Workspace scegli l'icona Collaboration (Collaborazione) dalla schermata di avvio o nella parte inferiore del pannello a sinistra.

   > **Note**
   >
   > Non visualizzerai il pannello Collaboration (Collaborazione) a meno che l'amministratore di Studio non ti abbia concesso l'autorizzazione per configurare la collaborazione per il WorkSpace. Per ulteriori informazioni, consulta Imposta la proprietà per la collaborazione di WorkSpace.

2. Assicurati che l'interruttore Allow WorkSpace collaboration (Consenti collaborazione di WorkSpace) sia abilitato. Quando si abilita la collaborazione, solo tu e i collaboratori aggiunti potete vedere il WorkSpace nell'elenco nella pagina WorkSpace di Studio.
3. Immetti un Collaborator name (Nome collaboratore). Il WorkSpace può disporre di un massimo di cinque collaboratori incluso te stesso. Un collaboratore può essere qualsiasi utente con

accesso al tuo EMR Studio. Se non accedi a un collaboratore, Workspace è un workspace privato accessibile solo all'utente.

Nella seguente tabella vengono specificati i valori del collaboratore applicabili da inserire in base al tipo di identità del proprietario.

> **Note**
>
> Un proprietario può invitare solo collaboratori con lo stesso tipo di identità. Ad esempio, un utente può aggiungere solo altri utenti e un utente IAM Identity Center può aggiungere solo altri utenti IAM Identity Center.

Modalità di autenticazione	Valore da inserire per Collaborator name (Nome collaboratore)
Autenticazione IAM	un nome utente. Questo è il nome che un utente vede quando effettua l'accesso alla AWS Management Console.
Federazione IAM	Il nome di un ruolo IAM e un nome di sessione facoltativo. Per aggiungere tutti gli utenti federati che devono assumere lo stesso ruolo IAM, specifica il nome di un ruolo IAM per la federazione. Per aggiungere un singolo utente come collaboratore, specifica un ruolo e un nome di sessione. Ad esempio, `MyRoleNam e:MySessionName` .
SSO	Un nome utente IAM Identity Center come `user@example.com`.

4. Scegli Aggiungi. Ora il collaboratore può vedere il WorkSpace sulla propria pagina WorkSpace di EMR Studio e avviarlo per utilizzarlo in tempo reale con te.

> **Note**
>
> Se disabiliti la collaborazione di WorkSpace, il WorkSpace torna allo stato condiviso e può essere visualizzato da tutti gli utenti di Studio. Allo stato condiviso, solo un utente di Studio alla volta può aprire e lavorare in un WorkSpace.

Esecuzione di un WorkSpace EMR Studio con un ruolo di runtime

> **Note**
>
> La funzionalità del ruolo di runtime descritta in questa pagina si applica solo ad Amazon EMR in esecuzione su Amazon EC2 e non si riferisce alla funzionalità del ruolo di runtime nelle applicazioni interattive EMR Serverless. Per ulteriori informazioni su come utilizzare i ruoli di runtime in EMR Serverless, consulta Ruoli di runtime del processo nella Guida per l'utente di Amazon EMR serverless.

Un ruolo di runtime è un ruolo AWS Identity and Access Management (IAM) che puoi specificare quando invii un lavoro o una query a un cluster Amazon EMR. Il job o la query che invii al tuo cluster EMR utilizza il ruolo di runtime per accedere alle AWS risorse, come gli oggetti in Amazon S3.

Quando colleghi un EMR Studio Workspace a un cluster EMR che utilizza Amazon EMR 6.11 o versioni successive, puoi selezionare un ruolo di runtime per il job o la query che invii da utilizzare quando accede alle risorse. AWS Tuttavia, se il cluster EMR non supporta i ruoli di runtime, il cluster EMR non assumerà il ruolo quando accede alle risorse. AWS

Per poter utilizzare un ruolo di runtime con un WorkSpace Amazon EMR Studio, un amministratore deve configurare le autorizzazioni utente in modo che l'utente Studio possa chiamare l'API `elasticmapreduce:GetClusterSessionCredentials` sul ruolo di runtime. Quindi, avvia un nuovo cluster con un ruolo di runtime che puoi utilizzare con il WorkSpace Amazon EMR Studio.

In questa pagina

- Configurazione delle autorizzazioni utente per il ruolo di runtime
- Avvio di un nuovo cluster con un ruolo di runtime
- Utilizzo del cluster EMR con un ruolo di runtime nei WorkSpace
- Considerazioni

Configurazione delle autorizzazioni utente per il ruolo di runtime

Configura le autorizzazioni utente in modo che l'utente Studio possa chiamare l'API `elasticmapreduce:GetClusterSessionCredentials` sul ruolo di runtime che l'utente desidera utilizzare. Devi inoltre configurare the section called “Autorizzazioni utente Studio (, EKS) EC2” prima che l'utente possa iniziare a utilizzare Studio.

> **Warning**
>
> Per concedere questa autorizzazione, crea una condizione basata sulla chiave di `elasticmapreduce:ExecutionRoleArn` contesto quando concedi a un chiamante l'accesso per chiamare il. `GetClusterSessionCredentials` APIs L'esempio seguente mostra come fare.

```
{
      "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"
              ]
          }
      }
  }
```

L'esempio seguente mostra come consentire a un principale IAM di utilizzare un ruolo IAM denominato `test-emr-demo3` come ruolo di runtime. Inoltre, il titolare della policy potrà accedere ai cluster Amazon EMR solo con l'ID cluster `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"
            ]
        }
    }
}
```

L'esempio seguente consente a un principale IAM di utilizzare qualsiasi ruolo IAM con un nome che inizia con la stringa `test-emr-demo4` come ruolo di runtime. Inoltre, il titolare della policy potrà accedere solo ai cluster Amazon EMR contrassegnati con la coppia chiave-valore `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*"
            ]
        }
    }
}
```

Avvio di un nuovo cluster con un ruolo di runtime

Ora che disponi delle autorizzazioni necessarie, avvia un nuovo cluster con un ruolo di runtime da utilizzare con il tuo WorkSpace Amazon EMR Studio.

Se hai già avviato un nuovo cluster con un ruolo di runtime, puoi passare alla sezione the section called “Utilizzo del cluster con il WorkSpace”.

1. Innanzitutto, completa i prerequisiti nella sezione Ruoli di runtime per le fasi di Amazon EMR.
2. Quindi, avvia un cluster con le seguenti impostazioni per utilizzare i ruoli di runtime con i WorkSpace Amazon EMR Studio. Per istruzioni sull'avvio del cluster, consulta la sezione Specificare una configurazione di sicurezza per un cluster Amazon EMR.
   - Scegli l'etichetta di release emr-6.11.0 o versioni successive.
   - Seleziona Spark, Livy e Jupyter Enterprise Gateway come applicazioni cluster.
   - Utilizza la configurazione di sicurezza creata nella fase precedente.
   - Facoltativamente, puoi abilitare Lake Formation per il cluster EMR. Per ulteriori informazioni, consulta Abilitazione di Amazon EMR con Lake Formation.

Dopo aver avviato il cluster, sei pronto per utilizzare il cluster con ruolo di runtime abilitato con un WorkSpace EMR Studio.

> **Note**
>
> Il ExecutionRoleArnvalore non è attualmente supportato dall'operazione StartNotebookExecutionAPI quando il `ExecutionEngineConfig.Type` valore è`EMR`.

Utilizzo del cluster EMR con un ruolo di runtime nei WorkSpace

Dopo aver configurato e avviato il cluster, puoi utilizzare il cluster abilitato al ruolo di runtime con il WorkSpace di EMR Studio.

1. Crea un nuovo workspace o avvia un workspace esistente. Per ulteriori informazioni, consulta Creazione di un WorkSpace EMR Studio.
2. Scegli la scheda Cluster EMR nella barra laterale sinistra del tuo Workspace aperto, espandi la sezione Tipo di calcolo e scegli il cluster dal menu Cluster EMR e il ruolo di runtime dal EC2 menu Ruolo Runtime.
3. Scegli Collega per collegare il cluster con ruolo di runtime al tuo WorkSpace.

> **Note**
>
> Quando scegli un ruolo di runtime, tieni presente che a esso possono essere associate politiche gestite sottostanti. Nella maggior parte dei casi consigliamo di scegliere risorse limitate, come notebook specifici. Se si sceglie un ruolo di runtime che include l'accesso a tutti i notebook, ad esempio, la policy gestita associata al ruolo fornisce l'accesso completo.

Considerazioni

Quando utilizzi un cluster abilitato al ruolo di runtime con il WorkSpace Amazon EMR Studio:

- Puoi selezionare un ruolo di runtime solo quando colleghi un WorkSpace EMR Studio a un cluster EMR che utilizza Amazon EMR rilascio 6.11 o successivi.
- La funzionalità del ruolo di runtime descritta in questa pagina è supportata solo con Amazon EMR in esecuzione su Amazon EC2 e non è supportata con le applicazioni interattive EMR Serverless. Per ulteriori informazioni sui ruoli di runtime per EMR Serverless, consulta Ruoli di runtime del processo nella Guida per l'utente di Amazon EMR serverless.
- Sebbene sia necessario configurare autorizzazioni aggiuntive per poter specificare un ruolo di runtime quando si invia un processo a un cluster, non sono necessarie autorizzazioni aggiuntive per accedere ai file generati da un WorkSpace EMR Studio. Le autorizzazioni per tali file sono le stesse dei file generati da cluster senza ruoli di runtime.
- Non è possibile utilizzare SQL Explorer in un WorkSpace EMR Studio con un cluster con un ruolo di runtime. Amazon EMR disabilita SQL Explorer nell'interfaccia utente quando un WorkSpace è collegato a un cluster EMR abilitato al ruolo di runtime.
- Non è possibile utilizzare la modalità di collaborazione in un WorkSpace EMR Studio con un cluster con un ruolo di runtime. Amazon EMR disabilita le funzionalità di collaborazione di un WorkSpace quando un WorkSpace è collegato a un cluster EMR abilitato al ruolo di runtime. Il WorkSpace rimarrà accessibile solo all'utente che lo ha collegato.
- Non è possibile utilizzare ruoli di runtime in uno Studio con la propagazione delle identità attendibili di IAM Identity Center abilitata.
- Potresti ricevere un avviso "La pagina potrebbe non essere sicura!" dall'interfaccia utente di Spark per un cluster con ruoli di runtime che utilizza Amazon EMR versione 7.4.0 e precedenti. Se ciò accade, ignora l'avviso per continuare a visualizzare l'interfaccia utente di Spark.

Esegui i notebook Amazon EMR Studio Workspace Workspace a livello di programmazione

> **Note**
>
> L'esecuzione programmatica dei notebook non è supportata con le applicazioni interattive di Amazon EMR serverless.

Puoi eseguire i tuoi notebook Amazon EMR Studio Workspace a livello di programmazione con uno script o sul AWS CLI. Per informazioni su come eseguire il notebook in modo programmatico, consulta Esempi di comandi programmatici per Notebooks EMR.

Sfoglia i dati con SQL Explorer per EMR Studio

> **Note**
>
> SQL Explorer per EMR Studio non è supportato con le applicazioni interattive Amazon EMR serverless o in uno Studio con la propagazione delle identità attendibili di IAM Identity Center abilitata.

In questo argomento vengono fornite informazioni utili per iniziare a utilizzare SQL Explorer in Amazon EMR Studio. SQL Explorer è uno strumento a pagina singola nel WorkSpace che aiuta a comprendere le origini dati nel catalogo dati del cluster EMR. Puoi utilizzare SQL Explorer per sfogliare i dati, eseguire query SQL per recuperare dati e scaricare i risultati delle query.

SQL Explorer supporta Presto. Prima di utilizzare SQL Explorer, assicurati di disporre di un cluster EMR che utilizzi Amazon EMR versione 5.34.0 o successive oppure versione 6.4.0 o successive con Presto installato. Amazon EMR Studio SQL Explorer non supporta i cluster Presto configurati con la crittografia in transito. Questo perché su questi cluster Presto viene eseguito in modalità TLS.

Sfogliare il catalogo dati del cluster

SQL Explorer fornisce un'interfaccia del browser di catalogo che è possibile utilizzare per esplorare e capire come sono organizzati i dati. Ad esempio, è possibile utilizzare il browser del catalogo dati per verificare i nomi di tabelle e colonne prima di scrivere una query SQL.

Sfogliare il catalogo dati

1. Apri SQL Explorer nel tuo WorkSpace.
2. Assicurati che il tuo spazio di lavoro sia collegato a un cluster EMR in esecuzione che utilizza EC2 Amazon EMR versione 6.4.0 o successiva con Presto installato. Puoi selezionare un cluster esistente o crearne uno nuovo. Per ulteriori informazioni, consulta Collegamento di un calcolo a un WorkSpace EMR Studio.
3. Seleziona un Database dall'elenco a discesa per iniziare a sfogliarlo.
4. Espandi una tabella nel database per visualizzare i nomi delle colonne della tabella. Puoi anche immettere una parola chiave nella barra di ricerca per filtrare i risultati delle tabelle.

Esecuzione di una query SQL per recuperare i dati

Recupero di dati con una query SQL e download dei risultati

1. Apri SQL Explorer nel tuo WorkSpace.
2. Assicurati che il tuo spazio di lavoro sia collegato a un cluster EMR in esecuzione con Presto e Spark EC2 installati. Puoi selezionare un cluster esistente o crearne uno nuovo. Per ulteriori informazioni, consulta Collegamento di un calcolo a un WorkSpace EMR Studio.
3. Seleziona Open editor (Apri editor) per aprire una nuova scheda dell'editor nel WorkSpace.
4. Componi la query SQL nella scheda dell'editor.
5. Seleziona Esegui.
6. Visualizza i risultati della query in Result preview (Anteprima dei risultati). SQL Explorer visualizza i primi 100 risultati per impostazione predefinita. Puoi scegliere un numero diverso di risultati da visualizzare (fino a 1000) utilizzando il menu a discesa Preview first 100 query results (Anteprima dei primi 100 risultati della query).
7. Scegli Download results (Scarica risultati) per scaricare i risultati in formato CSV. È possibile scaricare fino a 1000 righe di risultati.

Collegamento di un calcolo a un WorkSpace EMR Studio

Amazon EMR Studio esegue comandi notebook utilizzando un kernel su un cluster EMR. Prima di poter selezionare un kernel, devi collegare Workspace a un cluster che utilizza EC2 istanze Amazon, a un cluster Amazon EMR su EKS o a un'applicazione EMR Serverless. EMR Studio consente di

collegare WorkSpace a cluster nuovi o esistenti e offre la flessibilità necessaria per modificare i cluster senza chiudere il WorkSpace.

Questa sezione comprende i seguenti argomenti per informazioni su come lavorare con i cluster ed effettuare il relativo provisioning per EMR Studio:

- Collega un EC2 cluster Amazon a un EMR Studio Workspace
- Collegamento di un cluster Amazon EMR su EKS a un WorkSpace EMR Studio
- Collegamento di un'applicazione Amazon EMR serverless a un WorkSpace EMR Studio
- Creazione e collegamento di un nuovo cluster EMR a un Workspace EMR Studio
- Scollegamento di un calcolo da un WorkSpace EMR Studio

Collega un EC2 cluster Amazon a un EMR Studio Workspace

Puoi collegare un cluster EMR in esecuzione su Amazon EC2 a un Workspace quando crei il Workspace o collegare un cluster a un Workspace esistente. Se desideri creare e collegare un nuovo cluster, consulta Creazione e collegamento di un nuovo cluster EMR a un Workspace EMR Studio.

> **Note**
>
> Un workspace in uno Studio in cui è abilitata la propagazione delle identità attendibili di IAM Identity Center può collegarsi solo a un cluster EMR con una configurazione di sicurezza per cui è abilitato Identity Center.

On create

Collegamento di un cluster di calcolo Amazon EMR quando si crea un WorkSpace

1. Nella finestra di dialogo Create a WorkSpace (Crea un WorkSpace), verifica di aver già selezionato una sottorete per la nuova istanza WorkSpace. Espandi la sezione Advanced configuration (Configurazione avanzata).
2. Scegli Attach WorkSpace to an EMR cluster (Collega WorkSpace a un cluster EMR).
3. Nell'elenco a discesa Cluster EMR, seleziona un cluster EMR esistente per collegarlo al WorkSpace.

Dopo aver collegato un cluster, completa il processo creando il WorkSpace. Quando apri il nuovo WorkSpace per la prima volta e scegli il riquadro Cluster EMR, dovresti visualizzare il cluster selezionato collegato.

On launch

Collegamento di un cluster di calcolo Amazon EMR quando si avvia il WorkSpace

1. Vai all'elenco dei WorkSpace e seleziona la riga relativa al WorkSpace che desideri avviare. Quindi, seleziona Avvia Workspace > Avvia con opzioni.
2. Scegli un cluster EMR da collegare al tuo WorkSpace.

Dopo aver collegato un cluster, completa il processo creando il WorkSpace. Quando apri il nuovo WorkSpace per la prima volta e scegli il riquadro Cluster EMR, dovresti visualizzare il cluster selezionato collegato.

In JupyterLab

Collega un workspace a un cluster di calcolo Amazon EMR in JupyterLab

1. Seleziona il tuo WorkSpace, quindi seleziona Avvia Workspace > Avvio rapido.
2. All'interno JupyterLab, apri la scheda Cluster nella barra laterale sinistra.
3. Seleziona il menu a discesa EMR su EC2 cluster o seleziona un cluster Amazon EMR su EKS.
4. Seleziona Collega per collegare il cluster al tuo WorkSpace.

Dopo aver collegato un cluster, completa il processo creando il WorkSpace. Quando apri il nuovo WorkSpace per la prima volta e scegli il riquadro Cluster EMR, dovresti visualizzare il cluster selezionato collegato.

In the Workspace UI

Collegamento di un WorkSpace a un cluster di calcolo Amazon EMR dall'interfaccia utente del WorkSpace

1. Nel WorkSpace che desideri collegare a un cluster, scegli l'icona Cluster EMR dalla barra laterale sinistra per aprire il riquadro Cluster.
2. In Tipo di cluster, espandi il menu a discesa e seleziona Cluster EMR attivo. EC2

3. Scegli un cluster dall'elenco a discesa. Potrebbe essere necessario scollegare prima un cluster esistente per abilitare l'elenco a discesa di selezione del cluster.
4. Scegli Collega. Quando il cluster è collegato, viene visualizzato un messaggio di esito positivo.

Collegamento di un cluster Amazon EMR su EKS a un WorkSpace EMR Studio

Oltre a utilizzare i cluster Amazon EMR in esecuzione su Amazon EC2, puoi collegare un Workspace a un cluster Amazon EMR su EKS per eseguire il codice dei notebook. Per ulteriori informazioni su Amazon EMR su EKS, consulta Che cos'è Amazon EMR su EKS.

Per poter connettere un WorkSpace a un cluster Amazon EMR su EKS, l'amministratore di Studio deve concedere le autorizzazioni di accesso.

> **Note**
>
> Non è possibile avviare un cluster Amazon EMR su EKS in un EMR Studio che utilizza la propagazione delle identità attendibili di IAM Identity Center.

On create

Per collegare un cluster Amazon EMR su EKS quando si crea un WorkSpace

1. Nella finestra di dialogo Create a Workspace (Creazione di un WorkSpace), espandere la sezione Advanced configuration (Configurazione avanzata).
2. Scegli Collega il WorkSpace a un cluster Amazon EMR su EKS.
3. In Cluster Amazon EMR su EKS, scegli un cluster dall'elenco a discesa.
4. In Select an endpoint (Seleziona un endpoint), scegli un endpoint gestito da collegare all'istanza WorkSpace. Un endpoint gestito è un gateway che consente a EMR Studio di comunicare con il cluster scelto.
5. Scegli Crea un WorkSpace per completare il processo di creazione del WorkSpace e collegare il cluster selezionato.

Dopo aver collegato un cluster, è possibile completare il processo di creazione del WorkSpace. Quando apri il nuovo WorkSpace per la prima volta e selezioni il riquadro Cluster EMR, dovresti visualizzare il cluster selezionato collegato.

In the Workspace UI

Per collegare un cluster Amazon EMR su EKS dall'interfaccia utente del WorkSpace

1. Nel WorkSpace che desideri collegare a un cluster, scegli l'icona Cluster EMR dalla barra laterale sinistra per aprire il riquadro Cluster.
2. Espandi il menu a discesa Tipo di cluster e scegli Cluster EMR su EKS.
3. In Cluster EMR su EKS, scegli un cluster dall'elenco a discesa.
4. In Endpoint, scegli un endpoint gestito da collegare all'istanza WorkSpace. Un endpoint gestito è un gateway che consente a EMR Studio di comunicare con il cluster scelto.
5. Scegli Collega. Quando il cluster è collegato, viene visualizzato un messaggio di esito positivo.

Collegamento di un'applicazione Amazon EMR serverless a un WorkSpace EMR Studio

È possibile collegare un WorkSpace a un'applicazione EMR Serverless per eseguire carichi di lavoro interattivi. Per ulteriori informazioni, consulta Utilizzo dei notebook per eseguire carichi di lavoro interattivi con EMR Serverless tramite EMR Studio.

> **Note**
>
> Non è possibile collegare un'applicazione EMR serverless a un EMR Studio che utilizza la propagazione delle identità attendibili di IAM Identity Center.

Example Collega un workspace a un'applicazione EMR Serverless in JupyterLab

Prima di poter connettere un'istanza WorkSpace a un'applicazione Serverless, l'amministratore dell'account deve concedere le autorizzazioni di accesso come descritto in Autorizzazioni richieste per i carichi di lavoro interattivi.

1. Vai su EMR Studio e seleziona il tuo WorkSpace, quindi seleziona Avvia Workspace > Avvio rapido.
2. All'interno JupyterLab, apri la scheda Cluster nella barra laterale sinistra.
3. Seleziona EMR Serverless come opzione di calcolo, quindi seleziona un'applicazione EMR Serverless e un ruolo runtime.

4. Per collegare il cluster al tuo WorkSpace, scegli Collega.

Ora, quando apri questo WorkSpace, dovresti vedere l'applicazione selezionata collegata.

Creazione e collegamento di un nuovo cluster EMR a un Workspace EMR Studio

Gli utenti avanzati di EMR Studio possono effettuare il provisioning di nuovi cluster EMR in esecuzione su EC2 Amazon da utilizzare con un Workspace. Il nuovo cluster dispone di tutte le applicazioni Big Data necessarie per EMR Studio installate per impostazione predefinita.

Per creare cluster, l'amministratore dello Studio deve prima concedere l'autorizzazione a utilizzare una policy di sessione. Per ulteriori informazioni, consulta Creazione di policy di autorizzazione per gli utenti di EMR Studio.

È possibile creare un nuovo cluster nella finestra di dialogo Create a WorkSpace (Crea un WorkSpace) o nel riquadro Cluster dell'interfaccia utente di WorkSpace. In entrambi i casi, sono disponibili due opzioni di creazione del cluster:

1. Crea un cluster EMR: crea un cluster EMR scegliendo il tipo e il numero di EC2 istanze Amazon.
2. Use a cluster template (Utilizza un modello di cluster): provisioning rapido di un cluster selezionando un modello di cluster predefinito. Questa opzione è disponibile se si dispone dell'autorizzazione per utilizzare i modelli di cluster.

> **Note**
>
> Se hai abilitato la propagazione delle identità attendibili con IAM Identity Center per il tuo Studio, devi utilizzare un modello per creare un cluster.

Creazione di un cluster EMR fornendo una configurazione del cluster

1. Scegli un orario di inizio.

A…	Esegui questa operazione…
Crea il cluster durante la creazione di un WorkSpace tramite la finestra di dialogo Create a WorkSpace (Crea un WorkSpace).	Espandi la sezione Advanced configuration (Configurazione avanzata) nella finestra di dialogo Create a WorkSpace (Crea un

A…	Esegui questa operazione…
	WorkSpace) e seleziona Create an EMR cluster (Crea un cluster EMR).
Crea il cluster dal riquadro Cluster EMR nell'interfaccia utente del WorkSpace dopo aver creato un WorkSpace.	Scegli la scheda Cluster EMR nella barra laterale sinistra di un WorkSpace aperto, espandi la sezione Configurazione avanzata e scegli Crea cluster.

2. Immetti un Cluster name (Nome cluster). La denominazione del cluster consente di individuarlo successivamente nell'elenco Clusters (Cluster) di EMR Studio.
3. Per Rilascio di Amazon EMR, scegli una versione di rilascio di Amazon EMR per il cluster.
4. Ad esempio, seleziona il tipo e il numero di EC2 istanze Amazon per il cluster. Per ulteriori informazioni sulla selezione dei tipi di istanza, consulta Configurazione dei tipi di EC2 istanze Amazon da utilizzare con Amazon EMR. Un'istanza sarà utilizzata come nodo primario.
5. Selezionare una Subnet (Sottorete) dove EMR Studio può lanciare il nuovo cluster. Ogni opzione di sottorete è pre-approvata dall'amministratore di Studio, pertanto WorkSpace dovrebbe essere in grado di connettersi a un cluster in qualsiasi sottorete elencata.
6. Scegli un S3 URI for log storage (URI S3 per l'archiviazione dei log).
7. Scegli Create EMR cluster (Crea cluster EMR) per eseguire il provisioning del cluster. Se utilizzi la finestra di dialogo Crea un WorkSpace, scegli Crea un WorkSpace per creare il WorkSpace ed eseguire il provisioning del cluster. Dopo che EMR Studio esegue il provisioning del nuovo cluster, il cluster viene collegato automaticamente all'istanza WorkSpace.

Creazione di un cluster tramite un modello di cluster

1. Scegli un orario di inizio.

A…	Esegui questa operazione…
Crea il cluster durante la creazione di un WorkSpace tramite la finestra di dialogo Create a WorkSpace (Crea un WorkSpace).	Espandi la sezione Advanced configuration (Configurazione avanzata) nella finestra di dialogo Create a WorkSpace (Crea un WorkSpace) e seleziona Use a cluster template (Utilizza un modello di cluster).

A…	Esegui questa operazione…
Crea il cluster dal riquadro Cluster EMR nell'interfaccia utente del WorkSpace.	Scegli la scheda Cluster EMR nella barra laterale sinistra di un WorkSpace aperto, espandi la sezione Configurazione avanzata e quindi scegli Modello di cluster.

2. Seleziona un modello di cluster dall'elenco a discesa. Ogni modello di cluster disponibile include una breve descrizione che ti aiuta a effettuare una selezione.
3. Il modello di cluster scelto potrebbe avere parametri aggiuntivi, ad esempio la versione di Amazon EMR o il nome del cluster. È possibile scegliere o inserire valori oppure utilizzare i valori predefiniti selezionati dall'amministratore.
4. Selezionare una Subnet (Sottorete) dove EMR Studio può lanciare il nuovo cluster. Ogni opzione di sottorete è pre-approvata dall'amministratore di Studio e WorkSpace dovrebbe essere in grado di connettersi a un cluster in qualsiasi sottorete.
5. Scegli Use cluster template (Utilizza modello cluster) per eseguire il provisioning del cluster e collegarlo all'istanza WorkSpace. La creazione del cluster da parte di EMR Studio richiederà alcuni minuti. Se utilizzi la finestra di dialogo Crea un WorkSpace, scegli Crea un WorkSpace per creare il WorkSpace ed eseguire il provisioning del cluster. Dopo che EMR Studio esegue il provisioning del nuovo cluster, il cluster viene collegato automaticamente all'istanza WorkSpace.

Scollegamento di un calcolo da un WorkSpace EMR Studio

Per scambiare il cluster collegato a un'istanza WorkSpace, è possibile scollegare un cluster dall'interfaccia utente dell'istanza WorkSpace.

Scollegamento di un cluster da un'istanza WorkSpace

1. Nel WorkSpace che desideri scollegare da un cluster, scegli l'icona Cluster EMR dalla barra laterale sinistra per aprire il riquadro Cluster.
2. In Select cluster (Seleziona cluster), scegli Detach (Scollega) e attendi che EMR Studio scolleghi il cluster. Quando il cluster viene scollegato, visualizzerai un messaggio di esito positivo.

Scollegamento di un'applicazione EMR Serverless da un WorkSpace EMR Studio

Per scambiare il calcolo collegato a un'istanza WorkSpace, è possibile scollegare l'applicazione dall'interfaccia utente dell'istanza WorkSpace.

1. Nel WorkSpace che desideri scollegare da un cluster, scegli l'icona Calcolo Amazon EMR dalla barra laterale sinistra per aprire il riquadro Calcolo.
2. In Seleziona calcolo, scegli Scollega e attendi che EMR Studio scolleghi l'applicazione. Quando l'applicazione viene scollegata, visualizzerai un messaggio di esito positivo.

Collegamento di repository basati su Git a un WorkSpace EMR Studio

Associa fino a tre repository basati su Git a un'istanza WorkSpace di Amazon EMR Studio per salvare e condividere i tuoi file notebook.

Informazioni sui repository Git per EMR Studio

È possibile associare un massimo di tre repository Git a un WorkSpace EMR Studio. Per impostazione predefinita, ogni Workspace consente di scegliere da un elenco di repository Git associati allo stesso AWS account di Studio. Puoi anche creare un nuovo repository Git come risorsa per un'istanza WorkSpace.

Puoi eseguire comandi Git come il seguente utilizzando un comando terminale mentre sei connesso al nodo primario di un cluster.

```
!git pull origin <branch-name>
```

In alternativa, puoi utilizzare l'estensione jupyterlab-git. Aprila dalla barra laterale sinistra scegliendo l'icona Git. Per informazioni sull'estensione jupyterlab-git per, vedi jupyterlab-git. JupyterLab

Prerequisiti

- Per associare un repository Git a un'istanza WorkSpace, il tuo Studio deve essere configurato per consentire il collegamento del repository Git. L'amministratore dello Studio dovrebbe adottare le misure necessarie per Definizione di accesso e autorizzazioni per i repository basati su Git.
- Se si utilizza un CodeCommit repository, è necessario utilizzare credenziali Git e HTTPS. Le chiavi SSH e HTTPS con l'helper delle AWS Command Line Interface credenziali non sono supportate. CodeCommit inoltre non supporta i token di accesso personali (). PATs Per ulteriori informazioni, consulta Using IAM with CodeCommit nella IAM User Guide e Setup for HTTPS users using Git Credentials nella AWS CodeCommit User Guide.

Istruzioni

Collegamento di un repository Git associato a un'istanza WorkSpace

1. Apri l'istanza WorkSpace da collegare a un repository dall'elenco WorkSpaces (istanze WorkSpace) nel Studio.
2. Nella barra laterale sinistra, seleziona l'icona Repository Git Amazon EMR per aprire il pannello degli strumenti Repository Git.
3. In Git repositories (Repository Git), espandi l'elenco a discesa e seleziona un massimo di tre repository diversi da collegare all'istanza WorkSpace. EMR Studio registra la selezione e inizia a collegare ogni repository.

Il completamento del processo di collegamento potrebbe richiedere un po' di tempo. È possibile visualizzare lo stato di ogni repository selezionato nel pannello degli strumenti Git repository (Repository Git). Dopo che EMR Studio collega un repository all'istanza WorkSpace, i file che appartengono a tale repository dovrebbero essere visualizzati nel riquadro File browser (Browser di file).

Aggiunta di un nuovo repository Git all'istanza WorkSpace come risorsa

1. Apri il WorkSpace da collegare a un repository dall'elenco WorkSpaces (WorkSpace) nel tuo Studio.
2. Nella barra laterale sinistra, seleziona l'icona Repository Git Amazon EMR per aprire il pannello degli strumenti Repository Git.
3. Scegli Add new Git repository (Aggiungi nuovo repository Git).
4. Per Repository name (Nome repository), immetti un nome descrittivo da utilizzare per il repository in EMR Studio. I nomi possono contenere solo caratteri alfanumerici, trattini alti e trattini bassi.
5. Per URL repository Git), immetti l'URL del repository. Quando usi un CodeCommit repository, questo è l'URL che viene copiato quando scegli Clona URL e poi Clona HTTPS. Ad esempio, `https://git-codecommit.us-west-2.amazonaws.com/v1/repos/[MyCodeCommitRepoName]`.
6. Per Branch (Ramo), immetti il nome di un ramo esistente da controllare.
7. Per Git credentials (Credenziali Git), scegli un'opzione in base alle seguenti linee guida. EMR Studio accede alle credenziali Git utilizzando i segreti archiviati in Secrets Manager.

> **Note**
>
> Se utilizzi un GitHub repository, ti consigliamo di utilizzare un token di accesso personale (PAT) per l'autenticazione. A partire dal 13 agosto 2021, GitHub richiederà l'autenticazione basata su token e non accetterà più password per l'autenticazione delle operazioni Git. Per ulteriori informazioni, consulta il post sui requisiti di autenticazione dei token per le operazioni Git nel GitHub blog.

Opzione	Descrizione
Crea un nuovo segreto	Scegli questa opzione per associare le credenziali Git esistenti a un nuovo segreto che verrà creato AWS Secrets Manager per te. Esegui una delle seguenti operazioni in base alle credenziali Git utilizzate per il repository. Se utilizzi un nome utente e una password Git per accedere al repository, seleziona Nome utente e password, immetti il Nome segreto da utilizzare in Secrets Manager e quindi il Nome utente e la Password da associare al segreto. –OPPURE– Se utilizzi un token di accesso personale per accedere al repository, seleziona Personal access token (PAT) (Token di accesso personale (PAT)), immetti il Secret name (Nome segreto) da utilizzare in Secrets Manager e, in seguito, immetti il tuo personal access token (token di accesso personale). Per ulteriori informazioni, consulta Creazione di un token di accesso personale per la riga di comando GitHub e Token di accesso personali per Bitbucket. CodeCommit i repository non supportano questa opzione.
Utilizza un repository pubblico senza credenziali	Scegli questa opzione per accedere a un repository pubblico.

Opzione	Descrizione
Usa un segreto esistente AWS	Scegli questa opzione se le credenziali sono già state salvate come segreto in Secrets Manager, quindi seleziona il nome segreto dall'elenco. Se selezioni un segreto associato a un nome utente e una password Git, il segreto deve essere nel formato `{"gitUsername": "MyUserName ", "gitPassword": "MyPassword "}`.

8. Scegli Add repository (Aggiungi repository) per creare il nuovo repository. Dopo che EMR Studio ha creato il nuovo repository, visualizzerai un messaggio di esito positivo. Il nuovo repository viene visualizzato nell'elenco a discesa in Git repositories (Repository Git).
9. Per collegare il nuovo repository alla tua istanza WorkSpace, selezionalo dall'elenco a discesa in Git repositories (Repository Git).

Il completamento del processo di collegamento potrebbe richiedere un po' di tempo. Dopo che EMR Studio collega il nuovo repository all'istanza WorkSpace, verrà visualizzata una nuova cartella con lo stesso nome del repository nel riquadro File Browser (Browser di file).

Per aprire un repository collegato differente, passare alla relativa cartella nel File browser (Browser di file).

Utilizzo dell'editor Amazon Athena SQL in EMR Studio

Panoramica

Puoi usare Amazon EMR Studio per sviluppare ed eseguire query interattive su Amazon Athena. Ciò significa che puoi eseguire l'analisi SQL su Athena dalla stessa interfaccia EMR Studio utilizzata per eseguire Spark, Scala e altri carichi di lavoro. Con questa integrazione, puoi utilizzare il completamento automatico per sviluppare query rapidamente, sfogliare i dati nel tuo AWS Glue Data Catalog, creare query salvate, visualizzare la cronologia delle query e altro ancora.

Per ulteriori informazioni sull'utilizzo di Amazon Athena, consulta la sezione Utilizzo di Athena SQL nella Guida per l'utente di Amazon Athena.

Utilizzo dell'editor Athena SQL in EMR Studio

Utilizza i seguenti passaggi per sviluppare ed eseguire query interattive su Amazon Athena da EMR Studio:

1. Aggiungi le autorizzazioni necessarie al ruolo utente per gli utenti che accedono ai workspace in questo Studio. Le autorizzazioni sono elencate nella tabella AWS Identity and Access Management autorizzazioni per gli utenti di EMR Studio nella colonna Accedi all'editor Amazon Athena SQL dal tuo EMR Studio. In alternativa, puoi scegliere di copiare il contenuto della policy Avanzata da Esempi di policy utente per concedere agli utenti le autorizzazioni complete per le funzionalità di EMR Studio, inclusa questa.
2. Configura e crea un EMR Studio.
3. Vai al tuo Studio e seleziona Editor di query dalla barra laterale.

Viene visualizzata l'interfaccia utente nota dell'editor Athena. Per informazioni su come iniziare e usare Athena SQL per eseguire query interattive, consulta Guida introduttiva e Utilizzo di Athena SQL nella Guida per l'utente di Amazon Athena.

> **Note**
>
> Se hai abilitato la propagazione delle identità attendibili tramite IAM Identity Center per il tuo EMR Studio, devi utilizzare i gruppi di lavoro Athena per controllare l'accesso alle query e il gruppo di lavoro che utilizzi deve utilizzare anche la propagazione delle identità attendibili. Per i passaggi per configurare Identity Center e abilitare la propagazione delle identità attendibili per il tuo gruppo di lavoro, consulta Utilizzo dei gruppi di lavoro Athena abilitati per IAM Identity Center nella Guida per l'utente di Amazon Athena.

Considerazioni per l'utilizzo dell'editor Athena SQL in EMR Studio

- L'integrazione con Athena è disponibile in tutte le regioni commerciali in cui sono disponibili EMR Studio e Athena.
- Le seguenti funzionalità di Athena non sono disponibili in EMR Studio:
  - Funzionalità di amministrazione come creazione o aggiornamento di gruppi di lavoro Athena, origini dati o prenotazioni della capacità
  - Athena per Spark o notebook Spark

- DataZone Integrazione con Amazon
- Cost Based Optimizer (CBO)
- Funzioni passo-passo

CodeWhisperer Integrazione di Amazon con EMR Studio Workspaces

Panoramica

Puoi usare Amazon CodeWhisperer con Amazon EMR Studio per ottenere consigli in tempo reale durante la scrittura del codice. JupyterLab CodeWhispererpuoi completare i tuoi commenti, completare singole righe di codice, line-by-line formulare consigli e generare funzioni complete.

> **Note**
>
> Quando utilizzi Amazon EMR Studio, AWS potresti archiviare dati sull'utilizzo e sui contenuti per scopi di miglioramento del servizio. Per ulteriori informazioni e istruzioni su come disattivare la condivisione dei dati, consulta Sharing your data with AWS nella Amazon CodeWhisperer User Guide.

Considerazioni sull'utilizzo CodeWhisperer con Workspaces

- CodeWhisperer l'integrazione è disponibile nello stesso Regioni AWS luogo in cui è disponibile EMR Studio, come documentato nelle considerazioni su EMR Studio.
- Amazon EMR Studio utilizza automaticamente l' CodeWhisperer endpoint negli Stati Uniti orientali (Virginia settentrionale) (us-east-1) per i consigli, indipendentemente dalla regione in cui si trova lo studio.
- CodeWhisperer supporta solo il linguaggio Python per la codifica degli script ETL per i lavori Spark in EMR Studio.
- Un'opzione di telemetria lato client quantifica l'utilizzo di. CodeWhisperer Questa funzionalità non è supportata con EMR Studio.

Autorizzazioni richieste per CodeWhisperer

Per CodeWhisperer utilizzarla, devi allegare la seguente policy al tuo ruolo utente IAM per Amazon EMR Studio:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "CodeWhispererPermissions",
      "Effect": "Allow",
      "Action": [
        "codewhisperer:GenerateRecommendations"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Utilizza CodeWhisperer con Workspaces

Per visualizzare il log in CodeWhisperer di riferimento JupyterLab, apri il CodeWhispererpannello nella parte inferiore della JupyterLab finestra e scegli Open Code Reference Log.

L'elenco seguente contiene scorciatoie che puoi usare per interagire con CodeWhisperer i suggerimenti:

- Metti in pausa i consigli: utilizza i suggerimenti automatici di pausa dalle impostazioni. CodeWhisperer
- Accetta un suggerimento: premi Tab sulla tastiera.
- Rifiuta un suggerimento: premi Esc sulla tastiera.
- Passa da un suggerimento all'altro: usa le frecce Su e Giù sulla tastiera.
- Richiamo manuale: premi Alt e C sulla tastiera. Se usi un Mac, premi Cmd e C.

Puoi anche utilizzarlo CodeWhisperer per modificare impostazioni come il livello di registro e ottenere suggerimenti per i riferimenti al codice. Per ulteriori informazioni, consulta Configurazione CodeWhisperer con JupyterLab e funzionalità nella Amazon CodeWhisperer User Guide.

Debug di applicazioni e processi con EMR Studio

Con Amazon EMR Studio, è possibile avviare interfacce di applicazioni dati per analizzare le applicazioni e le esecuzioni di processo nel browser.

Puoi anche avviare le interfacce utente persistenti e fuori dal cluster per Amazon EMR in esecuzione su EC2 cluster dalla console Amazon EMR. Per ulteriori informazioni, consulta Visualizza le interfacce utente persistenti delle applicazioni in Amazon EMR.

> **Note**
>
> A seconda delle impostazioni del browser, potrebbe essere necessario abilitare i popup per l'apertura dell'interfaccia utente di un'applicazione.

Per informazioni sulla configurazione e sull'utilizzo delle interfacce delle applicazioni, consulta Il Timeline Server di YARN, Monitoraggio e strumentazione o Panoramica dell'interfaccia utente Tez.

Esegui il debug di Amazon EMR in esecuzione su Amazon Jobs EC2

Workspace UI

Avvio di un'interfaccia utente su cluster da un file notebook

Quando si utilizza Amazon EMR versione 5.33.0 e successive, è possibile avviare l'interfaccia utente Web Spark (l'interfaccia utente Spark o Spark History Server) da un notebook nel WorkSpace.

Sul cluster UIs funziona con i kernel PySpark, Spark o SparkR. La dimensione massima del file visualizzabile per i log eventi o i log del container di Spark è di 10 MB. Se i file di log superano i 10 MB, si consiglia di utilizzare lo Spark History Server persistente anziché l'interfaccia utente Spark su cluster per eseguire il debug dei processi.

> **Important**
>
> Affinché EMR Studio possa avviare le interfacce utente delle applicazioni su cluster da un Workspace, il cluster deve essere in grado di comunicare con il Gateway Amazon API. È necessario configurare il cluster EMR per consentire il traffico di rete in uscita verso Amazon API Gateway e assicurarsi che Amazon API Gateway sia raggiungibile dal cluster.

L'interfaccia utente Spark accede ai log del container risolvendo i nomi host. Se si utilizza un nome di dominio personalizzato, è necessario assicurarsi che i nomi host dei nodi del cluster possano essere risolti da Amazon DNS o dal server DNS specificato. A tale scopo, imposta le opzioni DHCP (Dynamic Host Configuration Protocol) per l'Amazon Virtual Private Cloud (VPC) associato al cluster. Per ulteriori informazioni sulle opzioni DHCP, consulta Set di opzioni DHCP nella Guida per l'utente di Amazon Virtual Private Cloud.

1. Nel tuo EMR Studio, apri l'area di lavoro che desideri utilizzare e assicurati che sia collegata a un cluster Amazon EMR su cui è in esecuzione. EC2 Per istruzioni, consultare Collegamento di un calcolo a un WorkSpace EMR Studio.
2. Apri un file notebook e usa il kernel PySpark, Spark o SparkR. Per selezionare un kernel, scegli il nome del kernel in alto a destra della barra degli strumenti del notebook per aprire la finestra di dialogo Select Kernel (Seleziona kernel). Il nome viene visualizzato come No Kernel! (Nessun kernel!) se non è stato selezionato alcun kernel.
3. Esegui il codice del notebook. Quando avvii il contesto Spark, viene visualizzato come output nel notebook. Potrebbero essere necessari alcuni secondi prima di visualizzarlo. Se hai avviato il contesto Spark, è possibile eseguire il comando `%%info` per accedere a un collegamento all'interfaccia utente Spark in qualsiasi momento.

   > **Note**
   >
   > Se i collegamenti dell'interfaccia utente Spark non funzionano o non vengono visualizzati dopo alcuni secondi, crea una nuova cella del notebook ed esegui il comando `%%info` per rigenerare i collegamenti.

4. Per avviare l'interfaccia utente Spark, seleziona Link (Collegamento) in Spark UI (Interfaccia utente Spark). Se l'applicazione Spark è in esecuzione, l'interfaccia utente Spark si apre in una nuova scheda. Se l'applicazione è stata completata, si apre Spark History Server.

   Dopo aver avviato l'interfaccia utente di Spark, puoi modificare l'URL nel browser per aprire YARN ResourceManager o Yarn Timeline Server. Aggiungi uno dei percorsi seguenti dopo `amazonaws.com`.

Interfaccia utente Web	Path	Esempio di URL modificato
FILATO ResourceM anager	/rm	https://*j-examplebby5ij* .emrappui-prod. *eu-west-1* .amazonaws.com */rm*
Timeline Server di Yarn	/yts	https://.emrappui-prod. *j-examplebby5ij eu-west-1* .amazonaws.com */yts*
Spark History Server	/shs	https://.emrappui-prod. *j-examplebby5ij eu-west-1* .amazonaws.com */shs*

Studio UI

Avvio dell'interfaccia utente persistente del Timeline Server di YARN, di Spark History Server o di Tez dall'interfacia utente di EMR Studio

1. Nel tuo EMR Studio, seleziona Amazon EMR EC2 sul lato sinistro della pagina per aprire l'elenco di Amazon EMR sui cluster. EC2
2. Filtra l'elenco dei cluster per name (nome), state (stato) oppure ID immettendo valori nella casella di ricerca. Puoi anche effettuare una ricerca per time range (intervallo temporale) di creazione.
3. Seleziona un cluster, quindi scegli Avvia applicazione UIs per selezionare l'interfaccia utente dell'applicazione. L'interfaccia utente dell'applicazione si apre in una nuova scheda del browser e potrebbe richiedere del tempo per il caricamento.

Esegui il debug di EMR Studio in esecuzione su EMR Serverless

Analogamente ad Amazon EMR in esecuzione su Amazon EC2, puoi utilizzare l'interfaccia utente Workspace per analizzare le tue applicazioni EMR Serverless. Dall'interfaccia utente WorkSpaces, quando utilizzi le versioni 6.14.0 e successive di Amazon EMR, è possibile avviare l'interfaccia utente Web Spark (l'interfaccia utente Spark o Spark History Server) da un notebook nel WorkSpace. Per comodità, forniamo anche un collegamento al log dei driver per accedere rapidamente ai log dei driver Spark.

Esecuzione del debug delle esecuzioni di processo Amazon EMR su EKS con Spark History Server

Quando invii un processo eseguito a un cluster Amazon EMR su EKS, puoi accedere ai registri per quel processo eseguito utilizzando Spark History Server. Spark History Server fornisce strumenti per il monitoraggio delle applicazioni Spark, come un elenco di fasi e processi di pianificazione, un riepilogo delle dimensioni RDD e dell'utilizzo della memoria e informazioni ambientali. È possibile avviare Spark History Server per Amazon EMR sulle esecuzioni del processo EKS nei seguenti modi:

- Quando invii l'esecuzione di un processo utilizzando EMR Studio con un endpoint gestito da Amazon EMR su EKS, puoi avviare Spark History Server da un file notebook nel tuo Workspace.
- Quando invii un job eseguito utilizzando AWS CLI o AWS SDK per Amazon EMR su EKS, puoi avviare Spark History Server dall'interfaccia utente di EMR Studio.

Per informazioni su come utilizzare Spark History Server, consulta Monitoraggio e strumentazionenella documentazione di Apache Spark. Per ulteriori informazioni sulle esecuzioni di processo, consulta Concetti e componenti nella Guida allo sviluppo di Amazon EMR su EKS.

Per l'avvio dello Spark History Server persistente da un file notebook nella tua istanza WorkSpace di EMR Studio

1. Aprire un'istanza WorkSpace connessa a un cluster Amazon EMR su EKS.
2. Seleziona e apri il file del notebook nell'istanza WorkSpace.
3. Scegli Spark UI (Interfaccia utente Spark) nella parte superiore di un file notebook per aprire lo Spark History Server persistente in una nuova scheda.

Per avviare Spark History Server dall'interfaccia utente di EMR Studio

> **Note**
>
> L'elenco dei lavori nell'interfaccia utente di EMR Studio mostra solo le esecuzioni di job inviate utilizzando AWS CLI o l' AWS SDK per Amazon EMR su EKS.

1. In EMR Studio, seleziona Amazon EMR su EKS a sinistra della pagina.
2. Cerca il cluster virtuale Amazon EMR su EKS utilizzato per inviare l'esecuzione del processo. Puoi filtrare l'elenco dei cluster per status (stato) o ID immettendo valori nella casella di ricerca.

3. Seleziona il cluster per aprire la relativa pagina dei dettagli. Nella pagina dei dettagli vengono visualizzate informazioni sul cluster, ad esempio ID, spazio dei nomi e stato. La pagina mostra anche un elenco di tutti i processi eseguiti inviati a quel cluster.
4. Dalla pagina dei dettagli del cluster, seleziona un processo da sottoporre a debug.
5. In alto a destra dell'elenco Jobs (Processi), seleziona Launch Spark History Server (Avvia Spark History Server) per aprire l'interfaccia dell'applicazione in una nuova scheda del browser.

Installazione di kernel e librerie in un'istanza WorkSpace di EMR Studio

Ogni istanza WorkSpace di Amazon EMR Studio viene fornito con un set di librerie e kernel preinstallati.

Kernel e librerie su cluster eseguiti su Amazon EC2

Puoi anche personalizzare l'ambiente per EMR Studio nei seguenti modi quando utilizzi cluster EMR in esecuzione su Amazon: EC2

- Installazione dei kernel Jupyter Notebook e delle librerie Python su un nodo primario del cluster: quando si installano le librerie utilizzando questa opzione, tutti i WorkSpace collegati allo stesso cluster condividono quelle librerie. È possibile installare kernel o librerie all'interno di una cella del notebook o mentre si è connessi tramite SSH al nodo primario di un cluster.
- Utilizzare librerie con ambito notebook: quando gli utenti delle istanze WorkSpace installano e utilizzano le librerie all'interno di una cella del notebook, tali librerie sono disponibili solo per quel notebook. Questa opzione consente a diversi notebook utilizzando lo stesso cluster di lavorare senza preoccuparsi di versioni di libreria in conflitto.

I WorkSpace EMR Studio hanno la stessa architettura sottostante dei notebook EMR. È possibile installare e utilizzare i kernel Jupyter Notebook e le librerie Python con EMR Studio allo stesso modo dei notebook EMR. Per istruzioni, consultare Installazione e utilizzo di kernel e librerie in EMR Studio.

Kernel e librerie sui cluster Amazon EMR su EKS

I cluster Amazon EMR su EKS includono i kernel Python 3.7 e PySpark Python con un set di librerie preinstallate. Amazon EMR su EKS non supporta l'installazione di librerie o cluster aggiuntivi.

Ogni cluster Amazon EMR su EKS viene fornito con i seguenti Python e le seguenti librerie installate: PySpark

- 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

Kernel e librerie sulle applicazioni EMR Serverless

Ogni applicazione EMR Serverless viene fornita con i seguenti Python e le seguenti librerie installate: PySpark

- Python – ggplot, matplotlib, numpy, pandas, plotly, bokeh, scikit-learn, scipy, seaborn
- PySpark – ggplot, matplotlib,numpy, pandas, plotly, bokeh, scikit-learn, scipy, seaborn

Migliora i kernel con magic i comandi in EMR Studio

Panoramica

EMR Studio ed EMR Notebooks supportano i comandi. magic Magici comandi, o magics, sono miglioramenti forniti dal IPython kernel per aiutarti a eseguire e analizzare i dati. IPythonè un ambiente shell interattivo creato con Python.

Amazon EMR supporta Sparkmagic anche un pacchetto che fornisce comandi magic specifici ai kernel relativi a Spark (PySpark, SparkR e Scala) e che utilizza Livy sul cluster per inviare lavori Spark.

È possibile utilizzare i comandi magic purché si disponga di un kernel Python nel notebook EMR. Analogamente, qualsiasi kernel relativo a Spark supporta i comandi Sparkmagic.

Comandi Magic, chiamati anche magic, sono disponibili in due varietà:

- Line magic: questi comandi magic sono denotati da un singolo prefisso `%` e funzionano su una singola riga di codice
- Cell magic: questi comandi magic sono indicati con un doppio prefisso `%%` e funzionano su più righe di codice

Per tutti i magic disponibili, consulta Elenco magic e comandi Sparkmagic.

Considerazioni e limitazioni

- EMR Serverless non supporta il `%%sh` per eseguire `spark-submit`. Non supporta i magic di EMR Notebooks.
- I cluster Amazon EMR su EKS non supportano i comandi Sparkmagic per EMR Studio. Questo perché i kernel Spark utilizzati con gli endpoint gestiti sono integrati in Kubernetes e non sono supportati da Sparkmagic e Livy. È possibile impostare la configurazione Spark direttamente nell' SparkContext oggetto come soluzione alternativa, come dimostra l'esempio seguente.

  ```
  spark.conf.set("spark.driver.maxResultSize", '6g')
  ```

- I seguenti magic comandi e azioni sono proibiti da: AWS
  - `%alias`
  - `%alias_magic`
  - `%automagic`
  - `%macro`
  - Modifica `proxy_user` con `%configure`
  - Modifica `KERNEL_USERNAME` con `%env` o `%set_env`

Elenco magic e comandi Sparkmagic

Utilizza i comandi seguenti per visualizzare un elenco di tutti i comandi magic disponibili:

- `%lsmagic` elenca tutte le funzioni magic attualmente disponibili.
- `%%help` elenca le funzioni magic attualmente disponibili su Spark fornite dal pacchetto Sparkmagic.

Utilizza **`%%configure`** per configurare Spark

Uno dei comandi Sparkmagic più utili è il comando `%%configure`, che configura i parametri di creazione della sessione. Utilizzando le impostazioni `conf`, è possibile configurare qualsiasi configurazione Spark menzionata nella documentazione di configurazione di Apache Spark.

Example Aggiungi un file JAR esterno ai EMR Notebooks dal repository Maven o da Amazon S3

È possibile utilizzare il seguente approccio per aggiungere una dipendenza da file JAR esterno a qualsiasi kernel relativo a Spark supportato da 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 : Configura Hudi

È quindi possibile utilizzare l'editor del notebook per configurare EMR Notebooks per l'utilizzo di 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"
     }
}
```

Utilizza `%%sh` per eseguire `spark-submit`

Il `%%sh` magic esegue comandi shell (interprete di comandi) in un sottoprocesso su un'istanza del cluster collegato. In genere, si utilizza uno dei kernel relativi a Spark per eseguire applicazioni Spark sul cluster collegato. Tuttavia, se si desidera utilizzare un kernel Python per inviare un'applicazione Spark, è possibile utilizzare la seguente magic, sostituendo il nome del bucket con il nome del bucket in minuscolo.

```
%%sh
spark-submit --master yarn --deploy-mode cluster s3://amzn-s3-demo-bucket/test.py
```

In questo esempio, il cluster deve accedere alla posizione di `s3://amzn-s3-demo-bucket/test.py` o il comando avrà esito negativo.

È possibile utilizzare qualsiasi comando Linux con `%%sh` magic. Per eseguire qualsiasi comando Spark o YARN, utilizza una delle seguenti opzioni per creare un Utente Hadoop `emr-notebook` e concedi all'utente le autorizzazioni per eseguire i comandi:

- Puoi creare esplicitamente un nuovo utente eseguendo i seguenti comandi.

```
hadoop fs -mkdir /user/emr-notebook
hadoop fs -chown emr-notebook /user/emr-notebook
```

- Puoi attivare la rappresentazione dell'utente in Livy, che crea automaticamente l'utente. Per ulteriori informazioni, consulta Abilitazione della rappresentazione utente per monitorare l'attività dell'utente e dei processi Spark.

Utilizza `%%display` per visualizzare i dataframe Spark

Puoi usare `%%display` magic per visualizzare un dataframe Spark. Per usare questo magic, eseguire il seguente comando.

```
%%display df
```

Scegli di visualizzare i risultati in formato tabella, come illustrato nell'immagine seguente.

È inoltre possibile scegliere di visualizzare i dati con cinque tipi di grafici. Le opzioni includono grafici a torta, a dispersione, a linee, ad area e a barre.

Utilizza i magic di EMR Notebooks

Amazon EMR fornisce i seguenti magic di EMR Notebooks che possono essere utilizzati con i kernel basati su Python3 e Spark:

- `%mount_workspace_dir` - Monta la directory WorkSpace sul cluster in modo da poter importare ed eseguire codice da altri file nel WorkSpace

> **Note**
>
> con `%mount_workspace_dir`, solo il kernel Python 3 può accedere ai file system locali. Gli executor Spark non avranno accesso alla directory montata con questo kernel.

- `%umount_workspace_dir` - Smonta la directory WorkSpace dal cluster
- `%generate_s3_download_url` - Genera un link per il download temporaneo nell'output del notebook per un oggetto Amazon S3

Prerequisiti

Prima di installare i magic di EMR Notebooks, completare i seguenti processi:

- Assicurarsi che il proprio Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2 disponga dell'accesso in lettura per Amazon S3. Il `EMR_EC2_DefaultRole` con la policy gestita `AmazonElasticMapReduceforEC2Role` soddisfa questo requisito. Se si utilizza un ruolo o una policy personalizzati, assicurarsi che dispongano delle autorizzazioni S3 necessarie.

  > **Note**
  >
  > magicEMR Notebooks viene eseguito su un cluster come utente del notebook e EC2 utilizza il profilo dell'istanza per interagire con Amazon S3. Quando si monta una directory WorkSpace su un cluster EMR, tutti i notebook WorkSpace e EMR con il permesso di collegarsi a tale cluster possono accedere alla directory montata.
  > Le directory sono montate in sola lettura per impostazione predefinita. Mentre `s3fs-fuse` e `goofys` permettono il supporto in lettura-scrittura, consigliamo vivamente di non modificare i parametri di montaggio per montare le directory in modalità di lettura-scrittura. Se si permette l'accesso in scrittura, le modifiche apportate alla directory vengono scritte nel bucket S3. Per evitare l'eliminazione o la sovrascrittura accidentale, è possibile abilitare il controllo delle versioni per il bucket S3. Per ulteriori informazioni, consulta Utilizzo della funzione controllo delle versioni nei bucket S3.

- Eseguire uno dei seguenti script sul cluster per installare le dipendenze per i magic di EMR Notebooks. Per eseguire uno script, è possibile eseguire Utilizzo di operazioni di bootstrap personalizzate oppure segui le istruzioni in Run commands and scripts on an Amazon EMR cluster (Esegui comandi e script su un cluster Amazon EMR) quando si dispone già di un cluster in esecuzione.

  È possibile scegliere quale dipendenza installare. Entrambi s3fs-fuse e goofys sono strumenti FUSE (Filesystem in Userspace) che consentono di montare un bucket Amazon S3 come file system locale su un cluster. Lo strumento `s3fs` offre un'esperienza simile a POSIX. Lo strumento `goofys` è una buona scelta quando si preferiscono le prestazioni rispetto a un file system conforme a POSIX.

  La serie Amazon EMR 7.x utilizza Amazon Linux 2023, che non supporta i repository EPEL. Se utilizzi Amazon EMR 7.x, segui le istruzioni di s3fs-fuse per l'installazione GitHub. `s3fs-fuse` Se usi la serie 5.x o 6.x, usa i seguenti comandi per l'installazione. `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
```

OPPURE

```
#!/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
```

Installare i magic di EMR Notebooks

> **Note**
>
> Con i rilasci di Amazon EMR da 6.0 a 6.9.0 e da 5.0 a 5.36.0, solo il pacchetto `emr-notebooks-magics` versione 0.2.0 e successive supporta `%mount_workspace_dir` magic.

Per installare i magic di EMR Notebooks, completare i seguenti passaggi.

1. Nel notebook, esegui i comandi seguenti per installare il pacchetto `emr-notebooks-magics`.

   ```
   %pip install boto3 --upgrade
   %pip install botocore --upgrade
   %pip install emr-notebooks-magics --upgrade
   ```

2. Riavviare il kernel per caricare i magic di EMR Notebooks.
3. Verificare l'installazione con il seguente comando, che dovrebbe visualizzare il testo della guida di output per `%mount_workspace_dir`.

   ```
   %mount_workspace_dir?
   ```

Montare una directory WorkSpace con `%mount_workspace_dir`

Il `%mount_workspace_dir` magic consente di montare la directory WorkSpace sul cluster EMR in modo da poter importare ed eseguire altri file, moduli o pacchetti archiviati nella directory.

L'esempio seguente monta l'intera directory di WorkSpace su un cluster e specifica l'argomento opzionale *<--fuse-type>* affinché vengano utilizzati goofys per il montaggio della directory.

```
%mount_workspace_dir . <--fuse-type goofys>
```

Per verificare che la directory WorkSpace sia montata, utilizzare l'esempio seguente per visualizzare l'attuale directory di lavoro con il comando `ls`. L'output dovrebbe visualizzare tutti i file nel WorkSpace.

```
%%sh
ls
```

Una volta che sono state apportate le modifiche nel WorkSpace, puoi smontare la directory del WorkSpace con il seguente comando:

> **Note**
>
> La directory WorkSpace rimane montata sul cluster anche quando il WorkSpace viene arrestato o distaccato. È necessario smontare esplicitamente la directory di WorkSpace.

```
%umount_workspace_dir
```

Scaricare un oggetto Amazon S3 con `%generate_s3_download_url`

Il comando `generate_s3_download_url` crea un URL prefirmato per un oggetto archiviato in Amazon S3. È possibile utilizzare l'URL preimpostato per scaricare l'oggetto sul computer locale. Ad esempio si potrebbe eseguire `generate_s3_download_url` per scaricare il risultato di una query SQL che il codice scrive su Amazon S3.

L'URL prefirmato è valido per 60 minuti per impostazione predefinita. È possibile modificare il tempo di scadenza specificando un numero di secondi per il flag `--expires-in`. Ad esempio, `--expires-in 1800` crea un URL valido per 30 minuti.

L'esempio seguente genera un collegamento per il download di un oggetto specificando il percorso completo di Amazon S3: *s3://EXAMPLE-DOC-BUCKET/path/to/my/object*.

```
%generate_s3_download_url s3://EXAMPLE-DOC-BUCKET/path/to/my/object
```

Per ulteriori informazioni sull'uso di `generate_s3_download_url`, eseguire il comando di seguito per visualizzare il testo della guida.

```
%generate_s3_download_url?
```

Gestione di un notebook in modalità headless con **`%execute_notebook`**

Con `%execute_notebook` magic, puoi eseguire un altro notebook in modalità headless e visualizzare l'output di ogni cella che hai utilizzato. Ciò magic richiede autorizzazioni aggiuntive per il ruolo dell'istanza condiviso da Amazon EMR e Amazon EC2 . Per ulteriori dettagli su come concedere autorizzazioni aggiuntive, esegui il comando `%execute_notebook?`.

Durante un processo di lunga durata, il sistema potrebbe andare in modalità di sospensione a causa dell'inattività o perdere temporaneamente la connettività Internet. Ciò potrebbe interrompere la connessione tra il browser e il server Jupyter. In questo caso, potresti perdere l'output delle celle che hai eseguito e inviato dal server Jupyter.

Se utilizzi il notebook in modalità headless con `%execute_notebook` magic, EMR Notebooks acquisisce l'output dalle celle che hai eseguito, anche in caso di interruzione della rete locale. EMR Notebooks salva l'output in modo incrementale in un nuovo notebook con lo stesso nome del notebook che hai utilizzato. EMR Notebooks inserisce quindi il notebook in una nuova cartella all'interno del workspace. Le esecuzioni headless avvengono sullo stesso cluster e utilizzano il ruolo di servizio `EMR_Notebook_DefaultRole`, ma argomenti aggiuntivi possono modificare i valori predefiniti.

Per eseguire un notebook in modalità headless, usa il seguente comando:

```
%execute_notebook <relative-file-path>
```

Per specificare un ID cluster e il ruolo di servizio per un'esecuzione headless, usa il seguente comando:

```
%execute_notebook <notebook_name>.ipynb --cluster-id <emr-cluster-id> --service-role
 <emr-notebook-service-role>
```

Quando Amazon EMR e Amazon EC2 condividono un ruolo di istanza, il ruolo richiede le seguenti autorizzazioni aggiuntive:

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**
>
> Per utilizzare `%execute_notebook` magic, installa il pacchetto `emr-notebooks-magics`, versione 0.2.3 o successive.

Usare notebook multilingue con i kernel Spark

Ogni kernel notebook di Jupyter ha un linguaggio di default. Ad esempio, il linguaggio predefinito del kernel Spark è Scala e il linguaggio predefinito del PySpark kernel è Python. Con Amazon EMR 6.4.0 e versioni successive, EMR Studio supporta notebook multilingue. Ciò significa che ogni kernel di EMR Studio può supportare le seguenti linguaggi oltre al linguaggio di default: Python, Spark, R e Spark SQL.

Per attivare questa funzionalità, specificare uno dei seguenti comandi magic all'inizio di qualsiasi cella.

Lingua	Comando
Python	`%%pyspark`
Scala	`%%scalaspark`
R	`%%rspark` Non è supportato per i carichi di lavoro interattivi con EMR Serverless.
Spark SQL	`%%sql`

Quando vengono invocati, questi comandi eseguono l'intera cella all'interno della stessa sessione Spark utilizzando l'interprete del linguaggio corrispondente.

La `%%pyspark` cella magic consente agli utenti di scrivere PySpark codice in tutti i kernel Spark.

```
%%pyspark
a = 1
```

La cella `%%sql` magic consente agli utenti di eseguire codice Spark-SQL in tutti i kernel Spark.

```
%%sql
SHOW TABLES
```

La cella `%%rspark` magic consente agli utenti di eseguire codice SparkR in tutti i kernel Spark.

```
%%rspark
a <- 1
```

La cella `%%scalaspark` magic consente agli utenti di eseguire il codice Spark Scala in tutti i kernel Spark.

```
%%scalaspark
val a = 1
```

Condividere i dati tra interpreti del linguaggio utilizzando tabelle temporanee

È inoltre possibile condividere dati tra interpreti del linguaggio utilizzando tabelle temporanee. Nell'esempio seguente viene utilizzato `%%pyspark` in una cella per creare una tabella temporanea in Python e utilizza `%%scalaspark` nella cella seguente per leggere i dati di quella tabella in 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)
```

Panoramica di Notebook Amazon EMR

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Puoi utilizzare Amazon EMR Notebooks insieme ai cluster Amazon EMR che eseguono Apache Spark per creare e aprire Jupyter Notebook e interfacce all'interno della console Amazon EMR. JupyterLab Un notebook EMR è un notebook "serverless" che puoi utilizzare per eseguire query e codice. A differenza di un notebook tradizionale, i contenuti di un notebook EMR, ossia le equazioni, le query, i modelli, il codice e il testo narrativo all'interno delle celle del notebook, vengono eseguiti in un client. I comandi vengono eseguiti utilizzando un kernel sul cluster EMR. I contenuti del notebook vengono salvati in Amazon S3 separatamente dai dati del cluster per una maggiore durata e un riutilizzo flessibile.

È possibile avviare un cluster, collegare un notebook EMR per l'analisi e infine terminare il cluster. Inoltre, puoi chiudere un notebook collegato a un cluster in esecuzione e passare a un altro. Più utenti possono collegare notebook allo stesso cluster contemporaneamente e condividere i file dei notebook tra loro in Amazon S3. Queste funzionalità consentono di eseguire cluster on demand per risparmiare sui costi e ridurre il tempo richiesto per riconfigurare i notebook per cluster e set di dati diversi.

Inoltre, puoi eseguire un notebook EMR a livello di programmazione utilizzando l'API Amazon EMR, senza la necessità di interagire con la console Amazon EMR ("esecuzione headless"). Nel notebook EMR, dovrai includere una cella contenente un tag di parametri. Questa cella consente a uno script di passare nuovi valori di input al notebook. I notebook parametrizzati possono essere riutilizzati con diversi set di valori di input. Non è necessario creare copie dello stesso notebook per modificarlo ed eseguirlo con nuovi valori di input. Amazon EMR crea e salva il notebook di output su S3 per ogni esecuzione del notebook parametrizzato. Per visualizzare alcuni esempi di codice API per notebook EMR, consulta Esempi di comandi programmatici per Notebooks EMR.

> **Important**
>
> La funzionalità Notebook EMR supporta cluster che utilizzano i rilasci 5.18.0 e successivi di Amazon EMR. Consigliamo di utilizzare Notebook EMR con cluster che utilizzano la versione più recente di Amazon EMR o almeno 5.30.0, 5.32.0 o 6.2.0. Con questi rilasci, i kernel Jupyter vengono eseguiti sul cluster collegato anziché su un'istanza Jupyter. Ciò migliora le prestazioni e aumenta la possibilità di personalizzare kernel e librerie. Per ulteriori informazioni, consulta Differenze nelle funzionalità in base alla versione del cluster.

Si applicano i costi previsti per l'archiviazione Amazon S3 e per i cluster Amazon EMR.

I notebook Amazon EMR sono disponibili come aree di lavoro Amazon EMR Studio nella console.

Transizione da Notebook EMR ai Workspace

Nella nuova console Amazon EMR, abbiamo unito Notebook EMR ai Workspace Amazon EMR Studio in un'unica esperienza. Quando utilizzi un EMR Studio, puoi creare e configurare diversi Workspace per organizzare ed eseguire notebook. Se avevi notebook Amazon EMR nella vecchia console, sono disponibili come EMR Studio Workspaces nella console.

Amazon EMR ha creato questi nuovi Workspace EMR Studio per tuo conto. Il numero di Studio che abbiamo creato corrisponde al numero di Notebooks VPCs distinti utilizzati da EMR Notebooks. Ad esempio, se ti connetti ai cluster EMR in due notebook diversi da VPCs EMR, abbiamo creato due nuovi EMR Studios. I tuoi notebook vengono distribuiti tra i nuovi Studio.

> **Important**
>
> Abbiamo disattivato la possibilità di creare nuovi notebook nella vecchia console Amazon EMR. Utilizza invece il pulsante Crea Workspace nella nuova console Amazon EMR.

Per ulteriori informazioni sui Workspace Amazon EMR Studio, consulta Scopri gli spazi di lavoro di EMR Studio. Per una panoramica concettuale di EMR Studio, consulta WorkSpace nella pagina Come funziona Amazon EMR Studio.

Che cosa occorre fare?

Sebbene sia ancora possibile utilizzare i notebook esistenti nella vecchia console, consigliamo di utilizzare invece Amazon EMR Studio Workspaces nella console. Devi configurare le autorizzazioni di ruolo aggiuntive per attivare le funzionalità di EMR Studio che non sono disponibili in Notebook EMR.

> **Note**
>
> Come minimo, per visualizzare i notebook EMR esistenti come Workspace EMR Studio e per crearne di nuovi, gli utenti devono disporre delle autorizzazioni `elasticmapreduce:ListStudios` e `elasticmapreduce:CreateStudioPresignedUrl` relative ai propri ruoli. Per accedere a tutte le funzionalità di EMR Studio, consulta Abilitazione delle funzionalità di EMR Studio per gli utenti di Notebook EMR per l'elenco completo delle autorizzazioni aggiuntive necessarie agli utenti di Notebook EMR.

Funzionalità avanzate in EMR Studio oltre Notebook EMR

Con Amazon EMR Studio, puoi configurare e utilizzare le seguenti funzionalità che non sono disponibili con Notebook EMR:

- Sfogliare e collegare i cluster EMR dall'interno di Jupyterlab
- Sfogliare e collegare i cluster virtuali di Notebook EMR dall'interno di Jupyterlab
- Connettersi ai repository Git dall'interno di Jupyterlab
- Collaborare con altri membri del team per scrivere ed eseguire il codice del notebook
- Sfogliare i dati con SQL Explorer
- Eseguire il provisioning di cluster EMR con Service Catalog

Per un elenco completo delle funzionalità di Amazon EMR Studio, consulta Funzionalità principali di EMR Studio.

Abilitazione delle funzionalità di EMR Studio per gli utenti di Notebook EMR

Le nuove istanze EMR Studio che creeremo nell'ambito di questa unione utilizzano il ruolo IAM `EMR_Notebooks_DefaultRole` esistente come ruolo di servizio EMR Studio.

Gli utenti che effettuano il passaggio da Notebook EMR a EMR Studio e desiderano utilizzare le funzionalità aggiuntive di EMR Studio devono richiedere nuove autorizzazioni di ruolo. Aggiungi le seguenti autorizzazioni ai ruoli degli utenti di Notebook EMR che intendono utilizzare EMR Studio.

> **Note**
>
> Come minimo, per visualizzare i notebook EMR esistenti come Workspace EMR Studio e per crearne di nuovi, gli utenti devono disporre delle autorizzazioni `elasticmapreduce:ListStudios` e `elasticmapreduce:CreateStudioPresignedUrl` relative ai propri ruoli. Per utilizzare tutte le funzionalità di EMR Studio, aggiungi tutte le autorizzazioni elencate di seguito. Gli utenti amministratori devono inoltre disporre dell'autorizzazione per creare e gestire un EMR Studio. Per ulteriori informazioni, consulta Autorizzazioni di amministratore per creare e gestire 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"
```

Le seguenti autorizzazioni sono necessarie anche per utilizzare le funzionalità di collaborazione in EMR Studio, ma non erano richieste con i notebook EMR.

```
"sso-directory:SearchUsers",
"iam:GetUser",
"iam:GetRole",
"iam:ListUsers",
"iam:ListRoles",
"sso:GetManagedApplicationInstance"
```

Requisiti, differenze nelle versioni di rilascio e sicurezza per i notebook EMR

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Considera i seguenti requisiti, le differenze nelle versioni di rilascio, le informazioni sulla sicurezza e altre considerazioni quando crei cluster e sviluppi soluzioni utilizzando il notebook EMR.

Requisiti del cluster

- Attivazione del blocco dell'accesso pubblico Amazon EMR: l'accesso in ingresso a un cluster consente agli utenti del cluster di eseguire i kernel dei notebook. Assicurati che solo gli utenti autorizzati possano accedere al cluster. Consigliamo vivamente di lasciare abilitato il blocco dell'accesso pubblico e di limitare il traffico SSH in ingresso solo a origini affidabili. Per ulteriori informazioni, consultare Utilizzo del blocco dell'accesso pubblico di Amazon EMR e Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.
- Utilizzo di un cluster compatibile: un cluster collegato a un notebook deve soddisfare i seguenti requisiti:

- Sono supportati solo i cluster creati utilizzando Amazon EMR. È possibile creare un cluster in modo indipendente all'interno di Amazon EMR e, successivamente, collegare un notebook EMR, oppure è possibile creare un cluster compatibile durante la creazione di un notebook EMR.
- Solo i cluster creati utilizzando Amazon EMR versione 5.18.0 e successive sono supportati. Consultare the section called “Differenze nelle funzionalità in base alla versione del cluster”.
- I cluster creati utilizzando EC2 istanze Amazon con processori AMD EPYC, ad esempio i tipi di istanza m5a.* e r5a.*, non sono supportati.
- Notebook EMR funziona solo con cluster creati con `VisibleToAllUsers` impostato su `true`. `VisibleToAllUsers` è `true` per impostazione predefinita.
- Il cluster deve essere avviato all'interno di un EC2 -VPC. Sono supportate sottoreti pubbliche e private. La piattaforma EC2 -Classic non è supportata.
- I cluster devono essere avviati con Hadoop, Spark e Livy installati. Possono essere installate altre applicazioni, ma attualmente Notebook EMR supporta solo i cluster Spark.

  > **Important**
  >
  > Per le versioni di Amazon EMR 5.32.0 e successive, o 6.2.0 e successive, il cluster deve eseguire anche l'applicazione Jupyter Enterprise Gateway per poter lavorare con Notebook EMR.

- I cluster che utilizzano l'autenticazione Kerberos non sono supportati.
- I cluster integrati AWS Lake Formation supportano solo l'installazione di librerie con ambito notebook. L'installazione di kernel e librerie nel cluster non è supportata.
- I cluster con più nodi primari non sono supportati.
- I cluster che utilizzano EC2 istanze Amazon basate su AWS Graviton2 non sono supportati.

Differenze nelle funzionalità in base alla versione del cluster

Consigliamo di utilizzare Notebook EMR con cluster creati utilizzando Amazon EMR versione 5.30.0, 5.32.0 o successive oppure 6.2.0 o successive. Con queste versioni, Notebook EMR esegue i kernel sul cluster Amazon EMR collegato. I kernel e le librerie possono essere installati direttamente sul nodo primario del cluster. L'uso di EMR Notebooks con queste versioni del cluster presenta i seguenti vantaggi:

- Prestazioni migliorate: i kernel dei notebook vengono eseguiti su cluster con tipi di istanze selezionati dall'utente. EC2 Le versioni precedenti eseguono i kernel su un'istanza specializzata che non è ridimensionabile, accessibile o personalizzabile.
- Possibilità di aggiungere e personalizzare i kernel: è possibile connettersi al cluster per installare i pacchetti kernel utilizzando `conda` e `pip`. Inoltre, l'installazione `pip` è supportata utilizzando i comandi del terminale all'interno delle celle di notebook. Nelle versioni precedenti, erano disponibili solo kernel preinstallati (Python PySpark, Spark e SparkR). Per ulteriori informazioni, consulta Installazione di kernel e librerie Python su un nodo primario del cluster.
- Possibilità di installare librerie Python: è possibile installare librerie Python sul nodo primario del cluster utilizzando `conda` e `pip`. Consigliamo l'uso di `conda`. Nelle versioni precedenti, sono supportate solo le librerie con ambito notebook per. PySpark

Funzionalità EMR Notebooks supportate dalla versione del cluster

Versione di rilascio del cluster	Librerie con ambito notebook per PySpark	Installazione del kernel sul cluster	Installazione della libreria Python sul nodo primario
Precedente a 5.18.0	Notebook EMR non supportato		
5.18.0-5.25.0	No	No	No
5.26.0–5.29.0	Sì	No	No
5.30.0	Sì	Sì	Sì
6.0.0	No	No	No
5.32.0 e versioni successive e 6.2.0 e versioni successive	Sì	Sì	Sì

Limiti di notebook EMR collegati contemporaneamente

Quando crei un cluster che supporta i notebook, considera il tipo di istanza del nodo primario del EC2 cluster. I vincoli di memoria di questa EC2 istanza determinano il numero di notebook che possono essere pronti contemporaneamente per eseguire codice e query sul cluster.

EC2 Tipo di istanza del nodo primario	Numero di notebook EMR
*.medium	2
*.large	4
*.xlarge	8
*.2xlarge	16
*.4xlarge	24
*.8xlarge	24
*.16xlarge	24

Versioni Jupyter Notebook e Python

EMR Notebooks esegue Jupyter Notebook versione 6.0.2 e Python 3.6.5 a prescindere dalla versione Amazon EMR del cluster collegato.

Considerazioni relative alla sicurezza

Utilizzo di posizioni S3 crittografate

Se si specifica un percorso crittografato in Amazon S3 per archiviare i file del notebook, è necessario impostare Ruolo di servizio per EMR Notebooks come un utente chiave. Il ruolo di servizio predefinito è `EMR_Notebooks_DefaultRole`. Se utilizzi una AWS KMS chiave per la crittografia, consulta Using key policy in AWS KMS nella AWS Key Management Service Developer Guide e l'articolo di supporto per l'aggiunta di utenti chiave.

Utilizzo dei cookie con domini di hosting

Per aumentare la sicurezza delle applicazioni off-console che potresti utilizzare con Amazon EMR, i domini di hosting delle applicazioni sono registrati nella Public Suffix List (PSL). Alcuni esempi di questi domini di hosting includono: `emrstudio-prod.us-east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Per maggiore sicurezza, se hai bisogno di impostare cookie sensibili nel nome di dominio predefinito, consigliamo di utilizzare i cookie con un prefisso `__Host-`. Questa pratica ti aiuterà a difendere il tuo dominio dai tentativi CSRF (cross-site

request forgery). Per ulteriori informazioni, consulta la pagina Set-Cookie in Mozilla Developer Network.

Creare un notebook in EMR Studio

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Per creare un notebook EMR, puoi utilizzare la vecchia console Amazon EMR. La creazione di notebook utilizzando o AWS CLI l'API Amazon EMR non è supportata.

Creazione di un notebook EMR

1. Apri la console di Amazon EMR all'indirizzo https://console.aws.amazon.com/elasticmapreduce/.
2. Scegli Notebook, Crea un notebook.
3. Immetti un Nome notebook e una Descrizione notebook facoltativa.
4. Se disponi di un cluster attivo a cui desideri collegare il notebook, lascia selezionata l'impostazione Scegli un cluster esistente predefinita, fai clic su Scegli, seleziona un cluster dall'elenco, quindi fai clic su Scegli cluster. Per informazioni sui requisiti del cluster per Notebook EMR, consulta Requisiti, differenze nelle versioni di rilascio e sicurezza per i notebook EMR.

   —oppure—

   Scegli Crea un cluster, immetti un Nome cluster e scegli le opzioni in base alle linee guida riportate di seguito. Il cluster viene creato nel VPC predefinito per l'account utilizzando istanze on demand.

Impostazione	Descrizione
Nome cluster	Nome descrittivo utilizzato per identificare il cluster.

Impostazione	Descrizione
Versione	Non si può modificare. Il valore predefinito è l'ultima versione di Amazon EMR (5.36.2).
Applicazioni	Non si può modificare. Elenca le applicazioni installate nel cluster.
Istanza	Inserisci il numero di istanze e seleziona il tipo di istanza. EC2 Per il nodo primario viene utilizzata un'istanza. Le altre vengono utilizzate per i nodi principali. Il tipo di istanza determina il numero di notebook che possono si possono collegare simultane amente al cluster. Per ulteriori informazioni, consulta Limiti di notebook EMR collegati contemporaneamente.
Ruolo EMR	Lascia l'impostazione predefinita o scegli il collegamento per specificare un ruolo di servizio personalizzato per Amazon EMR. Per ulteriori informazioni, consulta Ruolo di servizio per Amazon EMR (ruolo EMR).
EC2 profilo dell'istanza	Lascia il valore predefinito o scegli il link per specificare un ruolo di servizio personalizzato per le EC2 istanze. Per ulteriori informazioni, consulta Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2.
EC2 coppia di chiavi	Scegli una EC2 key pair per connetterti alle istanze del cluster. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

Impostazione	Descrizione
Terminazione automatica	La terminazione automatica è supportata per Amazon EMR versione 5.30.0, 6.1.0 e successive. Seleziona la casella di spunta per abilitare la terminazione automatica, quindi specifica il tempo di inattività dopo il quale il cluster dovrebbe spegnersi in automatico. Per ulteriori informazioni, consulta Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR.

5. Per Gruppo di sicurezza, scegli Usa gruppi di sicurezza predefiniti. In alternativa, fai clic su Scegli gruppi di sicurezza e seleziona i gruppi di sicurezza personalizzati disponibili nel VPC del cluster. Selezionane uno per l'istanza primaria e un altro per l'istanza client del notebook. Per ulteriori informazioni, consulta the section called “Gruppi di sicurezza per EMR Notebooks”.
6. Per Ruolo di servizio AWS , lascia l'impostazione predefinita o scegli un ruolo personalizzato dall'elenco. L'istanza client per il notebook utilizza questo ruolo. Per ulteriori informazioni, consulta Ruolo di servizio per EMR Notebooks.
7. Per Notebook location (Percorso notebook) scegli il percorso Amazon S3 in cui è salvato il file del notebook, oppure specifica un percorso. Se il bucket e la cartella non esistono, Amazon EMR li crea.

   Amazon EMR crea una cartella con Notebook ID (ID notebook) come nome della cartella e salva il notebook in un file denominato *`NotebookName`*`.ipynb`. Ad esempio, se si specifica il percorso Amazon S3 `s3://amzn-s3-demo-bucket/MyNotebooks` per un notebook denominato `MyFirstEMRManagedNotebook`, il file del notebook viene salvato in `s3://amzn-s3-demo-bucket/MyNotebooks/`*`NotebookID`*`/MyFirstEMRManagedNotebook.ipynb`.

   Se si specifica un percorso crittografato in Amazon S3, è necessario impostare Ruolo di servizio per EMR Notebooks come utente chiave. Il ruolo di servizio predefinito è `EMR_Notebooks_DefaultRole`. Se utilizzi una AWS KMS chiave per la crittografia, consulta Uso delle politiche chiave in AWS KMS nella Guida per gli AWS Key Management Service sviluppatori e l'articolo di supporto per l'aggiunta di utenti chiave.

8. Facoltativamente, se hai aggiunto ad Amazon EMR un repository basato su Git che desideri associare a questo notebook, scegli Repository Git, fai clic su Scegli repository e seleziona un repository dall'elenco. Per ulteriori informazioni, consulta Associazione di repository basati su Git con EMR Notebooks.
9. Facoltativamente, puoi scegliere Tag e quindi aggiungere ulteriori tag chiave-valore per il notebook.

   > **Important**
   >
   > Ai fini dell'accesso vengono applicati un tag predefinito con la stringa Chiave impostata su `creatorUserID` e il valore impostato per l'ID dell'utente IAM. Consigliamo di non modificare o rimuovere questo tag perché può essere utilizzato per controllare l'accesso. Per ulteriori informazioni, consulta Utilizzo di tag di cluster e notebook con policy IAM per il controllo degli accessi.

10. Scegli Crea un notebook.

Lavorare con EMR Notebooks

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Dopo aver creato un notebook EMR, questo si avvia poco dopo. Lo Stato nell'elenco Notebook mostra Avvio in corso. È possibile aprire un notebook quando il suo stato è Pronto. Potrebbe essere necessario più tempo a un notebook per raggiungere lo stato Pronto se hai creato un cluster insieme a esso.

> **Tip**
>
> Aggiorna il browser oppure scegli l'icona di aggiornamento sopra l'elenco dei notebook per aggiornare lo stato.

Comprensione dello stato dei notebook

Un notebook EMR può avere il seguente Status (Stato) nell'elenco Notebooks (Notebook).

Stato	Significato
Pronto	Puoi aprire il notebook utilizzando l'editor di notebook. Quando un notebook è in stato Pronto, è possibile arrestarlo o eliminarlo. Per modificare i cluster, devi prima arrestare il notebook. Un notebook in stato Pronto che rimane inattivo a lungo viene arrestato in automatico.
Avvio in corso	Il notebook viene creato e collegato al cluster. Mentre un notebook è in fase di avvio, non è possibile aprire l'editor di notebook, arrestarlo, eliminarlo o modificare i cluster.
In attesa	Il notebook è stato creato ed è in attesa dell'integrazione con il cluster per il completamento. È possibile che il cluster stia ancora eseguendo il provisioning delle risorse o rispondendo ad altre richieste. Puoi aprire l'editor di notebook con il notebook in modalità locale. Il codice che si affida a processi di cluster non viene eseguito e dà esito negativo.
In arresto	Il notebook è in fase di arresto o il cluster a cui è collegato il notebook è in fase di terminazione. Mentre un notebook è in fase di arresto,

Stato	Significato
	non è possibile aprire l'editor di notebook, arrestarlo, eliminarlo o modificare i cluster.
Arrestato	Il notebook è stato arrestato. È possibile avviare il notebook sullo stesso cluster, purché quest'ultimo sia ancora in esecuzione. Puoi modificare i cluster ed eliminare il cluster.
Eliminazione in corso	Il cluster è in fase di rimozione dall'elenco dei cluster disponibili. Il file del notebook, *NotebookName* `.ipynb` , resta in Amazon S3 e continua ad accumulare addebiti di archiviazione applicabili.

Utilizzo dell'editor di notebook

Un vantaggio dell'utilizzo di un notebook EMR è che è possibile avviare il notebook in Jupyter o JupyterLab direttamente dalla console.

Con EMR Notebooks, l'editor di notebook a cui accedi dalla console Amazon EMR è il familiare editor open source Jupyter Notebook o. JupyterLab Poiché l'editor di notebook viene avviato dalla console di Amazon EMR, la configurazione dell'accesso è più efficiente di quanto non sia con un notebook ospitato in un cluster Amazon EMR. Non è necessario configurare un client dell'utente per avere l'accesso Web attraverso SSH, le regole per i gruppi di sicurezza e le configurazioni del proxy. Se un utente dispone di autorizzazioni sufficienti, basta semplicemente aprire l'editor di notebook nella console di Amazon EMR.

EMR Notebooks può essere aperto da un solo utente alla volta da Amazon EMR. Se un altro utente cerca di aprire un notebook EMR già aperto, si verifica un errore.

> **Important**
>
> Amazon EMR crea un URL prefirmato univoco per ogni sessione di editor notebook, valido solo per un breve periodo di tempo. Consigliamo di non condividere l'URL dell'editor di notebook. Ciò comporta un rischio per la sicurezza, perché i destinatari dell'URL adottano le autorizzazioni per modificare il notebook ed eseguire il codice del notebook per tutta la durata

dell'URL. Se altri utenti necessitano di accedere a un notebook, fornisci loro le autorizzazioni attraverso le policy di autorizzazione e assicurati che il ruolo di servizio di Notebook EMR disponga dell'accesso al percorso di Amazon S3. Per ulteriori informazioni, consultare the section called “Sicurezza” e Ruolo di servizio per EMR Notebooks.

Apertura dell'editor di notebook per un notebook EMR

1. Seleziona un notebook con lo Stato su Pronto o In attesa dall'elenco Notebook.
2. Scegli Apri JupyterLab in o Apri in Jupyter.

   Si apre una nuova scheda del browser nell'editor JupyterLab o Jupyter Notebook.
3. Dal menu Kernel, scegli Cambia kernel, quindi seleziona il kernel per il tuo linguaggio di programmazione.

   Ora è tutto pronto per scrivere ed eseguire il codice dall'interno dell'editor di notebook.

Salvataggio dei contenuti di un notebook

Quando utilizzi l'editor di notebook, i contenuti delle celle di notebook e l'output vengono salvati in automatico nel file del notebook in Amazon S3 con cadenza periodica. Un notebook che non ha avuto modifiche dall'ultima volta che è una cella stata modificata mostra la dicitura (salvato in automatico) accanto al nome del notebook nell'editor. Se le modifiche non sono state ancora salvate, viene visualizzato modifiche non salvate.

È possibile salvare un notebook manualmente. Dal menu File, scegli Salva ed esegui il checkpoint o premi CTRL+S. In questo modo viene creato un file denominato *NotebookName*`.ipynb` in una cartella checkpoints all'interno della cartella del notebook in Amazon S3. Ad esempio, `s3://`*amzn-s3-demo-bucket*`/`*MyNotebookFolder*`/`*NotebookID*`/checkpoints/`*NotebookName*`.ipynb`. Solo i file di checkpoint più recenti vengono salvati in questa posizione.

Modifica dei cluster

È possibile modificare il cluster a cui è collegato un notebook EMR senza modificare i contenuti del notebook stesso. È possibile modificare i cluster solo per i notebook che hanno lo stato Arrestato.

Modifica del cluster di un notebook EMR

1. Se il notebook che desideri modificare è in esecuzione, selezionalo dall'elenco Notebook e scegli Arresta.
2. Quando lo stato del notebook è Arrestato, seleziona il notebook dall'elenco Notebook, quindi scegli Visualizza dettagli.
3. Seleziona Modifica cluster.
4. Se disponi di un cluster attivo che esegue Hadoop, Spark e Livy a cui desideri collegare il notebook, lascia l'impostazione predefinita e seleziona un cluster dall'elenco. Sono elencati solo i cluster che soddisfano i requisiti.

   oppure

   Seleziona Crea un cluster e quindi scegli le opzioni del cluster. Per ulteriori informazioni, consulta Requisiti del cluster.
5. Scegli un'opzione per i Gruppi di sicurezza, quindi scegli Modifica il cluster e avvia il notebook.

Eliminazione dei notebook e dei relativi file

Quando si elimina un notebook EMR mediante la console di Amazon EMR, il notebook viene eliminato dall'elenco dei notebook disponibili. Tuttavia, i file del notebook restano in Amazon S3 e continuano ad accumulare costi di archiviazione.

Eliminazione di un notebook e rimozione dei file associati

1. Apri la console di Amazon EMR all'indirizzo https://console.aws.amazon.com/elasticmapreduce/.
2. Scegli Notebook, seleziona il notebook dall'elenco e quindi scegli Visualizza dettagli.
3. Scegli l'icona della cartella accanto a Posizione del notebook e copia l'URL, che si trova nel pattern `s3://MyNotebookLocationPath/NotebookID/`.
4. Scegliere Delete (Elimina).

   Il notebook viene rimosso dall'elenco e i dettagli del notebook non possono più essere visualizzati.
5. Per istruzioni, consulta Come eliminare cartelle da un bucket S3 nella Guida per l'utente di Amazon Simple Storage Service. Passa al bucket e alla cartella della fase 3.

   oppure

Se lo hai AWS CLI installato, apri un prompt dei comandi e digita il comando alla fine di questo paragrafo. Sostituisci il percorso Amazon S3 con quello copiato in precedenza. Assicurati che AWS CLI sia configurato con le chiavi di accesso di un utente con le autorizzazioni per eliminare la posizione Amazon S3. Per ulteriori informazioni, consulta Configurazione della AWS CLI nella Guida per l'utente di AWS Command Line Interface .

```
aws s3 rm s3://MyNotebookLocationPath/NotebookID
```

Condivisione di file del notebook

Ogni notebook EMR viene salvato su Amazon S3 come file denominato *NotebookName*.ipynb. Finché un file del notebook è compatibile con la stessa versione del notebook Jupyter su cui è basato EMR Notebooks, è possibile aprire il notebook come un notebook EMR.

Il modo più semplice per aprire un file di notebook di un altro utente consiste nel salvare il file*.ipynb di un altro utente nel file system locale, quindi utilizzare la funzione di caricamento in Jupyter e negli editor. JupyterLab

È possibile utilizzare questo processo per l'uso di EMR Notebooks condivisi da altri utenti, notebook condivisi nella community Jupyter o per ripristinare un notebook che è stato eliminato dalla console quando si dispone ancora del file del notebook.

Uso di un altro file del notebook come base per un notebook EMR

1. Prima di procedere, chiudi l'editor di notebook per tutti notebook da utilizzare, quindi arresta il notebook se è un notebook EMR.
2. Crea un notebook EMR e assegnagli un nome. Il nome assegnato al notebook sarà il nome del file da sostituire. Il nuovo nome del file deve corrispondere esattamente a questo.
3. Annota il percorso scelto per il notebook in Amazon S3. Il file sostituito si trova in una cartella con un percorso e un nome di file simile al pattern seguente: s3://*MyNotebookLocation*/*NotebookID*/*MyNotebookName*.ipynb.
4. Arresta il notebook.
5. Sostituisci il vecchio file del notebook nel percorso Amazon S3 con il nuovo file, utilizzando esattamente lo stesso nome.

Il seguente AWS CLI comando per Amazon S3 sostituisce un file salvato su un computer locale chiamato per `SharedNotebook.ipynb` un notebook EMR con il nome MyNotebook, un ID di `e-12A3BCDEFJHIJKLMNO45PQRST` e creato con specificato in `amzn-s3-demo-bucket/MyNotebooksFolder` Amazon S3. Per informazioni sull'uso della console di Amazon S3 per copiare e sostituire i file, consulta Caricamento, download e gestione di oggetti nella Guida per l'utente di Amazon Simple Storage Service.

```
aws s3 cp SharedNotebook.ipynb s3://amzn-s3-demo-bucket/
MyNotebooksFolder/-12A3BCDEFJHIJKLMNO45PQRST/MyNotebook.ipynb
```

Esempi di comandi programmatici per Notebooks EMR

Panoramica

È possibile eseguire notebook EMR con esecuzione APIs da uno script o dalla riga di comando. Quando si avvia, si interrompe, si elencano e si descrivono le esecuzioni dei notebook EMR al di fuori della AWS console, è possibile controllare a livello di programmazione un notebook EMR. Puoi trasmettere diversi valori di parametro a un notebook con una cella di notebook parametrizzata. Questa opzione elimina la necessità di creare una copia del notebook per ogni nuovo set di valori di parametro. Per ulteriori informazioni, consulta le operazioni dell'API di Amazon EMR.

Puoi pianificare o raggruppare in batch le esecuzioni di notebook EMR con Amazon CloudWatch Events e. AWS Lambda Per ulteriori informazioni, consulta Using AWS Lambda with Amazon CloudWatch Events.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Autorizzazioni di ruolo per l'esecuzione a livello di programmazione

Per utilizzare l'esecuzione a livello di programmazione con Notebook EMR, è necessario configurare le autorizzazioni utente con le policy seguenti:

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"
      ]
    }
  ]
}
```

Quando esegui Notebook EMR a livello di programmazione su un cluster Notebook EMR, devi aggiungere queste autorizzazioni supplementari:

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"
      ]
    }
  ]
}
```

Limitazioni relative all'esecuzione a livello di programmazione

- Sono supportate un massimo di 100 esecuzioni simultanee per account. Regione AWS
- Un'esecuzione viene terminata se dura più di 30 giorni.

- L'esecuzione programmatica dei notebook non è supportata con le applicazioni interattive di Amazon EMR serverless.

Esempi di esecuzione di notebook EMR a livello di programmazione

Le seguenti sezioni forniscono diversi esempi di esecuzione programmatica di notebook EMR con AWS CLI Boto3 SDK (Python) e Ruby:

- Esempi di comandi CLI per notebook in EMR Studio
- Esempi di Python per un notebook EMR
- Esempi di Ruby per un notebook EMR

Puoi anche eseguire notebook parametrizzati nell'ambito dei flussi di lavoro pianificati con uno strumento di orchestrazione come Apache Airflow o Amazon Managed Workflows per Apache Airflow (MWAA). Per ulteriori informazioni, consulta Orchestrazione dei processi di analisi su Notebook EMR con MWAA nel blog AWS Big Data.

Esempi di comandi CLI per notebook in EMR Studio

Questo argomento mostra esempi di comandi CLI per un notebook EMR. L'esempio utilizza il notebook demo della console EMR Notebooks. Per individuare il notebook, utilizza il percorso di file relativo alla home directory. In questo esempio, ci sono due file di notebook che puoi eseguire: `demo_pyspark.ipynb` e `my_folder/python3.ipynb`.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Il percorso relativo del file `demo_pyspark.ipynb` è `demo_pyspark.ipynb`, mostrato di seguito.

Il percorso relativo per `python3.ipynb` è `my_folder/python3.ipynb`, mostrato di seguito.

Per ulteriori informazioni sulle operazioni `NotebookExecution` dell'API di Amazon EMR, consulta Operazioni dell'API di Amazon EMR.

Esecuzione di un notebook

Puoi usare il AWS CLI per far funzionare il tuo notebook con l'`start-notebook-execution`azione, come dimostrano gli esempi seguenti.

Example — Esecuzione di un notebook EMR in un ambiente di lavoro EMR Studio con un cluster Amazon EMR (in esecuzione su 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 - Esecuzione di un notebook EMR in un Workspace EMR Studio con un cluster Notebook 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 - Esecuzione di un notebook EMR specificando la sua posizione 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"}'
```

Output del notebook

Ecco l'output di un notebook campione. La cella 3 mostra i valori di parametro appena inseriti.

Descrizione di un notebook

Puoi utilizzare l'operazione `describe-notebook-execution` per accedere alle informazioni sull'esecuzione di un notebook specifico.

```
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": []
    }
}
```

Arrestare un notebook

Se il notebook sta eseguendo un'esecuzione che desideri arrestare, puoi farlo con il 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": []
    }
```

```
}
```

Elencazione delle esecuzioni di un notebook per ora di inizio

Puoi trasmettere un parametro `--from` a `list-notebook-executions` per elencare le esecuzioni del notebook in base all'ora di inizio.

```
# 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
        }
    ]
}
```

Elencazione delle esecuzioni di un notebook per ora di inizio e stato

Il comando `list-notebook-executions` può anche filtrare i risultati in base a un parametro `--status`.

```
# 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
        }
    ]
}
```

Esempi di Python per un notebook EMR

Questo argomento contiene un file di comandi di esempio. L'esempio di codice è un file SDK for Python (Boto3) chiamato. `demo.py` Mostra l'esecuzione del notebook. APIs

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Per ulteriori informazioni sulle operazioni `NotebookExecution` dell'API di Amazon EMR, consulta Operazioni dell'API di 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")
```

Quello mostrato è l'output dall'esecuzione di `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}}
```

Esempi di Ruby per un notebook EMR

Questo argomento contiene un esempio di Ruby che illustra la funzionalità dei notebook.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori

autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

I seguenti esempi di codice Ruby dimostrano l'utilizzo dell'API di esecuzione dei notebook.

```
# prepare an Amazon EMR client

emr = Aws::EMR::Client.new(
  region: 'us-east-1',
  access_key_id: 'AKIA...JKPKA',
  secret_access_key: 'rLMeu...vU0OLrAC1',
)
```

Avvio dell'esecuzione di notebook e ottenimento dell'ID di esecuzione

In questo esempio, l'editor Amazon S3 e il notebook EMR sono `s3://amzn-s3-demo-bucket/notebooks/e-EA8VGAA429FEQTC8HC9ZHWISK/test.ipynb`.

Per ulteriori informazioni sulle operazioni `NotebookExecution` dell'API di Amazon EMR, consulta Operazioni dell'API di 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
```

Descrizione dell'esecuzione di notebook e stampa dei dettagli

```
describe_resp = emr.describe_notebook_execution({
    notebook_execution_id: notebook_execution_id
})
puts describe_resp.notebook_execution
```

L'output dei comandi precedenti sarà il seguente.

```
{
: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=>[]
}
```

Filtri del notebook

```
"EditorId": "e-XXXX",           [Optional]
"From" : "1593400000.000",    [Optional]
"To" :
```

Arresto dell'esecuzione di notebook

```
stop_resp = emr.stop_notebook_execution({
    notebook_execution_id: notebook_execution_id
})
```

Abilitazione della rappresentazione utente per monitorare l'attività dell'utente e dei processi Spark

EMR Notebooks consente di configurare la rappresentazione degli utenti in un cluster Spark. Questa funzionalità consente di monitorare le attività dei processi avviati dall'interno dell'editor di notebook. Inoltre, Notebook EMR dispone di un widget di notebook Jupyter integrato per visualizzare i dettagli del processo Spark insieme all'output delle query nell'editor di notebook. Il widget è disponibile per

impostazione predefinita e non richiede alcuna configurazione speciale. Tuttavia, per visualizzare i server della cronologia, il client deve essere configurato per visualizzare le interfacce Web Amazon EMR ospitate nel nodo primario.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Impostazione della rappresentazione utente Spark

Per impostazione predefinita, i processi Spark inviati dagli utenti mediante l'editor di notebook risultano provenienti da un'identità utente `livy` indistinta. È possibile configurare la rappresentazione utente per il cluster in modo che tali processi siano associati piuttosto all'identità utente che ha eseguito il codice. Le directory utente HDFS nel nodo primario vengono create per ciascuna identità utente che esegue codice nel notebook. Ad esempio, se l'utente `NbUser1` esegue il codice dall'editor di notebook, puoi collegarti al nodo primario e vedere che `hadoop fs -ls /user` mostra la directory `/user/user_NbUser1`.

Puoi abilitare questa funzionalità impostando le proprietà nelle classificazioni di configurazione `core-site` e `livy-conf`. Questa funzionalità non è disponibile per impostazione predefinita quando Amazon EMR crea un cluster insieme a un notebook. Per ulteriori informazioni su come utilizzare le classificazioni di configurazione per personalizzare le applicazioni, consulta Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR.

Utilizza i seguenti valori e classificazioni di configurazione per abilitare la rappresentazione degli utenti per EMR Notebooks:

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

```
    },
    {
        "Classification": "livy-conf",
        "Properties": {
          "livy.impersonation.enabled": "true"
        }
    }
]
```

Utilizzo del widget di monitoraggio dei processi Spark

Quando esegui codice nell'editor di notebook che esegue i processi Spark nel cluster EMR, l'output include un widget di notebook Jupyter per il monitoraggio dei processi Spark. Il widget fornisce i dettagli del processo e collegamenti utili alla pagina dei server della cronologia di Spark e alla pagina della cronologia di Hadoop, insieme a collegamenti ai log dei processi in Amazon S3 per tutti i processi non riuscito.

Per visualizzare le pagine dei server della cronologia nel nodo primario del cluster, devi configurare un client SSH e un proxy adeguati. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR. Per visualizzare i log nel cluster Amazon S3 la registrazione deve essere abilitata, il che corrisponde all'impostazione predefinita per i nuovi cluster. Per ulteriori informazioni, consulta Visualizzazione dei file di log archiviati in Amazon S3.

Di seguito è riportato un esempio del monitoraggio dei processi Spark.

Sicurezza e controllo dell'accesso a EMR Notebooks

Sono disponibili alcune caratteristiche per aiutarti a personalizzare l'assetto di sicurezza di EMR Notebooks. Ciò contribuisce a garantire che solo gli utenti autorizzati possano accedere a un notebook EMR, lavorare con i notebook e utilizzare l'editor di notebook per l'esecuzione del codice nel cluster. Queste funzionalità si accompagnano alle funzionalità di sicurezza disponibili per Amazon EMR e per i cluster Amazon EMR. Per ulteriori informazioni, consulta Sicurezza in Amazon EMR.

- Puoi utilizzare le dichiarazioni AWS Identity and Access Management politiche insieme ai tag dei notebook per limitare l'accesso. Per ulteriori informazioni, consultare Funzionamento di Amazon EMR con IAM e Dichiarazioni di policy basate su identità di esempio per EMR Notebooks.
- I gruppi EC2 di sicurezza di Amazon agiscono come firewall virtuali che controllano il traffico di rete tra l'istanza principale del cluster e l'editor del notebook. Puoi utilizzare le impostazioni predefinite

o personalizzare questi gruppi di sicurezza. Per ulteriori informazioni, consulta Specificazione dei gruppi EC2 di sicurezza per i notebook EMR.

- Si specifica un ruolo AWS di servizio che determina le autorizzazioni di cui dispone un notebook EMR quando interagisce con altri servizi. AWS Per ulteriori informazioni, consulta Ruolo di servizio per EMR Notebooks.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Installazione e utilizzo di kernel e librerie in EMR Studio

Ogni notebook EMR viene fornito con un set di librerie e kernel preinstallati. Puoi installare librerie e kernel aggiuntivi in un cluster EMR se il cluster dispone dell'accesso al repository in cui si trovano i kernel e le librerie. Ad esempio, per i cluster in sottoreti private, potrebbe essere necessario configurare Network Address Translation (NAT) e fornire un percorso per il cluster per accedere al repository PyPI pubblico per installare una libreria. Per ulteriori informazioni sulla configurazione dell'accesso esterno per diverse configurazioni di rete, consulta Scenari ed esempi nella Guida per l'utente di Amazon VPC.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Le applicazioni EMR Serverless sono dotate delle seguenti librerie preinstallate per Python e: PySpark

- Librerie Python – ggplot, matplotlib, numpy, pandas, plotly, bokeh, scikit-learn, scipy, scipy
- PySpark librerie —ggplot,,matplotlib,,numpy,,pandas, plotly bokeh scikit-learn scipy scipy

Installazione di kernel e librerie Python su un nodo primario del cluster

Con Amazon EMR versione 5.30.0 e successive, esclusa la versione 6.0.0, puoi installare librerie e kernel Python aggiuntivi sul nodo primario del cluster. Dopo l'installazione, questi kernel e librerie sono disponibili a ogni utente che esegue un notebook EMR collegato al cluster. Le librerie Python installate in questo modo sono disponibili solo per i processi in esecuzione sul nodo primario. Le librerie non sono installate nei nodi principali o attività e non sono disponibili per gli executor in esecuzione su tali nodi.

> **Note**
>
> Per le versioni Amazon EMR 5.30.1, 5.31.0 e 6.1.0, devi eseguire ulteriori passaggi per installare kernel e librerie nel nodo primario di un cluster.
> Per abilitare questa funzione, procedi come segue:
>
> 1. Assicurati che le policy di autorizzazione associate al ruolo di servizio per EMR Notebooks consentano l'operazione seguente:
>
>   `elasticmapreduce:ListSteps`
>
>   Per ulteriori informazioni, consulta il Ruolo di servizio per EMR Notebooks.
> 2. Utilizzare AWS CLI per eseguire un passaggio sul cluster che configura EMR Notebooks, come illustrato nell'esempio seguente. È necessario utilizzare il nome della fase `EMRNotebooksSetup`. Sostituisci *`us-east-1`* con la regione in cui risiede il cluster. Per ulteriori informazioni, consulta Aggiunta di fasi a un cluster con 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"]
>    ```

Puoi installare kernel e librerie eseguendo `pip` o `conda` nella directory `/emr/notebook-env/bin` sul nodo primario.

Example - Installazione di librerie Python

Dal kernel Python3, esegui il `%pip` magic come un comando dall'interno di una cella del notebook per installare librerie Python.

```
%pip install pmdarima
```

Potrebbe essere necessario riavviare il kernel per utilizzare i pacchetti aggiornati. Puoi anche utilizzare `%%sh` magic Spark per invocare `pip`.

```
%%sh
/emr/notebook-env/bin/pip install -U matplotlib
/emr/notebook-env/bin/pip install -U pmdarima
```

Quando si utilizza un PySpark kernel, è possibile installare le librerie sul cluster utilizzando `pip` i comandi oppure utilizzare librerie con ambito notebook dall'interno di un notebook. PySpark

Per eseguire comandi `pip` sul cluster dal terminale, prima esegui la connessione al nodo primario utilizzando SSH, come illustrato nei seguenti comandi.

```
sudo pip3 install -U matplotlib
sudo pip3 install -U pmdarima
```

In alternativa, puoi utilizzare librerie con ambito notebook. Con le librerie con ambito notebook, l'installazione della libreria è limitata all'ambito della sessione e si verifica su tutti gli executor Spark. Per ulteriori informazioni, consulta Utilizzo di librerie con ambito notebook.

Se vuoi impacchettare più librerie Python all'interno di un PySpark kernel, puoi anche creare un ambiente virtuale Python isolato. Per esempi di utilizzo, consulta Utilizzo di Virtualenv.

Per creare un ambiente virtuale Python in una sessione, utilizza la proprietà Spark `spark.yarn.dist.archives` dal comando `%%configure` magic nella prima cella di un notebook, come illustrato nell'esempio seguente.

```
%%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"
   }
}
```

Puoi creare allo stesso modo un ambiente executor 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"
   }
}
```

Puoi utilizzare anche `conda` per installare librerie Python. Per utilizzare `conda` non è necessario l'accesso sudo. Devi connetterti al nodo primario con SSH e quindi eseguire `conda` dal terminale. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

Example - Installazione di kernel

L'esempio seguente illustra l'installazione del kernel Kotlin utilizzando un comando del terminale mentre è connesso al nodo primario di un cluster:

```
sudo /emr/notebook-env/bin/conda install kotlin-jupyter-kernel -c jetbrains
```

> **Note**
>
> Queste istruzioni non prevedono l'installazione delle dipendenze del kernel. Se il kernel ha dipendenze di terze parti, potrebbe essere necessario eseguire ulteriori passaggi di configurazione prima di poterlo utilizzare con il notebook.

Considerazioni e limitazioni relative alle librerie con ambito notebook

Quando utilizzi le librerie con ambito notebook, tieni presente quanto segue:

- Le librerie con ambito notebook sono disponibili con i cluster creati utilizzando le versioni 5.26.0 e successive di Amazon EMR.
- Le librerie con ambito notebook sono pensate per essere utilizzate solo con il kernel. PySpark
- Qualsiasi utente può installare ulteriori librerie con ambito notebook dall'interno di una cella di notebook. Queste librerie sono disponibili solo per tale utente del notebook durante una singola sessione di notebook. Se altri utenti richiedono le stesse librerie o lo stesso utente richiede le stesse librerie in una sessione diversa, è necessario reinstallare la libreria.
- È possibile disinstallare solo le librerie che sono state installate con l'API `install_pypi_package`. Non è possibile disinstallare le librerie preinstallate nel cluster.
- Se nel cluster sono installate versioni diverse delle stesse librerie come librerie con ambito notebook, la versione della libreria con ambito notebook sostituisce la versione della libreria del cluster.

Lavorare con le librerie con ambito notebook

Per installare le librerie, il cluster Amazon EMR deve avere accesso al repository PyPI in cui queste si trovano.

Gli esempi seguenti mostrano semplici comandi per elencare, installare e disinstallare le librerie dall'interno di una cella del notebook utilizzando il kernel e. PySpark APIs Per altri esempi, consulta il post Installare le librerie Python su un cluster in esecuzione con EMR Notebooks sul blog Big Data. AWS

Example - Elenco delle librerie correnti

Il comando riportato di seguito elenca i pacchetti Python disponibili per la sessione Spark corrente del notebook. Questo comando elenca le librerie installate nel cluster e le librerie con ambito notebook.

```
sc.list_packages()
```

Example - Installazione della libreria Celery

Il comando seguente installa la libreria Celery come libreria con ambito notebook.

```
sc.install_pypi_package("celery")
```

Dopo l'installazione della libreria, il comando seguente conferma che la libreria è disponibile sul driver e sugli executor Spark.

```
import celery
sc.range(1,10000,1,100).map(lambda x: celery.__version__).collect()
```

Example - Installazione della libreria Arrow, specifica della versione e del repository

Il comando seguente installa la libreria Arrow come libreria con ambito notebook, con una specifica dell'URL del repository e della versione della libreria.

```
sc.install_pypi_package("arrow==0.14.0", "https://pypi.org/simple")
```

Example - Disinstallazione di una libreria

Il comando seguente disinstalla la libreria Arrow rimuovendola come libreria con ambito notebook dalla sessione corrente.

```
sc.uninstall_package("arrow")
```

Associazione di repository basati su Git con EMR Notebooks

È possibile associare i repository basati su Git ai notebook Amazon EMR per salvare i notebook in un ambiente controllato dalla versione. È possibile associare fino a tre repository a un notebook. Sono supportati i seguenti servizi basati su Git:

- AWS CodeCommit
- GitHub
- Bitbucket
- GitLab

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

L'associazione di repository basati su GIT al notebook presenta i seguenti vantaggi.

- Controllo della versione: è possibile registrare le modifiche al codice in un sistema di controllo della versione in modo da poter rivedere la cronologia delle modifiche e annullarle selettivamente.
- Collaborazione: i colleghi che lavorano in notebook diversi possono condividere il codice tramite i repository basati su Git. I notebook possono clonare o unire il codice proveniente da repository remoti e inviare nuovamente le modifiche agli stessi.
- Riutilizzo del codice: molti notebook Jupyter che dimostrano tecniche di analisi dei dati o di apprendimento automatico sono disponibili in repository ospitati pubblicamente, come. GitHub Puoi associare i notebook a un repository per riutilizzare i notebook Jupyter contenuti in un repository.

Per utilizzare i repository basati su Git con EMR Notebooks, aggiungi i repository come risorse nella console di Amazon EMR, associa le credenziali per i repository che richiedono l'autenticazione e collegali ai notebook. È possibile visualizzare un elenco di repository che sono archiviati nell'account e informazioni dettagliate su ogni repository nella console di Amazon EMR. Puoi associare un repository basato su Git esistente a un notebook quando lo crei.

Argomenti

- Prerequisiti e considerazioni per l'integrazione di un notebook EMR con un repository
- Aggiunta di un repository basato su Git ad Amazon EMR
- Aggiorna o elimina un repository basato su Git da un EMR Studio Workspace
- Collega o scollega un repository basato su Git in EMR Studio
- Crea un nuovo Notebook con un repository Git associato in EMR Studio
- Utilizzare gli archivi Git in un notebook EMR Studio

Prerequisiti e considerazioni per l'integrazione di un notebook EMR con un repository

Prendi in considerazione le seguenti best practice relative a commit, autorizzazioni e hosting quando pianifichi di integrare un repository basato su Git con EMR Notebooks.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere

ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

AWS CodeCommit

Se usi un CodeCommit repository, devi usare le credenziali Git e HTTPS con. CodeCommit Le chiavi SSH e l'HTTPS con l'helper per le AWS CLI credenziali non sono supportati. CodeCommit non supporta i token di accesso personali (). PATs Per ulteriori informazioni, consulta Utilizzo di IAM con CodeCommit: credenziali Git, chiavi SSH e chiavi di AWS accesso nella Guida utente IAM e Configurazione per utenti HTTPS che utilizzano credenziali Git nella Guida per l'AWS CodeCommit utente.

Considerazioni su accesso e autorizzazione

Prima di associare un repository al notebook, assicurati che il cluster, il ruolo IAM per EMR Notebooks e i gruppi di sicurezza dispongano delle impostazioni e delle autorizzazioni corrette. È inoltre possibile configurare i repository basati su Git ospitati in una rete privata seguendo le istruzioni riportate in Configurazione di un repository Git ospitato privatamente per EMR Notebooks.

- Accesso a Internet del cluster: l'interfaccia di rete avviata dispone di solo un indirizzo IP privato. Ciò significa che il cluster a cui il notebook si connette deve trovarsi in una sottorete privata con un gateway Network Address Translation (NAT) o deve essere in grado di accedere a Internet attraverso un gateway privato virtuale. Per ulteriori informazioni, consulta la sezione relativa alle Opzioni di Amazon VPC.

  I gruppi di sicurezza del notebook devono includere una regola in uscita che consenta al notebook di instradare il traffico a Internet dal cluster. È consigliabile creare gruppi di sicurezza personali. Per ulteriori informazioni, vedere Specificazione dei gruppi EC2 di sicurezza per i notebook EMR.

  > **Important**
  >
  > Se l'interfaccia di rete viene avviata in una sottorete pubblica, non sarà in grado di comunicare con Internet tramite un gateway Internet (IGW).

- Autorizzazioni per AWS Secrets Manager: se si utilizza Secrets Manager per archiviare segreti utilizzati per accedere a un repository, è the section called “Ruolo EMR

Notebooks" necessario allegare una politica di autorizzazioni che consenta l'azione. `secretsmanager:GetSecretValue`

Configurazione di un repository Git ospitato privatamente per EMR Notebooks

Utilizza le istruzioni seguenti per configurare repository ospitati privatamente per EMR Notebooks. È necessario fornire un file di configurazione con informazioni sui server DNS e Git. Amazon EMR utilizza queste informazioni per configurare EMR Notebooks in grado di instradare il traffico ai repository ospitati privatamente.

Prerequisiti

Prima di configurare un repository Git ospitato privatamente per EMR Notebooks, devi disporre di quanto segue:

- Una Amazon S3 Control posizione in cui verranno salvati i file del notebook EMR.

Configurazione di uno o più repository Git ospitati privatamente per EMR Notebooks

1. Crea un file di configurazione utilizzando il modello fornito. Includi i seguenti valori per ogni server Git che desideri specificare nella configurazione:

   - **`DnsServerIpV4`**- L' IPv4 indirizzo del tuo server DNS. Se si forniscono valori per `DnsServerIpV4` e `GitServerIpV4List`, il valore per `DnsServerIpV4` ha la precedenza e verrà utilizzato per risolvere il `GitServerDnsName`.

     > **Note**
     >
     > Per utilizzare repository Git ospitati privatamente, il server DNS deve consentire l'accesso in ingresso da EMR Notebooks. Si consiglia di proteggere il server DNS da altri accessi non autorizzati.

   - **`GitServerDnsName`**: il nome DNS del server Git. Ad esempio `"git.example.com"`.
   - **`GitServerIpV4List`**- Un elenco di IPv4 indirizzi che appartengono ai tuoi server 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. Salva il file di configurazione come `configuration.json`.
3. Carica il file di configurazione nel percorso di archiviazione Amazon S3 designato in una cartella denominata `life-cycle-configuration`. Ad esempio, se il percorso S3 predefinito è `s3://amzn-s3-demo-bucket/notebooks`, il file di configurazione dovrebbe trovarsi in `s3://amzn-s3-demo-bucket/notebooks/`**`life-cycle-configuration/configuration.json`**.

> **Important**
>
> Si consiglia fortemente di limitare l'accesso alla cartella `life-cycle-configuration` solo agli amministratori di EMR Notebooks e al ruolo di servizio per EMR Notebooks. Dovresti inoltre proteggere `configuration.json` contro l'accesso non autorizzato. Per istruzioni, consulta Controllo dell'accesso a un bucket con policy utente o Best practice di sicurezza per Amazon S3.

Per istruzioni sul caricamento, consulta Creazione di una cartella e Caricamento degli oggetti nella Guida per l'utente di Amazon Simple Storage Service.

Aggiunta di un repository basato su Git ad Amazon EMR

Fai riferimento alle seguenti sezioni per informazioni su come aggiungere un repository basato su Git a un notebook EMR nella vecchia console o a un EMR Studio Workspace nella console.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Console

Poiché i notebook EMR sono Workspace EMR Studio nella nuova console, puoi seguire le istruzioni riportate in Collegamento di repository basati su Git a un WorkSpace EMR Studio per associare fino a tre repository Git al tuo Workspace.

In alternativa, puoi utilizzare l'estensione JupyterLab Git. Scegli l'icona Git dalla barra laterale sinistra del notebook Jupyterlab per accedere all'estensione. Per informazioni sull'estensione, consulta il repository jupyterlab-git. GitHub

Per associare un repository Git a un Workspace, l'amministratore dello Studio deve adottare misure per configurare lo Studio in modo da consentire il collegamento del repository Git. Per ulteriori informazioni, consulta Definizione di accesso e autorizzazioni per i repository basati su Git.

Aggiorna o elimina un repository basato su Git da un EMR Studio Workspace

Fai riferimento alle seguenti sezioni per informazioni su come eliminare un repository basato su Git da un notebook EMR nella vecchia console o da un EMR Studio Workspace nella console.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere

ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Console

Poiché i Notebook EMR sono Workspace EMR Studio nella nuova console, puoi fare riferimento a Collegamento di repository basati su Git a un WorkSpace EMR Studio per avere ulteriori informazioni sull'uso dei repository Git nel tuo Workspace. Al momento, tuttavia, non puoi eliminare i repository Git dai Workspace.

Collega o scollega un repository basato su Git in EMR Studio

Usa i seguenti passaggi per collegare o scollegare un repository basato su Git a un notebook EMR nella vecchia console o a un EMR Studio Workspace nella console.

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Console

Poiché i Notebook EMR sono Workspace EMR Studio nella nuova console, puoi fare riferimento a Collegamento di repository basati su Git a un WorkSpace EMR Studio per avere ulteriori informazioni sull'uso dei repository Git nel tuo Workspace. Al momento, tuttavia, non puoi eliminare i repository Git dai Workspace.

Informazioni sullo stato del repository

Un repository Git può presentare uno dei seguenti stati nell'elenco dei repository. Per ulteriori informazioni sul collegamento dei EMR Notebooks con i repository Git, consulta Collega o scollega un repository basato su Git in EMR Studio.

Stato	Significato
Collegamento	Il repository Git è in fase di collegamento al notebook. Mentre il repository è nello stato Collegamento, non è possibile arrestarlo.
Collegato	Il repository Git è collegato al notebook. Quando è nello stato Collegato, il repository è collegato al repository remoto.
Collegamento non riuscito	Il repository Git non è riuscito a collegarsi al notebook. È possibile riprovare a collegarlo.
Scollegamento	Il repository Git è in fase di scollegamento dal notebook. Mentre il repository è nello stato Scollegamento, non è possibile arrestarlo. Lo scollegamento di un repository Git da un notebook lo disconnette solo dal repository remoto; non elimina alcun codice dal notebook.
Scollegamento non riuscito	Il repository Git non è riuscito a scollegarsi dal notebook. È possibile riprovare a scollegarlo.

Crea un nuovo Notebook con un repository Git associato in EMR Studio

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Creazione di un notebook e relativa associazione ai repository Git nella vecchia console Amazon EMR

1. Segui le istruzioni riportate in Creare un notebook in EMR Studio.
2. Per Gruppo di sicurezza, scegli Utilizza gruppo di sicurezza personale.

> **Note**
>
> I gruppi di sicurezza del notebook devono includere una regola in uscita per consentire al notebook di instradare il traffico a Internet attraverso il cluster. È consigliabile creare gruppi di sicurezza personali. Per ulteriori informazioni, vedere Specificazione dei gruppi EC2 di sicurezza per i notebook EMR.

3. Per Repository Git, seleziona Scegli repository per selezionare il repository da associare al notebook.
    1. Scegli un repository archiviato come risorsa nel tuo account e quindi Salva.
    2. Per aggiungere un nuovo repository come risorsa nell'account, scegli Aggiungi un nuovo repository. Completa il flusso di lavoro Aggiungi repository in una nuova finestra.

Utilizzare gli archivi Git in un notebook EMR Studio

> **Note**
>
> I notebook EMR sono disponibili come aree di lavoro EMR Studio nella console. Il pulsante Crea area di lavoro nella console consente di creare nuovi notebook. Per accedere ai Workspace o crearne di nuovi, gli utenti di Notebook EMR necessitano di ulteriori autorizzazioni per i ruoli IAM. Per ulteriori informazioni, consulta Amazon EMR Notebooks are Amazon EMR Studio Workspace nella console e nella console Amazon EMR.

Puoi scegliere di aprire JupyterLab o Apri in Jupyter quando apri un notebook.

Se scegli di aprire il notebook in Jupyter, viene visualizzato un elenco di file e cartelle espandibili all'interno del notebook. È possibile eseguire manualmente comandi Git come il seguente in una cella di notebook.

```
!git pull origin primary
```

Per aprire uno dei repository aggiuntivi, passa ad altre cartelle.

Se scegli di aprire il notebook con un' JupyterLab interfaccia, puoi usare l'estensione JupyterLab Git preinstallata. Per avere informazioni sull'estensione, consulta jupyterlab-git.

Pianifica, configura e avvia i cluster Amazon EMR

Questa sezione illustra le opzioni di configurazione e le istruzioni per la pianificazione, la configurazione e l'avvio di cluster tramite Amazon EMR. Prima di avviare un cluster, puoi definire le scelte riguardanti il sistema in base ai dati che stai elaborando e alle tue esigenze in termini di velocità, capacità, disponibilità, sicurezza e gestibilità. Le opzioni disponibili includono:

- La regione in cui eseguire un cluster, dove e come archiviare i dati e quale output preferire per i risultati. Consultare Configurazione della posizione e dello storage dei dati del cluster Amazon EMR.
- Sia che si eseguano i cluster Amazon EMR su Outposts o Local Zones. Vedi Cluster EMR attivi AWS Outposts o Cluster EMR su Local Zones AWS.
- Se un cluster è a lunga durata o transitoria e quale software vi viene eseguito. Consulta Configurazione di un cluster Amazon EMR per continuare o terminare dopo l'esecuzione delle fasi e Configura le applicazioni all'avvio del cluster Amazon EMR.
- Se un cluster dispone di un singolo nodo primario o tre nodi primari. Consultare Pianifica e configura i nodi primari nel tuo cluster Amazon EMR.
- Le opzioni hardware e di rete che ottimizzano costi, prestazioni e disponibilità per la tua applicazione. Consultare Configurazione dell'hardware e della rete del cluster Amazon EMR.
- Come configurare i cluster in modo da poter gestire più facilmente e monitorare le attività, le prestazioni e lo stato. Consulta Configurazione del logging e del debug dei cluster Amazon EMR e Etichetta e classifica le risorse del cluster Amazon EMR.
- Come autenticare e autorizzare l'accesso alle risorse del cluster e come crittografare i dati. Consultare Sicurezza in Amazon EMR.
- Come eseguire l'integrazione con altri software e servizi. Consultare Integrazione di driver e applicazioni di terze parti su Amazon EMR.

Avvia rapidamente un cluster Amazon EMR

Segui questi passaggi per avviare un cluster Amazon EMR in pochi minuti.

Per avviare rapidamente un cluster con la console

1. Accedi a e apri la console Amazon EMR nei https://console.aws.amazon.com/emr/ cluster. AWS Management Console

2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Nella pagina Create Cluster (Crea cluster), inserisci o seleziona i valori per i campi forniti. Il pannello di riepilogo persistente consente di visualizzare in tempo reale le opzioni del cluster attualmente selezionate. Seleziona un'intestazione nel pannello di riepilogo per accedere alla sezione corrispondente e apportare modifiche. Il nome del cluster non può contenere i caratteri <, >, $, | o `(backtick). È necessario completare tutte le configurazioni richieste prima di poter scegliere Create cluster (Crea cluster).
4. Scegli Create cluster (Crea cluster) per accettare la configurazione come mostrato.
5. Si apre la pagina dei dettagli del cluster. Cerca l'indicazione Stato accanto al nome del cluster. Lo stato dovrebbe cambiare da Starting (Avvio in corso) a Running (In esecuzione) a Waiting (In attesa) durante il processo di creazione del cluster. Potrebbe essere necessario scegliere l'icona di aggiornamento in alto a destra o aggiornare il browser per ricevere gli aggiornamenti.

   Quando lo stato cambia in Waiting (In attesa), il cluster è attivo, in esecuzione e pronto per accettare fasi e connessioni SSH.

Configurazione della posizione e dello storage dei dati del cluster Amazon EMR

Questa sezione descrive come configurare la regione per un cluster, i diversi file system disponibili per l'utilizzo con Amazon EMR e come utilizzarli. Descrive inoltre come preparare o caricare dati su Amazon EMR, se necessario, nonché come preparare un percorso di output per i file di log e qualsiasi file di dati di output configurato.

Argomenti

- Scegli una AWS regione per il tuo cluster Amazon EMR
- Utilizzo di sistemi di storage e file con Amazon EMR
- Prepara i dati di input per l'elaborazione con Amazon EMR
- Configurare una posizione per l'output del cluster Amazon EMR

Scegli una AWS regione per il tuo cluster Amazon EMR

Amazon Web Services viene eseguito su server in data center di tutto il mondo. I data center sono organizzati per Regione geografica. Quando si avvia un cluster Amazon EMR, occorre specificare

una Regione. Puoi scegliere una Regione per ridurre la latenza, minimizzare i costi o rispondere ai requisiti normativi. Per un elenco di tutte le regioni e di tutti gli endpoint supportati da Amazon EMR, consulta Regioni ed endpoint in Riferimenti generali di Amazon Web Services.

Per ottenere il massimo delle prestazioni, devi avviare il cluster nella stessa Regione in cui si trovano i dati. Ad esempio, se il bucket Amazon S3 di archiviazione dei dati di input si trova nella Regione Stati Uniti occidentali (Oregon), il cluster deve essere avviato nella Regione Stati Uniti occidentali (Oregon) per evitare i costi del trasferimento di dati tra Regioni. Se utilizzi un bucket Amazon S3 per ricevere l'output del cluster, è opportuno crearlo nella Regione Stati Uniti occidentali (Oregon).

Se prevedi di associare una coppia di EC2 chiavi Amazon al cluster (necessaria per utilizzare SSH per accedere al nodo master), la coppia di chiavi deve essere creata nella stessa regione del cluster. Analogamente, i gruppi di sicurezza creati da Amazon EMR per gestire il cluster si trovano nella stessa Regione del cluster.

Se ti sei registrato a partire dal 17 maggio 2017, la regione predefinita quando accedi a una risorsa dal AWS Management Console è Stati Uniti orientali (Ohio) (us-east-2); per gli account precedenti, la regione predefinita è Stati Uniti occidentali (Oregon) (us-west-2) o Stati Uniti orientali (Virginia settentrionale) (us-east-1). Account AWS Per ulteriori informazioni, consulta Regioni ed endpoint.

Alcune funzionalità sono disponibili solo in regioni limitate. AWS Ad esempio, le istanze Cluster Compute sono disponibili solo nella Regione Stati Uniti orientali (Virginia settentrionale), mentre la Regione Asia Pacifico (Sydney) supporta solo Hadoop rilascio 1.0.3 e successivi. Quando scegli una Regione, controlla che supporti le caratteristiche che desideri utilizzare.

Per prestazioni ottimali, utilizza la stessa regione per tutte le AWS risorse che verranno utilizzate con il cluster. La tabella seguente mappa i nomi delle Regioni tra i servizi. Per un elenco delle regioni di Amazon EMR, consulta la sezione Regioni AWS ed endpoint in Riferimenti generali di Amazon Web Services.

Scelta di una Regione con la console

La tua Regione predefinita viene visualizzata a sinistra delle informazioni del tuo account nella barra di navigazione. Per cambiare Regione, sia nella nuova console sia in quella vecchia, scegli il menu a discesa Region (Regione) e seleziona una nuova opzione.

Specificare una regione con AWS CLI

Specificare una regione predefinita AWS CLI utilizzando il aws configure comando o la variabile di `AWS_DEFAULT_REGION` ambiente. Per ulteriori informazioni, vedere Configurazione della AWS regione nella Guida per l'AWS Command Line Interface utente.

Scelta di una Regione con un SDK o l'API

Per scegliere una Regione utilizzando un SDK, configura l'applicazione per utilizzare l'endpoint di tale Regione. Se stai creando un'applicazione client utilizzando un SDK AWS , puoi modificare l'endpoint client chiamando `setEndpoint`, come mostrato nell'esempio seguente:

```
client.setEndpoint("elasticmapreduce.us-west-2.amazonaws.com");
```

Dopo che l'applicazione ha specificato una regione impostando l'endpoint, è possibile impostare la zona di disponibilità per le istanze del EC2 cluster. Le zone di disponibilità sono posizioni geografiche distinte sviluppate per essere isolate dai guasti in altre zone di disponibilità e offrono una connettività di rete economica a bassa latenza ad altre zone di disponibilità nella stessa Regione. Una Regione contiene una o più zone di disponibilità. Per ottimizzare le prestazioni e ridurre la latenza, tutte le risorse devono trovarsi nella stessa zona di disponibilità del cluster che le utilizza.

Utilizzo di sistemi di storage e file con Amazon EMR

Amazon EMR e Hadoop offrono un'ampia gamma di file system che puoi utilizzare durante l'elaborazione delle fasi del cluster. Puoi specificare il file system da utilizzare tramite il prefisso dell'URI di accesso ai dati. Ad esempio, `s3://amzn-s3-demo-bucket1/path` fa riferimento a un bucket Amazon S3 utilizzando S3A (dalla versione EMR-7.10.0). La tabella seguente elenca i file system disponibili, con suggerimenti riguardo all'utilizzo più appropriato per ciascuno.

Amazon EMR e Hadoop in genere usano due o più dei seguenti file system durante l'elaborazione di un cluster. HDFS e S3A sono i due file system principali utilizzati con Amazon EMR.

> **Important**
>
> A partire dalla versione 5.22.0 di Amazon EMR, Amazon EMR AWS utilizza la versione 4 di Signature esclusivamente per autenticare le richieste verso Amazon S3. Le versioni precedenti di Amazon EMR utilizzano in alcuni casi AWS Signature Version 2, a meno che le note di rilascio non indichino che viene utilizzata esclusivamente la versione Signature 4. Per ulteriori informazioni, consulta Authenticating Requests (AWS Signature Version 4) e

Authenticating Requests (AWS Signature Version 2) nella Amazon Simple Storage Service Developer Guide.

File system	Prefix	Descrizione
HDFS	`hdfs://` (o nessun prefisso)	HDFS è un file system distribuito, scalabile e portatile per Hadoop. HDFS ha il vantaggio di garantire la consapevolezza dei dati tra i nodi del cluster Hadoop che gestiscono i cluster e i nodi del cluster Hadoop che gestiscono le singole fasi. Per ulteriori informazioni, consulta la documentazione di Hadoop. HDFS viene utilizzato per i nodi master e principal i. Uno dei vantaggi è la sua rapidità, mentre ha lo svantaggio di essere uno storage temporaneo che viene recuperato quando il cluster termina. Trova il miglior utilizzo nel caching dei risultati prodotti dalle fasi intermedie del flusso di elaborazione.
S3A	`s3://`, `s3a://`, `s3n://`	Il file system Hadoop S3A è un connettore S3 open source che consente ad Apache Hadoop e al suo ecosistema di interagire direttamente con lo storage Amazon S3. Consente agli utenti di leggere e scrivere dati su bucket S3 utilizzando operazioni di file compatibili con Hadoop, fornendo una perfetta integrazione tra le applicazioni Hadoop e lo storage cloud. **Note** Prima di EMR-7.10.0, Amazon EMR utilizzava EMRFS per lo schema s3://e s3n://.
File system locale		Il file system locale fa riferimento a un disco con connessione locale. Quando viene creato un cluster Hadoop, ogni nodo viene creato da un' EC2 istanza

File system	Prefix	Descrizione
		dotata di un blocco preconfigurato di storage su disco precollegato chiamato instance store. I dati sui volumi dell'Instance Store persistono solo per la durata dell'istanza. EC2 I volumi instance store sono ideali per lo storage di dati temporanei in continua evoluzione, come buffer, cache, dati Scratch e altri contenuti temporanei. Per ulteriori informazioni, consulta Amazon EC2 instance storage. Il file system locale viene utilizzato da HDFS, ma Python viene eseguito anche dal file system locale ed è possibile scegliere di archiviare file di applicazione aggiuntivi sui volumi archivio istanza.
(Legacy) File system a blocchi Amazon S3	`s3bfs://`	Il file system a blocchi Amazon S3 è un file storage system legacy. e ne sconsigliamo caldamente l'utilizzo. **Important** Consigliamo di non utilizzare questo file system perché può attivare una race condition che può causare errori nel cluster. Tuttavia, può essere richiesto da applicazioni legacy.

Accesso ai file system

Puoi specificare il file system da utilizzare tramite il prefisso dell'URI (Uniform Resource Identifier) di accesso ai dati. Le seguenti procedure illustrano come fare riferimento a diversi tipi di file system.

Per accedere all'HDFS locale

- Specifica il prefisso `hdfs:///` nell'URI. Amazon EMR risolve i percorsi che non specificano un prefisso nell'URI dell'HDFS locale. Ad esempio, entrambe le seguenti situazioni URIs verrebbero risolte nella stessa posizione in HDFS.

```
hdfs:///path-to-data

/path-to-data
```

Per accedere all'HDFS remoto

- Includi l'indirizzo IP del nodo master nell'URI, come illustrato negli esempi seguenti.

```
hdfs://master-ip-address/path-to-data

master-ip-address/path-to-data
```

Accesso ad Amazon S3

- Utilizza il prefisso `s3://`.

```
s3://bucket-name/path-to-file-in-bucket
```

Accesso al file system a blocchi Amazon S3

- Utilizzalo solo per le applicazioni legacy che richiedono il file system a blocchi Amazon S3. Per accedere o archiviare i dati con questo file system, utilizza il prefisso `s3bfs://` nell'URI.

  Il file system a blocchi Amazon S3 è un file system legacy utilizzato per supportare caricamenti su Amazon S3 di dimensioni superiori a 5 GB. Con la funzionalità di caricamento multiparte fornita da Amazon EMR tramite AWS Java SDK, puoi caricare file di grandi dimensioni sul file system nativo di Amazon S3 e il file system a blocchi Amazon S3 è obsoleto. Per ulteriori informazioni sul caricamento multiparte per EMR, consulta Configurare il caricamento multiparte per Amazon S3. Per ulteriori informazioni sui limiti delle dimensioni degli oggetti e delle parti di S3, consulta i limiti di caricamento multiparte di Amazon S3 nella Guida per l'utente di Amazon Simple Storage Service.

> **Warning**
>
> Poiché questo file system legacy può creare race condition che possono danneggiare il file system, è opportuno evitarlo e utilizzare invece EMRFS.

```
s3bfs://bucket-name/path-to-file-in-bucket
```

Prepara i dati di input per l'elaborazione con Amazon EMR

La maggior parte dei cluster carica i dati di input e li elabora. Per caricare i dati, deve essere in una posizione alla quale il cluster possa accedere e in un formato che il cluster possa elaborare. Lo scenario più comune è quello di caricare i dati di input in Amazon S3. Amazon EMR fornisce strumenti per il cluster per importare o leggere dati da Amazon S3.

Il formato di input predefinito in Hadoop è quello dei file di testo, anche se è possibile personalizzare Hadoop e utilizzare strumenti per importare i dati memorizzati in altri formati.

Argomenti

- Tipi di input che Amazon EMR può accettare
- Diversi modi per trasferire dati in Amazon EMR

Tipi di input che Amazon EMR può accettare

Il formato di input predefinito per un cluster è un file di testo con ogni riga separata da un carattere di newline (\n), che è il formato di input più comunemente usato.

Se i dati di input sono in un formato diverso da quello predefinito dei file di testo, è possibile utilizzare l'interfaccia di Hadoop `InputFormat` per specificare altri tipi di input. Puoi persino creare una sottoclasse della classe `FileInputFormat` per gestire i tipi di dati personalizzati. Per ulteriori informazioni, consulta http://hadoop.apache. org/docs/current/api/org/apache/hadoop/mapred/InputFormat.html.

Se si utilizza Hive, è possibile utilizzare a serializer/deserializer (SerDe) per leggere i dati da un determinato formato in HDFS. Per ulteriori informazioni, consulta https://cwiki.apache. org/confluence/display/Hive/SerDe.

Diversi modi per trasferire dati in Amazon EMR

Amazon EMR fornisce diversi modi per ottenere dati su un cluster. Il modo più comune è quello di caricare i dati in Amazon S3 e utilizzare le funzionalità integrate di Amazon EMR per caricare i dati sul cluster. È inoltre possibile utilizzare la funzionalità DistributedCache di Hadoop per trasferire i file da un file system distribuito al file system locale. L'implementazione di Hive fornita da Amazon EMR (Hive versione 0.7.1.1 e successive) include funzionalità che è possibile utilizzare per importare ed esportare i dati tra DynamoDB e un cluster Amazon EMR. Se si dispone di grandi quantità di dati da elaborare, il servizio AWS Direct Connect può essere utile.

Argomenti

- Caricamento dei dati su Amazon S3
- Carica i dati con AWS DataSync
- Importazione di file con cache distribuita con Amazon EMR
- Rilevamento ed elaborazione di file compressi con Amazon EMR
- Importa dati DynamoDB in Hive con Amazon EMR
- Connettiti ai dati con AWS Direct Connect Amazon EMR
- Carica grandi quantità di dati per Amazon EMR con AWS Snowball Edge

Caricamento dei dati su Amazon S3

Per istruzioni su come caricare oggetti su Amazon S3, consulta Aggiunta di un oggetto a un bucket nella Guida per l'utente di Amazon Simple Storage Service. Per ulteriori informazioni sull'uso di Amazon S3 con Hadoop, consulta http://wiki.apache. org/hadoop/AmazonS3.

Argomenti

- Creazione e configurazione di un bucket Amazon S3
- Configurazione del caricamento in più parti per Amazon S3
- Best practice
- Caricamento dei dati in Amazon S3 Express One Zone

Creazione e configurazione di un bucket Amazon S3

Amazon EMR utilizza il AWS SDK per Java con Amazon S3 per archiviare dati di input, file di log e dati di output. Amazon S3 fa riferimento a questi percorsi di archiviazione come bucket. I bucket presentano determinate restrizioni e limitazioni in conformità con i requisiti di Amazon S3 e DNS. Per ulteriori informazioni, consulta Restrizioni e limitazioni dei bucket nella Guida per l'utente di Amazon Simple Storage Service.

Questa sezione mostra come utilizzare Amazon S3 per creare e quindi impostare AWS Management Console le autorizzazioni per un bucket Amazon S3. È anche possibile creare e impostare autorizzazioni per un bucket Amazon S3 utilizzando l'API Amazon S3 o la AWS CLI. È anche possibile utilizzare Curl insieme a una modifica per passare i parametri di autenticazione appropriati per Amazon S3.

Consulta le seguenti risorse:

- Per creare un bucket utilizzando la console, consultare Creare un bucket nella Guida per l'utente di Amazon S3.
- Per creare e lavorare con i bucket utilizzando il AWS CLI, consulta Usare i comandi S3 di alto livello con la AWS Command Line Interface nella Amazon S3 User Guide.
- Per creare un bucket tramite un SDK, consulta Esempi di creazione di un bucket nella Guida per l'utente di Amazon Simple Storage Service.
- Per utilizzare i bucket mediante Curl, consulta Strumento di autenticazione Amazon S3 per Curl.
- Per ulteriori informazioni sulla definizione di bucket specifici per Regione, consulta Accesso a un bucket nella Guida per l'utente di Amazon Simple Storage Service.
- Per lavorare con bucket utilizzando Amazon S3 Access Points, consulta Utilizzo di un alias in stile bucket per il punto di accesso nella Guida per l'utente di Amazon S3. Puoi utilizzare facilmente Amazon S3 Access Points con l'alias di Access Point di Amazon S3 anziché il nome del bucket Amazon S3. Puoi utilizzare l'alias di Access Point di Amazon S3 per applicazioni esistenti e nuove, tra cui Spark, Hive, Presto e altre.

> **Note**
>
> Se si abilita la registrazione per un bucket, saranno abilitati solo i log di accesso al bucket e non i log del cluster Amazon EMR.

Durante o dopo la creazione del bucket, è possibile impostare le autorizzazioni appropriate per accedere al bucket a seconda dell'applicazione. In genere, consenti a te stesso (il proprietario) l'accesso in lettura e scrittura e il solo accesso in lettura agli utenti autenticati.

Per poter creare un cluster, sono necessari i bucket Amazon S3 richiesti. È necessario caricare in Amazon S3 tutti gli script e i dati a cui viene fatto riferimento nel cluster.

Configurazione del caricamento in più parti per Amazon S3

Amazon EMR supporta il caricamento multiparte di Amazon S3 tramite l'SDK AWS for Java. Il caricamento in più parti consente di caricare un singolo oggetto come un insieme di parti. È possibile caricare queste parti dell'oggetto in modo indipendente e in qualsiasi ordine. Se la trasmissione di una parte non riesce, è possibile ritrasmettere tale parte senza influire sulle altre. Una volta caricate tutte le parti dell'oggetto, Amazon S3 le assembla e crea l'oggetto.

Per ulteriori informazioni, consulta Panoramica del caricamento in più parti nella Guida per l'utente di Amazon Simple Storage Service.

Inoltre, Amazon EMR include proprietà che consentono di controllare con maggiore precisione il cleanup delle parti di un caricamento in più parti.

Nella tabella seguente vengono descritte le proprietà di configurazione di Amazon EMR per il caricamento in più parti. Per la configurazione, utilizza la classificazione di configurazione `core-site`. Per ulteriori informazioni, consulta Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR.

Nome del parametro di configurazione	Valore predefinito	Descrizione
`fs.s3n.multipart.uploads.enabled`	`true`	Un tipo booleano che indica se abilitare il caricamento in più parti. Quando la visualizzazione coerente EMRFS è abilitata, i caricamenti in più parti sono abilitati per impostazione predefinita e l'impostazione di questo valore su `false` viene ignorata.
`fs.s3n.multipart.uploads.split.size`	`134217728`	Specifica la dimensione massima di una parte, in byte, prima che EMRFS avvii il caricamento di una nuova parte quando

Nome del parametro di configurazione	Valore predefinito	Descrizione
		è abilitato il caricamento in più parti. Il valore minimo è `5242880` (5 MB). Se viene specificato un valore più basso, viene utilizzato `5242880`. Il valore massimo è `5368709120` (5 GB). Se viene specificato un valore più alto, viene utilizzato `5368709120`. Se la crittografia lato client di EMRFS è disattivata e il committer ottimizzato Amazon S3 è disabilitato, questo valore controlla anche la dimensione massima di un file di dati che può aumentare finché EMRFS utilizza i caricamenti in più parti, anziché una richiesta `PutObject` per caricare il file. Per ulteriori informazioni, consulta
`fs.s3n.ssl.enabled`	`true`	Un tipo booleano che indica se utilizzare http o https.
`fs.s3.buckets.create.enabled`	`false`	Un tipo booleano che indica se è necessario creare un bucket in caso non esista. L'impostazione su `false` causa un'eccezione nelle operazioni `CreateBucket`.
`fs.s3.multipart.clean.enabled`	`false`	Un tipo booleano che indica se abilitare il cleanup periodico in background dei caricamenti in più parti incompleti.
`fs.s3.multipart.clean.age.threshold`	`604800`	Un tipo long che specifica la durata minima, in secondi, di un caricamento in più parti, prima che venga considerato idonea per il cleanup. Il valore predefinito è una settimana.

Nome del parametro di configurazione	Valore predefinito	Descrizione
`fs.s3.multipart.clean.jitter.max`	`10000`	Un numero intero che specifica il ritardo massimo, in secondi, nel jitter casuale aggiunto al ritardo fisso di 15 minuti prima di pianificare la fase successiva di cleanup.

Disabilitazione dei caricamenti in più parti

Console

Per disabilitare i caricamenti in più parti con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Software Settings (Modifica delle impostazioni), inserisci la seguente configurazione: `classification=core-site,properties=[fs.s3n.multipart.uploads.enabled=false]`.
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per disabilitare il caricamento in più parti utilizzando il AWS CLI

Questa procedura illustra come disabilitare il caricamento in più parti utilizzando la AWS CLI. Per disabilitare il caricamento in più parti, digitare il comando `create-cluster` con il parametri `--bootstrap-actions`.

1. Crea un file, `myConfig.json`, con il seguente contenuto e salvalo nella stessa directory in cui esegui il comando:

```
[
  {
```

```
    "Classification": "core-site",
    "Properties": {
      "fs.s3n.multipart.uploads.enabled": "false"
    }
  }
]
```

2. Digita il comando seguente e *myKey* sostituiscilo con il nome della tua EC2 key pair.

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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

Disabilitazione dei caricamento in più parti con l'API

- Per informazioni sull'utilizzo dei caricamenti in più parti di Amazon S3 a livello di programmazione, consulta Utilizzo dell'SDK AWS per Java per il caricamento in più parti nella Guida per l'utente di Amazon Simple Storage Service.

  Per ulteriori informazioni sull' AWS SDK per Java, AWS consulta SDK for Java.

Best practice

Di seguito sono riportate le raccomandazioni per l'utilizzo di bucket Amazon S3 con cluster EMR.

Abilita il controllo delle versioni

La funzione Versioni multiple è una configurazione consigliata per il bucket Amazon S3. Abilitando la funzione Versioni multiple, si garantisce che anche se i dati vengono eliminati o sovrascritti

involontariamente, possono essere ripristinati. Per ulteriori informazioni, consulta Utilizzo del controllo delle versioni nella Guida per l'utente di Amazon Simple Storage Service.

Eliminazione dei caricamenti in più parti non riusciti

Per impostazione predefinita, i componenti del cluster EMR utilizzano caricamenti multiparte tramite l'SDK per AWS Java con Amazon S3 per scrivere file di log e inviare dati in uscita su Amazon APIs S3. Per ulteriori informazioni sulla modifica delle proprietà relative a questa configurazione utilizzando Amazon EMR, consulta Configurazione del caricamento in più parti per Amazon S3. A volte il caricamento di un file di grandi dimensioni può risultare in un caricamento in più parti incompleto Amazon S3. Quando un caricamento in più parti non riesce a essere completato con successo, il caricamento in più parti in corso continua a occupare il bucket e comporta costi di storage. Per evitare lo storage eccessivo dei dati, consigliamo le seguenti opzioni:

- Per i bucket utilizzati con Amazon EMR, utilizza una regola di configurazione del ciclo di vita in Amazon S3 per rimuovere i caricamenti in più parti incompleti tre giorni dopo la data di inizio del caricamento. Le regole di configurazione del ciclo di vita consentono di controllare la classe di storage e la durata degli oggetti. Per ulteriori informazioni, consulta Gestione del ciclo di vita degli oggetti e Interruzione dei caricamenti in più parti incompleti utilizzando una policy per il ciclo di vita del bucket.
- Abilita la funzione di eliminazione in più parti di Amazon EMR impostando `fs.s3.multipart.clean.enabled` su `true` e l'ottimizzazione di altri parametri di eliminazione. Questa funzione è utile in caso di volume elevato, su larga scala e con cluster con tempi di attività limitati. In questo caso, il parametro `DaysAfterIntitiation` di una regola di configurazione del ciclo di vita potrebbe essere troppo lungo, anche se impostato sul valore minimo, causando picchi di archiviazione in Amazon S3. L'eliminazione multiparte di Amazon EMR consente un controllo più preciso. Per ulteriori informazioni, consulta Configurazione del caricamento in più parti per Amazon S3.

Gestione dei contrassegni di versione

Consigliamo di abilitare una regola di configurazione del ciclo di vita in Amazon S3 per rimuovere i contrassegni di eliminazione dell'oggetto scaduto per i bucket con versione utilizzati con Amazon EMR. Quando si elimina un oggetto in un bucket con versione, viene creato un contrassegno di eliminazione. Se tutte le versioni precedenti dell'oggetto scadono successivamente, nel bucket viene lasciato un contrassegno di eliminazione dell'oggetto scaduto. Sebbene non sia previsto alcun costo per questi contrassegni di eliminazione, la loro rimozione può migliorare le prestazioni delle richieste

di LIST. Per ulteriori informazioni, consulta Configurazione del ciclo di vita per un bucket con controllo delle versioni nella Guida per l'utente di Amazon Simple Storage Service.

Best practice sulle prestazioni

A seconda dei carichi di lavoro, tipi specifici di utilizzo dei cluster EMR e delle applicazioni su di essi possono portare a un elevato numero di richieste contro una bucket. Per maggiori informazioni, consulta Considerazioni sulla percentuale di richieste e sulle prestazioni nella Guida per l'utente di Amazon Simple Storage Service.

Caricamento dei dati in Amazon S3 Express One Zone

Panoramica

Con Amazon EMR 6.15.0 e versioni successive, puoi utilizzare Amazon EMR con Apache Spark insieme alla classe di storage Amazon S3 Express One Zone per migliorare le prestazioni nei tuoi processi Spark. Le versioni 7.2.0 e successive di Amazon EMR supportano HBase anche Flink e Hive, quindi puoi trarre vantaggio anche da S3 Express One Zone se utilizzi queste applicazioni. S3 Express One Zone è una classe di storage S3 per applicazioni che accedono frequentemente ai dati con centinaia di migliaia di richieste al secondo. Al momento del suo rilascio, S3 Express One Zone offre lo storage di oggetti cloud con la latenza più bassa e le prestazioni più elevate in Amazon S3.

Prerequisiti

- Autorizzazioni S3 Express One Zone: quando S3 Express One Zone richiama inizialmente un'operazione come `GET`, `LIST` o PUT su un oggetto S3, la classe di archiviazione chiama `CreateSession` per tuo conto. La policy IAM deve consentire l'autorizzazione `s3express:CreateSession`, in modo che il connettore S3A possa richiamare l'API `CreateSession`. Per un esempio di policy con questa autorizzazione, consulta Nozioni di base su Amazon S3 Express One Zone.
- Connettore S3A: per configurare il cluster Spark per accedere ai dati da un bucket Amazon S3 che utilizza la classe di storage S3 Express One Zone, devi utilizzare il connettore Apache Hadoop S3A. Per utilizzare il connettore, assicurati che tutti gli S3 utilizzino lo schema. URIs `s3a` In caso contrario, puoi modificare l'implementazione del file system che utilizzi per gli schemi `s3` e `s3n`.

Per modificare lo schema `s3`, specifica le seguenti configurazioni del cluster:

```
[
  {
    "Classification": "core-site",
```

```
    "Properties": {
      "fs.s3.impl": "org.apache.hadoop.fs.s3a.S3AFileSystem",
      "fs.AbstractFileSystem.s3.impl": "org.apache.hadoop.fs.s3a.S3A"
    }
  }
]
```

Per modificare lo schema `s3n`, specifica le seguenti configurazioni del cluster:

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "fs.s3n.impl": "org.apache.hadoop.fs.s3a.S3AFileSystem",
      "fs.AbstractFileSystem.s3n.impl": "org.apache.hadoop.fs.s3a.S3A"
    }
  }
]
```

Nozioni di base su Amazon S3 Express One Zone

Argomenti

- Creazione di una policy di autorizzazione
- Creazione e configurazione del cluster
- Panoramica delle configurazioni

Creazione di una policy di autorizzazione

Prima di poter creare un cluster che utilizza Amazon S3 Express One Zone, devi creare una policy IAM da collegare al profilo dell' EC2 istanza Amazon per il cluster. La policy deve disporre delle autorizzazioni per accedere alla classe di storage S3 Express One Zone. La policy di esempio seguente mostra come concedere l'autorizzazione richiesta. Dopo avere creato la policy, collegala al ruolo del profilo dell'istanza utilizzato per creare il cluster EMR, come descritto nella sezione Creazione e configurazione del cluster.

JSON

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

```
  "Statement": [
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3express:*:123456789012:bucket/example-s3-bucket"
      ],
      "Action": [
        "s3express:CreateSession"
      ],
      "Sid": "AllowS3EXPRESSCreatesession"
    }
  ]
}
```

Creazione e configurazione del cluster

Successivamente, crea un cluster che esegua Spark HBase, Flink o Hive con S3 Express One Zone. I passaggi seguenti descrivono una panoramica di alto livello per creare un cluster nella AWS Management Console:

1. Accedi alla console Amazon EMR e seleziona Cluster dalla barra laterale. Quindi, scegli Crea cluster.
2. Se usi Spark, seleziona la `emr-6.15.0` versione Amazon EMR o successiva. Se usi HBase Flink o Hive, seleziona o una versione successiva. `emr-7.2.0`
3. Seleziona le applicazioni che desideri includere nel cluster, come Spark o Flink HBase.
4. Per abilitare Amazon S3 Express One Zone, inserisci una configurazione simile all'esempio seguente nella sezione Impostazioni software. Le configurazioni e i valori consigliati sono descritti nella sezione Panoramica delle configurazioni che segue questa procedura.

```
[
  {
    "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. Nella sezione Profilo dell'EC2 istanza per Amazon EMR, scegli di utilizzare un ruolo esistente e utilizza un ruolo con la policy allegata che hai creato nella Creazione di una policy di autorizzazione sezione precedente.
6. Configura il resto delle impostazioni del cluster in base all'applicazione, quindi seleziona Crea cluster.

Panoramica delle configurazioni

Le tabelle seguenti descrivono le configurazioni e i valori suggeriti da specificare quando si configura un cluster che utilizza S3 Express One Zone con Amazon EMR, come descritto nella sezione Creazione e configurazione del cluster.

Configurazioni S3A

Parametro	Valore predefinito	Valore consigliato	Spiegazione
`fs.s3a.aws.credentials.provider`	Se non specificato, viene utilizzato `AWSCredentialProviderList` nel seguente ordine: `TemporaryAWSCredentialsProvider`, `SimpleAWSCredentialsProvider`, `EnvironmentVariabl`	`software.amazon.awssdk.auth.credentials.InstanceProfileCredentialsProvider`	Il ruolo del profilo dell'istanza Amazon EMR deve avere la policy che consente al filesystem S3A di chiamare `s3express:CreateSession`. Anche altri provider di credenziali funzionano se dispongono delle autorizzazioni S3 Express One Zone.

Parametro	Valore predefinito	Valore consigliato	Spiegazione
	`eCredentialsProvider`, `IAMInstanceCredentialsProvider`.		
`fs.s3a.endpoint.region`	null	Il Regione AWS luogo in cui hai creato il bucket.	La logica di risoluzione regionale non funziona con la classe di storage S3 Express One Zone.
`fs.s3a.select.enabled`	`true`	`false`	Amazon S3 `select` non è supportato con la classe di storage S3 Express One Zone.
`fs.s3a.change.detection.mode`	`server`	nessuno	Il rilevamento delle modifiche da parte di S3A funziona tramite verifica di `etags` basati su MD5. La classe di storage S3 Express One Zone non supporta i `checksums` MD5.

Configurazioni Spark

Parametro	Valore predefinito	Valore consigliato	Spiegazione
		`false`	

Parametro	Valore predefinito	Valore consigliato	Spiegazione
`spark.sql.sources.fastS3PartitionDiscovery.enabled`	`true`		L'ottimizzazione interna utilizza un parametro API S3 non supportato dala classe di storage S3 Express One Zone.

Configurazioni Hive

Parametro	Valore predefinito	Valore consigliato	Spiegazione
`hive.exec.fast.s3.partition.discovery.enabled`	`true`	`false`	L'ottimizzazione interna utilizza un parametro API S3 non supportato dala classe di storage S3 Express One Zone.

Considerazioni

Considera quanto segue quando integri Apache Spark su Amazon EMR con la classe di storage S3 Express One Zone:

- Il connettore S3A è necessario per utilizzare S3 Express One Zone con Amazon EMR. Solo S3A dispone delle funzionalità e delle classi di storage necessarie per interagire con S3 Express One Zone. Per le fasi di configurazione del connettore, consulta the section called “Prerequisiti”.
- La classe di storage Amazon S3 Express One Zone supporta SSE-S3 e SSE-KMS crittografa. Per ulteriori informazioni, consulta Crittografia lato server con Amazon S3.
- La classe di storage Amazon S3 Express One Zone non supporta le scritture con S3A `FileOutputCommitter`. Le scritture con S3A `FileOutputCommitter` su bucket S3 Express One Zone generano un errore: InvalidStorageClass: The storage class you specified is not valid.

- Amazon S3 Express One Zone è supportato con le versioni 6.15.0 e successive di Amazon EMR su EMR on. EC2 Inoltre, è supportato nelle versioni 7.2.0 e successive di Amazon EMR su EKS e su Amazon EMR Serverless.

Carica i dati con AWS DataSync

AWS DataSync è un servizio di trasferimento dati online che semplifica, automatizza e accelera il processo di spostamento dei dati tra i servizi di archiviazione e archiviazione locali o tra i servizi di storage. AWS AWS DataSync supporta una varietà di sistemi di storage locali come Hadoop Distributed File System (HDFS), file server NAS e storage di oggetti autogestito.

Il modo più comune per inserire dati in un cluster è quello di caricare i dati in Amazon S3 e utilizzare le funzionalità integrate di Amazon EMR per caricare i dati sul cluster.

DataSync può aiutarti a svolgere le seguenti attività:

- Replica HDFS sul tuo cluster Hadoop in Amazon S3 per la continuità aziendale
- Copia HDFS in Amazon S3 per popolare i data lake
- Trasferisci i dati tra HDFS del cluster Hadoop e Amazon S3 per l'analisi e l'elaborazione

Per caricare i dati nel tuo bucket S3, devi prima implementare uno o più DataSync agenti nella stessa rete dello storage locale. Un agent (agente) è una macchina virtuale (VM) utilizzata per leggere o scrivere dati in una posizione autogestita. Quindi attivi i tuoi agenti nel Account AWS e nel luogo in Regione AWS cui si trova il tuo bucket S3.

Dopo aver attivato l'agente, crei una posizione di origine per l'archiviazione on-premise, una posizione di destinazione per il bucket S3 e un processo. Un'attivitàè costituita da un set di due percorsi (origine e destinazione) e un set di opzioni predefinite che permettono di controllarne il comportamento.

Infine, esegui la tua DataSync attività per trasferire i dati dalla sorgente alla destinazione.

Per ulteriori informazioni, consulta la pagina Nozioni di base di AWS DataSync.

Importazione di file con cache distribuita con Amazon EMR

DistributedCache è una funzionalità Hadoop in grado di aumentare l'efficienza quando una mappa o un'attività di riduzione richiede l'accesso a dati comuni. Se il cluster dipende da applicazioni

esistenti o da file binari non installati al momento della creazione del cluster, è possibile utilizzare DistributedCache per importare questi file. Questa funzione consente a un nodo del cluster di leggere i file importati dal file system locale, invece di recuperare i file da altri nodi del cluster.

Per ulteriori informazioni, consulta http://hadoop.apache. org/docs/stable/api/org/apache/hadoop/filecache/DistributedCache.html.

È possibile richiamare DistributedCache quando si crea il cluster. I file vengono messi in cache poco prima dell'avvio del lavoro Hadoop e rimangono in cache per tutta la durata del lavoro. È possibile mettere in cache i file archiviati su qualsiasi file system compatibile con Hadoop, ad esempio HDFS o Amazon S3. La dimensione predefinita della cache dei file è 10 GB. Per modificare la dimensione della cache, riconfigurare il parametro Hadoop, `local.cache.size` usando l'operazione di bootstrap. Per ulteriori informazioni, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR.

Argomenti

- Tipi di file supportati
- Percorso dei file memorizzati nella cache
- Accesso ai file della cache dalle applicazioni in streaming
- Accesso ai file della cache dalle applicazioni in streaming

Tipi di file supportati

DistributedCache consente sia singoli file che archivi. I singoli file vengono memorizzati in cache come file di sola lettura. Gli eseguibili e i file binari hanno le autorizzazioni di esecuzione impostate.

Gli archivi sono uno o più file assemblati utilizzando un'utilità, ad esempio `gzip`. DistributedCache trasferisce i file compressi in ogni nodo principale e decomprime l'archivio come parte della memorizzazione nella cache. DistributedCache supporta i seguenti formati di compressione:

- zip
- tgz
- tar.gz
- tar
- jar

Percorso dei file memorizzati nella cache

DistributedCache copia i file solo nei nodi principali. Se non sono presenti nodi principali nel cluster, DistributedCache copia i file nel nodo primario.

DistributedCache associa i file della cache alla directory di lavoro corrente del mappatore e del riduttore utilizzando collegamenti simbolici. Un collegamento simbolico è un alias a una posizione del file, non la posizione effettiva del file. Il valore del parametro, `yarn.nodemanager.local-dirs` in `yarn-site.xml`, specifica il percorso dei file temporanei. Amazon EMR imposta il parametro su `/mnt/mapred` o alcune variazioni in base al tipo di istanza e alla versione EMR. Ad esempio, un'impostazione può avere `/mnt/mapred` e `/mnt1/mapred` perché il tipo di istanza ha due volumi effimeri. I file della cache si trovano in una sottodirectory della posizione del file temporaneo in `/mnt/mapred/taskTracker/archive`.

Se si memorizza nella cache un singolo file, DistributedCache inserisce il file nella directory `archive`. Se si memorizza nella cache un archivio, DistributedCache decomprime il file e crea una sottodirectory in `/archive` con lo stesso nome del file di archivio. I singoli file si trovano nella nuova sottodirectory.

Puoi utilizzare DistributedCache solo quando utilizzi lo streaming.

Accesso ai file della cache dalle applicazioni in streaming

Per accedere ai file memorizzati nella cache dalle applicazioni mappatore o riduttore, accertati di aver aggiunto la directory di lavoro corrente (. /) nel percorso dell'applicazione e di aver fatto riferimento ai file memorizzati nella cache come se fossero presenti nella directory di lavoro corrente.

Accesso ai file della cache dalle applicazioni in streaming

È possibile utilizzare AWS Management Console and the AWS CLI per creare cluster che utilizzano Distributed Cache.

Console

Specifica dei file della cache distribuita con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.

3. In Steps (Fasi), scegli Add step (Aggiungi fase). Si apre la finestra di dialogo Add step (Aggiungi fase). Nel campo Arguments (Argomenti), includi i file e gli archivi da salvare nella cache. La dimensione del file (o la dimensione totale dei file in un file di archivio) deve essere inferiore alla dimensione della cache assegnata.

   Se desideri aggiungere un singolo file alla cache distribuita, specifica `-cacheFile` seguito dal nome e dalla posizione del file, dal segno cancelletto (#) e quindi dal nome che desideri assegnare al file quando viene collocato nella cache locale. L'esempio seguente mostra come aggiungere un singolo file alla cache distribuita.

   ```
   -cacheFile \
   s3://amzn-s3-demo-bucket/file-name#cache-file-name
   ```

   Se desideri aggiungere un file di archivio alla cache distribuita, inserisci `-cacheArchive` seguito dal percorso dei file in Amazon S3, dal segno cancelletto (#) e quindi dal nome che desideri assegnare alla raccolta di file nella cache locale. L'esempio seguente mostra come aggiungere un file di archivio alla cache distribuita.

   ```
   -cacheArchive \
   s3://amzn-s3-demo-bucket/archive-name#cache-archive-name
   ```

   Inserisci i valori appropriati negli altri campi di dialogo. Le opzioni variano a seconda del tipo di fase. Per aggiungere la fase e uscire dalla finestra di dialogo, scegli Add step (Aggiungi fase).
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per specificare i file di cache distribuiti con AWS CLI

- Per inviare una fase di Streaming quando un cluster è stato creato, digita il comando `create-cluster` con il parametro `--steps`. Per specificare i file di cache distribuiti utilizzando il AWS CLI, specificate gli argomenti appropriati quando inviate un passaggio di Streaming.

Se desideri aggiungere un singolo file alla cache distribuita, specifica `-cacheFile` seguito dal nome e dalla posizione del file, dal segno cancelletto (#) e quindi dal nome che desideri assegnare al file quando viene collocato nella cache locale.

Se desideri aggiungere un file di archivio alla cache distribuita, inserisci `-cacheArchive` seguito dal percorso dei file in Amazon S3, dal segno cancelletto (#) e quindi dal nome che desideri assegnare alla raccolta di file nella cache locale. L'esempio seguente mostra come aggiungere un file di archivio alla cache distribuita.

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Example 1

Digita il comando seguente per avviare un cluster e invia una fase di Streaming che utilizza `-cacheFile` per aggiungere un file, `sample_dataset_cached.dat`, alla cache.

```
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"]
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

Se in precedenza non sono stati creati il ruolo del servizio EMR e il profilo di EC2 istanza predefiniti, digitare `aws emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Example 2

Il comando seguente mostra la creazione di un cluster di streaming e si avvale di `-cacheArchive` per aggiungere un archivio di file nella cache.

```
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"]
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

Se in precedenza non sono stati creati il ruolo del servizio EMR e il profilo di EC2 istanza predefiniti, digitare `aws emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Rilevamento ed elaborazione di file compressi con Amazon EMR

Hadoop controlla l'estensione del file per rilevare i file compressi. I tipi di compressione supportati da Hadoop sono: gzip, bzip2 e LZO. Non è necessario intraprendere alcuna azione aggiuntiva per estrarre i file utilizzando questo tipo di compressione; Hadoop lo gestisce per voi.

Per indicizzare i file LZO, puoi usare la libreria hadoop-lzo che può essere scaricata da hadoop-lzo. https://github.com/kevinweil/ Trattandosi di una libreria di terze parti, Amazon EMR non offre supporto agli sviluppatori su come utilizzare questo strumento. Per informazioni sull'utilizzo, consulta il file readme di hadoop-lzo.

Importa dati DynamoDB in Hive con Amazon EMR

L'implementazione di Hive fornita da Amazon EMR include funzionalità che è possibile utilizzare per importare ed esportare i dati tra DynamoDB e un cluster Amazon EMR. Questa funzione è utile se i dati di input vengono memorizzati in DynamoDB. Per ulteriori informazioni, consulta Esportazione, importazione, esecuzione di query e unione di tabelle in DynamoDB con Amazon EMR.

Connettiti ai dati con AWS Direct Connect Amazon EMR

AWS Direct Connect è un servizio che puoi utilizzare per stabilire una connessione di rete privata dedicata ad Amazon Web Services dal tuo data center, ufficio o ambiente di colocation. Se disponi di grandi quantità di dati di input, l'utilizzo AWS Direct Connect può ridurre i costi di rete, aumentare la velocità di trasmissione della larghezza di banda e fornire un'esperienza di rete più coerente rispetto alle connessioni basate su Internet. Per ulteriori informazioni, consulta la Guida per l'utente AWS Direct Connect.

Carica grandi quantità di dati per Amazon EMR con AWS Snowball Edge

AWS Snowball Edge è un servizio che puoi utilizzare per trasferire grandi quantità di dati tra Amazon Simple Storage Service (Amazon S3) e la tua posizione di archiviazione dati onsite a velocità elevate. faster-than-internet Snowball Edge supporta due tipi di processi: processi di importazione ed esportazione. I processi di importazione comportano un trasferimento di dati da un'origine on-premise a un bucket Amazon S3. I processi di esportazione comportano un trasferimento di dati da bucket Amazon S3 a un'origine on-premise. Per entrambi i tipi di lavoro, i dispositivi Snowball Edge proteggono e proteggono i dati, mentre i corrieri di spedizione regionali li trasportano tra Amazon S3 e la posizione di archiviazione dei dati in sede. I dispositivi Snowball Edge sono fisicamente robusti e protetti da (). AWS Key Management Service AWS KMS Per ulteriori informazioni, consulta la Guida per gli sviluppatori di AWS Snowball Edge Edge.

Configurare una posizione per l'output del cluster Amazon EMR

Il formato di output più comune di un cluster Amazon EMR sono i file di testo, compressi o decompressi. Di solito, tali file vengono scritti su un bucket Amazon S3. Occorre creare questo bucket prima dell'avvio del cluster. Specifica il bucket S3 come percorso di output al momento dell'avvio del cluster.

Per ulteriori informazioni, consulta i seguenti argomenti:

Argomenti

- Creazione e configurazione di un bucket Amazon S3
- Quali formati Amazon EMR è in grado di restituire?
- Come scrivere dati su un bucket Amazon S3 che non possiedi con Amazon EMR
- Modi per comprimere l'output del tuo cluster Amazon EMR

Creazione e configurazione di un bucket Amazon S3

Amazon EMR usa Amazon S3 per archiviare dati di input, file di log e dati di output. Amazon S3 fa riferimento a questi percorsi di archiviazione come bucket. I bucket presentano determinate restrizioni e limitazioni in conformità con i requisiti di Amazon S3 e DNS. Per ulteriori informazioni, consulta Restrizioni e limitazioni dei bucket nella Guida per gli sviluppatori di Amazon Simple Storage Service.

Per creare un bucket Amazon S3, segui le istruzioni nella pagina Creazione di un bucket della Guida per gli sviluppatori di Amazon Simple Storage Service.

> **Note**
>
> Se abiliti la registrazione nella procedura guidata Create a Bucket (Crea un bucket), sono abilitati solo i log di accesso al bucket e non i log del cluster.

> **Note**
>
> Per ulteriori informazioni sulla specificazione di bucket specifici per regione, consulta Bucket e regioni nella Amazon Simple Storage Service Developer Guide and Available Region Endpoints per. AWS SDKs

Dopo aver creato il bucket è possibile impostare le autorizzazioni appropriate su di esso. In genere, consenti a te stesso (il proprietario) l'accesso in lettura e scrittura. Si consiglia di seguire Best practice di sicurezza per Amazon S3 durante la configurazione del bucket.

Per poter creare un cluster, sono necessari i bucket Amazon S3 richiesti. È necessario caricare in Amazon S3 tutti gli script e i dati a cui viene fatto riferimento nel cluster. Nella seguente tabella vengono descritti dati, script e ubicazioni di file di log esempio.

Informazioni	Esempio di percorso su Amazon S3
script o programma	`s3://amzn-s3-demo-bucket1/script/MapperScript.py`
file di log	`s3://amzn-s3-demo-bucket1/logs`
dati di input	`s3://amzn-s3-demo-bucket1/input`
dati di output	`s3://amzn-s3-demo-bucket1/output`

Quali formati Amazon EMR è in grado di restituire?

Il formato di output predefinito per un cluster è il testo con coppie di chiavi e valori scritte su singole righe nei file di testo. Questo è il formato di output utilizzato più comunemente.

Se devi scrivere i dati di output in un formato diverso da quello predefinito dei file di testo, puoi utilizzare l'interfaccia di Hadoop `OutputFormat` per specificare altri tipi di output. Puoi persino creare una sottoclasse della classe `FileOutputFormat` per gestire i tipi di dati personalizzati. Per ulteriori informazioni, consulta http://hadoop.apache. org/docs/current/api/org/apache/hadoop/mapred/OutputFormat.html.

Se state lanciando un cluster Hive, potete usare a serializer/deserializer (SerDe) per inviare dati da HDFS in un determinato formato. Per ulteriori informazioni, consultate https://cwiki.apache. org/confluence/display/Hive/SerDe.

Come scrivere dati su un bucket Amazon S3 che non possiedi con Amazon EMR

Quando scrivi un file in un bucket Amazon Simple Storage Service (Amazon S3), per impostazione predefinita sei l'unico in grado di leggere tale file. Il presupposto è che scriverai file nel tuo bucket e questa impostazione predefinita protegge la privacy dei file.

Tuttavia, se stai utilizzando un cluster e desideri che l'output venga scritto nel bucket Amazon S3 di un altro AWS utente e desideri che l'altro AWS utente sia in grado di leggere quell'output, devi fare due cose:

- Chiedi all'altro AWS utente di concederti le autorizzazioni di scrittura per il suo bucket Amazon S3. Il cluster che avvii viene eseguito con AWS le tue credenziali, quindi tutti i cluster che avvii potranno anche scrivere nel bucket dell'altro utente. AWS
- Imposta le autorizzazioni di lettura per l'altro AWS utente sui file che tu o il cluster scrivete nel bucket Amazon S3. Il modo più semplice per impostare queste autorizzazioni di lettura consiste nell'utilizzare elenchi di controllo degli accessi predefiniti (ACLs), un insieme di politiche di accesso predefinite definite da Amazon S3.

Per informazioni su come l'altro AWS utente può concederti le autorizzazioni per scrivere file nel bucket Amazon S3 dell'altro utente, consulta Modifica delle autorizzazioni del bucket nella Amazon Simple Storage Service User Guide.

Affinché il cluster utilizzi «preimpostato» ACLs quando scrive file su Amazon S3, imposta `fs.s3.canned.acl` l'opzione di configurazione del cluster sull'ACL predefinito da utilizzare. La tabella seguente elenca i predefiniti attualmente. ACLs

ACL predefinita	Descrizione
`AuthenticatedRead`	Specifica che al proprietario viene concesso l'accesso `Permission.FullControl` e al gruppo `GroupGrantee.AuthenticatedUsers` assegnatario viene concesso l'accesso `Permission.Read`.
`BucketOwnerFullControl`	Specifica che al proprietario del bucket viene concesso l'accesso `Permission.FullControl`. Il proprietario del bucket non coincide necessariamente con il proprietario dell'oggetto.
`BucketOwnerRead`	Specifica che al proprietario del bucket viene concesso l'accesso `Permission.Read`. Il proprietario del bucket non coincide necessariamente con il proprietario dell'oggetto.
`LogDeliveryWrite`	Specifica che al proprietario viene concesso l'accesso `Permission.FullControl` e al gruppo `GroupGrantee.LogDelivery` assegnatario viene concesso l'accesso `Permission.Write`, in modo da poter recapitare i log di accesso.
`Private`	Specifica che al proprietario viene concesso l'accesso `Permission.FullControl`.
`PublicRead`	Specifica che al proprietario viene concesso l'accesso `Permission.FullControl` e al gruppo `GroupGrantee.AllUsers` assegnatario viene concesso l'accesso `Permission.Read`.
`PublicReadWrite`	Specifica che al proprietario viene concesso l'accesso `Permission.FullControl` e al gruppo `GroupGrantee.AllUsers` assegnatario viene concesso l'accesso `Permission.Read` e `Permission.Write`.

A seconda del tipo di cluster in esecuzione, esistono molti modi per impostare le opzioni di configurazione del cluster. Nelle procedure seguenti viene mostrato come impostare l'opzione per i casi comuni.

Per scrivere file utilizzando «preimpostato» in Hive ACLs

- Dal prompt dei comandi di Hive, imposta l'opzione di configurazione `fs.s3.canned.acl` per la lista di controllo accessi (ACL) predefinita che deve essere impostata dal cluster sui file scritti in Amazon S3. Per accedere al prompt dei comandi di Hive, esegui la connessione al nodo master tramite SSH e digita Hive al prompt dei comandi di Hadoop. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

  Nell'esempio seguente, l'opzione di configurazione `fs.s3.canned.acl` viene impostata su `BucketOwnerFullControl`. In questo modo, al proprietario del bucket Amazon S3 viene assegnato il controllo completo sul file. Si noti che il comando set rileva la distinzione tra maiuscole e minuscole e non contiene né virgolette né spazi.

  ```
  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;
  ```

  Nelle ultime due righe dell'esempio viene creata una tabella che è archiviata in Amazon S3 e vengono scritti dati nella tabella.

Per scrivere file usando «in scatola ACLs » in Pig

- Dal prompt dei comandi di Pig, imposta l'opzione di configurazione `fs.s3.canned.acl` per la lista di controllo accessi (ACL) predefinita che deve essere impostata dal cluster sui file scritti in Amazon S3. Per accedere al prompt dei comandi di Pig, esegui la connessione al nodo master utilizzando SSH e digita Pig al prompt dei comandi di Hadoop. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

  L'esempio seguente imposta l'opzione di `fs.s3.canned.acl` configurazione su BucketOwnerFullControl, che offre al proprietario del bucket Amazon S3 il controllo completo sul file. Si noti che il comando set include uno spazio prima del nome della lista di controllo accessi (ACL) predefinita e non contiene virgolette.

```
pig> set fs.s3.canned.acl BucketOwnerFullControl;
store some data into 's3://amzn-s3-demo-bucket/pig/acl';
```

Per scrivere file utilizzando «predefiniti» ACLs in un file JAR personalizzato

- Imposta l'opzione di configurazione `fs.s3.canned.acl` utilizzando Hadoop con il flag -D. Questo viene mostrato nell'esempio sottostante.

```
hadoop jar hadoop-examples.jar wordcount
-Dfs.s3.canned.acl=BucketOwnerFullControl s3://amzn-s3-demo-bucket/input s3://amzn-
s3-demo-bucket/output
```

Modi per comprimere l'output del tuo cluster Amazon EMR

Esistono diversi modi per comprimere l'output derivante dall'elaborazione dei dati. Gli strumenti di compressione utilizzati dipendono dalle proprietà dei dati. La compressione può migliorare le prestazioni quando si trasferiscono grandi quantità di dati.

Compressione dei dati di output

In questo modo si comprime l'output del processo Hadoop. Se si utilizza TextOutputFormat il risultato è un file di testo con gzip. Se stai scrivendo a, SequenceFiles il risultato è un file compresso SequenceFile internamente. Questo può essere abilitato impostando la configurazione mapred.output.compress su "true".

Se si sta eseguendo un processo in streaming, è possibile abilitarlo trasmettendo il processo in streaming a questi argomenti.

```
-jobconf mapred.output.compress=true
```

È anche possibile utilizzare un'operazione di bootstrap per comprimere automaticamente tutti gli output del lavoro. Ecco come procedere con il client Ruby.

```
--bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-hadoop \
--args "-s,mapred.output.compress=true"
```

Infine, se si sta scrivendo un Jar personalizzato, è possibile attivare la compressione dell'output con la seguente riga quando si crea un lavoro.

```
FileOutputFormat.setCompressOutput(conf, true);
```

Compressione dei dati intermedi

Se il processo mischia una quantità significativa di dati dai mappatori ai riduttori, è possibile vedere un miglioramento delle prestazioni grazie all'abilitazione della compressione intermedia. Comprime l'output della mappa e lo decomprime quando arriva sul nodo principale. L'impostazione di configurazione è mapred.compress.map.output. È possibile abilitarlo in modo simile alla compressione di output.

Quando si scrive un Jar personalizzato, utilizzare il seguente comando:

```
conf.setCompressMapOutput(true);
```

Utilizzo della libreria Snappy con Amazon EMR

Snappy è una libreria di compressione e decompressione ottimizzata per la velocità. È disponibile su Amazon EMR AMIs versione 2.0 e successive e viene utilizzato come impostazione predefinita per la compressione intermedia. Per ulteriori informazioni su Snappy, vai alla pagina http://code.google.com/p/snappy/.

Pianifica e configura i nodi primari nel tuo cluster Amazon EMR

Quando avvii un cluster Amazon EMR, puoi scegliere di avere uno o tre nodi primari nel cluster. L'elevata disponibilità, ad esempio, delle flotte è supportata dalle versioni di Amazon EMR 5.36.1, 5.36.2, 6.8.1, 6.9.1, 6.10.1, 6.11.1, 6.12.0 e successive. Per i gruppi di istanze, la disponibilità elevata

è supportata con Amazon EMR 5.23.0 e rilasci successivi. Per migliorare ulteriormente la disponibilità dei cluster, Amazon EMR può utilizzare i gruppi di EC2 posizionamento Amazon per garantire che i nodi primari siano posizionati su hardware sottostante distinto. Per ulteriori informazioni, consulta Integrazione di Amazon EMR con EC2 i gruppi di collocamento.

Un cluster Amazon EMR con più nodi primari fornisce i seguenti vantaggi:

- Il nodo primario non rappresenta più un singolo punto di errore. Se uno dei nodi primari riscontra errori, il cluster utilizza gli altri due nodi primari e viene eseguito senza interruzioni. Nel frattempo, Amazon EMR sostituisce automaticamente i nodi primari con errori con un nuovo nodo dotato delle stesse operazioni di configurazione e bootstrap.
- Amazon EMR abilita le funzionalità di alta disponibilità Hadoop di HDFS NameNode e YARN ResourceManager e supporta l'alta disponibilità per alcune altre applicazioni open source.

  Per ulteriori informazioni su come un cluster EMR con più nodi primari supporta le applicazioni open source e altre caratteristiche Amazon EMR, consulta la sezione Funzionalità che supportano l'elevata disponibilità in un cluster Amazon EMR e il loro funzionamento con applicazioni open source.

> **Note**
>
> I cluster possono trovarsi in una sola zona di disponibilità o sottorete.

Questa sezione fornisce informazioni sulle applicazioni e sulle caratteristiche supportate di un cluster Amazon EMR con più nodi primari nonché i dettagli di configurazione, le best practice e le considerazioni in merito all'avvio del cluster.

Argomenti

- Funzionalità che supportano l'elevata disponibilità in un cluster Amazon EMR e il loro funzionamento con applicazioni open source
- Avvio di un cluster Amazon EMR con più nodi primari
- Integrazione di Amazon EMR con EC2 i gruppi di collocamento
- Considerazioni e best practice per la creazione di un cluster Amazon EMR con più nodi primari

Funzionalità che supportano l'elevata disponibilità in un cluster Amazon EMR e il loro funzionamento con applicazioni open source

Questo argomento fornisce informazioni sulle funzionalità Hadoop ad alta disponibilità di HDFS NameNode e YARN in ResourceManager un cluster Amazon EMR e su come le funzionalità ad alta disponibilità funzionano con le applicazioni open source e altre funzionalità di Amazon EMR.

HDFS a disponibilità elevata

Un cluster Amazon EMR con più nodi primari abilita la funzionalità di disponibilità elevata HDFS NameNode in Hadoop. Per ulteriori informazioni, consulta l'argomento relativo alla Disponibilità elevata HDFS.

In un cluster Amazon EMR, due o più nodi separati sono configurati come. NameNodes Uno NameNode è in uno `active` stato e gli altri sono in uno `standby` stato. In caso di `active` NameNode guasto del nodo, Amazon EMR avvia un processo di failover HDFS automatico. Un nodo `standby` NameNode diventa `active` e assume il controllo di tutte le operazioni dei client nel cluster. Amazon EMR sostituisce il nodo con esito negativo con uno nuovo che quindi si ricollega come `standby`.

> **Note**
>
> Nelle versioni di Amazon EMR da 5.23.0 a 5.36.2, solo due dei tre nodi primari eseguono HDFS. NameNode
> Nelle versioni 6.x e successive di Amazon EMR, tutti e tre i nodi primari eseguono HDFS. NameNode

Se hai bisogno di scoprire qual NameNode è`active`, puoi usare SSH per connetterti a qualsiasi nodo primario del cluster ed eseguire il seguente comando:

```
hdfs haadmin -getAllServiceState
```

L'output elenca i nodi in cui NameNode è installato e il loro stato. Ad esempio,

```
ip-##-#-#-##1.ec2.internal:8020 active
ip-##-#-#-##2.ec2.internal:8020 standby
ip-##-#-#-##3.ec2.internal:8020 standby
```

YARN ad alta disponibilità ResourceManager

Un cluster Amazon EMR con più nodi primari abilita la funzionalità di ResourceManager alta disponibilità YARN in Hadoop. Per ulteriori informazioni, consulta la sezione Alta disponibilità. ResourceManager

In un cluster Amazon EMR con più nodi primari, YARN ResourceManager viene eseguito su tutti e tre i nodi primari. Uno ResourceManager è in `active` stato e gli altri due sono in `standby` stato. Se il nodo primario si `active` ResourceManager guasta, Amazon EMR avvia un processo di failover automatico. Un nodo primario con un `standby` ResourceManager si occupa di tutte le operazioni. Amazon EMR sostituisce il nodo primario guasto con uno nuovo, che quindi si ricongiunge al quorum come un. ResourceManager `standby`

Puoi connetterti a «http: //:8088/cluster*`master-public-dns-name`*» per qualsiasi nodo primario, che ti indirizza automaticamente al gestore delle risorse. `active` Per sapere quale gestore è `active`, utilizza SSH per connetterti a un nodo primario del cluster. Quindi esegui il comando seguente per ottenere un elenco dei tre nodi primari e il relativo stato:

```
yarn rmadmin -getAllServiceState
```

Applicazioni supportate in un cluster Amazon EMR con più nodi primari

È possibile installare ed eseguire le seguenti applicazioni in un cluster Amazon EMR con più nodi primari. Per ogni applicazione, il processo di failover del nodo primario varia.

Applicazione	Disponibilità durante il failover del nodo primario	Note
Flink	Disponibilità non interessata dal failover del nodo primario	I processi Flink su Amazon EMR vengono eseguiti come applicazioni YARN. Flink JobManagers funziona come YARN sui nodi principali. ApplicationMasters Non JobManage r è influenzato dal processo di failover del nodo primario. Se utilizzi Amazon EMR versione 5.27.0 o precedente, si JobManager tratta di un singolo punto di errore. In caso di JobManager errore,

Applicazione	Disponibilità durante il failover del nodo primario	Note
		perde tutti gli stati del processo e non riprende i processi in esecuzione. È possibile abilitare l' JobManager alta disponibilità configurando il conteggio dei tentativi dell'applicazione, il checkpoint e l'abilitazione ZooKeeper come storage di stato per Flink. Per ulteriori informazioni, consulta la sezione Configuring Flink on an Amazon EMR Cluster with multiple primary nodes (Configurazione di Flink su un cluster Amazon EMR con più nodi primari). A partire dalla versione 5.28.0 di Amazon EMR, non è necessaria alcuna configurazione manuale per consentire l'elevata disponibilità. JobManager
Ganglia	Disponibilità non interessata dal failover del nodo primario	Ganglia è disponibile su tutti i nodi primari, quindi può continuare a funzionare durante il processo di failover del nodo primario.
Hadoop	Elevata disponibilità	HDFS NameNode e YARN eseguono ResourceManager automaticamente il failover sul nodo di standby in caso di guasto del nodo primario attivo.
HBase	Elevata disponibilità	HBase esegue automaticamente il failover sul nodo di standby in caso di guasto del nodo primario attivo. Se ci si connette HBase tramite un server REST o Thrift, è necessario passare a un nodo primario diverso in caso di guasto del nodo primario attivo.

Applicazione	Disponibilità durante il failover del nodo primario	Note
HCatalog	Disponibilità non interessata dal failover del nodo primario	HCatalog è basato sul metastore Hive, che esiste all'esterno del cluster. HCatalog rimane disponibile durante il processo di failover del nodo primario.
JupyterHub	Elevata disponibilità	JupyterHub è installato su tutte e tre le istanze principali. Si consiglia vivamente di configura re la persistenza del notebook per evitare la perdita del notebook in caso di errore del nodo primario. Per ulteriori informazioni, consulta la sezione Configurazione della persistenza per i notebook in Amazon S3.
Livy	Elevata disponibilità	Livy è installato su tutti e tre i nodi primari. Quando il nodo primario attivo riscontra un errore, perdi l'accesso alla sessione Livy corrente e devi creare una nuova sessione Livy su un altro nodo primario o sul nuovo nodo sostitutivo.
Mahout	Disponibilità non interessata dal failover del nodo primario	Poiché Mahout non prevede il daemon, non viene interessato dal processo di failover del nodo primario.
MXNet	Disponibilità non interessata dal failover del nodo primario	Poiché non MXNet dispone di un demone, non è influenzato dal processo di failover del nodo primario.

Applicazione	Disponibilità durante il failover del nodo primario	Note
Phoenix	Elevata disponibilità	Phoenix' QueryServer viene eseguito solo su uno dei tre nodi primari. Phoenix su tutti e tre i master è configurato per connettere Phoenix. QueryServer È possibile trovare l'IP privato del server Query di Phoenix utilizzando il file `/etc/phoenix/conf/phoenix-env.sh`
Pig	Disponibilità non interessata dal failover del nodo primario	Poiché Pig non prevede il daemon, non viene interessato dal processo di failover del nodo primario.
Spark	Elevata disponibilità	Tutte le applicazioni Spark vengono eseguite in container YARN e possono reagire al failover del nodo primario allo stesso modo delle funzionalità YARN a disponibilità elevata.
Sqoop	Elevata disponibilità	Per impostazione predefinita, sqoop-job e sqoop-metastore memorizzano i dati (descrizi oni del lavoro) sul disco locale del master che esegue il comando; se si desidera salvare i dati del metastore su database esterno, fare riferimento alla documentazione di apache Sqoop
Tez	Elevata disponibilità	Poiché i container Tez funzionano su YARN, Tez si comporta allo stesso modo di YARN durante il processo di failover del nodo primario.
TensorFlow	Disponibilità non interessata dal failover del nodo primario	Poiché non TensorFlow dispone di un demone, non è influenzato dal processo di failover del nodo primario.

Applicazione	Disponibilità durante il failover del nodo primario	Note
Zeppelin	Elevata disponibilità	Zeppelin è installato su tutti e tre i nodi primari. Zeppelin memorizza note e configurazioni di interpreti in HDFS per impostazione predefini ta per evitare la perdita di dati. Le sessioni di interprete sono completamente isolate in tutte e tre le istanze primarie. I dati della sessione andranno persi in caso di errore principale. Si consiglia di non modificare la stessa nota contemporaneamente su istanze primarie diverse.
ZooKeeper	Elevata disponibilità	ZooKeeper è la base della funzionalità di failover automatico HDFS. ZooKeeper fornisce un servizio ad alta disponibilità per la gestione dei dati di coordinamento, la notifica ai clienti delle modifiche di tali dati e il monitoraggio dei client in caso di guasti. Per ulteriori informazi oni, consulta la sezione relativa al Failover automatico di HDFS.

Per eseguire le seguenti applicazioni in un cluster Amazon EMR con più nodi primari, è necessario configurare un database esterno. Il database esterno esiste al di fuori del cluster e rende i dati persistenti durante il processo di failover del nodo primario. Per le applicazioni seguenti, i componenti del servizio verranno ripristinati automaticamente durante il processo di failover del nodo primario, ma i processi attivi potrebbero non riuscire e dovranno essere riattivati.

Applicazione	Disponibilità durante il failover del nodo primario	Note
Hive	Disponibilità elevata solo per i componenti di servizio	È necessario un metastore esterno per Hive. Deve trattarsi di un metastore esterno MySQL, poiché PostgreSQL non è supportato per

Applicazione	Disponibilità durante il failover del nodo primario	Note
		cluster multi-master. Per ulteriori informazioni, consulta Configurazione di un metastore esterno per Hive.
Hue	Disponibilità elevata solo per i componenti di servizio	È necessario un database esterno per Hue. Per ulteriori informazioni, consulta Utilizzo di Hue con un database remoto in Amazon RDS.
Oozie	Disponibilità elevata solo per i componenti di servizio	È necessario un database esterno per Oozie. Per ulteriori informazioni, consulta Utilizzo di Oozie con un database remoto in Amazon RDS. Oozie-server e oozie-client sono installati su tutti e tre i nodi primari. I client oozie-sono configurati per connettersi al server oozie-corretto per impostazione predefinita.

Applicazione	Disponibilità durante il failover del nodo primario	Note
PrestoDB o PrestoSQL/ Trino	Disponibilità elevata solo per i componenti di servizio	È necessario un metastore Hive esterno per PrestoDB (PrestoSQL su Amazon EMR 6.1.0-6.3.0 o Trino su Amazon EMR 6.4.0 e versioni successive). Puoi usare Presto con il AWS Glue Data Catalog o usare un database MySQL esterno per Hive. La CLI di Presto è installata su tutti e tre i nodi primari, quindi è possibile utilizzarla per accedere a Presto Coordinator da uno qualsiasi dei nodi primari. Presto Coordinat or è installato su un solo nodo primario. Puoi trovare il nome DNS del nodo primario in cui è installato Presto Coordinator chiamando l'API `describe-cluster` di Amazon EMR e la lettura del valore restituito del campo `MasterPublicDnsName` nella risposta.

> **Note**
>
> Quando un nodo primario riscontra un errore, Java Database Connectivity (JDBC) o Open Database Connectivity (ODBC) termina la connessione al nodo primario. Puoi connetterti a uno dei nodi primari rimanenti per continuare il lavoro dal momento che il daemon metastore Hive viene eseguito su tutti i nodi primari. In alternativa, puoi attendere che il nodo primario con errori venga sostituito.

Funzionamento delle caratteristiche di Amazon EMR in un cluster con più nodi primari

Connessione ai nodi primari tramite SSH

È possibile connettersi a uno dei tre nodi primari in un cluster Amazon EMR utilizzando SSH nello stesso modo in cui ci si connette a un singolo nodo primario. Per ulteriori informazioni, consulta la sezione Connect to the primary node using SSH (Connessione al nodo primario tramite SSH).

Se un nodo primario riscontra un errore, la connessione SSH per quel nodo primario viene terminata. Per continuare il lavoro, puoi connetterti a uno degli altri due nodi primari. In alternativa, puoi accedere al nuovo nodo primario dopo che Amazon EMR sostituisce quello con errori con un nuovo nodo.

> **Note**
>
> L'indirizzo IP privato per il nodo primario sostitutivo rimane uguale a quello precedente. L'indirizzo IP pubblico per il nodo primario sostitutivo potrebbe cambiare. Puoi recuperare i nuovi indirizzi IP nella console o utilizzando il comando `describe-cluster` nella CLI AWS . NameNode funziona solo su due dei nodi primari. Tuttavia, è possibile eseguire i comandi della CLI `hdfs` e gestire i processi per accedere a HDFS su tutti e tre i nodi primari.

Utilizzo delle fasi in un cluster Amazon EMR con più nodi primari

Puoi inviare le fasi a un cluster Amazon EMR con più nodi primari nello stesso modo in cui utilizzi le fasi in un cluster con un singolo nodo primario. Per ulteriori informazioni, consulta Invio di lavoro a un cluster.

Di seguito sono riportate le considerazioni sull'utilizzo di fasi in un cluster Amazon EMR con più nodi primari:

- Se un nodo primario riscontra un errore, le fasi in esecuzione sul nodo primario vengono contrassegnate come FAILED. I dati scritti in locale andranno perduti. Tuttavia, lo stato FAILED potrebbe non riflettere lo stato reale delle fasi.
- Se una fase in esecuzione ha avviato un'applicazione YARN quando il nodo primario riscontra un errore, la fase può continuare e avere esito positivo a causa del failover automatico del nodo primario.
- È consigliabile controllare lo stato delle fasi consultando l'output dei processi. Ad esempio, i MapReduce processi utilizzano un `_SUCCESS` file per determinare se il processo viene completato correttamente.
- Si consiglia di impostare il ActionOnFailure parametro su CONTINUE o CANCEL_AND_WAIT, anziché TERMINATE_JOB_FLOW o TERMINATE_CLUSTER.

Protezione da cessazione automatica

Amazon EMR abilita automaticamente la protezione da terminazione per tutti i cluster con più nodi primari e sostituisce tutte le impostazioni di esecuzione di fasi fornite durante la creazione del cluster. È possibile disattivare la protezione da cessazione dopo l'avvio del cluster. Consultare Configurazione della protezione da cessazione per i cluster in esecuzione. Per chiudere un cluster con più nodi primari, è necessario modificare gli attributi del cluster per disabilitare la protezione da terminazione. Per istruzioni, consultare Terminazione di un cluster Amazon EMR con più nodi primari.

Per ulteriori informazioni sulla protezione da cessazione, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

Caratteristiche non supportate in un cluster Amazon EMR con più nodi primari

Le seguenti caratteristiche Amazon EMR non sono attualmente disponibili in un cluster EMR con più nodi primari:

- Notebook EMR
- Accesso con un clic al server della cronologia Spark persistente
- Interfacce utente delle applicazioni persistenti
- L'accesso con un clic alle interfacce utente persistenti delle applicazioni non è attualmente disponibile per i cluster Amazon EMR con più nodi primari o per i cluster Amazon EMR integrati con Lake Formation. AWS

> **Note**
>
> Per utilizzare l'autenticazione Kerberos nel cluster, è necessario configurare un KDC esterno. A partire dalla versione 5.27.0 di Amazon EMR, puoi configurare la crittografia trasparente HDFS su un cluster EMR con più nodi primari. Per ulteriori informazioni, consulta l'argomento relativo alla Crittografia trasparente in HDFS su Amazon EMR.

Avvio di un cluster Amazon EMR con più nodi primari

Questo argomento contiene i dettagli di configurazione e gli esempi per avviare un cluster Amazon EMR con più nodi primari.

> **Note**
>
> Amazon EMR abilita automaticamente la protezione da terminazione per tutti i cluster con più nodi primari e sostituisce tutte le impostazioni di terminazione automatica fornite durante la creazione del cluster. Per chiudere un cluster con più nodi primari, è necessario modificare gli attributi del cluster per disabilitare la protezione da terminazione. Per istruzioni, consultare Terminazione di un cluster Amazon EMR con più nodi primari.

Prerequisiti

- Puoi avviare un cluster Amazon EMR con più nodi primari in sottoreti VPC pubbliche e private. EC2-La versione classica non è supportata. Per avviare un cluster Amazon EMR con più nodi primari in una sottorete pubblica, devi abilitare le istanze in questa sottorete a ricevere un indirizzo IP pubblico selezionando Assegna automatica IPv4 nella console o eseguendo il comando seguente. Sostituiscilo con il tuo ID di sottorete. *22XXXX01*

  ```
  aws ec2 modify-subnet-attribute --subnet-id subnet-22XXXX01 --map-public-ip-on-launch
  ```

- Per eseguire Hive oppure Oozie in un cluster Amazon EMR con più nodi primari, è necessario creare un metastore esterno per Hive. Per ulteriori informazioni, consulta la sezione Configurazione di un metastore esterno per Hive, Utilizzo di Hue con un database remoto in Amazon RDS o Apache Oozie.
- Per utilizzare l'autenticazione Kerberos nel cluster, è necessario configurare un KDC esterno. Per ulteriori informazioni, consulta Configurazione di Kerberos su Amazon EMR.

Avvio di un cluster Amazon EMR con più nodi primari

Puoi avviare un cluster con più nodi primari quando utilizzi gruppi di istanze o parchi istanze. Quando utilizzi gruppi di istanze con più nodi primari, devi specificare un valore di conteggio delle istanze pari a `3` per il gruppo di istanze del nodo primario. Quando utilizzi parchi istanze con più nodi primari, devi specificare il valore `TargetOnDemandCapacity` di `3`, `TargetSpotCapacity` o `0` per il parco istanze primario e il valore `WeightedCapacity` di `1` per ogni tipo di istanza configurato per il parco istanze primario.

I seguenti esempi illustrano come avviare il cluster utilizzando l'AMI predefinita o un'AMI personalizzata con i gruppi di istanze e i parchi istanze:

> **Note**
>
> È necessario specificare l'ID della sottorete quando si avvia un cluster Amazon EMR con più nodi primari mediante la AWS CLI. Sostituisci *22XXXX01* e *22XXXX02* con il tuo ID di sottorete negli esempi seguenti.

Default AMI, instance groups

Example Esempio: avvio di un cluster del gruppo di istanze Amazon EMR con più nodi primari utilizzando un'AMI predefinita

```
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 Esempio: avvio di un cluster del parco istanze Amazon EMR con più nodi primari utilizzando un'AMI predefinita

```
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 Esempio: avvio di un cluster del gruppo di istanze Amazon EMR con più nodi primari utilizzando un'AMI personalizzata

```
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 Esempio: avvio di un cluster del parco istanze Amazon EMR con più nodi primari utilizzando un'AMI personalizzata

```
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
```

Terminazione di un cluster Amazon EMR con più nodi primari

Per terminare un cluster Amazon EMR con più nodi primari, è necessario prima disabilitare la protezione dalla terminazione, come dimostrato nel seguente esempio. Sostituisci *j-3KVTXXXXXX7UG* con l'ID del tuo cluster.

```
aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-
protected
aws emr terminate-clusters --cluster-id j-3KVTXXXXXX7UG
```

Integrazione di Amazon EMR con EC2 i gruppi di collocamento

Quando avvii un cluster Amazon EMR con più nodi primari su Amazon EC2, hai la possibilità di utilizzare strategie di gruppo di posizionamento per specificare come desideri che le istanze del nodo primario vengano distribuite per proteggerle da guasti hardware.

Le strategie dei gruppi di collocamento sono supportate a partire da Amazon EMR versione 5.23.0 come opzione per i cluster con più nodi primari. Attualmente, solo i tipi di nodi primari sono supportati dalla strategia del gruppo di collocamento e la strategia SPREAD viene applicata a tali nodi primari. La strategia SPREAD colloca un piccolo gruppo di istanze su hardware sottostante separato per evitare la perdita di più nodi primari in caso di guasto hardware. Una richiesta di avvio di un'istanza potrebbe non riuscire se l'hardware univoco è insufficiente per l'esecuzione della richiesta. Per ulteriori informazioni sulle strategie e le limitazioni di EC2 posizionamento, consulta i gruppi di collocamento nella Guida EC2 utente per le istanze Linux.

Amazon prevede un limite iniziale EC2 di 500 cluster abilitati alla strategia dei gruppi di collocamento che possono essere lanciati per regione. AWS Contatta l' AWS assistenza per richiedere un aumento del numero di gruppi di collocamento consentiti. Puoi identificare i gruppi di EC2 collocamento creati da Amazon EMR tracciando la coppia chiave-valore che Amazon EMR associa alla strategia

dei gruppi di collocamento di Amazon EMR. Per ulteriori informazioni sui tag delle istanze cluster, consulta EC2 . Visualizza le istanze di cluster in Amazon EC2

Allega la politica gestita dal gruppo di collocamento ad Amazon EMRrole

La strategia dei gruppi di collocamento richiede una policy gestita chiamataAmazonElasticMapReducePlacementGroupPolicy, che consente ad Amazon EMR di creare, eliminare e descrivere i gruppi di collocamento su Amazon. EC2 È necessario collegare AmazonElasticMapReducePlacementGroupPolicy al ruolo di servizio per Amazon EMR prima di avviare un cluster Amazon EMR con più nodi primari.

In alternativa, puoi collegare la policy AmazonEMRServicePolicy_v2 gestita al ruolo di servizio Amazon EMR anziché alla policy gestita dal gruppo di collocamento. AmazonEMRServicePolicy_v2consente lo stesso accesso ai gruppi di collocamento su Amazon EC2 diAmazonElasticMapReducePlacementGroupPolicy. Per ulteriori informazioni, consulta Ruolo di servizio per Amazon EMR (ruolo EMR).

La policy gestita da AmazonElasticMapReducePlacementGroupPolicy è il seguente testo JSON che viene creato e gestito da Amazon EMR.

> **Note**
>
> Poiché la politica AmazonElasticMapReducePlacementGroupPolicy gestita viene aggiornata automaticamente, la politica mostrata qui potrebbe essere out-of-date. Utilizza la console di AWS gestione per visualizzare la politica corrente.

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"
    }
  ]
}
```

Avvio di un cluster Amazon EMR con più nodi primari utilizzando una strategia per gruppi di posizionamento

Per avviare un cluster Amazon EMR con più nodi primari con una strategia per gruppi di posizionamento, collega la policy gestita dal gruppo di posizionamento `AmazonElasticMapReducePlacementGroupPolicy` al ruolo Amazon EMR. Per ulteriori informazioni, consulta Allega la politica gestita dal gruppo di collocamento ad Amazon EMRrole.

Ogni volta che utilizzi questo ruolo per avviare un cluster Amazon EMR con più nodi primari, Amazon EMR tenta di avviare un cluster con la strategia `SPREAD` applicata ai suoi nodi primari. Se utilizzi un ruolo che non dispone della policy gestita dal gruppo di posizionamento `AmazonElasticMapReducePlacementGroupPolicy` ad esso collegato, Amazon EMR tenta di avviare un cluster Amazon EMR con più nodi primari senza una strategia per gruppi di posizionamento.

Se avvii un cluster Amazon EMR con più nodi primari con il `placement-group-configs` parametro utilizzando Amazon EMRAPI o CLI, Amazon EMR avvia il cluster solo se EMRrole Amazon ha la policy gestita del gruppo di collocamento allegata. `AmazonElasticMapReducePlacementGroupPolicy` Se Amazon EMRrole non ha la policy allegata, l'avvio del cluster Amazon EMR con più nodi primari non funziona.

Amazon EMR API

Example Esempio: utilizza una strategia del gruppo di posizionamento per avviare un cluster del gruppo di istanze con più nodi primari dall'API Amazon EMR

Quando utilizzi l' RunJobFlow azione per creare un cluster Amazon EMR con più nodi primari, imposta la `PlacementGroupConfigs` proprietà su quanto segue. Attualmente, il ruolo dell'istanza `MASTER` utilizza automaticamente `SPREAD` come strategia del gruppo di collocamento.

```
{
   "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`*Sostituiscilo con il nome del tuo cluster ad alta disponibilità.
- *`subnet-22XXXX01`*Sostituiscilo con il tuo ID di sottorete.
- Sostituisci *`ec2_key_pair_name`* con il nome della tua EC2 key pair per questo cluster. EC2 key pair è facoltativo e richiesto solo se si desidera utilizzare SSH per accedere al cluster.

AWS CLI

Example Esempio: utilizzo di una strategia del gruppo di posizionamento per avviare un cluster del parco istanze con più nodi primari dalla AWS Command Line Interface

Quando utilizzi l' RunJobFlow azione per creare un cluster Amazon EMR con più nodi primari, imposta la `PlacementGroupConfigs` proprietà su quanto segue. Attualmente, il ruolo dell'istanza `MASTER` utilizza automaticamente `SPREAD` come strategia del gruppo di collocamento.

```
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*Sostituiscilo con il nome del tuo cluster ad alta disponibilità.
- Sostituisci *ec2_key_pair_name* con il nome della tua EC2 key pair per questo cluster. EC2 key pair è facoltativo e richiesto solo se si desidera utilizzare SSH per accedere al cluster.

- Sostituisci *subnet-22XXXX01* e *subnet-22XXXX02* con la tua IDs sottorete.

Avvio di un cluster con più nodi primari Amazon EMR senza strategia del gruppo di collocamento

Affinché un cluster con più nodi primari possa avviare nodi primari senza la strategia del gruppo di collocamento, dovrai effettuare una delle seguenti operazioni:

- Rimuovi la politica gestita dal gruppo di collocamento `AmazonElasticMapReducePlacementGroupPolicy` da Amazon EMRrole, oppure
- Avvia un cluster con più nodi primari Amazon EMR con il parametro `placement-group-configs` utilizzando l'API o la CLI di Amazon EMR oppure scegliendo `NONE` come strategia del gruppo di collocamento.

Amazon EMR API

Example - Avvio di un cluster con più nodi primari senza strategia del gruppo di collocamento utilizzando l'API di Amazon EMR.

Quando utilizzi l' RunJobFlow azione per creare un cluster con più nodi primari, imposta la `PlacementGroupConfigs` proprietà su quanto segue.

```
{
   "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"
}
```

- Sostituiscilo *ha-cluster* con il nome del cluster ad alta disponibilità.
- *subnet-22XXXX01*Sostituiscilo con il tuo ID di sottorete.
- Sostituisci *ec2_key_pair_name* con il nome della tua EC2 key pair per questo cluster. EC2 key pair è facoltativo e richiesto solo se si desidera utilizzare SSH per accedere al cluster.

Amazon EMR CLI

Example - Avvio di un cluster con più nodi primari senza strategia del gruppo di collocamento utilizzando la CLI Amazon EMR.

Quando utilizzi l' RunJobFlow azione per creare un cluster con più nodi primari, imposta la `PlacementGroupConfigs` proprietà su quanto segue.

```
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
```

- Sostituiscilo *ha-cluster* con il nome del cluster ad alta disponibilità.
- *subnet-22XXXX01*Sostituiscilo con il tuo ID di sottorete.
- Sostituisci *ec2_key_pair_name* con il nome della tua EC2 key pair per questo cluster. EC2 key pair è facoltativo e richiesto solo se si desidera utilizzare SSH per accedere al cluster.

Controllo della configurazione della strategia del gruppo di collocamento associata al cluster con più nodi primari di Amazon EMR

Puoi utilizzare l'API describe cluster di Amazon EMR per visualizzare la configurazione della strategia del gruppo di collocamento associata al cluster con più nodi primari di Amazon EMR.

Example

```
aws emr describe-cluster --cluster-id "j-xxxxx"
{
   "Cluster":{
      "Id":"j-xxxxx",
      ...
      ...
      "PlacementGroups":[
         {
            "InstanceRole":"MASTER",
            "PlacementStrategy":"SPREAD"
         }
      ]
   }
}
```

Considerazioni e best practice per la creazione di un cluster Amazon EMR con più nodi primari

Quando crei un cluster Amazon EMR con più nodi primari, considera quanto segue:

> **Important**
>
> Per avviare cluster EMR a disponibilità elevata con più nodi primari, si consiglia di utilizzare il rilascio più recente di Amazon EMR. Ciò garantisce il massimo livello di resilienza e stabilità per i cluster a disponibilità elevata.

- L'elevata disponibilità, ad esempio, delle flotte è supportata dalle versioni di Amazon EMR 5.36.1, 5.36.2, 6.8.1, 6.9.1, 6.10.1, 6.11.1, 6.12.0 e successive. Per i gruppi di istanze, la disponibilità elevata è supportata con Amazon EMR 5.23.0 e rilasci successivi. Per ulteriori informazioni, consulta Informazioni sui rilasci di Amazon EMR.

- Nei cluster a disponibilità elevata, Amazon EMR supporta solo l'avvio di nodi primari con istanze On Demand. Ciò garantisce la massima disponibilità per il cluster.
- È ancora possibile specificare più tipi di istanze per il parco istanze primario, ma tutti i nodi primari dei cluster a disponibilità elevata vengono avviati con lo stesso tipo di istanza, comprese le sostituzioni per i nodi primari non integri.
- Per continuare le operazioni, un cluster a disponibilità elevata con più nodi primari richiede che due nodi primari su tre siano integri. Di conseguenza, se due nodi primari riportano errori contemporaneamente, il cluster EMR avrà esito negativo.
- Tutti i cluster EMR, compresi i cluster a disponibilità elevata, vengono avviati in un'unica zona di disponibilità. Pertanto, non possono tollerare gli errori della zona di disponibilità. Nel caso di un'interruzione nella zona di disponibilità, perdi l'accesso al cluster.
- Se utilizzi un ruolo o una policy di servizio personalizzato quando avvii un cluster all'interno di una flotta di istanze, puoi aggiungere l'`ec2:DescribeInstanceTypeOfferings`autorizzazione in modo che Amazon EMR possa filtrare le zone di disponibilità (AZ) non supportate. Quando Amazon EMR filtra i nodi primari AZs che non supportano alcun tipo di istanza, Amazon EMR impedisce che l'avvio del cluster non riesca a causa di tipi di istanze primarie non supportati. Per ulteriori informazioni, consulta Tipo di istanza non supportato.
- Amazon EMR non garantisce la disponibilità elevata per le applicazioni open-source diverse da quelle specificate in Applicazioni supportate in un cluster Amazon EMR con più nodi primari.
- Nelle versioni di Amazon EMR da 5.23.0 a 5.36.2, vengono eseguiti solo due dei tre nodi primari per un cluster di gruppi di istanze. HDFS NameNode
- Nelle versioni 6.x e successive di Amazon EMR, vengono eseguiti tutti e tre i nodi primari di un gruppo di istanze. HDFS NameNode

Considerazioni per la configurazione della sottorete:

- Un cluster Amazon EMR con più nodi primari può trovarsi in una sola zona di disponibilità o sottorete. Amazon EMR non è in grado di sostituire un nodo primario con errori se la sottorete è completamente utilizzata o sovrascritta in caso di failover. Per evitare questo scenario, è opportuno dedicare un'intera sottorete a un cluster Amazon EMR. Inoltre, assicurati che nella sottorete siano disponibili sufficienti indirizzi IP privati.

Considerazioni per la configurazione dei nodi core:

- Per garantire che anche i nodi principali siano altamente disponibili, ti consigliamo di avviare almeno quattro nodi principali. Se decidi di avviare un cluster più piccolo con tre nodi principali (o un numero minore), imposta `dfs.replication parameter` almeno su 2 in modo che HDFS abbia una replica DFS sufficiente. Per ulteriori informazioni, consulta la sezione dedicata alla Configurazione di HDFS.

> **Warning**
>
> 1. L'impostazione di `dfs.replication` su 1 per i cluster con meno di quattro nodi può causare la perdita di dati HDFS in caso di disattivazione anche di un singolo nodo. Ti consigliamo di utilizzare un cluster con almeno quattro nodi principali per i carichi di lavoro di produzione.
> 2. Amazon EMR non consente ai cluster di dimensionare i nodi principali al di sotto di `dfs.replication`. Ad esempio, se `dfs.replication = 2`, il numero minimo di nodi principali è 2.
> 3. Quando utilizzi il dimensionamento gestito, il dimensionamento automatico o scegli di dimensionare manualmente il cluster, ti consigliamo di impostare `dfs.replication` su 2 o su un valore superiore.

Considerazioni per l'impostazione di allarmi sui parametri:

- Amazon EMR non fornisce parametri specifici dell'applicazione su HDFS o YARN. Ti consigliamo di impostare allarmi per monitorare il numero di istanze del nodo principale. Configura gli allarmi utilizzando i seguenti CloudWatch parametri di Amazon:`MultiMasterInstanceGroupNodesRunning`,`MultiMasterInstanceGroupNodesRunning`o. `MultiMasterInstanceGroupNodesRequested` CloudWatch ti avviserà in caso di guasto e sostituzione del nodo primario.
  - Se `MultiMasterInstanceGroupNodesRunningPercentage` è inferiore al 100% e superiore al 50%, il cluster potrebbe aver perso un nodo primario. In questo caso, Amazon EMR tenta di sostituire un nodo primario.
  - Se il valore `MultiMasterInstanceGroupNodesRunningPercentage` scende al di sotto del 50%, è possibile che due nodi primari abbiano avuto esito negativo. In questo caso, il quorum viene perso e il cluster non può essere recuperato. È necessario eseguire manualmente la migrazione dei dati al di fuori del cluster.

Per ulteriori informazioni, consulta Impostazione di allarmi per i parametri.

Cluster EMR attivi AWS Outposts

A partire da Amazon EMR 5.28.0, puoi creare ed eseguire cluster EMR su. AWS Outposts AWS Outposts abilita AWS servizi, infrastrutture e modelli operativi nativi in strutture locali. Negli AWS Outposts ambienti, puoi utilizzare gli stessi AWS APIs strumenti e la stessa infrastruttura che usi nel AWS cloud. Amazon EMR on AWS Outposts è ideale per carichi di lavoro a bassa latenza che devono essere eseguiti in prossimità di dati e applicazioni locali. Per ulteriori informazioni su AWS Outposts, consulta la Guida per l'utente.AWS Outposts

Prerequisiti

Di seguito sono indicati i prerequisiti per l'utilizzo di Amazon EMR in AWS Outposts:

- È necessario averlo installato e configurato AWS Outposts nel data center locale.
- È necessario disporre di una connessione di rete affidabile tra l'ambiente Outpost e una AWS regione.
- È necessario disporre di una capacità sufficiente per i tipi di istanze supportati da Amazon EMR disponibili nel tuo Outpost.

Limitazioni

Di seguito sono riportate le limitazioni dell'utilizzo di Amazon EMR su AWS Outposts:

- Le istanze on demand sono l'unica opzione supportata per le istanze Amazon EC2 . Le istanze Spot non sono disponibili per Amazon EMR in AWS Outposts.
- Se hai bisogno di volumi di storage Amazon EBS aggiuntivi, è supportato solo General Purpose SSD (GP2).
- Se lo utilizzi AWS Outposts con le versioni di Amazon EMR dalla 5.28 alla 6.x, puoi usare solo bucket S3 che archiviano oggetti in un ambiente da te specificato. Regione AWS Con Amazon EMR 7.0.0 e versioni successive, Amazon EMR on AWS Outposts è supportato anche con il client di file system, prefix. S3A `s3a://`
- Solo i seguenti tipi di istanza sono supportati da Amazon EMR in AWS Outposts:

Classe di istanza	Tipi di istanza
Uso generale	`m5.xlarge | m5.2xlarge | m5.4xlarge | m5.12xlarge | m5.24xlarge | m5d.xlarge | m5d.2xlarge | m5d.4xlarge | m5d.12xlarge | m5d.24xlarge`
Ottimizzato per il calcolo	`c5.xlarge | c5.2xlarge | c5.4xlarge | c5.18xlarge | c5d.xlarge | c5d.2xlarge | c5d.4xlarge | c5d.18xlarge`
Ottimizzato per la memoria	`r5.xlarge | r5.2xlarge | r5.4xlarge | r5.12xlarge | r5d.xlarge | r5d.2xlarge | r5d.4xlarge | r5d.12xlarge | r5d.24xlarge`
Ottimizzato per lo storage	`i3en.xlarge | i3en.2xlarge | i3en.3xlarge | i3en.6xlarge | i3en.12xlarge | i3en.24xlarge`

Considerazioni sulla connettività di rete

- Se la connettività di rete tra Outpost e la relativa AWS regione viene interrotta, i cluster continueranno a funzionare. Tuttavia, non potrai creare nuovi cluster o intraprendere nuove azioni nei cluster esistenti fino a quando la connettività non viene ripristinata. In caso di errori di istanza, l'istanza non verrà sostituita automaticamente. Inoltre, azioni come l'aggiunta di passaggi a un cluster in esecuzione, la verifica dello stato di esecuzione delle fasi e l'invio di CloudWatch metriche ed eventi verranno ritardate.
- Ti consigliamo di fornire una connettività di rete affidabile e ad alta disponibilità tra Outpost e la regione. AWS Se la connettività di rete tra Outpost e la relativa AWS regione viene interrotta per più di qualche ora, i cluster che hanno abilitato la protezione da terminazione continueranno a funzionare e i cluster che hanno disabilitato la protezione da terminazione potrebbero essere terminati.
- Se la connettività di rete subirà danni a causa della manutenzione di routine, suggeriamo di abilitare proattivamente una protezione dalle interruzioni. Più in generale, l'interruzione della connettività significa che eventuali dipendenze esterne che non sono locali per Outpost o la

rete clienti non saranno accessibili. Sono inclusi Amazon S3, DynamoDB utilizzato con la visualizzazione coerente EMRFS e Amazon RDS se un'istanza interna alla Regione è utilizzata per un cluster Amazon EMR con più nodi primari.

Creazione di un cluster Amazon EMR su AWS Outposts

La creazione di un cluster Amazon EMR su AWS Outposts è simile alla creazione di un cluster Amazon EMR nel cloud. AWS Quando crei un cluster Amazon EMR su AWS Outposts, devi specificare una EC2 sottorete Amazon associata al tuo Outpost.

Un Amazon VPC può estendersi su tutte le zone di disponibilità di una regione. AWS AWS Outposts sono estensioni delle zone di disponibilità e puoi estendere un Amazon VPC in un account per estenderlo a più zone di disponibilità e sedi Outpost associate. Quando si configura l'Outpost, si associa a esso un gruppo di sottoreti per estendere l'ambiente regionale del VPC alla struttura locale. Le istanze Outpost e i servizi correlati fanno parte del VPC regionale, in modo simile a una zona di disponibilità con le sottoreti associate. Per ulteriori informazioni, consulta la Guida per l'utente di AWS Outposts.

Console

Per creare un nuovo cluster Amazon EMR AWS Outposts con AWS Management Console, specifica una EC2 sottorete Amazon associata al tuo Outpost.

Console

Per creare un cluster con la console AWS Outposts

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Cluster configuration (Configurazione del cluster), seleziona Instance groups (Gruppi di istanze) oppure Instance fleets (Parchi istanze). Quindi, scegli un tipo di istanza dal menu a discesa Scegli il tipo di EC2 istanza o seleziona Azioni e scegli Aggiungi volumi EBS. Amazon EMR on AWS Outposts supporta volumi e tipi di istanze Amazon EBS limitati.
4. In Rete, seleziona una EC2 sottorete con un ID Outpost in questo formato: op-123456789.
5. Scegli qualsiasi altra opzione applicabile al cluster.

6. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per creare un cluster con AWS OutpostsAWS CLI

- Per creare un nuovo cluster Amazon EMR AWS Outposts con AWS CLI, specifica una EC2 sottorete associata al tuo Outpost, come nell'esempio seguente. *subnet-22XXXX01*Sostituiscilo con il tuo ID di EC2 sottorete 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
```

Cluster EMR su Local Zones AWS

A partire dalla versione 5.28.0 di Amazon EMR, puoi creare ed eseguire cluster Amazon EMR su una sottorete Local AWS Zones come estensione logica di una regione che supporta Local Zones. AWS Una zona locale consente alle funzionalità di Amazon EMR e a un sottoinsieme di AWS servizi, come i servizi di elaborazione e storage, di essere posizionati più vicino agli utenti per fornire un accesso a latenza molto bassa alle applicazioni eseguite localmente. Per un elenco delle zone locali disponibili, consulta Local Zones AWS. Per informazioni sull'accesso alle zone AWS locali disponibili, vedere Regioni, zone di disponibilità e zone locali.

Tipi di istanze supportati

I seguenti tipi di istanza sono disponibili per i cluster Amazon EMR sulle Local Zones. La disponibilità del tipo di istanza può variare in base all'area geografica.

Classe di istanza	Tipi di istanza
Uso generale	m5.xlarge | m5.2xlarge | m5.4xlarge | m5.12xlarge | m5.24xlarge | m5d.xlarge | m5d.2xlarge | m5d.4xlarge | m5d.12xlarge | m5d.24xlarge

Classe di istanza	Tipi di istanza
Ottimizzato per il calcolo	c5.xlarge | c5.2xlarge | c5.4xlarge | c5.9xlarge | c5.18xlarge | c5d.xlarge | c5d.2xlarge | c5d.4xlarge| c5d.9xlarge | c5d.18xlarge
Ottimizzato per la memoria	r5.xlarge | r5.2xlarge | r5.4xlarge | r5.12xlarge | r5d.xlarge | r5d.2xlarge | r5d.4xlarge | r5d.12xlarge | r5d.24xlarge
Ottimizzato per lo storage	i3en.xlarge | i3en.2xlarge | i3en.3xlarge | i3en.6xlarge | i3en.12xlarge | i3en.24xlarge

Creazione di un cluster Amazon EMR sulle Local Zones

Crea un cluster Amazon EMR su AWS Local Zones avviando il cluster Amazon EMR in una sottorete Amazon VPC associata a una Local Zone. È possibile accedere al cluster utilizzando il nome della Local Zone, ad esempio us-west-2-lax-1a nella console US West (Oregon) (Stati Uniti occidentali (Oregon)).

Le Local Zones attualmente non supportano i notebook Amazon EMR o le connessioni dirette ad Amazon EMR utilizzando l'interfaccia VPC endpoint ().AWS PrivateLink

Console

Per creare un cluster in una zona locale con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Rete, selezionare una EC2 sottorete con un ID di zona locale in questo formato: subnet 123abc | us-west-2-lax-1a.
4. Scegli un tipo di istanza o aggiungi volumi di archiviazione Amazon EBS per i gruppi di istanze uniformi o i parchi istanze.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per creare un cluster in una zona locale con AWS CLI

- Utilizzate il comando create-cluster, insieme al comando SubnetId for the Local Zone, come illustrato nell'esempio seguente. Sostituite subnet-22 XXXX1234567 con la Local Zone SubnetId e sostituite le altre opzioni, se necessario. Per ulteriori informazioni, consulta 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
```

Configura Docker per l'uso con i cluster Amazon EMR

Amazon EMR 6.x supporta Hadoop 3, che consente NodeManager a YARN di avviare container direttamente sul cluster Amazon EMR o all'interno di un container Docker. I contenitori Docker forniscono ambienti di esecuzione personalizzati in cui viene eseguito il codice dell'applicazione. L'ambiente di esecuzione personalizzato è isolato dall'ambiente di esecuzione di YARN e di altre applicazioni. NodeManager

I contenitori Docker possono includere librerie speciali utilizzate dall'applicazione e possono fornire diverse versioni di strumenti e librerie native, come R e Python. È possibile utilizzare gli strumenti Docker familiari per definire librerie e dipendenze di runtime per le applicazioni.

I cluster Amazon EMR 6.x sono configurati per impostazione predefinita per consentire alle applicazioni YARN, ad esempio Spark, di essere eseguite utilizzando container Docker. Per personalizzare la configurazione del container, modificare le opzioni di supporto Docker definite nei file `yarn-site.xml` e `container-executor.cfg` disponibili nella directory `/etc/hadoop/conf`. Per informazioni dettagliate su ciascuna opzione di configurazione e su come viene utilizzata, consulta Avvio di applicazioni utilizzando contenitori Docker.

È possibile scegliere di utilizzare Docker quando si invia un processo. Utilizzare le variabili seguenti per specificare il runtime Docker e l'immagine Docker.

- `YARN_CONTAINER_RUNTIME_TYPE=docker`

- `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`

Quando si utilizzano contenitori Docker per eseguire le applicazioni YARN, YARN scarica l'immagine Docker specificata quando si invia il lavoro. Per fare in modo che YARN risolva questa immagine Docker, è necessario configurarla con un Registro di sistema Docker. Le opzioni di configurazione per un Registro di sistema Docker dipendono dal fatto che si distribuisca il cluster utilizzando una sottorete pubblica o privata.

Registri Docker

Un registro Docker è un sistema di archiviazione e distribuzione per le immagini Docker. Per Amazon EMR ti consigliamo di usare Amazon ECR, che è un registro del container Docker completamente gestito, che consente di creare le proprie immagini personalizzate e ospitarle in un'architettura altamente disponibile e scalabile.

Considerazioni sulla distribuzione

I registri Docker richiedono l'accesso alla rete da ciascun host nel cluster. Questo perché ogni host scarica immagini dal Registro di sistema Docker quando l'applicazione YARN è in esecuzione nel cluster. Questi requisiti di connettività di rete possono limitare la scelta del Registro di sistema Docker, a seconda che si distribuisca il cluster Amazon EMR in una sottorete pubblica o privata.

Public subnet (Sottorete pubblica)

Quando i cluster EMR vengono distribuiti in una sottorete pubblica, i nodi che eseguono YARN NodeManager possono accedere direttamente a qualsiasi registro disponibile su Internet.

Sottorete privata

Quando i cluster EMR vengono distribuiti in una sottorete privata, i nodi che eseguono YARN NodeManager non hanno accesso diretto a Internet. Le immagini Docker possono essere ospitate in Amazon ECR e accessibili tramite. AWS PrivateLink

Per ulteriori informazioni su come utilizzare per consentire l'accesso AWS PrivateLink ad Amazon ECR in uno scenario di sottorete privata, consulta Configurazione per AWS PrivateLink Amazon ECS e Amazon ECR.

Configurazione dei registri Docker

Per utilizzare i registri Docker con Amazon EMR, è necessario configurare Docker per considerare attendibile il Registro di sistema specifico che si desidera utilizzare per risolvere le immagini Docker.

I registri di trust predefiniti sono locali (privati) e centos. Per utilizzare altri repository pubblici oppure Amazon ECR, è possibile ignorare le impostazioni `docker.trusted.registries` in `/etc/hadoop/conf/container-executor.cfg` utilizzando l'API di classificazione EMR con la chiave di classificazione `container-executor`.

Nell'esempio seguente viene illustrato come configurare il cluster per considerare attendibile sia un repository pubblico, denominato `your-public-repo`, sia un endpoint del Registro di sistema ECR, denominato `123456789123.dkr.ecr.us-east-1.amazonaws.com`. Se si utilizza ECR, sostituire questo endpoint con l'endpoint ECR specifico.

```
[
  {
    "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"
            }
        }
    ]
  }
]
```

Per avviare un cluster Amazon EMR 6.0.0 con questa configurazione utilizzando AWS Command Line Interface (AWS CLI), crea un file denominato `container-executor.json` con il contenuto della precedente configurazione JSON di ontainer-executor. Quindi, utilizzare i seguenti comandi per avviare il cluster.

```
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
```

Configurazione di YARN per accedere ad Amazon ECR su EMR 6.0.0 e versioni precedenti

Se sei nuovo su Amazon ECR, segui le istruzioni riportate in Guida introduttiva di Amazon ECR e verifica di avere accesso ad Amazon ECR da ogni istanza del cluster Amazon EMR.

Su EMR 6.0.0 e versioni successive, per accedere ad Amazon ECR utilizzando il comando Docker è necessario innanzitutto generare le credenziali. Per verificare se YARN può accedere alle immagini da Amazon ECR, utilizza la variabile di ambiente container `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` per passare un riferimento alle credenziali generate.

Eseguire il seguente comando su uno dei nodi principali per ottenere la riga di accesso per l'account ECR.

```
aws ecr get-login --region us-east-1 --no-include-email
```

Il comando `get-login` genera il comando CLI Docker corretto da eseguire per creare le credenziali. Copiare ed eseguire l'output da `get-login`.

```
sudo docker login -u AWS -p <password> https://<account-id>.dkr.ecr.us-
east-1.amazonaws.com
```

Questo comando genera un file `config.json` nella cartella `/root/.docker`. Copia questo file in HDFS in modo che i processi inviati al cluster possano utilizzarlo per autenticarsi in Amazon ECR.

Eseguire i comandi riportati di seguito per copiare il file `config.json` nella directory home.

```
mkdir -p ~/.docker
sudo cp /root/.docker/config.json ~/.docker/config.json
sudo chmod 644 ~/.docker/config.json
```

Eseguire i comandi riportati di seguito per inserire config.json in HDFS in modo che possa essere utilizzato dai processi in esecuzione nel cluster.

```
hadoop fs -put ~/.docker/config.json /user/hadoop/
```

YARN può accedere a ECR come registro immagini Docker ed estrarre contenitori durante l'esecuzione del processo.

Dopo aver configurato i registri Docker e YARN, è possibile eseguire applicazioni YARN utilizzando contenitori Docker. Per ulteriori informazioni, consulta Esecuzione di applicazioni Spark con Docker utilizzando Amazon EMR 6.0.0.

In EMR 6.1.0 e versioni successive, non è necessario configurare manualmente l'autenticazione su Amazon ECR. Se viene rilevato un registro Amazon ECR nella chiave di classificazione `container-executor`, viene attivata la funzione di autenticazione automatica Amazon ECR e YARN gestisce il processo di autenticazione quando si invia un processo Spark con un'immagine ECR. È possibile verificare se l'autenticazione automatica è abilitata controllando `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` in yarn-site. L'autenticazione automatica è abilitata e l'impostazione di autenticazione YARN è impostata su `true` se `docker.trusted.registries` contiene un URL del registro ECR.

Prerequisiti per l'utilizzo dell'autenticazione automatica per Amazon ECR

- EMR versione 6.1.0 o versioni successive
- Il registro ECR incluso nella configurazione si trova nella stessa Regione con il cluster
- Ruolo IAM con autorizzazioni per ottenere token di autorizzazione ed estrarre qualsiasi immagine

Per ulteriori informazioni, consulta Configurazione di Amazon ECR.

Come abilitare l'autenticazione automatica

Segui Configurazione dei registri Docker per impostare un registro Amazon ECR come registro attendibile e assicurarti che il repository Amazon ECR e il cluster si trovino nella stessa Regione.

Per attivare questa caratteristica anche quando il registro ECR non è impostato nel registro attendibile, utilizza la classificazione di configurazione per impostare `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` su `true`.

Come disabilitare l'autenticazione automatica

Per impostazione predefinita, l'autenticazione automatica viene disattivata se non viene rilevato alcun registro Amazon ECR nel registro attendibile.

Per disattivare l'autenticazione automatica anche quando il registro Amazon ECR è impostato nel registro attendibile, utilizza la classificazione di configurazione per impostare `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` su `false`.

Come verificare se l'autenticazione automatica è abilitata in un cluster

Sul nodo master, utilizza un editor di testo come `vi` per visualizzare il contenuto del file: `vi /etc/hadoop/conf.empty/yarn-site.xml`. Verifica il valore di `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled`.

Controlla la terminazione dei cluster Amazon EMR

Questa sezione descrive le opzioni per la arresta dei cluster Amazon EMR. Copre la protezione di terminazione automatica e terminazione e il modo in cui interagiscono con altre caratteristiche di Amazon EMR.

Puoi arrestare un cluster Amazon EMR nei seguenti modi:

- Terminazione dopo l'esecuzione dell'ultima fase: crea un cluster transitorio che si arresta dopo il completamento di tutti i fasi.
- Terminazione automatica (dopo inattività): crea un cluster con una policy di terminazione automatica che si arresta dopo un periodo di inattività specificato. Per ulteriori informazioni, consulta Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR.
- Terminazione manuale: crea un cluster di lunga durata che continua a essere eseguito finché non viene terminato deliberatamente. Per informazioni su come terminare un cluster manualmente, consulta Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

Puoi anche impostare la protezione dalla terminazione su un cluster per evitare di chiudere EC2 le istanze accidentalmente o per errore.

Quando Amazon EMR chiude il cluster, tutte le EC2 istanze Amazon del cluster si chiudono. I dati dell'istanza e i volumi EBS non sono più disponibili né recuperabili. La comprensione e la gestione della terminazione dei cluster è fondamentale per sviluppare una strategia di gestione e conservazione dei dati per le operazioni di scrittura su Amazon S3 e il bilanciamento dei costi.

Argomenti

- Configurazione di un cluster Amazon EMR per continuare o terminare dopo l'esecuzione delle fasi
- Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR
- Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali

Configurazione di un cluster Amazon EMR per continuare o terminare dopo l'esecuzione delle fasi

Questo argomento spiega le differenze tra l'utilizzo di un cluster di lunga durata e la creazione di un cluster transitorio che si arresta dopo l'esecuzione dell'ultimo passaggio. Viene inoltre descritto come configurare l'esecuzione delle fasi per un cluster.

Crea un cluster di lunga durata

Per impostazione predefinita, i cluster creati con la console o il sono di lunga durata. AWS CLI I cluster di lunga durata continuano a funzionare, accettano il lavoro e accumulano addebiti fino a quando non si interviene per arrestarli.

Un cluster di lunga durata è efficace nelle situazioni seguenti:

- Quando è necessario eseguire una query sull dati in modo interattivo o automatico.
- Quando è necessario interagire regolarmente con applicazioni Big Data ospitate sul cluster.
- Quando se si esegue un'elaborazione periodica su un set di dati di dimensioni talmente grandi che risulterebbe inefficace avviare nuovi cluster e caricare ogni volta i dati.

È inoltre possibile impostare la protezione dalla terminazione su un cluster a esecuzione prolungata per evitare la chiusura accidentale o per errore delle EC2 istanze. Per ulteriori informazioni, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

> **Note**
>
> Amazon EMR abilita automaticamente la protezione da terminazione per tutti i cluster con più nodi primari e sostituisce tutte le impostazioni di esecuzione di fasi fornite durante la creazione del cluster. È possibile disattivare la protezione da cessazione dopo l'avvio del

cluster. Consultare Configurazione della protezione da cessazione per i cluster in esecuzione. Per chiudere un cluster con più nodi primari, è necessario modificare gli attributi del cluster per disabilitare la protezione da terminazione. Per istruzioni, consultare Terminazione di un cluster Amazon EMR con più nodi primari.

Configurare un cluster da terminare dopo l'esecuzione della fase

Quando si configura la terminazione dopo l'esecuzione della fase, il cluster viene avviato, esegue le azioni bootstrap e quindi esegue le fasi specificate. Non appena viene completato l'ultimo passaggio, Amazon EMR termina le istanze Amazon del cluster. EC2 Per i cluster avviati utilizzando l'API di Amazon EMR, l'esecuzione delle fasi è abilitata per impostazione predefinita.

La terminazione dopo l'esecuzione della fase è efficace per i cluster che esegue un'operazione di elaborazione periodica, ad esempio un'esecuzione giornaliera dell'elaborazione dei dati. L'esecuzione delle fasi consente inoltre di garantire che viene fatturato solo il tempo necessario per elaborare i dati. Per ulteriori informazioni sulle fasi, consulta Invia il lavoro a un cluster Amazon EMR.

Console

Per attivare la terminazione dopo l'esecuzione della fase con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Steps (Fasi), scegli Add step (Aggiungi fase). Inserisci i valori appropriati nei campi della finestra di dialogo Add step (Aggiungi fase). Le opzioni variano a seconda del tipo di fase. Per aggiungere la fase e uscire dalla finestra di dialogo, scegli Add step (Aggiungi fase).
4. In Cluster termination (Terminazione del cluster), seleziona la casella di controllo Terminate cluster after last step completes (Termina il cluster dopo il completamento dell'ultima fase).
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per attivare la terminazione dopo l'esecuzione di una fase con AWS CLI

- Specifica il parametro `--auto-terminate` quando utilizzi il comando `create-cluster` per creare un cluster transitorio.

  L'esempio seguente illustra come utilizzare il parametro `--auto-terminate`. È possibile digitare il comando seguente e sostituirlo *myKey* con il nome della propria EC2 key pair.

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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

Per disattivare la terminazione dopo l'esecuzione di fasi con l'API Amazon EMR all'avvio del cluster

1. Quando utilizzi l'RunJobFlowazione per creare un cluster, imposta la KeepJobFlowAliveWhenNoStepsproprietà su. `false`
2. Per modificare la configurazione della terminazione dopo l'esecuzione delle fasi con l'API Amazon EMR dopo il lancio del cluster:

   Usa l' SetKeepJobFlowAliveWhenNoSteps azione.

Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR

Una policy di terminazione automatica consente di orchestrare la pulizia del cluster senza la necessità di monitorare e terminare manualmente i cluster inutilizzati. Quando si aggiunge una policy di terminazione automatica a un cluster, si specifica la quantità di tempo di inattività dopo il quale il cluster deve arrestarsi automaticamente.

A seconda della versione di rilascio, Amazon EMR utilizza criteri diverse per contrassegnare un cluster come inattivo. Nella tabella seguente viene illustrato come Amazon EMR determina l'inattività del cluster.

Quando utilizzi...	Un cluster è considerato inattivo quando...
Amazon EMR versione 5.34.0 e successive, e versione 6.4.0 e successive	• Non esistono applicazioni YARN attive • L'utilizzo dell'HDFS è inferiore al 10% • Non ci sono connessioni attive per EMR Notebooks o EMR Studio • Non sono in uso interfacce utente dell'applicazione on-cluster • Non ci sono passaggi in sospeso
Amazon EMR versioni 5.30.0 - 5.33.0 e 6.1.0 - 6.3.0	• Non esistono applicazioni YARN attive • Il cluster non ha processi Spark attivi **Note** Amazon EMR contrassegna un cluster come inattivo e potrebbe terminare automaticamente il cluster anche se si dispone di un kernel Python3

Quando utilizzi...	Un cluster è considerato inattivo quando...
	attivo. Questo perché l'esecuzione di un kernel Python3 non invia un processo Spark sul cluster. Per utilizzare la terminazione automatica con un kernel Python3, consigliamo di utilizzare Amazon EMR versione 6.4.0 o successive.

> **Note**
>
> Amazon EMR versione 6.4.0 e successive supportano un file su cluster per il rilevamento dell'attività sul nodo primario: `/emr/metricscollector/isbusy`. Quando si utilizza un cluster per eseguire script di shell o applicazioni non YARN, è possibile toccare o aggiornare periodicamente `isbusy` per indicare ad Amazon EMR che il cluster non è inattivo.

È possibile allegare una policy di terminazione automatica quando si crea un cluster o si aggiunge una policy a un cluster esistente. Per modificare o disabilitare la terminazione automatica, è possibile aggiornare o rimuovere la policy.

Considerazioni

Prima di utilizzare una policy di terminazione automatica, considera le seguenti caratteristiche e limitazioni:

- Di seguito Regioni AWS, la terminazione automatica di Amazon EMR è disponibile con Amazon EMR 6.14.0 e versioni successive:
  - Europa (Spagna) (eu-south-2)
- Di seguito Regioni AWS, la terminazione automatica di Amazon EMR è disponibile con Amazon EMR 5.30.0 e 6.1.0 e versioni successive:
  - Stati Uniti orientali (Virginia settentrionale) (us-east-1)
  - Stati Uniti orientali (Ohio) (us-east-2)
  - Stati Uniti occidentali (Oregon) (us-west-2)

- Stati Uniti occidentali (California settentrionale) (us-west-1)
- Africa (Città del Capo) (af-south-1)
- Asia Pacifico (Hong Kong) (ap-east-1)
- Asia Pacifico (Mumbai) (ap-south-1)
- Asia Pacifico (Hyderabad) (ap-south-2)
- Asia Pacifico (Seoul) (ap-northeast-2)
- Asia Pacifico (Osaka-Locale) (ap-northeast-3)
- Asia Pacifico (Singapore) (ap-southeast-1)
- Asia Pacifico (Sydney) (ap-southeast-2)
- Asia Pacific (Giacarta) (ap-southeast-3)
- Asia Pacifico (Tokyo) (ap-northeast-1)
- Canada (Centrale) (ca-central-1)
- Sud America (San Paolo) (sa-east-1)
- Europa (Francoforte) (eu-central-1)
- Europa (Zurigo) (eu-central-2)
- Europa (Irlanda) (eu-west-1)
- Europa (Londra) (eu-west-2)
- Europa (Milano) (eu-south-1)
- Europe (Parigi) (eu-west-3)
- Europa (Stoccolma) (eu-north-1)
- Israele (Tel Aviv) (il-central-1)
- Medio Oriente (EAU) (me-central-1)
- Cina (Pechino) cn-north-1
- Cina (Ningxia) cn-nordovest-1
- AWS GovCloud (Stati Uniti orientali) (-1) us-gov-east
- AWS GovCloud (Stati Uniti occidentali) (us-gov-west-1)

- Il timeout inattivo è predefinito di 60 minuti (un'ora) quando non si specifica un importo. È possibile specificare un timeout minimo di inattività di un minuto e un timeout massimo di 7 giorni.
- Con Amazon EMR versioni 6.4.0 e successive, la terminazione automatica è abilitata per impostazione predefinita quando si crea un nuovo cluster tramite la console Amazon EMR.

- Amazon EMR pubblica Amazon CloudWatch parametri ad alta risoluzione quando abiliti la terminazione automatica per un cluster. Puoi utilizzare queste metriche per monitorare l'attività e l'inattività del cluster. Per ulteriori informazioni, consulta Parametri della capacità del cluster.
- La terminazione automatica non è supportata quando utilizzi applicazioni non basate su Yarn come Presto, Trino o. HBase
- Per utilizzare la terminazione automatica in API Gateway, il processo di raccolta dei parametri deve essere in grado di connettersi all'endpoint API pubblico. Se utilizzi un nome DNS privato con Amazon Virtual Private Cloud, la terminazione automatica non funzionerà correttamente. Per garantire che la terminazione automatica funzioni, è consigliabile eseguire una delle seguenti operazioni:
  - Rimuovi l'endpoint VPC dell'interfaccia API Gateway dal tuo Amazon VPC.
  - Segui le istruzioni in Perché ricevo un errore HTTP 403 Forbidden quando mi connetto al mio API Gateway APIs da un VPC? per disabilitare l'impostazione del nome DNS privato.
  - In alternativa, avvia il cluster in una sottorete privata. Per ulteriori informazioni, consulta l'argomento in Sottoreti private.
- (EMR rilascio 5.30.0 e successivi) Se si rimuove la regola predefinita Allow All (Consenti tutto) in uscita su 0.0.0.0/ nel gruppo di sicurezza primario, è necessario aggiungere una regola per consentire la connettività TCP in uscita al gruppo di sicurezza di accesso al servizio sulla porta 9443. Inoltre, il gruppo di sicurezza di accesso al servizio deve consentire il traffico TCP in ingresso sulla porta 9443 dal gruppo di sicurezza primario. Per ulteriori informazioni sulla configurazione dei gruppi di sicurezza, consulta la sezione Amazon EMR-managed security group for the primary instance (private subnets) (Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria [sottoreti private]).

Autorizzazioni per l'utilizzo della terminazione automatica

Per poter applicare e gestire le policy di terminazione automatica per Amazon EMR, devi collegare le autorizzazioni elencate nell'esempio seguente sulle policy di autorizzazioni IAM alle risorse IAM che gestiscono il cluster EMR.

```
{
  "Version": "2012-10-17",
  "Statement": {
      "Sid": "AllowAutoTerminationPolicyActions",
      "Effect": "Allow",
      "Action": [
```

```
        "elasticmapreduce:PutAutoTerminationPolicy",
        "elasticmapreduce:GetAutoTerminationPolicy",
        "elasticmapreduce:RemoveAutoTerminationPolicy"
      ],
      "Resource": "<your-resources>"
    }
}
```

Allega, aggiorna o rimuovi una policy di terminazione automatica

Questa sezione include istruzioni per allegare, aggiornare o rimuovere una policy di terminazione automatica da un cluster Amazon EMR. Prima di lavorare con le policy di terminazione automatica, assicurati di disporre delle autorizzazioni IAM necessarie. Consultare Autorizzazioni per l'utilizzo della terminazione automatica.

Console

Per allegare una politica di terminazione automatica quando si crea un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Cluster termination (Terminazione del cluster), seleziona Terminate cluster after idle time (Termina il cluster dopo il tempo di inattività).
4. Specifica il numero di ore e minuti di inattività dopo i quali il cluster deve terminare automaticamente. Il tempo di inattività predefinito è 1 ora.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

Per allegare, aggiornare o rimuovere una politica di terminazione automatica su un cluster in esecuzione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.

3. Nella scheda Properties (Proprietà) della pagina dei dettagli del cluster, cerca Cluster termination (Terminazione del cluster) e seleziona Edit (Modifica).
4. Seleziona o deseleziona Enable auto-termination (Abilita la terminazione automatica) per attivare o disattivare la caratteristica. Se attivi la terminazione automatica, specifica il numero di ore e minuti di inattività dopo il quale il cluster termina automaticamente. Quindi seleziona Save changes (Salva modifiche) per confermare.

AWS CLI

Prima di iniziare

Prima di utilizzare le policy di terminazione automatica, si consiglia di eseguire l'aggiornamento alla versione più recente della AWS CLI. Per le istruzioni, consulta Installazione, aggiornamento e disinstallazione della AWS CLI.

Per allegare o aggiornare una policy di terminazione automatica utilizzando la AWS CLI

- Puoi utilizzare il comando `aws emr put-auto-termination-policy` per allegare o aggiornare una policy di terminazione automatica su un cluster.

  L'esempio seguente specifica 3600 secondi per. *`IdleTimeout`* Se non si specifica*`IdleTimeout`*, il valore predefinito è un'ora.

```
aws emr put-auto-termination-policy \
--cluster-id <your-cluster-id> \
--auto-termination-policy IdleTimeout=3600
```

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

  È anche possibile specificare un valore per `--auto-termination-policy` quando si utilizza il comando `aws emr create-cluster`. Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta il AWS CLI Command Reference.

Per rimuovere una politica di terminazione automatica con AWS CLI

- Utilizzo il comando `aws emr remove-auto-termination-policy` per rimuovere una policy di terminazione automatica da un cluster. Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta il AWS CLI Command Reference.

```
aws emr remove-auto-termination-policy --cluster-id <your-cluster-id>
```

Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali

La protezione dalla terminazione protegge i cluster dalla chiusura accidentale, il che può essere particolarmente utile per i cluster di lunga durata che elaborano carichi di lavoro critici. Quando la protezione da cessazione è abilitata su un cluster di lunga durata, è necessario rimuovere esplicitamente la protezione da cessazione dal cluster prima di poterlo terminare. Questo aiuta a garantire che EC2 le istanze non vengano chiuse a causa di un incidente o di un errore. È possibile abilitare la protezione da cessazione al momento della creazione di un cluster ed è possibile modificare l'impostazione su un cluster in esecuzione.

Se la protezione da cessazione è abilitata, l'operazione `TerminateJobFlows` nell'API di Amazon EMR non funziona. Gli utenti non sono in grado di terminare il cluster utilizzando questa API né il comando `terminate-clusters` di AWS CLI. L'API restituisce un errore e l'interfaccia a riga di comando (CLI) si chiude con un codice restituito diverso da zero. Se utilizzi la console Amazon EMR per terminare un cluster, ti viene richiesto di eseguire una fase supplementare per disattivare la protezione da terminazione.

> **Warning**
>
> La protezione dalla terminazione non garantisce che i dati vengano conservati in caso di errore umano o soluzione alternativa, ad esempio se un comando di riavvio viene emesso dalla riga di comando mentre si è connessi all'istanza tramite SSH, se un'applicazione o uno script in esecuzione sull'istanza emette un comando di riavvio o se l'API Amazon o EC2 Amazon EMR viene utilizzata per disabilitare la protezione dalla terminazione. Questo vale anche se utilizzi le versioni 7.1 e successive di Amazon EMR e un'istanza diventa non integra e irrecuperabile. Anche se la protezione da cessazione è abilitata, i dati salvati nell'archiviazione dell'istanza, inclusi i dati HDFS, possono andare persi. Scrivi l'output dei

dati nei percorsi Amazon S3 e crea strategie di backup appropriate per i tuoi requisiti di business continuity.

Questo tipo di protezione non pregiudica la possibilità di dimensionare le risorse del cluster mediante una delle seguenti operazioni:

- Ridimensionamento manuale di un cluster con o. AWS Management Console AWS CLI Per ulteriori informazioni, consulta Ridimensiona manualmente un cluster Amazon EMR in esecuzione.
- Rimozione delle istanze da un gruppo di istanze principali o di task utilizzando una policy di riduzione con scalabilità automatica. Per ulteriori informazioni, consulta Utilizzo del ridimensionamento automatico con una politica personalizzata, ad esempio gruppi in Amazon EMR.
- Rimozione delle istanze da un parco istanze mediante la riduzione della capacità di destinazione. Per ulteriori informazioni, consulta Opzioni del parco istanze.

Protezione dalla terminazione e Amazon EC2

L'impostazione di protezione dalla terminazione in un cluster Amazon EMR corrisponde all'attributo per tutte le istanze EC2 Amazon `DisableApiTermination` del cluster. Ad esempio, se abiliti la protezione dalle terminazioni in un cluster EMR, Amazon EMR `DisableApiTermination` imposta automaticamente su true per tutte EC2 le istanze all'interno del cluster EMR. Lo stesso vale se disabiliti la protezione dalla terminazione. Amazon EMR si imposta automaticamente su false `DisableApiTermination` per tutte le EC2 istanze all'interno del cluster EMR. Se interrompi o riduci un cluster da Amazon EMR e le impostazioni di EC2 Amazon sono in conflitto per EC2 un'istanza, Amazon EMR dà la priorità all'`DisableApiStop`impostazione Amazon EMR rispetto EC2 alle impostazioni and di Amazon e continua a `DisableApiTermination` terminare l'istanza. EC2

Ad esempio, puoi utilizzare la EC2 console Amazon per abilitare la protezione dalla terminazione su un' EC2 istanza Amazon in un cluster EMR con la protezione dalla terminazione disabilitata. Se interrompi o riduci il cluster con la console Amazon EMR, AWS CLI l'API Amazon EMR, Amazon EMR `DisableApiTermination` sovrascrive l'impostazione, la imposta su false e termina l'istanza insieme ad altre istanze.

Puoi anche utilizzare la EC2 console Amazon per abilitare la protezione dall'arresto su un' EC2 istanza Amazon in un cluster EMR con la protezione dalla terminazione disabilitata. Se interrompi o

riduci il cluster, Amazon EMR `DisableApiStop` imposta su false in EC2 Amazon e termina l'istanza insieme ad altre istanze.

Amazon EMR sostituisce l'`DisableApiStop`impostazione solo quando chiudi o riduci un cluster. Quando abiliti o disabiliti la protezione dalla terminazione in un cluster EMR, Amazon EMR non modifica `disableApiStop` l'impostazione per nessuna delle istanze EC2 nel rispettivo cluster EMR.

> **Important**
>
> Se crei un'istanza come parte di un cluster Amazon EMR con protezione dalla terminazione e utilizzi l' EC2 API o AWS CLI i comandi di Amazon per modificare l'istanza in modo che `DisableApiTermination` sia così`false`, e quindi l' EC2 API o AWS CLI i comandi Amazon eseguono l'`TerminateInstances`operazione, l'istanza Amazon EC2 viene terminata.

Protezione da cessazione e nodi YARN con stato Unhealthy (Non integro)

Amazon EMR controlla periodicamente lo stato di Apache Hadoop YARN dei nodi in esecuzione su istanze Amazon core e task in un cluster. EC2 Lo stato di salute viene segnalato dal servizio di controllo dello stato di salute. NodeManager Se un nodo segnala`UNHEALTHY`, il controller di istanza Amazon EMR aggiunge il nodo a una denylist e non gli assegna contenitori YARN finché non torna a funzionare. A seconda dello stato di protezione dalla terminazione, della sostituzione dei nodi non integri e della versione di rilascio di Amazon EMR, Amazon EMR sostituirà l'istanza non integra o interromperà l'allocazione dei controller all'istanza.

Protezione dalla terminazione e terminazione dopo l'esecuzione di una fase

Quando abiliti la terminazione dopo l'esecuzione della fase e abiliti anche la protezione dalla terminazione, Amazon EMR ignora la protezione dalla terminazione.

Quando si inviano fasi a un cluster, è possibile impostare la proprietà `ActionOnFailure` per stabilire cosa accede se la fase non può completare l'esecuzione a causa di un errore. I valori possibile per questa impostazione sono `TERMINATE_CLUSTER` (`TERMINATE_JOB_FLOW` con le versioni precedenti), `CANCEL_AND_WAIT` e `CONTINUE`. Per ulteriori informazioni, consulta Invia il lavoro a un cluster Amazon EMR.

Se una fase configurata con `ActionOnFailure` set to fallisce`CANCEL_AND_WAIT`, se è abilitata la terminazione dopo l'esecuzione della fase, il cluster si interrompe senza eseguire i passaggi successivi.

Se una fase, configurata con `ActionOnFailure` impostato su `TERMINATE_CLUSTER`, non riesce, è possibile utilizzare la tabella delle impostazioni riportata di seguito per determinare il risultato.

ActionOnFailure	Terminazione dopo l'esecuzione della fase	Termination protection (Protezione da cessazione)	Risultato
`TERMINATE_CLUSTER`	Abilitato	Disabilitato	Il cluster viene terminato
	Abilitato	Abilitato	Il cluster viene terminato
	Disabilitato	Abilitato	Il cluster non viene terminato
	Disabilitato	Disabilitato	Il cluster viene terminato

Protezione da cessazione e istanze Spot

La protezione dalla terminazione di Amazon EMR non impedisce la chiusura di un'istanza Amazon EC2 Spot quando il prezzo Spot supera il prezzo Spot massimo.

Configurazione della protezione da cessazione all'avvio di un cluster

Puoi abilitare o disabilitare la protezione dalla terminazione quando avvii un cluster utilizzando la console, l'API o l'API AWS CLI.

Per i cluster a nodo singolo, le impostazioni predefinite di protezione dalla terminazione sono le seguenti:

- Avvio di un cluster tramite Amazon EMR Console —Termination Protection è disabilitato per impostazione predefinita.
- L'avvio di un cluster tramite AWS CLI **`aws emr create-cluster`** —Termination Protection è disabilitato a meno che non sia specificato. `--termination-protected`
- L'avvio di un cluster tramite il comando RunJobFlowAPI Amazon EMR: Termination Protection è disabilitato a meno che `TerminationProtected` il valore booleano non sia impostato su. `true`

Per i cluster ad alta disponibilità, le impostazioni predefinite di protezione dalla terminazione sono le seguenti:

- Avvio di un cluster tramite Amazon EMR Console: la protezione dalle terminazioni è abilitata per impostazione predefinita.
- L'avvio di un cluster tramite AWS CLI **`aws emr create-cluster`** —Termination Protection è disabilitato a meno che non sia specificato. `--termination-protected`
- L'avvio di un cluster tramite il comando RunJobFlowAPI Amazon EMR: Termination Protection è disabilitato a meno che `TerminationProtected` il valore booleano non sia impostato su. `true`

Console

Per attivare o disattivare la protezione dalla terminazione quando crei un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Per EMR release version (Versione del rilascio EMR), scegli emr-6.6.0 o un rilascio successivo.
4. In Terminazione del cluster e sostituzione dei nodi, assicurati che l'opzione Usa protezione dalla terminazione sia preselezionata oppure deseleziona la selezione per disattivarla.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per attivare o disattivare la protezione dalla terminazione quando crei un cluster utilizzando AWS CLI

- Con AWS CLI, è possibile avviare un cluster con la protezione dalla terminazione abilitata con il `create-cluster` comando con il `--termination-protected` parametro. Per impostazione predefinita, la protezione da cessazione è disabilitata.

  Nell'esempio seguente viene creato un cluster con protezione da cessazione abilitata:

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per la leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

  ```
  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
  ```

  Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Configurazione della protezione da cessazione per i cluster in esecuzione

È possibile configurare la protezione da terminazione per un cluster in esecuzione con la console o la AWS CLI.

Console

Per attivare o disattivare la protezione dalla terminazione per un cluster in esecuzione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.

2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Properties (Proprietà) della pagina dei dettagli del cluster, cerca Cluster termination (Terminazione del cluster) e seleziona Edit (Modifica).
4. Seleziona o deseleziona la casella di controllo Use termination protection (Utilizza la protezione da terminazione) per attivare o disattivare la caratteristica. Quindi seleziona Save changes (Salva modifiche) per confermare.

AWS CLI

Per attivare o disattivare la protezione dalla terminazione per un cluster in esecuzione utilizzando il AWS CLI

- Per abilitare la protezione da terminazione su un cluster in esecuzione con la AWS CLI, utilizza il comando `modify-cluster-attributes` con il parametro `--termination-protected`. Per disabilitarla, utilizza il parametro `--no-termination-protected`.

  L'esempio seguente abilita la protezione dalla terminazione sul cluster con ID: *j-3KVTXXXXXX7UG*

  ```
  aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --termination-protected
  ```

  Nell'esempio seguente viene disabilitata la protezione da cessazione sullo stesso cluster:

  ```
  aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-protected
  ```

Sostituzione di nodi non integri con Amazon EMR

Amazon EMR utilizza periodicamente il servizio di controllo dello NodeManager stato di salute in Apache Hadoop per monitorare lo stato dei nodi principali nei cluster Amazon EMR su Amazon. EC2 Se un nodo non funziona in modo ottimale, viene contrassegnato come non integro e il responsabile dello stato lo segnala al controller Amazon EMR. Il controller Amazon EMR aggiunge il nodo a una lista di negati, impedendo al nodo di ricevere nuove applicazioni YARN fino a quando lo stato del nodo non migliora.

> **Note**
>
> Un motivo comune per cui un nodo non è integro è l'esaurimento dello spazio su disco. Per ulteriori informazioni su quando lo spazio su disco di un nodo principale è quasi esaurito, è utile il seguente articolo del Re:post Knowledge Center: Perché il nodo principale del mio cluster Amazon EMR sta esaurendo lo spazio su disco?

> **Note**
>
> Hadoop offre la possibilità di eseguire controlli personalizzati dello stato dei nodi. Questo è spiegato più dettagliatamente nella documentazione di Apache Hadoop all'indirizzo. NodeManager

Puoi scegliere se Amazon EMR deve terminare i nodi non integri o mantenerli nel cluster. Se disattivi la sostituzione dei nodi non sani, questi restano nell'elenco dei nodi non accettati e continuano a essere conteggiati ai fini della capacità del cluster. Puoi comunque connetterti alla tua istanza EC2 principale di Amazon per la configurazione e il ripristino, in modo da poter ridimensionare il cluster se desideri aggiungere capacità. Per ulteriori informazioni su come funzionano la sostituzione e la terminazione dei nodi, consulta Utilizzo della protezione dalla terminazione.

Se è attivata la sostituzione dei nodi non integri, Amazon EMR termina un nodo principale non integro ed effettua il provisioning di una nuova istanza, in base al numero di istanze nel gruppo di istanze o in base alla capacità target per le flotte di istanze. Se alcuni nodi non sono integri per più di 45 minuti, Amazon EMR sostituirà correttamente i nodi. Se la disattivazione regolare di un nodo non viene completata entro un'ora, il nodo viene chiuso forzatamente, a meno che la chiusura non porti il cluster al di sotto dei limiti del fattore di replica o della capacità HDFS.

> **Important**
>
> Tieni presente che il tempo necessario prima che un nodo venga disattivato o terminato correttamente può essere soggetto a modifiche.
> Sebbene una sostituzione non corretta dei nodi riduca in modo significativo la possibilità di perdita di dati, non elimina completamente il rischio. I dati HDFS possono andare persi definitivamente durante la sostituzione graduale di un'istanza principale non integra. Ti consigliamo di eseguire sempre il backup dei dati.

Per ulteriori informazioni sull'identificazione dei nodi non integri e sul ripristino, consulta Errori nelle risorse. Inoltre, per ulteriori best practice da seguire per mantenere l'integrità di un cluster, consulta la seguente documentazione sull'errore di risorsa Il cluster Amazon EMR termina con NO_SLAVE_LEFT e i nodi principali FAILED_BY_MASTER.

Amazon EMR pubblica CloudWatch Amazon Events per la sostituzione di nodi non integri, così puoi tenere traccia di ciò che accade con le tue istanze principali non integre. Per ulteriori informazioni, consulta Eventi di sostituzione dei nodi non integri.

Impostazioni predefinite per la sostituzione dei nodi e la protezione dalla terminazione

La sostituzione dei nodi non integri è disponibile per tutte le versioni di Amazon EMR, ma le impostazioni predefinite dipendono dall'etichetta di rilascio scelta. Puoi modificare qualsiasi di queste impostazioni configurando la sostituzione dei nodi non integri quando crei un nuovo cluster o accedendo alla configurazione del cluster in qualsiasi momento.

Se stai creando un cluster a nodo singolo o un cluster ad alta disponibilità che esegue Amazon EMR versione 7.0 o precedente, l'impostazione predefinita della sostituzione dei nodi non integri dipende dalla protezione dalla terminazione:

- L'attivazione della protezione dalla terminazione disabilita la sostituzione non corretta dei nodi.
- La disabilitazione della protezione dalla terminazione consente la sostituzione non corretta dei nodi.

Configurazione della sostituzione di nodi non integri all'avvio di un cluster

Puoi abilitare o disabilitare la sostituzione dei nodi non integri quando avvii un cluster utilizzando la console AWS CLI, l'o l'API.

L'impostazione predefinita per la sostituzione dei nodi non integri dipende da come si avvia il cluster:

- Console Amazon EMR: la sostituzione dei nodi non integri è abilitata per impostazione predefinita.
- AWS CLI `aws emr create-cluster`— la sostituzione dei nodi non integri è abilitata per impostazione predefinita, a meno che tu non lo specifichi. `--no-unhealthy-node-replacement`
- Comando RunJobFlow API Amazon EMR: la sostituzione dei nodi non integri è abilitata per impostazione predefinita a meno che non imposti il valore `UnhealthyNodeReplacement` booleano su o. `True` `False`

Console

Per attivare o disattivare la sostituzione dei nodi non integri quando crei un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Per la versione di rilascio EMR, scegli l'etichetta di rilascio di Amazon EMR che desideri.
4. In Terminazione del cluster e sostituzione del nodo, assicurati che Unhealthy node replacement (consigliato) sia preselezionato oppure deseleziona la selezione per disattivarla.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per attivare o disattivare la sostituzione dei nodi non integri quando crei un cluster utilizzando AWS CLI

- Con AWS CLI, è possibile avviare un cluster con la sostituzione dei nodi non integri abilitata con il `create-cluster` comando con il `--unhealthy-node-replacement` parametro. La sostituzione dei nodi non integri è attiva per impostazione predefinita.

  L'esempio seguente crea un cluster con la sostituzione dei nodi non integri abilitata:

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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
```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta Comandi Amazon EMR. AWS CLI

Configurazione della sostituzione non corretta dei nodi in un cluster in esecuzione

Puoi attivare o disattivare la sostituzione dei nodi non integri per un cluster in esecuzione utilizzando la console AWS CLI, l'o l'API.

Console

Per attivare o disattivare la sostituzione di nodi non integri per un cluster in esecuzione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Proprietà della pagina dei dettagli del cluster, trova la terminazione del cluster e la sostituzione dei nodi e seleziona Modifica.
4. Seleziona o deseleziona la casella di controllo Sostituzione dei nodi non integri per attivare o disattivare la funzionalità. Quindi seleziona Save changes (Salva modifiche) per confermare.

AWS CLI

Per attivare o disattivare la sostituzione di nodi non integri per un cluster in esecuzione, utilizzare il AWS CLI

- Per attivare la sostituzione di nodi non integri su un cluster in esecuzione con il AWS CLI, usa il `modify-cluster-attributes` comando con il `--unhealthy-node-replacement` parametro. Per disabilitarla, utilizza il parametro `--no-unhealthy-node-replacement`.

  L'esempio seguente attiva la sostituzione di nodi non integri nel cluster con ID: *`j-3KVTXXXXXX7UG`*

```
aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --unhealthy-node-replacement
```

L'esempio seguente disattiva la sostituzione di nodi non integri sullo stesso cluster:

```
aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-unhealthy-node-replacement
```

Lavorare con Amazon Linux AMIs in Amazon EMR

Immagini di macchine Amazon Linux Amazon (AMIs)

Amazon EMR utilizza un'Amazon Linux Amazon Machine Image (AMI) per inizializzare le istanze EC2 Amazon quando crei e avvii un cluster. L'AMI contiene il sistema operativo Amazon Linux, altro software e le configurazioni richieste per l'hosting su ogni istanza delle applicazioni del cluster.

Per impostazione predefinita, quando crei un cluster, Amazon EMR utilizza un'AMI Amazon Linux (AL) predefinita creata appositamente per la versione di rilascio di Amazon EMR che utilizzi. Per ulteriori informazioni sull'AMI Amazon Linux di default, consulta Utilizzo dell'AMI Amazon Linux predefinita per Amazon EMR. Se utilizzi Amazon EMR 5.7.0 o versioni successive, puoi scegliere di specificare un'AMI Amazon Linux personalizzata anziché l'AMI Amazon Linux predefinita per Amazon EMR. Un'AMI personalizzata consente di crittografare il volume dispositivo root e di personalizzare le applicazioni e le configurazioni come alternativa all'utilizzo delle operazioni di bootstrap. È possibile specificare un'AMI personalizzata per ogni tipo di istanza nella configurazione del gruppo di istanze o del parco istanze di un cluster Amazon EMR. Il supporto AMI personalizzato multiplo offre la flessibilità di utilizzare più tipi di architettura in un cluster. Consultare Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR.

Amazon EMR collega automaticamente un volume SSD Amazon EBS General Purpose come dispositivo root per tutti. AMIs Il supporto EBS migliora le prestazioni. AMIs Per ulteriori informazioni su Amazon Linux AMIs, consulta Amazon Machine Images (AMI). Per ulteriori informazioni sull'archiviazione dell'istanza di Amazon EMR, consulta Opzioni e comportamento di storage delle istanze in Amazon EMR.

Argomenti

- Utilizzo dell'AMI Amazon Linux predefinita per Amazon EMR

- Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR
- Modifica la versione di Amazon Linux quando crei un cluster EMR
- Personalizzazione del volume del dispositivo root Amazon EBS

Utilizzo dell'AMI Amazon Linux predefinita per Amazon EMR

Ogni versione di Amazon EMR utilizza un'AMI Amazon Linux predefinita per Amazon EMR a meno che non venga specificata un'AMI personalizzata. A partire dalle versioni di Amazon EMR 5.36, Amazon EMR 6.6 e Amazon EMR 7.0, il comportamento predefinito per l'aggiornamento di Amazon Linux 2 (AL2 per EMR 5.x e 6.x, 023 per EMR 7.x) in un'AMI Amazon AL2 EMR predefinita consiste nell'applicare automaticamente la versione più recente di Amazon Linux per l'AMI Amazon EMR predefinita.

Aggiornamenti automatici di Amazon Linux per i rilasci di Amazon EMR

Quando avvii un cluster con l'ultima versione di patch di Amazon EMR 7.0 o successivi, 6.6 o successivi o 5.36 o successivi, Amazon EMR utilizza l'ultimo rilascio di Amazon Linux per l'AMI Amazon EMR predefinita. Per esempio:

- Se è presente un rilascio `x.x.0` e un rilascio `x.x.1`, il rilascio `x.x.0` non riceve più gli aggiornamenti per l'AMI all'avvio di `x.x.1`.
- Allo stesso modo, `x.x.1` non riceve più gli aggiornamenti per l'AMI all'avvio di `x.x.2`.
- Successivamente, con i rilasci `x.y.0`, `x.x.`*`[latest]`* continua a ricevere gli aggiornamenti per l'AMI insieme a `x.y.`*`[latest]`*.

Per verificare se stai utilizzando l'ultimo rilascio della patch indicato dal numero dopo il secondo punto decimale (`6.8.`*`1`*) per un rilascio di Amazon EMR, consulta i rilasci disponibili nella Guida ai rilasci di Amazon EMR, controlla il menu a discesa dei rilasci di Amazon EMR quando crei un cluster nella console o utilizza l'API `ListReleaseLabels` o l'azione della CLI `list-release-labels`. Per ricevere aggiornamenti quando avvii un nuovo rilascio di Amazon EMR, iscriviti al feed RSS nella pagina Novità della Guida ai rilasci.

Se lo desideri, puoi scegliere di avviare il cluster con la versione di Amazon Linux fornita per la prima volta con il rilascio di Amazon EMR. Per informazioni su come specificare il rilascio di Amazon Linux per il cluster, consulta Modifica la versione di Amazon Linux quando crei un cluster EMR.

Versioni Amazon Linux predefinite

Argomenti

- Impostazione predefinita AMIs per Amazon EMR 7.0 e versioni successive
- Impostazione predefinita AMIs per Amazon EMR 6.6 e versioni successive
- Impostazione predefinita AMIs per Amazon EMR 5.x

Impostazione predefinita AMIs per Amazon EMR 7.0 e versioni successive

La tabella seguente elenca le informazioni di Amazon Linux per la versione patch più recente di Amazon EMR release 7.0 e successive.

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,82(50721,2	6.1,144-170,251.am zn2023	14 agosto 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

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• 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 • 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50623,1	6.1.141-155,222.am zn2023	21 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50609,0	6.1.140-154.222.am zn2023	14 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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-central-1 (7.1.0+)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50527,1	6,134-152.225. amzn2023	19 giugno 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50512,0	6,134-152.225. amzn2023	04 giugno 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50428,1	6,134-150.224. amzn2023	23 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50414,0	6,132-147.221.amzn2023	12 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,72(50331,0	6.1.131-143.221.am zn2023	18 aprile 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,62(50303.0	6.1.129-138.220amzn2023	27 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,6. 2025021 2	6.1.128-136.201.am zn2023	08 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20236.2(50211,0	6.1,127-135.201.am zn2023	26 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,6 20250124	6,124-134.200.amzn2023	27 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,62(50115,0	6.1.119-129.201.am zn2023	23 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,62(41121,0	6.1.115-126.197. amzn2023	12 dicembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(41031,0	6.1.112-124,190.am zn2023	15 novembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(41001,1	6,1109-118,189. amzn2023	4 ottobre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(40819,0	6.1.102-111.182.am zn2023	20 agosto 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(40730,0	6,197-104,177.amzn2023	2 agosto 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(40722,0	6,197-104,177.amzn2023	24 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,52(40708,0	6,196-102,177.amzn2023	23 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,32(40304,0	6,179-99,164. amzn2023	12 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20233,20 40219,0	6.1.77-99.164.amzn2023	1 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,32(40205.0	6,175-99,163. amzn2023	19 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20233,20 40122,0	6,172-96,166.amzn2023	5 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,32(40108,0	6,172-96,166.amzn2023	24 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2023,32(31211,4	6.1.66-91.160.amzn2023	19 dicembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

Impostazione predefinita AMIs per Amazon EMR 6.6 e versioni successive

La seguente tabella riporta informazioni su Amazon Linux per l'ultima versione della patch di Amazon EMR rilasci 6.6.x e successivi.

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2.0.2025 610,0	4,14,355-277,647.amzn2	14 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• eu-centra l-2 (6,7,0+) • ap-east-1 • ap-south-2 (6,7,0+) • ap-southeast-3 • ap-southe ast-5 (6.11.1 e 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 • ap-southe ast-7 (6.11.1 e 6.14.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-centra l-1 (6.7.0+) • ca-central-1 • ca-west-1 (6,9,0+) • us-gov-east-1 • us-gov-west-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `cn-north-1` • `cn-northwest-1` • `il-centra l-1` (6,7,0+) • `mx-central-1` (6.11.1 e 6.14.0+)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(27,1	4,14,355-277,647.amzn2	19 giugno 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-centra l-2 (6,7,0+) • ap-east-1 • ap-south-2 (6,7,0+) • ap-southeast-3 • ap-southe ast-5 (6.11.1 e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ap-southe ast-7 (6.11.1 e 6.14.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-centra l-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-centra l-1 (6,7,0+) • mx-central-1 (6.11.1 e 6.14.0+)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(12,0	4,14,355-277,643.amzn2	04 giugno 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-centra l-2 (6,7,0+) • ap-east-1 • ap-south-2 (6,7,0+) • ap-southeast-3 • ap-southe ast-5 (6.11.1 e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ap-southe ast-7 (6.11.1 e 6.14.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-centra l-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-centra l-1 (6,7,0+) • mx-central-1 (6.11.1 e 6.14.0+)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2025 428,0	4,14,355-276,639amzn2	23 maggio 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-centra l-2 (6,7,0+) • ap-east-1 • ap-south-2 (6,7,0+) • ap-southeast-3 • ap-southe ast-5 (6.11.1 e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ap-southe ast-7 (6.11.1 e 6.14.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-centra l-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-centra l-1 (6,7,0+) • mx-central-1 (6.11.1 e 6.14.0+)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(14,0	4,14,355-276,618.amzn2	12 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(21.0	4,14,355	09 aprile 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(05.0	4,14,355	18 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ e 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(20,0	4,14,355	08 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(01,0	4,14,355	28 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ e 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(23.4	4,14,355	27 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2025 116,0	4,14,355	23 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024' 17,0	4,14,355	8 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024' 01,0	4,1352	4 ottobre 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(16,0	4,14,350	21 agosto 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(09,0	4,14,349	20 agosto 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(19,0	4,14,348	25 luglio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(09,1	4,14,348	23 luglio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(23,0	4,14,336	8 marzo 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(31,0	4,14,336	14 febbraio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(24.0	4,14,336	7 febbraio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(09,0	4,1,334	24 gennaio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023' 18,0	4,14,330	2 gennaio 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023' 06,0	4,14,330	22 dicembre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023' 16,0	4,14,328	11 dicembre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023´ 01,0	4,14,327	17 novembre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023' 20,1	4,14,326	7 novembre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,02023' 12,1	4,14,326	26 ottobre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 926,0	4,14,322	19 ottobre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 8906,0	4,14,322	4 ottobre 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ 5.36.1)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(22,0	4,14,322	30 agosto 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.9+ 5.36.1)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(08,0	4,14,320	24 agosto 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.9+ 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 727,0	4,14,320	14 agosto 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.9+ 5.36.1)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 719,0	4,14,320	2 agosto 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+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ 5.36.1)

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 628,0	4,14,318	12 luglio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-south-1 • ca-central-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 612,0	4,14,314	23 giugno 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-south-1 • ca-central-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(04,1	4,14,313	16 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(18,0	4,14,311	3 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-south-1 • ca-central-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(04,1	4,14,311	18 aprile 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(04,0	4,14,311	10 aprile 2023	• `us-east-1` • `eu-west-3`

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(20,0	4,1409	30 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 307,0	4,14,305	15 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 207,0	4,14,304	3 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 119,1	4,14,31	9 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022 10,1	4,14,31	12 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022' 03,3	4,1496	5 dicembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022 04,0	4,1294	2 novembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022(12,1	4,14,291	7 ottobre 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
20,2022(05,0	4,14,287	30 agosto 2022	• us-west-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2022 719,0	4,14,287	10 agosto 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2022 426,0	4,1481	10 giugno 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2022 406,1	4,14,275	2 maggio 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

Impostazione predefinita AMIs per Amazon EMR 5.x

La seguente tabella riporta informazioni su Amazon Linux per l'ultima versione della patch di Amazon EMR 5.x rilasci 5.36 e successivi.

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2.0,2025 428,0	4,14,355-276,639amzn2	23 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(21,0	4,14,355	09 aprile 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(05.0	4,14,355	18 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ e 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(20,0	4,14,355	08 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(01,0	4,14,355	28 febbraio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2025(23.4	4,14,355	27 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2025 116,0	4,14,355	23 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024 17,0	4,14,355	8 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• ca-central-1 • il-central-1 • ca-west-1 (6.9.0+ e 5.36.1+) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(16,0	4,14,350	21 agosto 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(09,0	4,14,349	20 agosto 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(19,0	4,14,348	25 luglio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2024(09,1	4,14,348	23 luglio 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-centra l-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-southe ast-4 (6.8.1+ e 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
			• me-centra l-1 (6.10.1+) • me-south-1 • ca-central-1 • il-central-1 (6.8.1+ e 5.36.1) • ca-west-1 (6.9.1+ e 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(04,1	4,14,313	16 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(18,0	4,14,311	3 maggio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(04,1	4,14,311	18 aprile 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
20,2023(04,0	4,14,311	10 aprile 2023	• us-east-1 • eu-west-3

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2023(20,0	4,1409	30 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 307,0	4,14,305	15 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2023 207,0	4,14,304	3 marzo 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022' 10,1	4,14,31	12 gennaio 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022' 03,3	4,1496	5 dicembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022 04,0	4,1294	2 novembre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
20,2022(12,1	4,14,291	7 ottobre 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 (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2022 719,0	4,14,287	10 agosto 2022	• us-west-1 • eu-west-3 • eu-north-1 • eu-central-1 • ap-south-1 • me-south-1

OsRelea Label (versione AL)	Versione del kernel AL	Data di disponibilità	Regioni AWS
2,0,2022 426,0	4,1481	14 giugno 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

Considerazioni sull'aggiornamento del software

Prendi nota dei seguenti comportamenti di aggiornamento del software predefiniti:

Amazon EMR 7.x — Amazon Linux 2023

Le versioni 7.0 e successive di Amazon EMR funzionano su Amazon Linux 2023 (AL2023). I tuoi cluster contengono solo aggiornamenti di sicurezza disponibili nella versione dell' AL2AMI 023 che hai scelto al momento della creazione. Al momento del lancio, le istanze del cluster EMR non installeranno gli ultimi aggiornamenti di sicurezza dagli archivi di pacchetti abilitati. Per ricevere gli ultimi aggiornamenti di sicurezza, ti consigliamo di ricreare periodicamente il cluster. Per ulteriori informazioni su AL2 023, consulta Updating Amazon Linux 2023 nella Amazon Linux 2023 User Guide.

Amazon EMR 5.36+ e 6.6+ — Amazon Linux 2

Le versioni 5.36 e successive di Amazon EMR e 6.6 e successive vengono eseguite su Amazon Linux 2 (). AL2 I tuoi cluster contengono solo aggiornamenti di sicurezza disponibili nella versione dell' AL2 AMI che hai scelto al momento della creazione. Al momento del lancio, le istanze del cluster EMR non installeranno gli ultimi aggiornamenti di sicurezza dagli archivi di pacchetti abilitati. Per ricevere gli ultimi aggiornamenti di sicurezza, ti consigliamo di ricreare periodicamente il cluster. Per ulteriori informazioni AL2, consulta Gestisci il software sulla tua AL2 istanza nella Guida per l'utente di Amazon Linux 2.

Amazon EMR 3.x, 4.x, da 5.0.0 a 5.35.0 e da 6.0.0 a 6.5.0: AMI Amazon Linux bloccata nella versione di rilascio di Amazon EMR

Per le versioni di Amazon EMR 3.x, 4.x, da 5.0.0 a 5.35.0 e da 6.0.0 a 6.5.0, l'AMI predefinita si basa sulla maggior parte delle AMI Amazon Linux up-to-date disponibili al momento del rilascio di Amazon EMR. L'AMI è testata per verificarne la compatibilità con le applicazioni Big Data e le funzionalità di Amazon EMR incluse in quella versione di rilascio.

Quando un' EC2 istanza Amazon si avvia per la prima volta in un cluster basato su Amazon Linux (AL) o AL2 AMI predefiniti per Amazon EMR, verifica la presenza di aggiornamenti software che si applicano alla versione di rilascio negli archivi di pacchetti abilitati per AL e Amazon EMR. Come con altre EC2 istanze Amazon che eseguono AL or AL2 AMIs, gli aggiornamenti di sicurezza critici e importanti di questi repository vengono installati automaticamente.

Considera anche che, nella tua configurazione di rete, devi consentire l'uscita HTTP e HTTPS ai repository Amazon Linux in Amazon S3. In caso contrario, gli aggiornamenti di sicurezza avranno esito negativo. Per ulteriori informazioni, consulta Amazon Linux - Package repository nella Amazon EC2 User Guide. Per impostazione predefinita, altri pacchetti software e aggiornamenti del kernel che richiedono un riavvio, inclusi NVIDIA e CUDA, sono esclusi dal download automatico al primo avvio.

> **Important**
>
> Le versioni di Amazon EMR 3.x, 4.x, da 5.0.0 a 5.35.0 e da 6.0.0 a 6.5.0 utilizzano il comportamento predefinito di Amazon Linux e non scaricano e installano automaticamente aggiornamenti del kernel importanti e critici che richiedono un riavvio. Questo è lo stesso comportamento delle altre EC2 istanze Amazon che eseguono l'AL predefinito o AL2 AMIs. Se nuovi aggiornamenti software Amazon Linux che richiedono un riavvio (ad esempio, aggiornamenti del kernel, NVIDIA e CUDA) risultano disponibili dopo il rilascio di una versione di Amazon EMR, le istanze del cluster EMR che eseguono l'AMI predefinita non scaricano e installano automaticamente tali aggiornamenti. Per ottenere gli aggiornamenti del kernel, puoi utilizzare l'AMI Amazon Linux più recente come descritto in Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR.

Il cluster viene avviato con o senza aggiornamenti

Tieni presente che se gli aggiornamenti software non possono essere installati perché i repository dei pacchetti non sono raggiungibili al primo avvio del cluster, l'istanza del cluster continua a completare il suo avvio. Ad esempio, i repository potrebbero non essere raggiungibili perché S3 non è temporaneamente disponibile oppure è possibile che siano configurate regole VPC o firewall per bloccare l'accesso.

Non eseguire **`sudo yum update`**

Al momento della connessione a un'istanza cluster tramite SSH, nelle prime righe dell'output dello schermo vengono forniti un link alle note di rilascio per l'AMI Amazon Linux utilizzata dall'istanza, un avviso per indicare la versione AMI Amazon Linux più recente, un avviso indicante il numero di pacchetti disponibili per l'aggiornamento dai repository abilitati e una direttiva per l'esecuzione di `sudo yum update`.

> **Important**
>
> Consigliamo vivamente di non eseguire `sudo yum update` su istanze cluster, durante la connessione con SSH o mediante un'operazione di bootstrap. Questo potrebbe causare incompatibilità perché tutti i pacchetti vengono installati indistintamente.

Best practice per l'aggiornamento del software

Best practice per la gestione degli aggiornamenti software

- Se utilizzi una versione di rilascio precedente di Amazon EMR, prendi in considerazione e verifica la possibilità di effettuare una migrazione al rilascio più recente prima di aggiornare i pacchetti software.
- Se esegui la migrazione a una versione di rilascio più recente oppure aggiorni pacchetti software, verifica prima l'implementazione in un ambiente non di produzione. A tale scopo, è utile l'opzione di clonazione dei cluster con la console Amazon EMR.
- Valuta gli aggiornamenti software per le applicazioni in uso e per la versione dell'AMI Amazon Linux su base individuale. Verifica e installa i pacchetti solo negli ambienti di produzione dove sia assolutamente necessario per la sicurezza, la funzionalità o le prestazioni dell'applicazione.
- Verifica se sono disponibili aggiornamenti nel Centro di sicurezza Amazon Linux.
- Evita di installare pacchetti eseguendo il collegamento a singole istanze cluster mediante SSH. Utilizza invece un'operazione di bootstrap per installare e aggiornare i pacchetti su tutte le istanze cluster in base alle necessità. Per questa operazione è necessario chiudere e riavviare un cluster. Per ulteriori informazioni, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR.

Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR

Se utilizzi Amazon EMR 5.7.0 o versioni successive, puoi scegliere di specificare un'AMI Amazon Linux personalizzata anziché l'AMI Amazon Linux predefinita per Amazon EMR. Un'AMI personalizzata è utile se desideri effettuare le seguenti operazioni:

- Preinstallazione delle applicazioni ed esecuzione di altre personalizzazioni invece di utilizzare operazioni di bootstrap. In questo modo è possibile migliorare il tempo di avvio del cluster e semplificare il flusso di lavoro di startup. Per ulteriori informazioni e un esempio, consulta Creazione di un'AMI Amazon Linux personalizzata da un'istanza preconfigurata.
- Implementazione di configurazioni per cluster e nodo più sofisticate di quelle consentite dalle operazioni di bootstrap.
- Crittografa i volumi dei dispositivi root EBS (volumi di avvio) delle EC2 istanze nel cluster se utilizzi una versione di Amazon EMR precedente alla 5.24.0. Come per l'AMI predefinita, la dimensione minima del volume root per un'AMI personalizzata è 10 GiB per Amazon EMR rilascio

6.9 e precedenti e 15 GiB per Amazon EMR rilascio 6.10 e successivi. Per ulteriori informazioni, consulta Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato.

> **Note**
>
> A partire dalla versione 5.24.0 di Amazon EMR, puoi utilizzare un'opzione di configurazione di sicurezza per crittografare il dispositivo root e i volumi di storage EBS quando lo specifichi come provider di chiavi. AWS KMS Per ulteriori informazioni, consulta Crittografia del disco locale.

Un AMI personalizzato deve esistere nella stessa AWS regione in cui si crea il cluster. Deve inoltre corrispondere all'architettura dell' EC2 istanza. Ad esempio, un'istanza m5.xlarge ha un'architettura x86_64. Pertanto, per effettuare il provisioning di un m5.xlarge utilizzando un'AMI personalizzata, anche l'AMI personalizzata dovrebbe avere un'architettura x86_64. Allo stesso modo, per eseguire il provisioning di un'istanza m6g.xlarge, con architettura arm64, l'AMI personalizzata dovrebbe avere l'architettura arm64. Per ulteriori informazioni sull'identificazione di un'AMI Linux per il tuo tipo di istanza, consulta Find a Linux AMI nella Amazon EC2 User Guide.

> **Important**
>
> I cluster EMR che eseguono Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizzano il comportamento predefinito di Amazon Linux e non scaricano e installano automaticamente aggiornamenti del kernel importanti e critici che richiedono un riavvio. Questo è lo stesso comportamento delle altre EC2 istanze Amazon che eseguono l'AMI Amazon Linux predefinita. Se nuovi aggiornamenti software Amazon Linux che richiedono un riavvio (ad esempio, aggiornamenti del kernel, NVIDIA e CUDA) risultano disponibili dopo il rilascio di una versione di Amazon EMR, le istanze del cluster EMR che eseguono l'AMI predefinita non scaricano e installano automaticamente tali aggiornamenti. Per ottenere gli aggiornamenti del kernel, puoi personalizzare l'AMI di Amazon EMR per utilizzare l'AMI di Amazon Linux più recente.

Creazione di un'AMI Amazon Linux personalizzata da un'istanza preconfigurata

Di seguito è riportata la procedura di base per la preinstallazione di software e l'esecuzione di altre configurazioni per creare un'AMI Amazon Linux personalizzata per Amazon EMR:

- Avvia un'istanza dall'AMI Amazon Linux di base.
- Connessione all'istanza per l'installazione di software e l'esecuzione di altre personalizzazioni.
- Creazione di una nuova immagine (snapshot AMI) dell'istanza configurata.

Una volta creata l'immagine in base alla tua istanza personalizzata, puoi copiare tale immagine in una destinazione crittografata, come descritto in Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato.

Tutorial: creazione di un'AMI da un'istanza con software personalizzato installato

Per avviare un' EC2 istanza basata sull'AMI Amazon Linux più recente

1. Utilizzate il AWS CLI per eseguire il comando seguente, che crea un'istanza da un'AMI esistente. Sostituisci *MyKeyName* con la key pair che usi per connetterti all'istanza e *MyAmiId* con l'ID di un'AMI Amazon Linux appropriata. Per l'AMI più recente IDs, consulta AMI Amazon Linux.

   > **Note**
   >
   > I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

   ```
   aws ec2 run-instances --image-id MyAmiID \
   --count 1 --instance-type m5.xlarge \
   --key-name MyKeyName --region us-west-2
   ```

   Il valore di output `InstanceId` viene usato come *MyInstanceId* nella fase successiva.

2. Esegui il comando seguente:

   ```
   aws ec2 describe-instances --instance-ids MyInstanceId
   ```

   Il valore di output `PublicDnsName` viene usato per connettersi all'istanza nella fase successiva.

Per connettersi all'istanza e installare il software

1. Utilizza una connessione SSH che consente di eseguire comandi shell sull'istanza di Linux. Per ulteriori informazioni, consulta Connessione alla tua istanza Linux tramite SSH nella Amazon EC2 User Guide.
2. Esegui eventuali personalizzazioni necessarie. Ad esempio:

```
sudo yum install MySoftwarePackage
sudo pip install MySoftwarePackage
```

Per creare una snapshot dall'immagine personalizzata

- Dopo aver personalizzato l'istanza, utilizzare il comando `create-image` per creare un'AMI dall'istanza.

```
aws ec2 create-image --no-dry-run --instance-id MyInstanceId --name MyEmrCustomAmi
```

Il valore di output `imageID` viene utilizzato all'avvio del cluster o alla creazione di una snapshot crittografata. Per ulteriori informazioni, consultare Utilizzare un'AMI personalizzata singola in un cluster EMR e Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato.

Come utilizzare un'AMI personalizzata in un cluster Amazon EMR

Puoi utilizzare un'AMI personalizzata per eseguire il provisioning di un cluster Amazon EMR in due modi:

- Utilizza un'unica AMI personalizzata per tutte le EC2 istanze del cluster.
- Utilizza personalizzazioni diverse AMIs per i diversi tipi di EC2 istanze utilizzati nel cluster.

È possibile utilizzare solo una delle due opzioni durante il provisioning di un cluster EMR e non è possibile modificarlo una volta avviato il cluster.

Considerazioni sull'utilizzo di soluzioni personalizzate singole o multiple AMIs in un cluster Amazon EMR

Considerazione	AMI personalizzata singola	Molteplici personalizzazioni AMIs
Utilizza processori x86 e Graviton2 con personalizzazione AMIs nello stesso cluster	✗ Non supportato	✓ Supportato
La personalizzazione di AMI varia a seconda dei tipi di istanza	✗ Non supportato	✓ Supportato
Modifica la personalizzazione AMIs quando aggiungi una nuova istanza di attività. groups/fleets to a running cluster. Note: you cannot change the custom AMI of existing instance groups/fleets	✗ Non supportato	✓ Supportato
Usa AWS Console per avviare un cluster	✓ Supportato	✗ Non supportato
Usa AWS CloudFormation per avviare un cluster	✓ Supportato	✓ Supportato

Utilizzare un'AMI personalizzata singola in un cluster EMR

Per specificare un ID AMI personalizzato al momento della creazione di un cluster, utilizzare una delle seguenti opzioni:

- AWS Management Console
- AWS CLI
- Amazon EMR SDK
- API Amazon EMR RunJobFlow
- AWS CloudFormation (vedi la `CustomAmiID` proprietà in Cluster InstanceGroupConfig, Cluster InstanceTypeConfig InstanceGroupConfig, Resource o Resource InstanceFleetConfig -) InstanceTypeConfig

Amazon EMR console

Per specificare una singola AMI personalizzata dalla console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Name and applications (Nome e applicazioni), cerca Operating system options (Opzioni del sistema operativo). Scegli Custom AMI (AMI personalizzata) e inserisci l'ID della tua AMI nel campo Custom AMI (AMI personalizzata).
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per specificare una singola AMI personalizzata con AWS CLI

- Utilizza il parametro `--custom-ami-id` per specificare l'ID dell'AMI durante l'esecuzione del comando `aws emr create-cluster`.

  L'esempio seguente specifica un cluster che utilizza un'AMI personalizzata singola con 20 GiB di volume di avvio. Per ulteriori informazioni, consulta Personalizzazione del volume del dispositivo root Amazon EBS.

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per la leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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
```

Usa più opzioni personalizzate AMIs in un cluster Amazon EMR

Per creare un cluster utilizzando più elementi personalizzati AMIs, utilizza uno dei seguenti metodi:

- AWS CLI versione 1.20.21 o successiva
- AWS SDK
- Amazon EMR RunJobFlownella guida di riferimento all'API Amazon EMR
- AWS CloudFormation (vedi la `CustomAmiID` proprietà in Cluster InstanceGroupConfig, InstanceTypeConfigCluster, Resource o InstanceGroupConfig InstanceFleetConfigResource -) InstanceTypeConfig

La console di AWS gestione attualmente non supporta la creazione di un cluster utilizzando più elementi personalizzati AMIs.

Example - Utilizza la AWS CLI per creare un cluster di gruppi di istanze utilizzando più cluster personalizzati AMIs

Utilizzando la versione AWS CLI 1.20.21 o successiva, puoi assegnare una singola AMI personalizzata all'intero cluster oppure puoi assegnare più AMI personalizzate AMIs a ogni nodo di istanza del cluster.

L'esempio seguente mostra un cluster di gruppi di istanze uniformi creato con due tipi di istanze (m5.xlarge) utilizzate tra i tipi di nodi (primario, principale, attività). Ogni nodo ha più elementi personalizzati. AMIs L'esempio illustra diverse caratteristiche della configurazione AMI personalizzata multipla:

- Non è stata assegnata alcuna AMI personalizzata a livello di cluster. Questo serve a evitare conflitti tra più AMI personalizzate AMIs e una singola AMI personalizzata, che potrebbero causare il fallimento dell'avvio del cluster.
- Il cluster può avere più nodi di attività personalizzati AMIs tra nodi di attività primari, principali e individuali. Ciò consente personalizzazioni AMI individuali, come applicazioni preinstallate, configurazioni cluster sofisticate e volumi di dispositivi root Amazon EBS crittografati.
- Il nodo principale del gruppo di istanze può avere un solo tipo di istanza e una corrispondente AMI personalizzata. Analogamente, il nodo primario può avere un solo tipo di istanza e una corrispondente AMI personalizzata.

- Il cluster può avere più nodi attività.

```
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 - Utilizza la versione AWS CLI 1.20.21 o successiva per aggiungere un nodo di attività a un cluster di gruppi di istanze in esecuzione con più tipi di istanze e più opzioni personalizzate AMIs

Utilizzando la versione AWS CLI 1.20.21 o successiva, puoi aggiungere più elementi personalizzati AMIs a un gruppo di istanze da aggiungere a un cluster in esecuzione. L'argomento `CustomAmiId` può essere utilizzato con il comando `add-instance-groups` come mostrato nell'esempio seguente. Si noti che lo stesso ID AMI personalizzate multiple (ami-123456) viene utilizzato in più di 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 - Utilizza la versione AWS CLI 1.20.21 o successiva per creare un cluster di istanze, più tipi di istanze personalizzati AMIs, primari On-Demand, core On-Demand, core multipli e nodi task

```
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 - Utilizza la versione AWS CLI 1.20.21 o successiva per aggiungere nodi di attività a un cluster in esecuzione con più tipi di istanze e più opzioni personalizzate 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}']
```

Gestione degli aggiornamenti per i repository di pacchetti AMI

Al primo avvio, per impostazione predefinita, Amazon Linux si AMIs connette agli archivi di pacchetti per installare gli aggiornamenti di sicurezza prima dell'avvio di altri servizi. A seconda dei requisiti, puoi scegliere di disabilitare gli aggiornamenti quando specifichi un'AMI personalizzata per Amazon EMR. L'opzione di disattivazione di questa funzione è disponibile solo con le AMI personalizzate. Per impostazione predefinita, il kernel di Amazon Linux e altri pacchetti software che richiedono un riavvio non vengono aggiornati. Tieni presente che la tua configurazione di rete deve consentire l'uscita HTTP e HTTPS ai repository Amazon Linux in Amazon S3, altrimenti gli aggiornamenti della sicurezza non avranno esito positivo.

> **Warning**
>
> Specificando un'AMI personalizzata, consigliamo vivamente di scegliere di aggiornare tutti i pacchetti installati al riavvio. La scelta di non aggiornare pacchetti crea maggiori rischi per la sicurezza.

Con AWS Management Console, puoi selezionare l'opzione per disabilitare gli aggiornamenti quando scegli AMI personalizzata.

Con AWS CLI, è possibile specificare `--repo-upgrade-on-boot NONE` insieme a `--custom-ami-id` quando si utilizza il create-cluster comando.

Con l'API Amazon EMR, puoi specificare `NONE` il RepoUpgradeOnBootparametro.

Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato

Per crittografare il volume del dispositivo di root Amazon EBS di un'AMI Amazon Linux per Amazon EMR, copia un'immagine snapshot da un'AMI non crittografata a una destinazione crittografata. Per informazioni sulla creazione di volumi EBS crittografati, consulta Amazon EBS encryption nella Amazon EC2 User Guide. L'AMI di origine per lo snapshot può essere l'AMI Amazon Linux di base; in alternativa, poi copiare uno snapshot da un'AMI ottenuta da un'AMI Amazon Linux di base personalizzata.

> **Note**
>
> A partire dalla versione 5.24.0 di Amazon EMR, puoi utilizzare un'opzione di configurazione di sicurezza per crittografare il dispositivo root e i volumi di storage EBS quando lo specifichi come provider di chiavi. AWS KMS Per ulteriori informazioni, consulta Crittografia del disco locale.

Puoi utilizzare un provider di chiavi esterno o una chiave AWS KMS per crittografare il volume root EBS. Affinché Amazon EMR possa creare un cluster utilizzando l'AMI, il ruolo di servizio utilizzato da Amazon EMR (in genere il ruolo `EMR_DefaultRole` predefinito) deve essere autorizzato a crittografare e decrittare il volume. Quando si utilizza AWS KMS come fornitore di chiavi, ciò significa che devono essere consentite le seguenti azioni:

- `kms:encrypt`
- `kms:decrypt`
- `kms:ReEncrypt*`
- `kms:CreateGrant`
- `kms:GenerateDataKeyWithoutPlaintext"`
- `kms:DescribeKey"`

Il modo più semplice per farlo consiste nell'aggiungere il ruolo come utente chiave, come descritto nel seguente tutorial. L'esempio seguente di istruzione per policy viene fornito qualora fosse necessario personalizzare le policy di ruolo.

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: creazione di un'AMI personalizzata con un volume dispositivo di root crittografato tramite una chiave KMS

La prima cosa da fare in questo esempio è trovare l'ARN di una chiave KMS o crearne una nuova. Per ulteriori informazioni sulla creazione di chiavi, consulta Creazione di chiavi nella Guida per gli sviluppatori di AWS Key Management Service . La procedura riportata di seguito illustra come aggiungere il ruolo di servizio `EMR_DefaultRole` come utente chiave alla policy chiavi. Annota il valore ARN per la chiave quando la crei o la modifichi. L'ARN servirà in seguito, durante la creazione dell'AMI.

Per aggiungere il ruolo di servizio per Amazon EC2 all'elenco degli utenti con chiave di crittografia con la console

1. Accedi a AWS Management Console e apri la console AWS Key Management Service (AWS KMS) su https://console.aws.amazon.com/kms.
2. Per modificare il Regione AWS, usa il selettore della regione nell'angolo in alto a destra della pagina.
3. Scegliere l'alias della chiave KMS da utilizzare.
4. Nella pagina dei dettagli della chiave, in Key Users (Utenti di chiavi), scegli Add (Aggiungi).
5. Nella finestra di dialogo Attach (Collega), scegli il ruolo di servizio Amazon EMR. Il nome del ruolo di default è `EMR_DefaultRole`.
6. Scegli Collega.

Per creare un'AMI crittografata con AWS CLI

- Usa il `aws ec2 copy-image` comando di AWS CLI per creare un AMI con un volume del dispositivo root EBS crittografato e la chiave che hai modificato. Sostituisci il valore `--kms-key-id` specificato con l'ARN completo della chiave creata o modificata in precedenza.

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

```
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
```

L'output del comando fornisce l'ID dell'AMI creata, che è possibile specificare quando si crea un cluster. Per ulteriori informazioni, consulta Utilizzare un'AMI personalizzata singola in un cluster EMR. È anche possibile scegliere di personalizzare questa AMI installando software ed eseguendo altre

configurazioni. Per ulteriori informazioni, consulta Creazione di un'AMI Amazon Linux personalizzata da un'istanza preconfigurata.

Best practice e considerazioni

Quando crei un'AMI personalizzata per Amazon EMR, tieni in considerazione quanto segue:

- La serie Amazon EMR 7.x è basata su Amazon Linux 2023. Per queste versioni di Amazon EMR, è necessario utilizzare immagini basate su Amazon Linux 2023 per scopi personalizzati. AMIs Per trovare un'AMI personalizzata di base, consulta Ricerca di un'AMI Linux.
- Per le versioni di Amazon EMR precedenti alla 7.x, Amazon Linux 2023 AMIs non è supportato.
- Amazon EMR 5.30.0 e versioni successive e le serie Amazon EMR 6.x sono basate su Amazon Linux 2. Per queste versioni di Amazon EMR, è necessario utilizzare immagini basate su Amazon Linux 2 per scopi personalizzati. AMIs Per trovare un'AMI personalizzata di base, consulta Ricerca di un'AMI Linux.
- Per le versioni di Amazon EMR precedenti alla 5.30.0 e 6.x, Amazon Linux 2 non è supportato. AMIs
- È necessario utilizzare un'AMI Amazon Linux a 64 bit. Non è supportata un'AMI a 32 bit.
- Amazon Linux AMIs con più volumi Amazon EBS non è supportato.
- La personalizzazione deve basarsi sulla versione dell'AMI Amazon Linux più recente supportata da EBS. Per un elenco di Amazon Linux AMIs e delle AMI corrispondenti IDs, consulta Amazon Linux AMI.
- Non copiare una snapshot di un'istanza Amazon EMR esistente per creare un'AMI personalizzata, perché causa errori.
- Sono supportati solo il tipo di virtualizzazione HVM e le istanze compatibili con Amazon EMR. Assicurati di selezionare l'immagine HVM e un tipo di istanza compatibile con Amazon EMR durante il processo di personalizzazione dell'AMI. Per le istanze e i tipi di virtualizzazione compatibili, consulta Tipi di istanze supportati con Amazon EMR.
- Il tuo servizio ruolo deve disporre di autorizzazioni di avvio per l'AMI, perciò l'AMI deve essere pubblica oppure deve essere di tua proprietà o condivisa con te dal proprietario.
- Creare utenti dell'AMI con nome uguale a quello delle applicazioni causa errori (ad esempio, `hadoop`, `hdfs`, `yarn` o `spark`).
- I contenuti di `/tmp`, `/var` e `/emr` (se presenti sull'AMI) vengono spostati rispettivamente in `/mnt/tmp`, `/mnt/var` e `/mnt/emr` durante lo startup. I file vengono conservati ma, in presenza di una grande quantità di dati, lo startup potrebbe richiedere più tempo del previsto.

- Se si utilizza un'AMI Amazon Linux personalizzata basata su un'AMI Amazon Linux con una data di creazione 2018-08-11, il server Oozie non si avvia. Se si utilizza Oozie, è possibile creare un'AMI personalizzata basata su un ID AMI Amazon Linux con una data di creazione diversa. Puoi utilizzare il seguente AWS CLI comando per restituire un elenco di immagini IDs per tutti gli HVM Amazon Linux AMIs con una versione 2018.03, insieme alla data di rilascio, in modo da poter scegliere un'AMI Amazon Linux appropriata come base. MyRegion Sostituiscilo con il tuo identificativo regionale, ad esempio 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
```

- Nei casi in cui utilizzi un VPC con un nome di dominio e AmazonProvided DNS non standard, non dovresti usare l'`rotate`opzione nella configurazione DNS dei sistemi operativi.
- Se crei un'AMI personalizzata che include l'agente Amazon EC2 Systems Manager (SSM), l'agente SSM abilitato può causare un errore di provisioning sul cluster. Per evitare ciò, disabilita l'agente SSM quando utilizzi un'AMI personalizzata. Per fare ciò, quando scegli e avvii l' EC2 istanza Amazon, disabilita l'agente SSM prima di utilizzare l'istanza per creare un'AMI personalizzata e successivamente creare il tuo cluster EMR.

Per ulteriori informazioni, consulta Creazione di un'AMI Linux supportata da Amazon EBS nella Amazon EC2 User Guide.

Modifica la versione di Amazon Linux quando crei un cluster EMR

Quando avvii un cluster utilizzando Amazon EMR versione 6.6.0 o successive, viene utilizzata automaticamente la versione più recente di Amazon Linux 2 convalidata per l'AMI Amazon EMR predefinita. Puoi specificare un rilascio di Amazon Linux diverso per il cluster con la console Amazon EMR o la AWS CLI.

Amazon EMR console

Per modificare il rilascio di Amazon Linux durante la creazione di un cluster dalla console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.

3. Per Versione di EMR, scegli emr-6.6.0 o versione successiva.
4. In Opzioni del sistema operativo, scegli la versione di Amazon Linux, deseleziona la casella di controllo Applica automaticamente gli ultimi aggiornamenti di Amazon Linux e scegli la versione di Amazon Linux desiderata.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Modifica del rilascio di Amazon Linux durante la creazione di un cluster con la AWS CLI

- Utilizzo del parametro `--os-release-label` per specificare il parametro Versione Amazon Linux quando esegui il 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
```

Personalizzazione del volume del dispositivo root Amazon EBS

Puoi impostare il tipo di volume e altri attributi, a seconda del caso d'uso e dei requisiti di costo. È possibile accettare valori predefiniti o effettuare personalizzazioni.

Impostazioni predefinite del volume root EBS

Con Amazon EMR 4.x e rilasci successivi, puoi specificare la dimensione del volume root quando crei un cluster. Con Amazon EMR 6.15.0 e rilasci successivi, puoi anche specificare IOPS e velocità di trasmissione effettiva del volume root. Gli attributi si applicano solo al volume del dispositivo root Amazon EBS e si applicano a tutte le istanze del cluster. Gli attributi non si applicano ai volumi di storage, che vanno specificati separatamente per ogni tipo di istanza al momento della creazione del cluster.

- La dimensione predefinita del volume root è 15 GiB in Amazon EMR 6.10.0 e rilasci successivi. I rilasci precedenti avevano una dimensione predefinita del volume root di 10 GiB. È possibile regolare questo valore fino a 100 GiB.

- Il valore predefinito di IOPS del volume root è 3000. È possibile regolare questo valore fino a 16.000.
- La velocità effettiva predefinita del volume root è 125. MiB/s. You can adjust this up to 1000 Mib/s

> **Note**
>
> La dimensione e gli IOPS del volume root non possono avere un rapporto superiore a 1 volume per 500 IOPS (1:500), mentre gli IOPS e la velocità di trasmissione effettiva del volume root non possono avere un rapporto superiore a 1 IOPS per una velocità di trasmissione effettiva di 0,25 (1:0,25).

Per ulteriori informazioni su Amazon EBS, consulta Amazon EC2 root device volume.

Tipo di volume del dispositivo root con l'AMI predefinita

Quando utilizzi l'AMI predefinita, il tipo di volume del dispositivo root è determinato dal rilascio di Amazon EMR utilizzato.

- Con Amazon EMR rilascio 6.15.0 e successivi, Amazon EMR collega un SSD per uso generico (gp3) come tipo di volume del dispositivo root.
- Con rilasci di Amazon EMR precedenti al 6.15.0, Amazon EMR collega un SSD per uso generico (gp2) come tipo di volume del dispositivo root.

Tipo di volume del dispositivo root con AMI personalizzata

Un'AMI personalizzata può avere un altro tipo di volume del dispositivo root. Amazon EMR utilizza sempre il tipo di volume dell'AMI personalizzata.

- Con i rilasci 6.15.0 e successivi di Amazon EMR, puoi configurare la dimensione del volume root, gli IOPS e la velocità di trasmissione effettiva per le tue AMI personalizzate, a condizione che tali attributi siano applicabili al tipo di volume dell'AMI personalizzata.
- Con i rilasci di Amazon EMR precedenti al rilascio 6.15.0, puoi configurare solo la dimensione del volume root per la tua AMI personalizzata.

Se non configuri la dimensione del volume root, gli IOPS o la velocità di trasmissione effettiva quando crei il cluster, Amazon EMR utilizza i valori dell'AMI personalizzata, se applicabili. Se decidi

di configurare questi valori quando crei il cluster, Amazon EMR utilizza i valori specificati purché tali valori siano compatibili e supportati dal volume root dell'AMI personalizzata. Per ulteriori informazioni, consulta Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR.

Prezzi della dimensione del volume del dispositivo root

Il costo del volume dispositivo root EBS viene calcolato all'ora in base alle tariffe EBS mensili per il tipo di volume nella Regione in cui viene eseguito il cluster. Lo stesso vale per i volumi di storage. I costi sono in GB, ma le dimensioni del volume root vanno specificate in GiB: tienine conto nella tua stima (1 GB corrisponde a 0,931323 GiB).

Gli SSD per uso generico gp2 e gp3 vengono fatturati in modo diverso. Per stimare i costi associati ai volumi del dispositivo root EBS nel cluster, utilizza la seguente formula:

SSD per uso generico gp2

Il costo di gp2 include solo la dimensione del volume EBS in GB.

```
($EBS size in GB/month) * 0.931323 / 30 / 24 * EMR_EBSRootVolumesizeInGiB *
 InstanceCount
```

Prendiamo come esempio un cluster con un nodo primario e un nodo principale che utilizzi l'AMI Amazon Linux di base, con l'impostazione predefinita di 10 GiB per il volume del dispositivo root. Se il costo EBS nella regione è di 0,10 USD, diviso per 30 giorni, 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/month diviso per 24 ore, moltiplicato per 10 GB, moltiplicato per 2 istanze del cluster).

SSD per uso generico gp3

Il costo di gp3 include le dimensioni del volume EBS in GB, gli IOPS superiori a 3000 (3000 IOPS gratuiti) e il throughput superiore a 125 (gratuito). 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
```

Prendiamo come esempio un cluster con un nodo primario, un nodo principale e che utilizzi l'AMI Amazon Linux di base, con l'impostazione predefinita di 15 GiB per la dimensione del volume del dispositivo root, 4000 IOPS e una velocità di trasmissione effettiva di 140. Se il costo EBS nella regione è di 0,10 USD/oltre 125. GB/month, $0.005/provisioned IOPS/month over 3000, and $0.040/provisioned MB/s/month Il risultato sarà circa $ 0,009293 per istanza all'ora e $ 0,018586 all'ora per il cluster.

Definizione delle impostazioni di volume del dispositivo root personalizzate

> **Note**
>
> La dimensione e gli IOPS del volume root non possono avere un rapporto superiore a 1 volume per 500 IOPS (1:500), mentre gli IOPS e la velocità di trasmissione effettiva del volume root non possono avere un rapporto superiore a 1 IOPS per una velocità di trasmissione effettiva di 0,25 (1:0,25).

Console

Per specificare gli attributi del volume del dispositivo root Amazon EMR dalla relativa console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Seleziona Amazon EMR rilascio 6.15.0 o successivi.
4. In Configurazione del cluster, vai alla sezione Volume root EBS e inserisci un valore per qualsiasi attributo che desideri configurare.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per specificare gli attributi del volume del dispositivo root Amazon EBS con la AWS CLI

- Utilizza i parametri `--ebs-root-volume-size`, `--ebs-root-volume-iops` e `--ebs-root-volume-throughput` del comando create-cluster, come mostrato nell'esempio seguente.

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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
```

Configura le applicazioni all'avvio del cluster Amazon EMR

Quando selezioni un rilascio del software, Amazon EMR utilizza un'Amazon Machine Image (AMI) con Amazon Linux per installare il software che scegli quando avvii il cluster, ad esempio Hadoop, Spark e Hive. Amazon EMR fornisce regolarmente nuove versioni, aggiungendo nuove caratteristiche, nuove applicazioni e aggiornamenti generali. Se possibile, ti consigliamo di utilizzare la versione più recente per avviare il cluster. La versione più recente è l'opzione predefinita nel caso di avvio di un cluster dalla console.

Per ulteriori informazioni sui rilasci Amazon EMR e sulle versioni dei software disponibili con ciascun rilascio, consulta Guida ai rilasci di Amazon EMR. Per ulteriori informazioni su come modificare le configurazioni predefinite di applicazioni e software installati nel cluster, passa alla sezione relativa alla Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR. Alcune versioni dei componenti dell'ecosistema Hadoop e Spark open source incluse nei rilasci Amazon EMR dispongono di patch e miglioramenti che sono descritti nella Guida ai rilasci di Amazon EMR.

Oltre al software e alle applicazioni standard che sono disponibili per l'installazione sul cluster, puoi utilizzare operazioni di bootstrap per installare software personalizzato. Le operazioni di bootstrap sono script che vengono eseguiti sulle istanze all'avvio del cluster e su nuovi nodi che vengono aggiunti al cluster quando vengono creati. Le azioni Bootstrap sono utili anche per richiamare AWS CLI comandi su ciascun nodo per copiare oggetti da Amazon S3 a ciascun nodo del cluster.

> **Note**
>
> Le operazioni di bootstrap vengono utilizzate in maniera diversa in Amazon EMR versione 4.x e successive. Per ulteriori informazioni su queste differenze rispetto alle versioni 2.x e 3.x delle AMI Amazon EMR, consulta Differenze nella versione 4.x nella Guida ai rilasci di Amazon EMR.

Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR

Puoi utilizzare un'operazione di bootstrap per installare software aggiuntivo o personalizzare la configurazione delle istanze del cluster. Le operazioni di bootstrap sono script eseguiti sul cluster dopo che Amazon EMR ha avviato l'istanza utilizzando l'Amazon Machine Image (AMI) Amazon Linux. Le operazioni di bootstrap vengono eseguite prima che Amazon EMR installi le applicazioni specificate alla creazione del cluster e prima che i nodi del cluster inizino l'elaborazione dei dati. Aggiungendo nodi a un cluster in esecuzione, le operazioni di bootstrap vengono eseguite allo stesso modo anche su questi nodi. È possibile creare e specificare operazioni di bootstrap personalizzate al momento della creazione del cluster.

La maggior parte delle operazioni di bootstrap predefinite per le versioni 2.x e 3.x dell'AMI Amazon EMR non sono supportate nelle versioni 4.x di Amazon EMR. Ad esempio, `configure-Hadoop` e `configure-daemons` non sono supportati nel rilascio 4.x di Amazon EMR. Al contrario, il rilascio 4.x di Amazon EMR offre questa caratteristica in modo nativo. Per ulteriori informazioni su come eseguire la migrazione delle operazioni di bootstrap dalle versioni 2.x e 3.x dell'AMI Amazon EMR al rilascio 4.x di Amazon EMR, consulta Personalizzazione della configurazione di cluster e applicazioni con versioni AMI precedenti di Amazon EMR nella Guida ai rilasci di Amazon EMR.

Nozioni di base sulle operazioni di bootstrap

Per impostazione predefinita, le operazioni di bootstrap vengono eseguite come utente Hadoop. Puoi eseguire un'operazione di bootstrap con privilegi root utilizzando `sudo`.

Tutte le interfacce di gestione Amazon EMR supportano le operazioni di bootstrap. Puoi specificare fino a 16 azioni di bootstrap per cluster fornendo più `bootstrap-actions` parametri dalla console o dall'API AWS CLI.

Dalla console Amazon EMR, è possibile anche specificare un'operazione di bootstrap durante la creazione di un cluster.

Con l'interfaccia a riga di comando (CLI), puoi trasferire ad Amazon EMR i riferimenti agli script delle operazioni di bootstrap aggiungendo il parametro `--bootstrap-actions` quando crei il cluster con il comando `create-cluster`.

```
--bootstrap-actions Path="s3://amzn-s3-demo-bucket/filename",Args=[arg1,arg2]
```

Se l'operazione di bootstrap restituisce un codice errore diverso da zero, Amazon EMR lo considera un errore e termina l'istanza. Se si verificano errori nelle operazioni di bootstrap su troppe istanze, Amazon EMR termina il cluster. Se gli errori coinvolgono solo poche istanze, Amazon EMR cerca di riassegnarle e continuare. Utilizza il codice di errore `lastStateChangeReason` del cluster per identificare gli errori causati da un'operazione di bootstrap.

Esecuzione di un'operazione di bootstrap in modo condizionale

Per eseguire solo operazioni di bootstrap sul nodo principale, puoi utilizzare un'operazione di bootstrap personalizzata con una logica per determinare se il nodo è quello principale.

```
#!/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
```

Il seguente output verrà stampato da un nodo principale.

```
This is not master node, do nothing, exiting
```

Il seguente output verrà stampato da un nodo principale.

```
This is master, continuing to execute script
```

Per utilizzare questa logica, carica l'operazione di bootstrap, incluso il codice sopra, nel bucket Amazon S3. Sul AWS CLI, aggiungi il `--bootstrap-actions` parametro alla chiamata `aws emr create-cluster` API e specifica la posizione dello script di bootstrap come valore di. `Path`

Operazioni di arresto

Lo script di un'operazione di bootstrap può creare una o più operazioni di arresto scrivendo script nella directory `/mnt/var/lib/instance-controller/public/shutdown-actions/`. Alla chiusura di un cluster, tutti gli script della directory vengono eseguiti in parallelo. Ogni script deve essere eseguito e completato entro 60 secondi.

Se il nodo viene chiuso con un errore, l'esecuzione degli script per le operazioni di arresto non è garantita.

> **Note**
>
> Con le versioni 4.0 e successive di Amazon EMR, è necessario creare manualmente la directory `/mnt/var/lib/instance-controller/public/shutdown-actions/` nel nodo master. Non esiste di default ma, dopo la sua creazione, gli script nella directory vengono comunque eseguiti prima dell'arresto. Per ulteriori informazioni sulla connessione al nodo master per creare directory, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

Utilizzo di operazioni di bootstrap personalizzate

Puoi creare uno script personalizzato per eseguire un'operazione di bootstrap personalizzata. Alle operazioni di bootstrap personalizzate può fare riferimento qualsiasi interfaccia Amazon EMR.

> **Note**
>
> Per prestazioni ottimali, ti consigliamo di archiviare azioni di bootstrap personalizzate, script e altri file che desideri utilizzare con Amazon EMR in un bucket Amazon S3 che si trova nello stesso cluster. Regione AWS

Indice

- Aggiunta di operazioni di bootstrap personalizzate

- Utilizzo di un'operazione di bootstrap personalizzata per la copia di un oggetto da Amazon S3 su ogni nodo

Aggiunta di operazioni di bootstrap personalizzate

Console

Per creare un cluster con un'azione di bootstrap con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Bootstrap actions (Operazioni di bootstrap), scegli Add (Aggiungi) per specificare un nome, la posizione dello script e gli argomenti opzionali per l'operazione. Seleziona Add bootstrap action (Aggiungi operazione di bootstrap).
4. Se lo desideri, puoi aggiungere altre operazioni di bootstrap.
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per creare un cluster con un'azione di bootstrap personalizzata con AWS CLI

Quando si utilizza l'azione AWS CLI per includere un'azione bootstrap, specificare `Path` and `Args` come elenco separato da virgole. L'esempio seguente non utilizza un elenco di argomenti.

- Per avviare un cluster con un'azione bootstrap personalizzata, digita il seguente comando, sostituendolo *myKey* con il nome della tua EC2 key pair. Includi `--bootstrap-actions` come parametro e specifica la posizione dello script bootstrap come valore di `Path`.
  - Utenti Linux, UNIX e 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"
```

- Utenti 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"
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

> **Note**
>
> Se in precedenza non hai creato il ruolo e il profilo di EC2 istanza del servizio Amazon EMR predefiniti, digita `aws emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Utilizzo di un'operazione di bootstrap personalizzata per la copia di un oggetto da Amazon S3 su ogni nodo

Puoi utilizzare un'operazione di bootstrap per copiare oggetti da Amazon S3 su ogni nodo di un cluster prima di installare le tue applicazioni. AWS CLI Viene installato su ogni nodo di un cluster, in modo che l'azione di bootstrap possa richiamare AWS CLI comandi.

L'esempio seguente mostra il semplice script di un'operazione di bootstrap che copia un file, `myfile.jar`, da Amazon S3 in una cartella locale, `/mnt1/myfolder`, su ogni nodo del cluster. Lo script viene salvato su Amazon S3 con il nome di file `copymyfile.sh` e con i seguenti contenuti.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/myfilefolder/myfile.jar /mnt1/myfolder
```

All'avvio del cluster, è necessario specificare lo script. L' AWS CLI esempio seguente lo dimostra:

```
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"
```

Configurazione dell'hardware e della rete del cluster Amazon EMR

Una considerazione importante quando si crea un cluster Amazon EMR è il modo in cui si configurano EC2 le istanze Amazon e le opzioni di rete. In questo capitolo vengono descritte le seguenti opzioni e vengono illustrate le best practice e linee guida per tutte queste opzioni.

- Tipi di nodi: EC2 le istanze Amazon in un cluster EMR sono organizzate in tipi di nodi. Esistono tre tipi di nodi: nodi primari, nodi core e nodi attività. Ogni tipo di nodo esegue un set di ruoli definiti dalle applicazioni distribuite installate sul cluster. Durante un job Hadoop MapReduce o Spark, ad esempio, i componenti sui nodi core e task elaborano i dati, trasferiscono l'output su Amazon S3 o HDFS e forniscono i metadati di stato al nodo primario. Con un cluster a nodo singolo, tutti i componenti vengono eseguiti sul nodo primario. Per ulteriori informazioni, consulta Comprendi i tipi di nodi in Amazon EMR: nodi primari, core e task node.
- EC2 istanze: quando crei un cluster, fai delle scelte sulle EC2 istanze Amazon su cui verrà eseguito ogni tipo di nodo. Il tipo di EC2 istanza determina il profilo di elaborazione e archiviazione del nodo. La scelta dell' EC2 istanza Amazon per i tuoi nodi è importante perché determina il profilo prestazionale dei singoli tipi di nodi nel cluster. Per ulteriori informazioni, consulta Configurazione dei tipi di EC2 istanze Amazon da utilizzare con Amazon EMR.
- Reti: è possibile avviare il cluster Amazon EMR in un VPC utilizzando una sottorete pubblica, una sottorete privata o una sottorete condivisa. La configurazione di rete determina il modo in cui i clienti e i servizi possono connettersi ai cluster per eseguire il lavoro, il modo in cui i cluster si connettono agli archivi dati e ad altre risorse AWS e le opzioni disponibili per controllare il traffico su tali connessioni. Per ulteriori informazioni, consulta Configurazione della rete in un VPC per Amazon EMR.
- Raggruppamento di istanze: la raccolta di EC2 istanze che ospitano ciascun tipo di nodo viene denominata flotta di istanze o gruppo di istanze uniforme. La scelta se configurare o meno i gruppi di istanze viene fatta quando si crea un cluster, Questa scelta determina il modo in cui è possibile aggiungere nodi al cluster mentre è in esecuzione. La configurazione si applica a tutti i tipi di nodo. In seguito non può più essere modificata. Per ulteriori informazioni, consulta Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi.

> **Note**
>
> La configurazione dei parchi istanze è disponibile solo in Amazon EMR rilasci 4.8.0 e successivi, esclusi i rilasci 5.0.0 e 5.0.3.

Comprendi i tipi di nodi in Amazon EMR: nodi primari, core e task node

Utilizza questa sezione per scoprire il modo in cui Amazon EMR utilizza ognuno di questi tipi di nodo e come base per la pianificazione della capacità del cluster.

Nodo primario

Il nodo primario gestisce il cluster ed esegue in genere i componenti primari delle applicazioni distribuite. Ad esempio, il nodo primario esegue il ResourceManager servizio YARN per gestire le risorse per le applicazioni. Esegue inoltre il NameNode servizio HDFS, tiene traccia dello stato dei lavori inviati al cluster e monitora lo stato dei gruppi di istanze.

Per monitorare lo stato di avanzamento di un cluster e interagire direttamente con le applicazioni, puoi connetterti al nodo primario su SSH come utente Hadoop. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH. La connessione al nodo primario consente di accedere direttamente a directory e file, ad esempio i file di log Hadoop. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR. Puoi anche visualizzare le interfacce utente pubblicate dalle applicazioni come siti Web in esecuzione sul nodo primario. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

> **Note**
>
> Con Amazon EMR 5.23.0 e versioni successive, puoi avviare un cluster con tre nodi primari per supportare l'elevata disponibilità di applicazioni come YARN Resource Manager, HDFS, Spark, Hive e NameNode Ganglia. Con questa caratteristica, il nodo primario non rappresenta più un potenziale singolo punto di errore. Se uno dei nodi primari ha esito negativo, Amazon EMR esegue automaticamente il failover in un nodo primario in standby e sostituisce il nodo primario guasto con uno nuovo con le medesime operazioni di configurazione e di bootstrap. Per ulteriori informazioni, consulta la sezione Plan and Configure Primary Nodes (Pianificazione e configurazione dei nodi primari).

Nodi principali

I nodi core sono gestiti dal nodo primario. I nodi principali eseguono il daemon Data Node per coordinare lo storage dei dati come parte di Hadoop Distributed File System (HDFS). Inoltre, eseguono il daemon Task Tracker e altre attività di calcolo parallelo sui dati richieste dalle applicazioni installate. Ad esempio, un nodo principale esegue daemon YARNNodeManager , task Hadoop ed esecutori Spark. MapReduce

Esiste un solo gruppo di istanze principale o una flotta di istanze per cluster, ma possono esserci più nodi in esecuzione su più istanze Amazon nel gruppo di EC2 istanze o nel parco di istanze. Con i gruppi di istanze, puoi aggiungere e rimuovere EC2 istanze Amazon mentre il cluster è in esecuzione. È inoltre possibile impostare la scalabilità automatica per aggiungere istanze in base al valore di un parametro. Per ulteriori informazioni sull'aggiunta e la rimozione di EC2 istanze Amazon con la configurazione dei gruppi di istanze, consultaUsa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione.

Con i parchi istanze, puoi aggiungere e rimuovere agevolmente istanze modificando le capacità target del parco istanze su on demand e Spot di conseguenza. Per ulteriori informazioni sulle capacità target, consulta Opzioni del parco istanze.

> **Warning**
>
> La rimozione dei daemon HDFS da un nodo principale in esecuzione o la terminazione di nodi principali comporta il rischio di perdita dei dati. Fai attenzione quando configuri i nodi principali per l'utilizzo delle istanze Spot. Per ulteriori informazioni, consulta Quando occorre utilizzare le istanze Spot?.

Nodi attività

È possibile utilizzare i nodi task per aggiungere potenza per eseguire attività di calcolo parallele sui dati, come le attività Hadoop e gli esecutori MapReduce Spark. I nodi di task non eseguono il daemon Data Node, né archiviano dati in HDFS. Come per i nodi principali, puoi aggiungere nodi di attività a un cluster aggiungendo istanze Amazon a un gruppo di EC2 istanze uniforme esistente o modificando le capacità target per un parco di istanze di attività.

Con la configurazione del gruppo di istanze uniforme puoi avere un totale di 48 gruppi di istanze attività. La possibilità di aggiungere gruppi di istanze in questo modo ti consente di combinare tipi di

EC2 istanze Amazon e opzioni di prezzo, come istanze On-Demand e istanze Spot. Questo consente di rispondere ai requisiti di carico di lavoro in modo conveniente.

Con la configurazione del parco istanze, la possibilità di combinare tipi di istanze e opzioni di acquisto è integrata, perciò esiste un solo parco istanze attività.

Poiché le istanze Spot vengono spesso utilizzate per eseguire nodi attività, Amazon EMR dispone delle caratteristiche predefinite per la pianificazione dei processi YARN in modo che i processi in esecuzione non abbiano esito negativo quando i nodi attività in esecuzione su istanze Spot vengono terminati. Amazon EMR esegue questa operazione consentendo ai processi master delle applicazioni di funzionare solo sui nodi principali. Il processo master dell'applicazione controlla i processi in esecuzione e deve rimanere attivo per tutta la durata del processo.

Amazon EMR rilascio 5.19.0 e successivi utilizzano la caratteristica integrata etichette nodo YARN per questo scopo. (Le versioni precedenti utilizzavano una patch di codice). Le proprietà nelle classificazioni di configurazione `yarn-site` e `capacity-scheduler` sono configurate per impostazione predefinita in modo che capacity-scheduler e fair-scheduler YARN sfruttino le etichette dei nodi. Amazon EMR etichetta in automatico i nodi principali con l'etichetta `CORE` e imposta le proprietà in modo che i master dell'applicazione siano pianificati solo sui nodi con l'etichetta CORE. La modifica manuale delle proprietà correlate nelle classificazioni di configurazione del sito di YARN e del pianificatore di capacità o direttamente nei file XML associati potrebbe interrompere o alterare questa funzionalità.

A partire dalla serie di rilascio Amazon EMR 6.x, la funzione etichette nodo YARN è disabilitata per impostazione predefinita. Per impostazione predefinita, i processi primari dell'applicazione possono essere eseguiti sia sui nodi core sia su quelli attività. È possibile abilitare la caratteristica etichette nodo YARN configurando le seguenti proprietà:

- `yarn.node-labels.enabled: true`
- `yarn.node-labels.am.default-node-label-expression: 'CORE'`

A partire dalla serie di release di Amazon EMR 7.x, Amazon EMR assegna le etichette dei nodi YARN alle istanze in base al tipo di mercato, ad esempio On-Demand o Spot. Puoi abilitare le etichette dei nodi e limitare i processi applicativi a ON_DEMAND configurando le seguenti proprietà:

```
yarn.node-labels.enabled: true
yarn.node-labels.am.default-node-label-expression: 'ON_DEMAND'
```

Se utilizzi Amazon EMR 7.0 o versioni successive, puoi limitare il processo di applicazione ai nodi con l'`CORE`etichetta utilizzando la seguente configurazione:

```
yarn.node-labels.enabled: true
yarn.node-labels.am.default-node-label-expression: 'CORE'
```

Per le versioni 7.2 e successive di Amazon EMR, se il cluster utilizza la scalabilità gestita con etichette dei nodi, Amazon EMR cercherà di scalare il cluster in base al processo applicativo e alla domanda dell'esecutore in modo indipendente.

Ad esempio, se utilizzi le release 7.2 o successive di Amazon EMR e limiti il processo applicativo ai `ON_DEMAND` nodi, la scalabilità gestita aumenta la scalabilità dei `ON_DEMAND` nodi se la domanda del processo applicativo aumenta. Allo stesso modo, se si limita il processo di applicazione ai `CORE` nodi, la scalabilità gestita aumenta la scalabilità dei nodi se la domanda del processo applicativo `CORE` aumenta.

Per informazioni su proprietà specifiche, consulta Impostazioni di Amazon EMR per impedire gli errori nei processi a causa dell'interruzione delle istanze Spot nei nodi attività.

Configurazione dei tipi di EC2 istanze Amazon da utilizzare con Amazon EMR

EC2 le istanze sono disponibili in diverse configurazioni note come tipi di istanza. I tipi di istanza dispongono di diverse capacità di CPU, input/output e archiviazione. Oltre al tipo di istanza, puoi scegliere diverse opzioni di acquisto per le EC2 istanze Amazon. Puoi specificare diversi tipi di istanze e opzioni di acquisto all'interno di gruppi di istanze o parchi istanze uniformi. Per ulteriori informazioni, consulta Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi. Per indicazioni sulla scelta di tipi di istanza e opzioni di acquisto per l'applicazione, consulta Configurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot.

> **Important**
>
> Quando si sceglie un tipo di istanza utilizzando il AWS Management Console, il numero di vCPU mostrato per ogni tipo di istanza è il numero di vcore YARN per quel tipo di istanza, non il numero di EC2 v CPUs per quel tipo di istanza. Per ulteriori informazioni sul numero di v CPUs per ogni tipo di istanza, consulta Amazon EC2 Instance Types.

Argomenti

- Tipi di istanze supportati con Amazon EMR
- Configurazione della rete in un VPC per Amazon EMR
- Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi

Tipi di istanze supportati con Amazon EMR

Questa sezione descrive i tipi di istanza supportati da Amazon EMR, organizzate per Regione AWS. Per ulteriori informazioni sui tipi di istanze, consulta la matrice dei tipi di EC2 istanze Amazon AMI e Amazon Linux AMI.

Non tutti i tipi di istanza sono disponibili in tutte le Regioni e la disponibilità dell'istanza è soggetta alla disponibilità e alla domanda nella Regione e nella zona di disponibilità specificate. La zona di disponibilità di un'istanza è determinata dalla sottorete utilizzata per avviare il cluster.

Considerazioni

Considera quanto segue quando scegli i tipi di istanza per il tuo cluster Amazon EMR.

> **Important**
>
> Quando si sceglie un tipo di istanza utilizzando il AWS Management Console, il numero di vCPU mostrato per ogni tipo di istanza è il numero di vcore YARN per quel tipo di istanza, non il numero di EC2 v CPUs per quel tipo di istanza. Per ulteriori informazioni sul numero di v CPUs per ogni tipo di istanza, consulta Amazon EC2 Instance Types.

- Se crei un cluster utilizzando un tipo di istanza che non è disponibile nella regione specificata in una zona di disponibilità, è possibile che il cluster non riesca a effettuare il provisioning o che il provisioning si blocchi. Per informazioni sulla disponibilità delle istanze, consulta Pagina dei prezzi di Amazon EMR o consulta le tabelle Tipi di istanze supportati da Regione AWS in questa pagina.
- A partire dal rilascio di Amazon EMR versione 5.13.0, tutte le istanze utilizzano la virtualizzazione HVM e l'archiviazione supportata da EBS per i volumi root. Quando si utilizzano versioni del rilascio di Amazon EMR precedenti a 5.13.0, alcune istanze di generazione precedenti utilizzano la virtualizzazione PVM. Per ulteriori informazioni, consulta Tipi di virtualizzazione delle AMI Linux.
- A causa della mancanza di supporto hardware e di impostazioni predefinite che possono portare a un sottoutilizzo della memoria e dei core, non è consigliabile utilizzare i tipi di istanza`c7a`,,,,,, `c7i` `m7i` `m7i-flex` `r7a` `r7i` `r7izi4i.12xlarge`, `i4i.24xlarge` se esegui versioni di Amazon

EMR precedenti alla 5.36.1 e 6.10.0. Se esegui questi tipi di istanze in quelle versioni, potresti riscontrare prestazioni inferiori e non vedrai i vantaggi attesi dai tipi di istanze più recenti, come vs. `c7i` `c6i` Per un utilizzo ottimale delle risorse e prestazioni con questi tipi di prestazioni, è consigliabile eseguire la versione 5.36.1 e versioni successive o la versione 6.10.0 e versioni successive per massimizzarne le capacità.

- Alcuni tipi di istanze supportano caratteristiche di rete avanzate. Per ulteriori informazioni, consulta Reti avanzate su Linux.
- I driver NVIDIA e CUDA sono installati su tipi di istanze GPU per impostazione predefinita.

Tipi di istanze supportati da Regione AWS

Le tabelle seguenti elencano i tipi di EC2 istanze Amazon supportati da Amazon EMR, organizzati per. Regione AWS Le tabelle elencano anche i primi rilasci di Amazon EMR nelle serie 5.x, 6.x e 7.x che supportano ciascun tipo di istanza.

Stati Uniti orientali (Virginia settentrionale), us-east-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Stati Uniti orientali (Ohio): us-east-2

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Stati Uniti occidentali (California settentrionale) - us-west-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Stati Uniti occidentali (Oregon): us-west-2

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Stati Uniti occidentali) - -1 us-gov-west

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Stati Uniti orientali) - -1 us-gov-east

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Africa (Città del Capo) - af-south-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Hong Kong) - ap-east-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Giacarta) - ap-southeast-3

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Melbourne) - ap-southeast-4

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Malesia) - ap-southeast-5

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Mumbai) - ap-south-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Hyderabad) (ap-south-2)

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Osaka) - ap-northeast-3

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Seoul) - ap-northeast-2

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Singapore) - ap-southeast-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Sydney): ap-southeast-2

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 Pacifico (Tokyo), ap-northeast-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Canada (Centrale) - ca-central-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Canada occidentale (Calgary) - ca-west-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la memoria	r5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Cina (Ningxia) - cn-nordovest-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Cina (Pechino) - cn-north-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Francoforte) - eu-central-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Zurigo) (eu-central-2)

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Londra) - eu-west-2

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Milano) (eu-south-1)

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Europa (Spagna) (eu-south-2)

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Parigi) - eu-west-3

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Stoccolma) - eu-north-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Israele (Tel Aviv) - il-central-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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 (Bahrein) - me-south-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Regione Medio Oriente (EAU) (me-central-1)

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Sud America (San Paolo) - sa-east-1

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Uso generale	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di Amazon EMR (5.x, 6.x, 7.x)
Ottimizzazione per il calcolo	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Calcolo accelerato	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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Ottimizzazione per la 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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
Storage ottimizzato	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Classe di istanza	Tipo di istanza	Versione minima supportata di 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

Istanze di generazioni precedenti

Amazon EMR supporta le istanze di generazioni precedenti per supportare le applicazioni ottimizzate per tali istanze e non ancora aggiornate. Per ulteriori informazioni su questi tipi di istanze e percorsi di aggiornamento, consulta Istanze di generazioni precedenti.

Classe di istanza	Tipi di istanza
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] Utilizza l'AMI di virtualizzazione PVM con Amazon EMR in versione precedente alla 5.13.0. Per ulteriori informazioni, consulta Tipi di virtualizzazione delle AMI Linux.

[2] Non supportata nella versione di rilascio 5.15.0.

Opzioni di acquisto di istanze in Amazon EMR

Quando configuri un cluster, scegli un'opzione di acquisto per EC2 le istanze Amazon. Puoi scegliere istanze on demand, istanze Spot o entrambe. I prezzi variano in base al tipo di istanza e alla Regione. Il prezzo di Amazon EMR si aggiunge al prezzo di Amazon (il EC2 prezzo per i server sottostanti) e al prezzo di Amazon EBS (se si allegano volumi Amazon EBS). Per i prezzi correnti, consulta Prezzi di Amazon EMR.

La scelta di utilizzare gruppi di istanze o parchi istanze nel cluster determina la modalità di modifica delle opzioni di acquisto dell'istanza mentre un cluster è in esecuzione. Se scegli gruppi di istanze

uniformi, puoi specificare l'opzione di acquisto per un gruppo di istanze solo al momento della creazione e il tipo di istanza e l'opzione di acquisto si applicano a tutte le EC2 istanze Amazon in ogni gruppo di istanze. Se scegli parchi istanze, puoi modificare le opzioni di acquisto dopo aver creato il parco istanze e combinare le opzioni di acquisto per raggiungere la capacità target specificata. Per ulteriori informazioni su queste configurazioni, consulta Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi.

Istanze on demand

Con istanze On-Demand, paghi per capacità di elaborazione a ore. In alternativa, puoi configurare le istanze on demand in modo da utilizzare le opzioni di acquisto Istanza riservata o Istanza dedicata. Con le istanze riservate, puoi effettuare un pagamento una tantum per un'istanza per prenotare la capacità. Le istanze dedicate sono fisicamente isolate a livello di hardware host dalle istanze che appartengono ad altri account. AWS Per ulteriori informazioni sulle opzioni di acquisto, consulta Instance Purchasing Options nella Amazon EC2 User Guide.

Uso di istanze riservate

Per utilizzare le istanze riservate in Amazon EMR, usi EC2 Amazon per acquistare l'istanza riservata e specificare i parametri della prenotazione, incluso l'ambito della prenotazione applicabile a una regione o a una zona di disponibilità. Per ulteriori informazioni, consulta Amazon EC2 Reserved Instances e Buying Reserved Instances nella Amazon EC2 User Guide. Quando acquisti un'istanza riservata, se tutte le condizioni seguenti sono vere, Amazon EMR utilizza l'istanza riservata all'avvio di un cluster:

- Un'istanza on demand viene specificata nella configurazione cluster che corrisponde alla specifica dell'istanza riservata.
- Il cluster viene avviato nell'ambito della prenotazione dell'istanza (la zona di disponibilità o Regione).
- La capacità dell'istanza riservata è ancora disponibile

Ad esempio, supponiamo di acquistare un'istanza riservata `m5.xlarge` con la prenotazione dell'istanza calibrata sulla Regione Stati Uniti orientali. Viene quindi avviato un cluster Amazon EMR negli Stati Uniti orientali che utilizza due istanze `m5.xlarge`. La prima istanza viene fatturata alla tariffa dell'istanza riservata e l'altra viene fatturata alla tariffa on demand. La capacità dell'istanza riservata viene utilizzata prima della creazione di qualsiasi istanza on demand.

Uso di istanze dedicate

Per utilizzare le istanze dedicate, è necessario acquistare istanze dedicate utilizzando Amazon EC2 e quindi creare un VPC con l'attributo di tenancy dedicato. All'interno di Amazon EMR, specifica quindi che un cluster deve essere avviato in questo VPC. Qualsiasi istanza on demand nel cluster che corrisponde alle specifiche dell'istanza dedicata utilizza istanze dedicate disponibili all'avvio del cluster.

> **Note**
>
> Amazon EMR non supporta l'impostazione dell'attributo `dedicated` su singole istanze.

Spot Instances

Le istanze Spot in Amazon EMR offrono la possibilità di acquistare la capacità di istanze EC2 Amazon a un costo ridotto rispetto all'acquisto on demand. Lo svantaggio dell'utilizzo di istanze Spot è che le istanze possono terminare se la capacità Spot diventa indisponibile per il tipo di istanza in esecuzione. Per ulteriori informazioni su quando l'utilizzo di istanze Spot può risultate appropriato per l'applicazione, consulta Quando occorre utilizzare le istanze Spot?.

Quando Amazon EC2 dispone di capacità inutilizzata, offre EC2 istanze a un costo ridotto, denominato prezzo Spot. Questo prezzo varia in base alla disponibilità e alla domanda e viene stabilito in base alla Regione e alla zona di disponibilità. Quando scegli le istanze Spot, specifichi il prezzo Spot massimo che sei disposto a pagare per ogni EC2 tipo di istanza. Quando il prezzo Spot nella zona di disponibilità del cluster è inferiore al prezzo Spot massimo specificato per tale tipo di istanza, le istanze vengono avviate. Mentre le istanze sono in esecuzione, ti viene addebitato il prezzo Spot corrente non il prezzo Spot massimo.

> **Note**
>
> Le istanze spot con una durata definita (note anche come blocchi Spot) non sono più disponibili per i nuovi clienti a partire dal 1° luglio 2021. Per i clienti che hanno già utilizzato la funzione, continueremo a supportare le istanze spot con una durata definita fino al 31 dicembre 2022.

Per i prezzi correnti, consulta i prezzi di Amazon EC2 Spot Instances. Per ulteriori informazioni, consulta le istanze Spot nella Amazon EC2 User Guide. Quando crei e configuri un cluster, devi

specificare le opzioni di rete che determinano in ultima analisi la zona di disponibilità in cui il cluster viene avviato. Per ulteriori informazioni, consulta Configurazione della rete in un VPC per Amazon EMR.

> **Tip**
>
> Puoi visualizzare il prezzo Spot in tempo reale nella console passando il mouse sul suggerimento informazioni accanto all'opzione di acquisto Spot quando crei un cluster utilizzando Advanced Options (Opzioni avanzate). Vengono visualizzati i prezzi per ogni zona di disponibilità nella Regione selezionata. I prezzi più bassi si trovano nelle righe di colore verde. A causa della fluttuazione dei prezzi Spot tra zone di disponibilità, è possibile che selezionando la zona di disponibilità con il prezzo iniziale più basso tale prezzo non venga mantenuto per l'intera durata del cluster. Per risultati ottimali, occorre studiare la cronologia dei prezzi della zona di disponibilità prima di scegliere. Per ulteriori informazioni, consulta la Cronologia dei prezzi delle istanze Spot nella Amazon EC2 User Guide.

Le opzioni Istanza Spot variano a seconda che nella configurazione del cluster si utilizzano gruppi di istanze o parchi istanze uniformi.

Istanze Spot in gruppi di istanze uniformi

Quando utilizzi istanze Spot in un gruppo di istanze uniformi, tutte le istanze in un gruppo di istanze devono essere istanze Spot. Specifica una sottorete o zona di disponibilità singola per il cluster. Per ogni gruppo di istanze, specifica un'istanza Spot singola e un prezzo Spot massimo. Le istanze Spot di questo tipo vengono avviate se il prezzo Spot nella Regione e nella zona di disponibilità del cluster è inferiore al prezzo Spot massimo. Le istanze terminano se il prezzo Spot è superiore al prezzo Spot massimo. Puoi importare il prezzo Spot massimo solo quando configuri un gruppo di istanze. In seguito non può più essere modificata. Per ulteriori informazioni, consulta Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi.

Istanze Spot in parchi istanze

Quando utilizzi la configurazione dei parchi istanze, opzioni aggiuntive offrono un maggiore controllo sulla modalità di avvio e di terminazione delle istanze Spot. Sostanzialmente, i parchi istanze utilizzano un metodo diverso rispetto a gruppi di istanze uniformi per avviare le istanze. In pratica, definisci una capacità target per istanze Spot (e istanze on demand) e un massimo di cinque tipi di istanze. Puoi anche specificare una capacità ponderata per ogni tipo di istanza o utilizzare il vCPU (vcore YARN) del tipo di istanza come capacità ponderata. Questa capacità ponderata viene

conteggiata per il raggiungimento della capacità target quando si esegue il provisioning di un'istanza di tale tipo. Amazon EMR esegue il provisioning di istanze con entrambe le opzioni di acquisto finché non viene raggiunta la capacità target per ogni target. Inoltre, puoi definire un intervallo di zone di disponibilità per Amazon EMR da cui scegliere quando si avviano istanze. Puoi anche fornire opzioni Spot aggiuntive per ciascun parco istanze, incluso un timeout di provisioning. Per ulteriori informazioni, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR.

Opzioni e comportamento di storage delle istanze in Amazon EMR

Panoramica

Archivio istanza e archiviazione di volumi Amazon EBS vengono utilizzati per dati HDFS, nonché buffer, cache, dati grezzi e altri contenuti temporanei che alcune applicazioni possono "riversare" nel file system locale.

Amazon EBS funziona in modo diverso all'interno di Amazon EMR rispetto alle normali istanze Amazon EC2 . I volumi Amazon EBS collegati a cluster Amazon EMR sono temporanei: i volumi vengono eliminati al termine del cluster e dell'istanza (ad esempio, durante la riduzione di gruppi di istanze), pertanto non bisogna aspettarsi che i dati siano persistenti. Anche se sono temporanei, i dati in HDFS possono essere replicati a seconda del numero e della specializzazione dei nodi nel cluster. Quando si aggiungono volumi di stogare Amazon EBS, questi vengono montati come volumi aggiuntivi. Non fanno parte del volume di avvio. YARN è configurato per utilizzare tutti i volumi aggiuntivi, ma l'utente è responsabile dell'allocazione di volumi aggiuntivi come storage locale (ad esempio, per file di log locali).

Considerazioni

Tieni conto di queste considerazioni aggiuntive quando utilizzi Amazon EBS con cluster EMR:

- Non puoi eseguire lo snapshot di un volume Amazon EBS per poi ripristinarlo all'interno di Amazon EMR. Per creare configurazioni personalizzate riutilizzabili, utilizza un'AMI personalizzata (disponibile in Amazon EMR versione 5.7.0 e successive). Per ulteriori informazioni, consulta Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR.
- Un volume dispositivo root Amazon EBS crittografato è supportato solo utilizzando un'AMI personalizzata. Per ulteriori informazioni, consulta Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato.
- Se applichi tag utilizzando l'API di Amazon EMR, tali operazioni vengono applicate a volumi EBS.

- Esiste un limite di 25 volumi per istanza.
- I volumi Amazon EBS sui nodi principali non possono essere inferiori a 5 GB.
- Amazon EBS ha un limite fisso di 2.500 volumi EBS per richiesta di avvio dell'istanza. Questo limite si applica anche ad Amazon EMR su EC2 cluster. Ti consigliamo di avviare cluster con il numero totale di volumi EBS entro questo limite e quindi scalare manualmente il cluster o con la scalabilità gestita di Amazon EMR, se necessario. Per ulteriori informazioni sul limite di volume EBS, consulta Service quotas.

Archiviazione Amazon EBS di default per istanze

Per EC2 le istanze con storage solo EBS, Amazon EMR alloca i volumi di storage Amazon EBS gp2 o gp3 alle istanze. Quando crei un cluster utilizzando Amazon EMR rilascio 5.22.0 e successivi , la quantità predefinita di spazio di archiviazione di Amazon EBS aumenta in base alle dimensioni dell'istanza.

Suddividiamo l'eventuale spazio di archiviazione aggiuntivo su più volumi. Ciò offre migliori prestazioni IOPS e, di conseguenza, migliori prestazioni per alcuni carichi di lavoro standardizzati. Se desideri utilizzare una diversa configurazione dell'archiviazione delle istanze di Amazon EBS, puoi specificarla al momento della creazione di un cluster EMR o quando aggiungi nodi a un cluster esistente. Puoi utilizzare solo volumi gp2 o gp3 di Amazon EBS come volumi root e aggiungere volumi gp2 o gp3 come volumi supplementari. Per ulteriori informazioni, consulta Specifica di volumi di archiviazione EBS aggiuntivi.

La tabella seguente identifica il numero predefinito di volumi di archiviazione gp2 di Amazon EBS, le dimensioni e le dimensioni totali per tipo di istanza. Per informazioni sui volumi gp2 rispetto ai volumi gp3, consulta Confronto tra i tipi di volume gp2 e gp3 di Amazon EBS.

Volumi di archiviazione gp2 di Amazon EBS predefiniti e dimensioni per tipo di istanza per Amazon EMR 5.22.0 e successivi

Dimensioni istanza	Numero di volumi	Dimensioni del volume (GiB)	Dimensione totale (GiB)
*.large	1	32	32
*.xlarge	2	32	64
*.2xlarge	4	32	128

Dimensioni istanza	Numero di volumi	Dimensioni del volume (GiB)	Dimensione totale (GiB)
*.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

Volume root di Amazon EBS predefinito per le istanze

Con le release 6.15 e successive di Amazon EMR, Amazon EMR collega automaticamente un SSD Amazon EBS General Purpose (gp3) come dispositivo root per migliorare le prestazioni. AMIs Con i rilasci precedenti, Amazon EMR collega un SSD per uso generico EBS (gp2) come dispositivo root.

	6.15 e successivi	6.14 e precedenti
Tipo di volume root predefinito		
Dimensioni predefinite		
IOPS predefiniti		
Velocità di trasmissione effettiva predefinita		

Per informazioni su come personalizzare il volume del dispositivo root di Amazon EBS, consulta Specifica di volumi di archiviazione EBS aggiuntivi.

Specifica di volumi di archiviazione EBS aggiuntivi

Quando configuri tipi di istanze in Amazon EMR, puoi specificare volumi EBS aggiuntivi per aggiungere capacità oltre l'archivio istanza (se presente) e il volume EBS predefinito. Amazon EBS fornisce i seguenti tipi di volume: per scopo generico (SSD), IOPS con provisioning (SSD), velocità effettiva ottimizzata (HDD), Cold (HDD) e magnetici. Si differenziano per caratteristiche di prestazioni e prezzo, perciò puoi personalizzare il tuo spazio di archiviazione in base alle esigenze analitiche e aziendali delle applicazioni. Ad esempio, per alcune applicazioni potrebbe essere necessario riversare su disco, mentre altre possono funzionare in modo sicuro in memoria o con Amazon S3.

Puoi collegare volumi Amazon EBS a istanze solo al momento dello startup del cluster e quando aggiungi un gruppo di istanze del nodo attività aggiuntivo. Se un'istanza in un cluster Amazon EMR non va a buon fine, l'istanza e i volumi Amazon EBS collegati vengono sostituiti con nuovi volumi. Di conseguenza, se si scollega manualmente un volume Amazon EBS, questa operazione viene considerata da Amazon EMR come un errore e sostituisce l'archiviazione dell'istanza (se applicabile) e gli store di volumi.

Amazon EMR non consente di modificare il tipo di volume da gp2 a gp3 per un cluster EMR esistente. Per utilizzare gp3 per i tuoi carichi di lavoro/, devi avviare un nuovo cluster EMR. Inoltre, non è consigliabile aggiornare la velocità di trasmissione effettiva e gli IOPS su un cluster in uso o allocazione, poiché Amazon EMR utilizza i valori di velocità di trasmissione effettiva e IOPS specificati al momento dell'avvio del cluster per ogni nuova istanza aggiunta durante il dimensionamento del cluster. Per ulteriori informazioni, consultare Confronto tra i tipi di volume gp2 e gp3 di Amazon EBS e Selezione di IOPS e velocità effettiva durante la migrazione a tipi di volume Amazon EBS gp3.

> **Important**
>
> Per utilizzare un volume gp3 con il cluster EMR, devi avviare un nuovo cluster.

Confronto tra i tipi di volume gp2 e gp3 di Amazon EBS

Ecco un confronto tra i costi dei volumi gp2 e gp3 nella regione Stati Uniti orientali (Virginia settentrionale). Per informazioni più aggiornate, consulta la pagina del prodotto dedicata ai volumi per uso generico di Amazon EBS e la pagina dei prezzi di Amazon EBS.

Tipo di volume	gp3	gp2
Volume size (Dimensione dei volumi)	Da 1 GiB a 16 TiB	Da 1 GiB a 16 TiB
IOPS predefiniti/di base	3000	Da 3 IOPS/GiB (minimo 100 IOPS) a un massimo di 16.000 IOPS. I volumi inferiori a 1 TiB possono anche espandersi fino a 3.000 IOPS.
IOPS massimi per volume	16,000	16,000
Velocità di trasmissione effettiva predefinita/base	125 MiB/s	Il limite di throughput è compreso tra 128 MiB/s e 250 MiB/s, a seconda delle dimensioni del volume.
Velocità di trasmissione effettiva massima per volume	1.000 MiB/s	250 MiB/s
Prezzo	$0,08/gib al mese (3.000 IOPS gratuiti) e 0,005 USD/mese (IOPS) oltre 3.000 USD; 125 USD gratuiti e 0,04 USD al mese (forniti) MiB/s MiB/s-month over 125MiB/s	0,10 USD per GiB al mese

Selezione di IOPS e velocità effettiva durante la migrazione a tipi di volume Amazon EBS gp3

Per allocare un volume gp2, è necessario calcolare le dimensioni del volume al fine di ottenere IOPS e velocità di trasmissione effettiva proporzionali. Con gp3, non è necessario allocare un volume di dimensioni maggiori per ottenere prestazioni più elevate. Puoi scegliere le dimensioni e le prestazioni desiderate in base alle esigenze dell'applicazione. La selezione delle dimensioni e dei parametri prestazionali corretti (IOPS e velocità di trasmissione effettiva) consente di ridurre al massimo i costi senza influire sulle prestazioni.

Ecco una tabella per aiutarti a selezionare le opzioni di configurazione di gp3:

Volume size (Dimensione dei volumi)	IOPS	Prestazioni
Da 1 a 170 GiB	3000	125 MiB/s
Da 170 a 334 GiB	3000	125 MiB/s se il tipo di istanza scelto EC2 supporta 125 *. MiB/s or less, use higher as per usage, Max 250 MiB/s
Da 334 a 1.000 GiB	3000	125 MiB/s se il tipo di EC2 istanza scelto supporta 125 MiB/s or less, Use higher as per usage, Max 250 MiB/s *.
1.000 GiB e oltre	Corrisponde agli IOPS di gp2 (dimensioni in GiB x 3) o agli IOPS massimi forniti dal volume gp2 corrente	125 MiB/s se il tipo di EC2 istanza scelto supporta 125 MiB/s or less, Use higher as per usage, Max 250 MiB/s *.

*Gp3 è in grado di fornire un throughput fino a 1000MiB/s. Since gp2 provides a maximum of 250MiB/s, potrebbe non essere necessario superare questo limite quando si utilizza gp3.

Configurazione della rete in un VPC per Amazon EMR

La maggior parte dei cluster viene avviata in una rete virtuale utilizzando Amazon Virtual Private Cloud (Amazon VPC). Un VPC è una rete virtuale isolata all'interno AWS che è logicamente isolata all'interno del tuo account. AWS È possibile configurare aspetti quali gli intervalli di indirizzi IP privati, le sottoreti, le tabelle di routing e i gateway di rete. Per ulteriori informazioni, consulta la Guida per l'utente di Amazon VPC.

VPC offre le seguenti caratteristiche:

- Elaborazione di dati sensibili

  L'avvio di un cluster in un VPC è simile all'avvio del cluster in una rete privata con strumenti aggiuntivi, come tabelle di routing e ACLs rete, per definire chi ha accesso alla rete. Se si stanno elaborando dati sensibili nel cluster, potrebbe essere necessario il controllo degli accessi

aggiuntivo fornito dall'avvio del cluster in un VPC. Inoltre, puoi scegliere di avviare le risorse in una sottorete privata in cui nessuna di tali risorse dispone di una connessione a Internet diretta.

- Accesso alle risorse su una rete interna

  Se la fonte dei dati si trova in una rete privata, potrebbe essere poco pratico o indesiderabile caricare tali dati AWS per l'importazione in Amazon EMR, a causa della quantità di dati da trasferire o della natura sensibile dei dati. Invece, puoi avviare il cluster in un VPC e collegare il data center al VPC tramite una connessione VPN, consentendo al cluster di accedere a risorse sulla rete interna. Ad esempio, se nel data center è disponibile un database Oracle, l'avvio del cluster in un VPC connesso a tale rete tramite VPN consente al cluster di accedere al database Oracle.

Sottoreti pubbliche e private

Puoi avviare cluster Amazon EMR in sottoreti VPC pubbliche e private. Ciò significa che non è necessaria la connettività Internet per gestire un cluster Amazon EMR; tuttavia, potrebbe essere necessario configurare la traduzione degli indirizzi di rete (NAT) e i gateway VPN per accedere a servizi o risorse che si trovano al di fuori del VPC, ad esempio in una intranet aziendale o in endpoint di servizio pubblico come. AWS AWS Key Management Service

> **Important**
>
> Amazon EMR supporta solo l'avvio di cluster in sottoreti private nelle versioni 4.2 e successive.

Per ulteriori informazioni su Amazon VPC, consulta la Guida per l'utente di Amazon VPC.

Argomenti

- Opzioni Amazon VPC all'avvio di un cluster
- Configura un VPC per ospitare cluster Amazon EMR
- Avvia i cluster in un VPC con Amazon EMR
- Politiche di esempio per sottoreti private che accedono ad Amazon S3
- Altre risorse per saperne di più VPCs

Opzioni Amazon VPC all'avvio di un cluster

Quando avvii un cluster Amazon EMR all'interno di un VPC, è possibile avviarlo all'interno di una sottorete pubblica, privata o condivisa. A seconda del tipo di sottorete scelta per un cluster, esistono leggere ma significative differenze di configurazione.

Sottoreti pubbliche

I cluster EMR in una sottorete pubblica richiedono un gateway Internet connesso. Questo perché i cluster Amazon EMR devono accedere ai servizi AWS e ad Amazon EMR. Se un servizio, ad esempio Amazon S3, offre la possibilità di creare un endpoint VPC, puoi accedere a tali servizi utilizzando l'endpoint anziché accedere a un endpoint pubblico tramite un gateway Internet. Inoltre, Amazon EMR non è in grado di comunicare con i cluster in sottoreti pubbliche tramite un dispositivo NAT (Network Address Translation). A questo scopo è richiesto un gateway Internet, ma puoi continuare a utilizzare un'istanza NAT o gateway per altro traffico in scenari più complessi.

Tutte le istanze in un cluster si connettono ad Amazon S3 tramite un endpoint VPC o un gateway Internet. Altri AWS servizi che attualmente non supportano gli endpoint VPC utilizzano solo un gateway Internet.

Se disponi di AWS risorse aggiuntive che non desideri collegare al gateway Internet, puoi avviare tali componenti in una sottorete privata creata all'interno del tuo VPC.

I cluster in esecuzione in una sottorete pubblica utilizzano due gruppi di sicurezza: uno per il nodo primario e l'altro per i nodi core e attività. Per ulteriori informazioni, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Nel seguente diagramma viene mostrato in che modo un cluster Amazon EMR viene eseguito in un VPC utilizzando una sottorete pubblica. Il cluster è in grado di connettersi ad altre AWS risorse, come i bucket Amazon S3, tramite il gateway Internet.

Nel seguente diagramma viene mostrato come configurare un VPC per consentire a un cluster nel VPC di accedere a risorse nella propria rete, ad esempio un database Oracle.

Sottoreti private

Una sottorete privata consente di avviare AWS risorse senza richiedere alla sottorete di disporre di un gateway Internet collegato. Amazon EMR supporta l'avvio di cluster in sottoreti private con versioni 4.2.0 o successive.

> **Note**
>
> Quando si configura un cluster Amazon EMR in una sottorete privata, si consiglia di configurare anche gli endpoint VPC per Simple Storage Service (Amazon S3). Se il cluster EMR si trova in una sottorete privata senza endpoint VPC per Simple Storage Service (Amazon S3), verranno addebitati costi aggiuntivi del gateway NAT associati al traffico S3 perché il traffico tra il cluster EMR e S3 non rimarrà all'interno del VPC.

Le sottoreti private differiscono dalle sottoreti pubbliche nei modi seguenti:

- Per accedere ai AWS servizi che non forniscono un endpoint VPC, devi comunque utilizzare un'istanza NAT o un gateway Internet.
- Ti consigliamo di fornire almeno un percorso ai bucket di log del servizio Amazon EMR e al repository Amazon Linux in Amazon S3. Per ulteriori informazioni, consulta Politiche di esempio per sottoreti private che accedono ad Amazon S3
- Se utilizzi caratteristiche EMRFS, devi disporre di un endpoint VPC Amazon S3 e un percorso dalla sottorete privata a DynamoDB.
- Il debug funziona solo se fornisci un percorso dalla sottorete privata a un endpoint Amazon SQS pubblico.
- La creazione di una configurazione della sottorete privata con un'istanza NAT o un gateway in una sottorete pubblica è supportata solo utilizzando AWS Management Console. Il modo più semplice per aggiungere e configurare istanze NAT ed endpoint VPC Amazon S3 per cluster Amazon EMR è utilizzare la pagina VPC Subnets List (Elenco di sottoreti VPC) nella console Amazon EMR. Per configurare gateway NAT, consulta Gateway NAT nella Guida per l'utente di Amazon VPC.
- Non puoi modificare una sottorete con un cluster Amazon EMR esistente da pubblica a privata o viceversa. Per individuare un cluster Amazon EMR all'interno di una sottorete privata, il cluster deve essere avviato in tale sottorete privata.

Amazon EMR crea e utilizza diversi gruppi di sicurezza predefiniti per i cluster in una sottorete privata: ElasticMapReduce-Master-Private, ElasticMapReduce -Slave-Private e -. ElasticMapReduce ServiceAccess Per ulteriori informazioni, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Per un elenco completo NACLs del tuo cluster, scegli Gruppi di sicurezza per Primary e Gruppi di sicurezza per Core & Task nella pagina Dettagli del cluster della console Amazon EMR.

Nella seguente immagine viene mostrato in che modo un cluster Amazon EMR viene configurato all'interno di una sottorete privata. La sola comunicazione all'esterno della sottorete è verso Amazon EMR.

Nell'immagine seguente viene mostrata una configurazione di esempio per un cluster Amazon EMR all'interno di una sottorete privata connessa a un'istanza NAT che si trova in una sottorete pubblica.

Sottoreti condivise

La condivisione VPC consente ai clienti di condividere sottoreti con altri AWS account all'interno della stessa organizzazione. AWS Puoi avviare i cluster Amazon EMR in sottoreti condivise pubbliche e private, con le seguenti avvertenze.

Il proprietario della sottorete deve condividere una sottorete prima di poter avviare un cluster Amazon EMR. Tuttavia, le sottoreti condivise possono essere successivamente non condivise. Per ulteriori informazioni, consulta Working with Shared. VPCs Quando un cluster viene avviato in una sottorete condivisa e tale sottorete condivisa viene quindi non condivisa, è possibile osservare comportamenti specifici basati sullo stato del cluster Amazon EMR quando la sottorete non è condivisa.

- La condivisione della sottorete viene annullata prima che il cluster venga avviato: se il proprietario interrompe la condivisione dell'Amazon VPC o della sottorete mentre il partecipante avvia un cluster, il cluster potrebbe non avviarsi o essere parzialmente inizializzato senza eseguire il provisioning di tutte le istanze richieste.
- La condivisione della sottorete viene annullata dopo il cluster è stato lanciato: quando il proprietario interrompe la condivisione di una sottorete o dell'Amazon VPC con il partecipante, i cluster del partecipante non saranno in grado di eseguire il dimensionamento per aggiungere nuove istanze o sostituire istanze non integre.

Quando avvii un cluster Amazon EMR, vengono creati più i gruppi di sicurezza. In una sottorete condivisa, il partecipante controlla questi gruppi di sicurezza. Il proprietario della sottorete può visualizzare i gruppi di sicurezza, ma non è in grado di eseguire azioni. Se il proprietario della sottorete desidera rimuovere o modificare il gruppo di sicurezza, il partecipante che ha creato il gruppo di sicurezza deve eseguire l'azione.

Controllo delle autorizzazioni VPC con IAM

Per impostazione predefinita, tutti gli utenti possono vedere tutte le sottoreti per l'account e qualsiasi utente può avviare un cluster in qualsiasi sottorete.

Quando avvii un cluster in un VPC, puoi utilizzare AWS Identity and Access Management (IAM) per controllare l'accesso ai cluster e limitare le azioni utilizzando le policy, proprio come faresti con i cluster lanciati in Amazon Classic. EC2 Per ulteriori informazioni su IAM, consulta la Guida per l'utente IAM.

Puoi anche utilizzare IAM per controllare chi può creare e amministrare le sottoreti. Ad esempio, puoi creare un ruolo IAM per amministrare le sottoreti e un secondo ruolo che può avviare i cluster ma non può modificare le impostazioni di Amazon VPC. Per ulteriori informazioni sull'amministrazione delle politiche e delle azioni in Amazon EC2 e Amazon VPC, consulta IAM Policies for Amazon nella EC2 EC2 Amazon User Guide.

Configura un VPC per ospitare cluster Amazon EMR

Prima di poter avviare i cluster in un VPC, devi creare un VPC e una sottorete. Per sottoreti pubbliche, devi creare un gateway Internet e collegarlo alla sottorete. Le seguenti istruzioni descrivono come creare un VPC in grado di ospitare cluster Amazon EMR.

Creazione di un VPC con sottoretI per un cluster Amazon EMR

1. Apri la console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nella parte in alto a destra della pagina, scegli la Regione AWS per il VPC.
3. Seleziona Crea VPC.
4. Nella pagina VPC settings (Impostazioni VPC), scegli VPC and more (VPC e altro).
5. In Name tag auto-generation (Generazione automatica del tag nome), abilita Auto-generate (Generazione automatica) e inserisci un nome per il VPC. Ciò consente di identificare il VPC e la sottorete nella console Amazon VPC dopo che sono stati creati.
6. Nel campo blocco IPv4 CIDR, inserisci uno spazio di indirizzi IP privato per il tuo VPC per garantire la corretta risoluzione del nome host DNS; in caso contrario, potresti riscontrare errori del cluster Amazon EMR. Ciò include i seguenti intervalli di indirizzi 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. In Numero di zone di disponibilità (AZs), scegli il numero di zone di disponibilità in cui desideri avviare le sottoreti.
8. In Number of public subnets (Numero di sottoreti pubbliche), scegli una singola sottorete pubblica che vuoi aggiungere al tuo VPC. Se i dati utilizzati nel cluster sono disponibili su Internet (ad esempio, in Amazon S3 o Amazon RDS), ti basta utilizzare una sottorete pubblica senza bisogno di aggiungere una sottorete privata.
9. Per Number of private subnets (Numero di sottoreti private), scegli il numero di sottoreti private che vuoi aggiungere al tuo VPC. Selezionane una o più se i dati per l'applicazione sono archiviati nella rete in uso (ad esempio, in un database Oracle). Per un VPC in una sottorete privata, tutte le EC2 istanze Amazon devono avere almeno un percorso verso Amazon EMR tramite l'interfaccia di rete elastica. Nella console, questo viene configurato automaticamente.
10. In NAT gateways (Gateway NAT), puoi facoltativamente scegliere di aggiungere gateway NAT. Sono necessari solo se disponi di sottoreti private che devono comunicare con Internet.
11. In VPC endpoints (Endpoint VPC), puoi facoltativamente scegliere di aggiungere endpoint per Amazon S3 alle tue sottoreti.
12. Verifica che i campi Enable DNS hostnames (Abilita hostname DNS) e Enable DNS resolution (Abilita risoluzione DNS) siano selezionati. Per ulteriori informazioni, consulta Utilizzo del DNS con il tuo VPC.
13. Seleziona Crea VPC.
14. Una finestra di stato mostra lo stato di avanzamento del lavoro. Al termine del lavoro, scegli Visualizza VPC per accedere alla pagina Il VPCs tuo, che mostra il tuo VPC predefinito e il VPC appena creato. Il VPC creato è un VPC non predefinito, di conseguenza la colonna Default VPC (VPC predefinito) visualizza No.
15. Se desideri associare il VPC a una voce DNS che non include un nome di dominio, vai su DHCP option sets (Set di opzioni DHCP), scegli Create DHCP options set (Crea set di opzioni DHCP) e ometti un nome di dominio. Dopo aver creato il set di opzioni, accedi al nuovo VPC, scegli Edit DHCP options set (Modifica set di opzioni DHCP) nel menu Actions (Operazioni) e seleziona il nuovo set di opzioni. Non puoi modificare il nome di dominio utilizzando la console dopo che il set di opzioni DNS è stato creato.

    Con Hadoop e applicazioni correlate è una best practice garantire la risoluzione del nome di dominio completo (FQDN) per i nodi. Per garantire la risoluzione DNS corretta, configura un VPC che include un set di opzioni DHCP i cui parametri sono impostati sui seguenti valori:

    - domain-name = **`ec2.internal`**

Utilizza **`ec2.internal`** se la Regione è Stati Uniti orientali (Virginia settentrionale). Per altre regioni, usa. *`region-name`* **`.compute.internal`** Per gli esempi in `us-west-2`, utilizza **`us-west-2.compute.internal`**. Per la regione AWS GovCloud (Stati Uniti occidentali), usa**`us-gov-west-1.compute.internal`**.

- domain-name-servers = **`AmazonProvidedDNS`**

Per ulteriori informazioni, consulta Set opzioni DHCP nella Guida per l'utente di Amazon VPC.

16. Dopo che il VPC è stato creato, vai alla pagina Subnets (Sottoreti) e prendi nota del Subnet ID (ID sottorete) di una delle sottoreti del nuovo VPC. Utilizza queste informazioni quando avvii il cluster Amazon EMR nel VPC.

Avvia i cluster in un VPC con Amazon EMR

Quando disponi di una sottorete configurata per ospitare cluster Amazon EMR, avvia il cluster in tale sottorete specificando l'identificatore di sottorete associato durante la creazione del cluster.

> **Note**
>
> Amazon EMR supporta sottoreti private nelle versioni finali 4.2 e successive.

Quando il cluster viene avviato, Amazon EMR aggiunge gruppi di sicurezza a seconda che il cluster venga avviato in sottoreti private o pubbliche VPC. Tutti i gruppi di sicurezza consentono l'ingresso alla porta 8443 per la comunicazione con il servizio Amazon EMR, ma gli intervalli di indirizzi IP variano per sottoreti pubbliche e private. Amazon EMR gestisce tutti questi gruppi di sicurezza e nel tempo potrebbe dover aggiungere ulteriori indirizzi IP all' AWS intervallo. Per ulteriori informazioni, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Per gestire il cluster su un VPC, Amazon EMR collega un dispositivo di rete al nodo primario e lo gestisce tramite questo dispositivo. Puoi visualizzare questo dispositivo utilizzando l'azione Amazon EC2 API `DescribeInstances`. Se modifichi questo dispositivo in qualsiasi modo, il cluster potrebbe non riuscire.

Console

Per avviare un cluster in un VPC con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Networking (Reti), vai al campo Virtual private cloud (VPC) (Cloud privato virtuale [VPC]). Inserisci il nome del tuo VPC o scegli Browse (Sfoglia) per selezionarlo. In alternativa, scegli Create VPC (Crea VPC) per creare un VPC da utilizzare per il cluster.
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per avviare un cluster in un VPC con AWS CLI

> **Note**
>
> AWS CLI non fornisce un modo per creare automaticamente un'istanza NAT e connetterla alla sottorete privata. Tuttavia, per creare un endpoint S3 nella sottorete puoi utilizzare i comandi della CLI di Amazon VPC. Utilizza la console per creare istanze NAT e avviare cluster in una sottorete privata.

Dopo che il VPC è stato configurato, puoi avviare cluster Amazon EMR al suo interno utilizzando il sottocomando `create-cluster` con il parametro `--ec2-attributes`. Utilizza il parametro `--ec2-attributes` per specificare la sottorete VPC per il cluster.

- Per creare un cluster in una sottorete specifica, digita il seguente comando, sostituiscilo *myKey* con il nome della tua coppia di EC2 chiavi Amazon e sostituiscilo *77XXXX03* con il tuo ID di sottorete.

```
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
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzano il tipo di istanza specificato nel comando.

> **Note**
>
> Se in precedenza non hai creato il ruolo e il profilo di EC2 istanza del servizio Amazon EMR predefiniti, digita `aws emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Garantire la disponibilità degli indirizzi IP per un cluster EMR su EC2

Per garantire che una sottorete con un numero sufficiente di indirizzi IP liberi sia disponibile al momento dell'avvio, la selezione della EC2 sottorete verifica la disponibilità degli IP. Se il processo di creazione utilizza una sottorete con il numero necessario di indirizzi IP per avviare i nodi principali, primari e task come richiesto, anche se al momento della creazione iniziale vengono creati solo i nodi principali per il cluster. EMR verifica il numero di indirizzi IP necessari per avviare i nodi primari e i task node durante la creazione, oltre a calcolare separatamente il numero di indirizzi IP necessari per avviare i nodi principali. Il numero minimo di istanze o nodi primari e di task richiesti viene determinato automaticamente da Amazon EMR.

> **Important**
>
> Se nessuna sottorete nel VPC dispone di una quantità sufficiente IPs per ospitare i nodi essenziali, viene restituito un errore e il cluster non viene creato.

Nella maggior parte dei casi di implementazione, esiste una differenza di orario tra ogni avvio dei nodi principali, primari e task. Inoltre, è possibile che più cluster condividano una sottorete. In questi casi, la disponibilità degli indirizzi IP può variare e i successivi lanci di task-node, ad esempio, possono essere limitati dagli indirizzi IP disponibili.

Politiche di esempio per sottoreti private che accedono ad Amazon S3

Per le sottoreti private, devi offrire ad Amazon EMR almeno la possibilità di accedere ai repository Amazon Linux. Questa policy della sottorete privata fa parte delle policy endpoint VPC per accedere ad Amazon S3.

Con Amazon EMR 5.25.0 o versioni successive, per abilitare l'accesso con un clic a Spark History Server persistente, devi consentire ad Amazon EMR di accedere al bucket di sistema che raccoglie i log di eventi Spark. Se abiliti la registrazione, fornisci le autorizzazioni PUT al seguente bucket:

```
aws157-logs-${AWS::Region}/*
```

Per ulteriori informazioni, consulta Accesso con un clic a Spark History Server persistente.

Spetta a te determinare le restrizioni della policy che soddisfano le esigenze aziendali. La seguente policy di esempio fornisce le autorizzazioni per accedere ai repository Amazon Linux e al bucket di sistema Amazon EMR per la raccolta dei log di eventi Spark. Mostra alcuni esempi di nomi di risorse per i bucket.

Per ulteriori informazioni sull'utilizzo delle policy IAM con gli endpoint Amazon VPC, consulta Policy dell'endpoint per Amazon S3.

Il seguente esempio di policy contiene risorse di esempio nella regione 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/*"
      ]
    }
  ]
}
```

La seguente policy di esempio fornisce le autorizzazioni necessarie per accedere ai repository Amazon Linux 2 nella regione 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/*"
           ]
       }
   ]
}
```

La seguente policy di esempio fornisce le autorizzazioni necessarie per accedere ai repository Amazon Linux 2023 nella regione us-east-1.

```
{
    "Statement": [
        {
            "Sid": "AmazonLinux2023AMIRepositoryAccess",
            "Effect": "Allow",
            "Principal": "*",
```

```
            "Action": "s3:GetObject",
            "Resource": [
                "arn:aws:s3:::al2023-repos-us-east-1-de612dc2/*"
            ]
        }
    ]
}
```

Regioni disponibili

La tabella seguente contiene un elenco di bucket per regione e include sia un Amazon Resource Name (ARN) per il repository sia una stringa che rappresenta l'ARN per. `appinfo.src` L'ARN, o Amazon Resource Name, è una stringa che identifica in modo univoco una risorsa. AWS

Regione	Bucket di repository	AppInfo secchio
Stati Uniti orientali (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.e mr.amazonaws.com/*»	«arn:aws:s3: :prod.us-east-2.ap pinfo.src/*»
Stati Uniti orientali (Virginia settentrionale)	«arn:aws:s3::: packages.us-east-1 .amazonaws.com/», "arn:aws:s3::: repo.us-east-1.amazonaws.com/», "arn:aws:s3: ::repo.us-east-1.e mr.amazonaws.com/*»	«arn:aws:s3: :prod.us-east-1.ap pinfo.src/*»
Stati Uniti occidentali (California settentrionale)	«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.e mr.amazonaws.com/*»	«arn:aws:s3: :prod.us-west-1.ap pinfo.src/*»
Stati Uniti occidentali (Oregon)	«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.e mr.amazonaws.com/*»	«arn:aws:s3: :prod.us-west-2.ap pinfo.src/*»

Regione	Bucket di repository	AppInfo secchio
Africa (Città del Capo)	«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/*»
Africa (Città del Capo)	«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 Pacific (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/*»
Asia Pacifico (Giacarta)	«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 Pacifico (Malesia)	«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 Pacifico (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/*»

Regione	Bucket di repository	AppInfo secchio
Asia Pacifico (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 Pacifico (Osaka-Locale)	«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 Pacifico (Seoul)	«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/*»
Asia Pacifico (Singapore)	«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 Pacifico (Sydney)	«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 Pacifico (Tokyo)	«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/*»

Regione	Bucket di repository	AppInfo secchio
Canada (Centrale)	«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/*»
Canada occidentale (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 (Francoforte)	«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/*»
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 (Londra)	«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 (Milano)	«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/*»

Regione	Bucket di repository	AppInfo secchio
Europa (Parigi)	«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.e mr.amazonaws.com/*»	«arn:aws:s3: :prod.eu-west-3.ap pinfo.src/*»
Europa (Spagna)	«arn:aws:s3::: packages.eu-south-2.amazonaws.com/», "arn:aws: s3::: repo.eu-south-2.amazonaws.c om/», "arn:aws:s3: ::repo.eu-south-2. emr.amazonaws.com/*»	«arn:aws:s3: :prod.eu-south-2.a ppinfo.src/*»
Europa (Stoccolma)	«arn:aws:s3::: packages.eu-north-1.amazonaws.com/», "arn:aws: s3::: repo.eu-north-1.amazonaws.c om/», "arn:aws:s3: ::repo.eu-north-1. emr.amazonaws.com/*»	«arn:aws:s3: :prod.eu-north-1.a ppinfo.src/*»
Europa (Zurigo)	«arn:aws:s3::: packages.eu-centra l-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/*»
Israele (Tel Aviv)	«arn:aws:s3::: packages.il-centra l-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 (Bahrein)	«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/*»

Regione	Bucket di repository	AppInfo secchio
Medio Oriente (Emirati Arabi Uniti)	«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/*»
Sud America (San Paolo)	«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/*»
AWS GovCloud (Stati Uniti orientali)	«arn:aws:s3: :pacchetti. 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 (Stati Uniti occidentali)	«arn:aws:s3: :pacchetti. 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/*»

Altre risorse per saperne di più VPCs

Usa i seguenti argomenti per saperne di più sulle VPCs sottoreti.

- Sottoreti private in un VPC
  - Scenario 2: VPC con sottoreti pubbliche e private (NAT)
  - Istanze NAT
  - Elevata disponibilità per istanze NAT Amazon VPC: un esempio
- Sottoreti pubbliche in un VPC
  - Scenario 1: VPC con una sottorete pubblica singola

- Informazioni VPC generali
  - Guida per l'utente di Amazon VPC
  - Peering di VPC
  - Utilizzo di interfacce di rete elastiche con il tuo VPC
  - Connessione sicura alle istanze Linux in esecuzione in un VPC privato

Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi

Quando si crea un cluster e si specifica la configurazione del nodo primario, dei nodi core e dei nodi attività, sono disponibili due opzioni di configurazione. Puoi utilizzare parchi istanze o gruppi di istanze uniformi. L'opzione di configurazione scelta si applica a tutti i nodi, vale per la durata del cluster e i parchi istanze e gruppi di istanze non possono coesistere in un cluster. La configurazione di parchi istanze è disponibile in Amazon EMR versione 4.8.0 e successive, escluse le versioni 5.0.x.

Puoi utilizzare la console Amazon EMR AWS CLI, o l'API Amazon EMR per creare cluster con entrambe le configurazioni. Quando utilizzi il comando `create-cluster` di AWS CLI, puoi utilizzare i parametri `--instance-fleets` per creare il cluster utilizzando parchi istanze o, in alternativa, puoi utilizzare i parametri `--instance-groups` per crearlo utilizzando gruppi di istanze uniformi.

Lo stesso vale utilizzando l'API Amazon EMR. Utilizza la configurazione `InstanceGroups` per specificare una matrice di oggetti `InstanceGroupConfig`, oppure utilizza la configurazione `InstanceFleets` per specificare una matrice di oggetti `InstanceFleetConfig`.

Nella nuova console Amazon EMR, puoi scegliere di utilizzare gruppi di istanze o parchi istanze quando crei un cluster e hai la possibilità di utilizzare le istanze spot con ciascuno di essi. Nella vecchia console Amazon EMR, se utilizzi le impostazioni Quick Options (Opzioni rapide) predefinite quando crei un cluster, Amazon EMR applica al cluster la configurazione dei gruppi di istanze uniformi e utilizza le istanze on demand. Per utilizzare istanze Spot con gruppi di istanze uniformi o per configurare parchi istanze e altre personalizzazioni, scegli Advanced Options (Opzioni avanzate).

Parchi istanze

La configurazione delle flotte di istanze offre la più ampia varietà di opzioni di provisioning per le istanze Amazon. EC2 Ogni tipo di nodo dispone di un singolo parco istanze e l'utilizzo di un parco istanze dell'attività è opzionale. Puoi specificare fino a cinque tipi di EC2 istanze per flotta o 30 tipi di EC2 istanze per flotta quando crei un cluster utilizzando l'API AWS CLI o Amazon EMR e una strategia di allocazione per istanze On-Demand e Spot. Per i parchi istanze principali e attività, puoi assegnare una capacità target per istanze on demand e un'altra per istanze Spot. Amazon

EMR sceglie qualsiasi combinazione dei tipi di istanze specificati per raggiungere le capacità target, eseguendo il provisioning sia delle istanze on demand che Spot.

Per il tipo di nodo primario, Amazon EMR consente di scegliere un singolo tipo di istanza dall'elenco e di specificare se è stato effettuato il provisioning come un'istanza on demand o spot. I parchi istanze offrono anche opzioni aggiuntive per gli acquisti di istanze Spot e on demand. Le opzioni di istanza Spot includono un timeout che specifica un'azione da eseguire se non è possibile assegnare la capacità Spot e una strategia di allocazione preferita (ottimizzata per la capacità) per avviare i parchi di istanze Spot. I parchi di istanze on demand possono anche essere avviati utilizzando la strategia di allocazione (prezzo più basso). Se si utilizza un ruolo di servizio che non è il ruolo di servizio predefinito EMR o si utilizza una policy gestita da EMR nel ruolo del servizio, è necessario aggiungere autorizzazioni aggiuntive al ruolo del servizio cluster personalizzato per abilitare la strategia di allocazione. Per ulteriori informazioni, consulta Ruolo di servizio per Amazon EMR (ruolo EMR).

Per ulteriori informazioni sulla configurazione dei parchi istanze, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR.

Gruppi di istanze uniformi

I gruppi di istanze uniformi offrono una configurazione più semplice rispetto ai parchi istanze. Ogni cluster Amazon EMR può includere fino a 50 gruppi di istanze: un gruppo di istanze primario che contiene un' EC2 istanza Amazon, un gruppo di istanze core che contiene una o più EC2 istanze e fino a 48 gruppi di istanze di task opzionali. Ogni gruppo di istanze core e task può contenere un numero qualsiasi di EC2 istanze Amazon. Puoi scalare ogni gruppo di istanze aggiungendo e rimuovendo EC2 istanze Amazon manualmente oppure puoi configurare il ridimensionamento automatico. Per informazioni sull'aggiunta e la rimozione di istanze, consulta Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione.

Per ulteriori informazioni sulla configurazione di gruppi di istanze uniformi, consulta Configura gruppi di istanze uniformi per il tuo cluster Amazon EMR.

Utilizzo di parchi istanze e gruppi di istanze

Argomenti

- Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR
- Riconfigurazione delle flotte di istanze per il tuo cluster Amazon EMR
- Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR

- Configura gruppi di istanze uniformi per il tuo cluster Amazon EMR
- Flessibilità della zona di disponibilità per un cluster Amazon EMR
- Configurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot

Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR

> **Note**
>
> La configurazione dei parchi istanze è disponibile solo in Amazon EMR rilasci 4.8.0 e successivi, esclusi i rilasci 5.0.0 e 5.0.3.

La configurazione della flotta di istanze per i cluster Amazon EMR ti consente di selezionare un'ampia varietà di opzioni di provisioning per le EC2 istanze Amazon e ti aiuta a sviluppare una strategia di risorse flessibile ed elastica per ogni tipo di nodo del cluster.

Nella configurazione di un parco istanze, specifica una target capacity (capacità target) per le On-Demand Instances (istanze on demand) e le Spot Instances (istanze Spot) all'interno di ogni parco istanze. All'avvio del cluster, Amazon EMR effettua il provisioning di istanze fino al raggiungimento delle capacità target. Quando Amazon EC2 recupera un'istanza Spot in un cluster in esecuzione a causa di un aumento di prezzo o di un errore dell'istanza, Amazon EMR tenta di sostituire l'istanza con uno dei tipi di istanza specificati. In questo modo è più semplice riguadagnare la capacità durante un picco dei prezzi Spot.

Puoi specificare un massimo di cinque tipi di EC2 istanze Amazon per flotta da utilizzare con Amazon EMR per raggiungere gli obiettivi o un massimo di 30 tipi di EC2 istanze Amazon per flotta quando crei un cluster utilizzando l'API o AWS CLI Amazon EMR e una strategia di allocazione per istanze On-Demand e Spot.

Puoi anche selezionare più sottoreti per zone di disponibilità differenti. Quando Amazon EMR avvia il cluster, viene eseguita una ricerca in tali sottoreti per individuare le istanze e le opzioni di acquisto specificate. Se Amazon EMR rileva un evento AWS su larga scala in una o più zone di disponibilità, Amazon EMR tenta automaticamente di indirizzare il traffico lontano dalle zone di disponibilità interessate e tenta di lanciare nuovi cluster che crei in zone di disponibilità alternative in base alle tue selezioni. La selezione della zona di disponibilità del cluster avviene solo alla creazione del cluster. I nodi del cluster esistenti non vengono riavviati automaticamente in una nuova zona di disponibilità in caso di interruzione della zona di disponibilità.

Considerazioni sull'utilizzo di flotte di istanze

Considera gli elementi seguenti quando utilizzi il parco istanze con Amazon EMR.

- Puoi avere uno e un solo parco istanze per tipo di nodo (primario, principale, attività). Puoi specificare fino a cinque tipi di EC2 istanze Amazon per ogni parco istanze AWS Management Console (o un massimo di 30 tipi per flotta di istanze quando crei un cluster utilizzando l'API AWS CLI o Amazon EMR e una Strategia di allocazione per parchi istanze).
- Amazon EMR sceglie uno o tutti i tipi di EC2 istanze Amazon specificati per fornire opzioni di acquisto Spot e On-Demand.
- Puoi definire capacità target per istanze Spot e on demand per il parco istanze principale e il parco istanze attività. Usa vCPU o un'unità generica assegnata a ciascuna EC2 istanza Amazon che conta ai fini degli obiettivi. Amazon EMR effettua il provisioning di istanze finché ogni capacità target non viene completamente raggiunta. Per il parco istanze primarie, il target è sempre uno.
- Puoi scegliere una sottorete (zona di disponibilità) o un intervallo. Se selezioni un intervallo, Amazon EMR effettua il provisioning della capacità nella zona di disponibilità più adatta.
- Quando si specifica una capacità target per istanze Spot:
  - Per ogni tipo di istanza, specifica un prezzo Spot massimo. Amazon EMR effettua il provisioning delle istanze Spot se il prezzo Spot è inferiore al prezzo Spot massimo. L'utente paga il prezzo Spot, non necessariamente il prezzo Spot massimo.
  - Per ogni parco istanze, definire un periodo di timeout per il provisioning delle istanze Spot. Se Amazon EMR non è in grado di eseguire il provisioning della capacità Spot, è possibile terminare il cluster o passare al provisioning della capacità on demand. Ciò vale solo per il provisioning dei cluster, non per il loro ridimensionamento. Se il periodo di timeout termina durante il processo di ridimensionamento del cluster, le richieste Spot non sottoposte a provisioning verranno annullate senza il trasferimento alla capacità on demand.
- Per ogni parco istanze, puoi specificare una delle seguenti strategie di allocazione per le tue istanze Spot: ottimizzazione del rapporto prezzo/capacità, ottimizzazione della capacità, prezzo più basso o diversificata su tutti i pool capacity-optimized-prioritized.
- Per ogni flotta, puoi applicare le seguenti strategie di allocazione per le tue istanze on demand: la strategia con il prezzo più basso o la strategia prioritaria.
- Per ogni flotta con istanze On-Demand, puoi scegliere di applicare le opzioni di prenotazione della capacità.
- Se si utilizza la strategia di allocazione per flotte di esempio, quando si scelgono le sottoreti per il cluster EMR si applicano le seguenti considerazioni:

  - Quando Amazon EMR effettua il provisioning di un cluster con una task fleet, filtra le sottoreti prive di indirizzi IP disponibili sufficienti per effettuare il provisioning di tutte le istanze del cluster EMR richiesto. Ciò include gli indirizzi IP necessari per le flotte di istanze primarie, principali e task durante l'avvio del cluster. Amazon EMR sfrutta quindi la propria strategia di allocazione per determinare il pool di istanze, in base al tipo di istanza e alle sottoreti rimanenti con indirizzi IP sufficienti, per avviare il cluster.
  - Se Amazon EMR non è in grado di avviare l'intero cluster a causa di indirizzi IP disponibili insufficienti, tenterà di identificare le sottoreti con un numero sufficiente di indirizzi IP liberi per avviare le flotte di istanze essenziali (principali e primarie). In tali scenari, il parco istanze di attività passerà a uno stato sospeso, anziché chiudere il cluster con un errore.
  - Se nessuna delle sottoreti specificate contiene indirizzi IP sufficienti per effettuare il provisioning delle flotte di istanze principali e primarie essenziali, l'avvio del cluster avrà esito negativo e restituirà un VALIDATION_ERROR. Ciò attiva un evento di terminazione del cluster con gravità CRITICA, che notifica all'utente che il cluster non può essere avviato. Per evitare questo problema, consigliamo di aumentare il numero di indirizzi IP nelle sottoreti.
- Se esegui la versione di Amazon EMR emr-7.7.0 e successive e utilizzi la strategia di allocazione per flotte di istanze, puoi scalare il cluster fino a 4000 istanze e 14000 volumi EBS per flotta di EC2 istanze. Per le versioni di rilascio precedenti a emr-7.7.0, il cluster può essere scalato solo fino a 2000 istanze e 7000 volumi EBS per flotta di istanze. EC2
- Quando avvii le istanze On-Demand, puoi utilizzare prenotazioni di capacità aperte o mirate per i nodi primari, principali e di attività dei tuoi account. Potresti riscontrare una capacità insufficiente con le istanze on demand con strategia di allocazione, ad esempio le flotte. Ti consigliamo di specificare più tipi di istanze per diversificare e ridurre la possibilità che si verifichi una capacità insufficiente. Per ulteriori informazioni, consulta the section called “Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR”.

Opzioni del parco istanze

Utilizza le seguenti linee guida per comprendere le opzioni del parco istanze.

Argomenti

- Impostazione delle capacità target
- Opzioni di avvio
- Opzioni di sottorete (zone di disponibilità) multiple
- Configurazione del nodo master

Impostazione delle capacità target

Specifica le capacità target desiderate per il parco istanze core e il parco istanze di attività. Una volta fatto, ciò determina il numero di istanze on demand e istanze Spot di cui Amazon EMR effettua il provisioning. Durante la specifica di un'istanza è l'utente che decide la misura di conteggio di ogni istanza per il raggiungimento del target. Quando si effettua il provisioning di un'istanza on demand, esegue il conteggio fino al raggiungimento della capacità target on demand. Lo stesso vale per istanze Spot. A differenza dei parchi istanze core e attività, il parco istanze primarie è sempre un'unica istanza. Pertanto, la capacità target di questo parco istanze è sempre uno.

Quando usi la console, per impostazione predefinita vengono utilizzate le v CPUs del tipo di EC2 istanza Amazon come conteggio per le capacità target. Puoi cambiarlo in Unità generiche e quindi specificare il conteggio per ogni tipo di EC2 istanza. Quando si utilizza il AWS CLI, si assegnano manualmente unità generiche per ogni tipo di istanza.

> **Important**
>
> Quando si sceglie un tipo di istanza utilizzando il AWS Management Console, il numero di vCPU mostrato per ogni tipo di istanza è il numero di vcore YARN per quel tipo di istanza, non il numero di EC2 v CPUs per quel tipo di istanza. Per ulteriori informazioni sul numero di v CPUs per ogni tipo di istanza, consulta Amazon EC2 Instance Types.

Per ogni flotta, specifichi fino a cinque tipi di EC2 istanze Amazon. Se utilizzi Strategia di allocazione per parchi istanze e crei un cluster utilizzando l' AWS CLI API Amazon EMR, puoi specificare fino a 30 tipi di istanze per flotta di EC2 istanze. Amazon EMR sceglie qualsiasi combinazione di questi tipi di EC2 istanze per soddisfare le capacità target. Poiché l'obiettivo di Amazon EMR è raggiungere la capacità target completa, si possono verificare costi aggiuntivi. Ad esempio, se la capacità di due unità non viene raggiunta e Amazon EMR è in grado di effettuare il provisioning di un'istanza con un conteggio di cinque unità, il provisioning dell'istanza viene ancora effettuato, ossia la capacità target viene superata di tre unità.

Se riduci la capacità target per ridimensionare un cluster in esecuzione, Amazon EMR tenta di completare le attività dell'applicazione e termina le istanze per soddisfare il nuovo target. Per ulteriori informazioni, consulta Terminazione al completamento dell'attività.

Opzioni di avvio

Per le istanze Spot, puoi specificare un Maximum Spot price (Prezzo Spot massimo) per ogni tipo di istanza in un parco istanze. Puoi impostare questo prezzo come una percentuale del prezzo on demand o come un importo in dollari specifico. Amazon EMR effettua il provisioning di istanze Spot se il prezzo Spot corrente in una zona di disponibilità è inferiore al prezzo Spot massimo. L'utente paga il prezzo Spot, non necessariamente il prezzo Spot massimo.

> **Note**
>
> Le istanze spot con una durata definita (note anche come blocchi Spot) non sono più disponibili per i nuovi clienti a partire dal 1° luglio 2021. Per i clienti che hanno già utilizzato la funzione, continueremo a supportare le istanze spot con una durata definita fino al 31 dicembre 2022.

In Amazon EMR 5.12.1 e versioni successive hai la possibilità di avviare parchi di istanze Spot e on demand con allocazione della capacità ottimizzata. Questa opzione di strategia di allocazione può essere impostata nella versione precedente AWS Management Console o utilizzando l'API. `RunJobFlow` Tieni presente che non è possibile personalizzare la strategia di allocazione nella nuova console. L'utilizzo della strategia di allocazione richiede autorizzazioni aggiuntive per il ruolo di servizio. Se utilizzi il ruolo di servizio Amazon EMR predefinito e la policy gestita (`EMR_DefaultRole` e `AmazonEMRServicePolicy_v2`) per il cluster, le autorizzazioni per l'opzione della strategia di allocazione sono già incluse. Se non si utilizza il ruolo di servizio Amazon EMR e la policy gestita predefiniti, è necessario aggiungerli per utilizzare questa opzione. Consultare Ruolo di servizio per Amazon EMR (ruolo EMR).

Per ulteriori informazioni sulle istanze Spot, consulta le istanze Spot nella Amazon EC2 User Guide. Per ulteriori informazioni sulle istanze On-Demand, consulta On-Demand Instances nella Amazon User Guide. EC2

Se si sceglie di avviare parchi istanze on demand con la strategia di allocazione del prezzo più basso, è possibile utilizzare le prenotazioni di capacità. Le opzioni di prenotazione della capacità possono essere impostate utilizzando l'API `RunJobFlow` di Amazon EMR. Le prenotazioni della capacità richiedono autorizzazioni aggiuntive per il ruolo di servizio che è necessario aggiungere per utilizzare queste opzioni. Consultare Autorizzazioni della strategia di allocazione. Tieni presente che non è possibile personalizzare le prenotazioni di capacità nella nuova console.

Opzioni di sottorete (zone di disponibilità) multiple

Quando utilizzi flotte di istanze, puoi specificare più EC2 sottoreti Amazon all'interno di un VPC, ognuna corrispondente a una zona di disponibilità diversa. Se usi EC2 -Classic, specifichi le zone di disponibilità in modo esplicito. Amazon EMR identifica la migliore zona di disponibilità per avviare istanze in base alle specifiche del parco istanze. Il provisioning delle istanze viene sempre effettuato in una sola zona di disponibilità. Puoi selezionare sottoreti private o pubbliche, ma non puoi combinare i due tipi. Inoltre, le sottoreti che specifichi devono trovarsi all'interno dello stesso VPC.

Configurazione del nodo master

Poiché il parco istanze primarie è solo un'istanza singola, la sua configurazione è leggermente diversa dai parchi istanze core e attività. Per il parco istanze primarie, puoi selezionare solo on demand o spot perché è costituito da una sola istanza. Se utilizzi la console per creare il parco istanze, la capacità target per l'opzione di acquisto selezionata è impostata su 1. Se usi il AWS CLI, imposta sempre uno `TargetSpotCapacity` o su 1, `TargetOnDemandCapacity` a seconda dei casi. È comunque possibile scegliere fino a 5 tipi di istanza per il parco istanze primarie (o un massimo di 30 quando si utilizza l'opzione di allocazione delle istanze on demand o spot). Tuttavia, a differenza del parco istanze core e attività, in cui Amazon EMR può effettuare il provisioning di più istanze di diversi tipi, Amazon EMR seleziona un singolo tipo di istanza per effettuare il provisioning per il parco istanze primarie.

Strategia di allocazione per parchi istanze

Con Amazon EMR versione 5.12.1 e successive, puoi utilizzare la strategia di allocazione con istanze on demand e Spot per ogni nodo del cluster. Quando crei un cluster utilizzando l' AWS CLI API Amazon EMR o la console Amazon EMR con una strategia di allocazione, puoi specificare fino a 30 tipi di istanze Amazon EC2 per flotta. Con la configurazione predefinita del parco istanze del cluster Amazon EMR, puoi avere fino a 5 tipi di istanze per parco istanze. Si consiglia di utilizzare la strategia di allocazione per un provisioning dei cluster più rapido, un'allocazione più accurata delle istanze Spot e un minor numero di interruzioni delle istanze Spot.

Argomenti

- Strategia di allocazione con istanze On-Demand
- Strategia di allocazione con istanze Spot
- Autorizzazioni della strategia di allocazione
- Autorizzazioni IAM necessarie per una strategia di allocazione

Strategia di allocazione con istanze On-Demand

Le seguenti strategie di allocazione sono disponibili per le tue istanze On-Demand:

`lowest-price`(impostazione predefinita)

La strategia di allocazione del prezzo più basso lancia istanze On-Demand dal pool con il prezzo più basso con capacità disponibile. Se il pool con il prezzo più basso non dispone di capacità disponibile, le istanze on demand provengono dal successivo pool con il prezzo più basso con capacità disponibile.

`prioritized`

La strategia di allocazione prioritaria consente di specificare un valore di priorità per ogni tipo di istanza del parco istanze. Amazon EMR lancia le tue istanze on demand con la massima priorità. Se utilizzi questa strategia, devi configurare la priorità per almeno un tipo di istanza. Se non configuri il valore di priorità per un tipo di istanza, Amazon EMR assegna la priorità più bassa a quel tipo di istanza. Ogni flotta di istanze (principale, principale o task) in un cluster può avere un valore di priorità diverso per un determinato tipo di istanza.

> **Note**
>
> Se utilizzi la strategia di allocazione capacity-optimized-prioritizedSpot, Amazon EMR applica le stesse priorità sia alle istanze On-Demand che alle istanze Spot quando stabilisci le priorità.

Strategia di allocazione con istanze Spot

Per Istanze Spot puoi scegliere una delle seguenti strategie di allocazione:

`price-capacity-optimized` (consigliato)

La strategia di allocazione ottimizzata per prezzo e capacità avvia le istanze Spot dai pool di istanze Spot che hanno la capacità più elevata disponibile e il prezzo più basso per il numero di istanze in fase di avvio. Di conseguenza, la strategia ottimizzata tra prezzo e capacità in genere ha maggiori possibilità di ottenere capacità Spot e offre tassi di interruzione inferiori. Questa è la strategia predefinita per le versioni 6.10.0 e successive di Amazon EMR.

`capacity-optimized`

La strategia di allocazione ottimizzata per capacità avvia le istanze Spot nei pool più disponibili con le minori possibilità di interruzione nel breve termine. Questa è una buona opzione per i carichi di lavoro che potrebbero avere un costo di interruzione più elevato associato al riavvio del lavoro. Questa è la strategia predefinita per Amazon EMR rilasci 6.9.0 e precedenti.

`capacity-optimized-prioritized`

La strategia di capacity-optimized-prioritized allocazione consente di specificare un valore di priorità per ogni tipo di istanza nella flotta di istanze. Amazon EMR ottimizza innanzitutto la capacità, ma rispetta le priorità del tipo di istanza al massimo, ad esempio se la priorità non influisce in modo significativo sulla capacità della flotta di fornire una capacità ottimale. Consigliamo questa opzione se hai carichi di lavoro che devono subire interruzioni minime e che richiedono comunque determinati tipi di istanze. Se utilizzi questa strategia, devi configurare la priorità per almeno un tipo di istanza. Se non configuri una priorità per nessun tipo di istanza, Amazon EMR assegna il valore di priorità più basso a quel tipo di istanza. Ogni flotta di istanze (principale, principale o task) in un cluster può avere un valore di priorità diverso per un determinato tipo di istanza.

> **Note**
>
> Se utilizzi la strategia di allocazione On-Demand con priorità, Amazon EMR applica lo stesso valore di priorità sia alle istanze On-Demand che alle istanze Spot quando stabilisci le priorità.

`diversified`

Con la strategia di allocazione diversificata, Amazon EC2 distribuisce le istanze Spot in tutti i pool di capacità Spot.

`lowest-price`

La strategia di allocazione del prezzo più basso avvia le istanze Spot dal pool con il prezzo più basso con capacità disponibile. Se il pool con il prezzo più basso non ha capacità disponibile, le istanze Spot provengono dal successivo pool con il prezzo più basso che ha capacità disponibile. Se un pool esaurisce la capacità prima che soddisfi la capacità richiesta, la EC2 flotta Amazon attinge dal pool successivo con il prezzo più basso per continuare a soddisfare la tua richiesta. Per accertarti che la capacità desiderata sia soddisfatta, potresti ricevere istanze spot da vari

pool. Poiché questa strategia considera solo il prezzo dell'istanza e non la capacità disponibile, potrebbe comportare tassi di interruzione elevati.

Autorizzazioni della strategia di allocazione

L'opzione della strategia di allocazione richiede diverse autorizzazioni IAM incluse automaticamente nel ruolo di servizio Amazon EMR predefinito e nella policy gestita da Amazon EMR (`EMR_DefaultRole` e `AmazonEMRServicePolicy_v2`). Se si utilizza un ruolo di servizio o una policy gestita personalizzati per il cluster, è necessario aggiungere queste autorizzazioni prima di creare il cluster. Per ulteriori informazioni, consulta Autorizzazioni della strategia di allocazione.

Le prenotazioni opzionali di capacità On-Demand (ODCRs) sono disponibili quando utilizzi l'opzione della strategia di allocazione On-Demand. Le opzioni di prenotazione della capacità ti consentono di specificare una preferenza per utilizzare prima la capacità riservata per i cluster Amazon EMR. Puoi utilizzarla per assicurarti che i tuoi carichi di lavoro critici utilizzino la capacità che hai già riservato utilizzando open o targeting. ODCRs Per i carichi di lavoro non critici, le preferenze di prenotazione della capacità consentono di specificare se utilizzare la capacità riservata.

Le prenotazioni della capacità possono essere utilizzate solo da istanze che dispongono di attributi corrispondenti (tipo di istanza, piattaforma e zona di disponibilità). Per impostazione predefinita, le prenotazioni di capacità aperta vengono automaticamente utilizzate da Amazon EMR durante il provisioning di istanze on demand che corrispondono agli attributi dell'istanza. Se non disponi di istanze in esecuzione che corrispondono agli attributi delle prenotazioni di capacità, queste rimangono inutilizzate finché non avvii un'istanza con attributi corrispondenti. Se non si desidera utilizzare alcuna prenotazione di capacità durante l'avvio del cluster, è necessario impostare la preferenza di prenotazione di capacità su none (nessuna) nelle opzioni di avvio.

Tuttavia, puoi anche utilizzare una prenotazione di capacità per carichi di lavoro specifici. In questo modo puoi controllare in modo esplicito quali istanze possono essere eseguite in quella capacità riservata. Per ulteriori informazioni sulle prenotazioni di capacità On-Demand, consulta Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR.

Autorizzazioni IAM necessarie per una strategia di allocazione

Il Ruolo di servizio per Amazon EMR (ruolo EMR) richiede autorizzazioni aggiuntive per creare un cluster che utilizzi la strategia di allocazione per parchi istanze on demand o Spot.

Queste autorizzazioni sono incluse automaticamente nel ruolo di servizio Amazon EMR `EMR_DefaultRole` e nella policy gestita da Amazon EMR `AmazonEMRServicePolicy_v2`.

Se utilizzi un ruolo di servizio o una policy gestita personalizzati per il cluster, devi aggiungere le autorizzazioni seguenti:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DeleteLaunchTemplate",
        "ec2:CreateLaunchTemplate",
        "ec2:DescribeLaunchTemplates",
        "ec2:CreateLaunchTemplateVersion",
        "ec2:CreateFleet"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Deletelaunchtemplate"
    }
  ]
}
```

Di seguito sono riportate le autorizzazioni del ruolo di servizio necessarie per creare un cluster che utilizza prenotazioni di capacità aperte o mirate. È necessario includere queste autorizzazioni oltre alle autorizzazioni necessarie per l'utilizzo della strategia di allocazione.

Example Documento della policy per le prenotazioni di capacità del ruolo di servizio

Per utilizzare le prenotazioni di capacità aperte, è necessario includere le seguenti autorizzazioni aggiuntive.

JSON

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

```
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeCapacityReservations",
        "ec2:DescribeLaunchTemplateVersions",
        "ec2:DeleteLaunchTemplateVersions"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describecapacityreservations"
    }
  ]
}
```

Example

Per utilizzare le prenotazioni di capacità mirate, è necessario includere le seguenti autorizzazioni aggiuntive.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeCapacityReservations",
        "ec2:DescribeLaunchTemplateVersions",
        "ec2:DeleteLaunchTemplateVersions",
        "resource-groups:ListGroupResources"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describecapacityreservations"
    }
  ]
}
```

Configurazione di parchi istanze per il cluster

Console

Per creare un cluster con flotte di istanze con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e scegli Crea cluster.
3. In Cluster configuration (Configurazione del cluster), scegli Instance fleets (Parchi istanze).
4. Per ogni Node group (Gruppo di nodi), seleziona Add instance type (Aggiungi tipo di istanza) e scegli fino a 5 tipi di istanze per i parchi istanze primarie e core e fino a 15 tipi di istanze per i parchi istanze attività. Amazon EMR può effettuare il provisioning di qualsiasi combinazione di questi tipi di istanze quando avvia il cluster.
5. In ogni tipo di gruppo di nodi, scegli il menu a discesa Actions (Operazioni) accanto a ciascuna istanza per modificare queste impostazioni:

   Aggiunta di volumi EBS

   Specifica i volumi EBS da collegare al tipo di istanza dopo che Amazon EMR ne effettua il provisioning.

   Modifica della capacità ponderata

   Per il gruppo di nodi core, modifica questo valore con un numero qualsiasi di unità adatto alle tue applicazioni. Il numero di vCore YARN per ogni tipo di istanza del parco istanze viene utilizzato come relativa capacità ponderata predefinita. Non è possibile modificare la capacità ponderata del nodo primario.

   Modifica del prezzo spot massimo

   Per ogni tipo di istanza all'interno di un parco, specifica un prezzo spot massimo. Puoi impostare questo prezzo come una percentuale del prezzo on demand o come un importo in dollari specifico. Se il prezzo spot corrente in una zona di disponibilità è inferiore al prezzo spot massimo, Amazon EMR effettua il provisioning delle istanze spot. L'utente paga il prezzo Spot, non necessariamente il prezzo Spot massimo.
6. Facoltativamente, per aggiungere gruppi di sicurezza per i tuoi nodi, espandi i gruppi EC2 di sicurezza (firewall) nella sezione Rete e seleziona il gruppo di sicurezza per ogni tipo di nodo.

7. Facoltativamente, seleziona la casella di controllo accanto a Applica strategia di allocazione se desideri utilizzare tale opzione e seleziona la strategia di allocazione che desideri specificare per le istanze Spot. Se il tuo ruolo di servizio Amazon EMR non dispone delle autorizzazioni richieste, è consigliabile non selezionare questa opzione. Per ulteriori informazioni, consulta Strategia di allocazione per parchi istanze.
8. Scegli qualsiasi altra opzione applicabile al cluster.
9. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per creare e avviare un cluster con flotte di istanze con AWS CLI, segui queste linee guida:

- Per creare e avviare un cluster con parchi istanze, utilizza il comando `create-cluster` insieme ai parametri `--instance-fleet`.
- Per ottenere i dettagli sulla configurazione dei parchi istanze in un cluster, utilizza il comando `list-instance-fleets`.
- Per aggiungere più Amazon Linux personalizzati AMIs a un cluster che stai creando, usa l'`CustomAmiId`opzione con ogni `InstanceType` specifica. Puoi configurare nodi del parco istanze con più tipi di istanze e più nodi personalizzati AMIs in base alle tue esigenze. Consultare Esempi: creazione di un cluster con la configurazione di parchi istanze.
- Per apportare modifiche alla capacità target per un parco istanze, utilizza il comando `modify-instance-fleet`.
- Per aggiungere un parco istanze dell'attività a un cluster che non ne possiede già uno, utilizza il comando `add-instance-fleet`.
- AMIs È possibile aggiungere più istanze personalizzate al parco istanze di attività utilizzando l'CustomAmiId argomento con il add-instance-fleet comando. Consultare Esempi: creazione di un cluster con la configurazione di parchi istanze.
- Per utilizzare l'opzione di strategia di allocazione durante la creazione di un parco istanze, aggiornare il ruolo del servizio per includere il documento del policy di esempio nella sezione seguente.
- Per utilizzare le opzioni di riserva di capacità durante la creazione di un parco istanze con strategia di allocazione On-Demand, aggiornare il ruolo del servizio per includere il documento del policy di esempio nella sezione seguente.
- I parchi istanze vengono inclusi automaticamente nel ruolo di servizio EMR e nella policy gestita da Amazon EMR predefiniti (`EMR_DefaultRole` e `AmazonEMRServicePolicy_v2`). Se

si utilizza un ruolo di servizio personalizzato o una policy gestita personalizzata per il cluster, è necessario aggiungere le nuove autorizzazioni per la strategia di allocazione nella sezione seguente.

Esempi: creazione di un cluster con la configurazione di parchi istanze

Negli esempi seguenti vengono illustrati i comandi `create-cluster` con opzioni diverse che è possibile combinare.

> **Note**
>
> Se in precedenza non hai creato il ruolo e il profilo di EC2 istanza del servizio Amazon EMR predefiniti, usali `aws emr create-default-roles` per crearli prima di utilizzare il `create-cluster` comando.

Example Esempio: primaria on demand, core on demand con tipo di istanza singola, VPC predefinito

```
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 Esempio: primaria spot, core spot con tipo di istanza singola, VPC predefinito

```
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 Esempio: core misto primario on-demand con tipo di istanza singola, sottorete singola 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 Esempio: sistema primario On-Demand, core spot con più tipi di istanza ponderati, Timeout for Spot, Range of Subnet 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 Esempio: istanze primarie on demand, core miste e task con più tipi di istanze ponderate, timeout per le istanze Spot principali, intervallo di sottoreti 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 Esempio: primaria spot, nessuna core o attività, configurazione Amazon EBS, VPC predefinito

```
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 Esempio: più tipi di istanze personalizzate AMIs, più tipi di istanza, primarie on-demand, core on-demand

```
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 Esempio: aggiungere un nodo di attività a un cluster in esecuzione con più tipi di istanze e più istanze personalizzate 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 Esempio: utilizzo di un file di configurazione JSON

Puoi configurare i parametri del parco istanze in un file JSON, quindi fare riferimento al file JSON come l'unico parametro per i parchi istanze. Ad esempio, il comando seguente fa riferimento a un file di configurazione 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
```

Il *my-fleet-config.json* file specifica le flotte di istanze primarie, principali e task, come illustrato nell'esempio seguente. Il parco istanze principali utilizza un prezzo Spot massimo (`BidPrice`) come percentuale di On-Demand, mentre i parchi di istanze task e primarie utilizzano un prezzo Spot massimo (BidPriceAsPercentageofOnDemandPrice) come stringa in 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"
            }
        ]
    }
]
```

Modifica delle capacità target per un parco istanze

Per specificare nuove capacità target per un parco istanze, utilizza il comando `modify-instance-fleet`. Devi specificare l'ID cluster e l'ID del parco istanze. Utilizzate il `list-instance-fleets` comando per recuperare la flotta di istanze. IDs

```
aws emr modify-instance-fleet --cluster-id <cluster-id> \
  --instance-fleet \
    InstanceFleetId='<instance-fleet-id>',TargetOnDemandCapacity=1,TargetSpotCapacity=1
```

Aggiunta di un parco istanze attività a un cluster

Se un cluster contiene solo parchi istanze primarie e core, puoi utilizzare il comando `add-instance-fleet` per aggiungere un parco istanze attività. Puoi utilizzare questo comando solo per aggiungere parchi istanze dell'attività.

```
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}']
```

Ottenimento dei dettagli di configurazione per parchi istanze in un cluster

Per ottenere i dettagli di configurazione dei parchi istanze in un cluster, utilizza il comando `list-instance-fleets`. Il comando richiede un ID cluster come input. Nell'esempio seguente viene illustrato il comando e il relativo output per un cluster che contiene un gruppo di istanze attività primarie e un gruppo di istanze attività core. Per la sintassi di risposta completa, ListInstanceFleetsconsulta Amazon EMR API Reference.

```
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"
        }
    ]
}
```

Riconfigurazione delle flotte di istanze per il tuo cluster Amazon EMR

Con Amazon EMR versione 5.21.0 e successive, puoi riconfigurare le applicazioni cluster e specificare classificazioni di configurazione aggiuntive per ogni flotta di istanze in un cluster in esecuzione. A tale scopo, puoi utilizzare l'interfaccia a riga di AWS comando (AWS CLI) o l'SDK. AWS

È possibile tenere traccia dello stato di una flotta di istanze visualizzando gli CloudWatch eventi. Per ulteriori informazioni, consulta Eventi di riconfigurazione del parco istanze.

> **Note**
>
> È possibile sovrascrivere solo l'oggetto Configurazioni del cluster specificato durante la creazione del cluster. Per ulteriori informazioni sugli oggetti Configurations, consulta RunJobFlow la sintassi della richiesta. Se ci sono differenze tra la configurazione esistente e il file fornito, Amazon EMR reimposta le configurazioni modificate manualmente, ad esempio le configurazioni modificate durante la connessione al cluster tramite SSH, ai valori predefiniti del cluster per la flotta di istanze specificata.

Quando invii una richiesta di riconfigurazione utilizzando la console Amazon EMR, AWS l'interfaccia a riga di comando AWS CLI() o AWS l'SDK, Amazon EMR verifica il file di configurazione esistente sul cluster. Se ci sono differenze tra la configurazione esistente e il file fornito, Amazon EMR avvia azioni di riconfigurazione, riavvia alcune applicazioni e ripristina tutte le configurazioni modificate manualmente, ad esempio le configurazioni modificate durante la connessione al cluster tramite SSH, ai valori predefiniti del cluster per la flotta di istanze specificata.

Comportamenti di riconfigurazione

La riconfigurazione sovrascrive la configurazione sul cluster con il set di configurazione appena inviato e può sovrascrivere le modifiche alla configurazione apportate al di fuori dell'API di riconfigurazione.

Amazon EMR segue un processo continuo per riconfigurare le istanze nel parco istanze Task e Core. Solo una percentuale delle istanze per un singolo tipo di istanza viene modificata e riavviata alla volta. Se il tuo parco istanze ha più configurazioni di tipi di istanze diverse, queste verranno riconfigurate in parallelo.

Le riconfigurazioni vengono dichiarate a livello. InstanceTypeConfig Per un esempio visivo, fare riferimento a. Riconfigurazione di un parco di istanze È possibile inviare richieste di riconfigurazione che contengono impostazioni di configurazione aggiornate per uno o più tipi di istanza all'interno di una singola richiesta. È necessario includere tutti i tipi di istanze che fanno parte del parco istanze nella richiesta di modifica; tuttavia, i tipi di istanze con campi di configurazione compilati verranno riconfigurati, mentre le altre `InstanceTypeConfig` istanze del parco istanze rimarranno invariate. Una riconfigurazione è considerata riuscita solo quando tutte le istanze dei tipi di istanze specificati completano la riconfigurazione. Se una delle istanze non riesce a riconfigurarsi, l'intero parco istanze torna automaticamente all'ultima configurazione stabile conosciuta.

Limitazioni

Quando riconfigurate un parco istanze in un cluster in esecuzione, considerate le seguenti limitazioni:

- Le applicazioni non YARN possono avere esito negativo durante il riavvio o causare problemi al cluster, soprattutto se le applicazioni non sono configurate correttamente. I cluster che si avvicinano al massimo utilizzo della memoria e della CPU potrebbero essere interessati da problemi dopo il processo di riavvio. Ciò è particolarmente vero per la flotta di istanze principali. Consulta la Risolvi i problemi di riconfigurazione del parco istanze sezione.
- Le operazioni di ridimensionamento e riconfigurazione non avvengono in parallelo. Le richieste di riconfigurazione attenderanno un ridimensionamento continuo e viceversa.
- Le operazioni di ridimensionamento e riconfigurazione non avvengono in parallelo. Le richieste di riconfigurazione attenderanno un ridimensionamento continuo e viceversa.
- Dopo aver riconfigurato una flotta di istanze, Amazon EMR riavvia le applicazioni per consentire alle nuove configurazioni di avere effetto. Errori del processo o altri comportamenti dell'applicazione imprevisti possono verificarsi se le applicazioni sono in uso durante la riconfigurazione.

- Se una riconfigurazione per qualsiasi tipo di configurazione di istanze in un parco istanze fallisce, Amazon EMR inverte i parametri di configurazione alla versione funzionante precedente per l'intero parco istanze, oltre a emettere eventi e aggiornare i dettagli sullo stato. Se anche il processo di reversione fallisce, devi inviare una nuova `ModifyInstanceFleet` richiesta per ripristinare la flotta di istanze dallo stato. `ARRESTED` Gli errori di reversione provocano eventi di riconfigurazione del parco istanze e cambiamenti di stato.
- Le richieste di riconfigurazione per le classificazioni di configurazione Phoenix sono supportate solo in Amazon EMR versione 5.23.0 e successive e non in Amazon EMR versione 5.21.0 o 5.22.0.
- Le richieste di HBase riconfigurazione per le classificazioni di configurazione sono supportate solo in Amazon EMR versione 5.30.0 e successive e non sono supportate nelle versioni di Amazon EMR da 5.23.0 a 5.29.0.
- La riconfigurazione della classificazione hdfs-encryption-zones o di una qualsiasi classificazione di configurazione Hadoop KMS non è supportata su un cluster Amazon EMR con più nodi primari.
- Amazon EMR attualmente non supporta alcune richieste di riconfigurazione per lo scheduler di capacità YARN che richiedono il riavvio di YARN. ResourceManager Ad esempio, non è possibile rimuovere completamente una coda.
- Quando YARN deve essere riavviato, tutti i job YARN in esecuzione vengono in genere terminati e persi. Ciò potrebbe causare ritardi nell'elaborazione dei dati. Per eseguire i job YARN durante un riavvio di YARN, puoi creare un cluster Amazon EMR con più nodi primari o impostare yarn.resourcemanager.recovery.enabled su nella tua classificazione di configurazione yarn-site. `true` Per ulteriori informazioni sull'utilizzo di più nodi master, consulta YARN ad alta disponibilità. ResourceManager

Riconfigurazione di un parco di istanze

Using the AWS CLI

Utilizza il `modify-instance-fleet` comando per specificare una nuova configurazione per un parco istanze in un cluster in esecuzione.

> **Note**
>
> Negli esempi seguenti, sostituisci j-2 AL4 XXXXXX5 T9 con il tuo ID del cluster e sostituisci if-1xxxxxxx9 con l'ID del tuo parco istanze.

Esempio: sostituisci una configurazione per un parco di istanze

> **Warning**
>
> Specificate tutti `InstanceTypeConfig` i campi che avete usato al momento del lancio. La mancata inclusione dei campi può comportare la sovrascrittura delle specifiche dichiarate al momento del lancio. InstanceTypeConfigPer un elenco, fare riferimento a.

L'esempio seguente fa riferimento a un file JSON di configurazione chiamato InstanceFleet.json per modificare la proprietà del controllore dello stato del disco YARN NodeManager per un parco di istanze.

Modifica della flotta di istanze (JSON)

1. Prepara la classificazione della configurazione e salvala come InstanceFleet.json nella stessa directory in cui eseguirai il 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. Esegui il comando seguente.

```
aws emr modify-instance-fleet \
--cluster-id j-2AL4XXXXXX5T9 \
--region us-west-2 \
--instance-fleet instanceFleet.json
```

Esempio: aggiungi una configurazione a un parco istanze

Se desideri aggiungere una configurazione a un tipo di istanza, devi includere tutte le configurazioni specificate in precedenza per quel tipo di istanza nella nuova `ModifyInstanceFleet` richiesta. In caso contrario, le configurazioni specificate in precedenza vengono rimosse.

L'esempio seguente aggiunge una proprietà per il controllo della memoria NodeManager virtuale YARN. La configurazione include anche i valori specificati in precedenza per il controllo dello stato del NodeManager disco YARN in modo che i valori non vengano sovrascritti.

1. Prepara i seguenti contenuti in InstanceFleet.json e salvalo nella stessa directory in cui eseguirai il 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. Esegui il comando seguente.

```
aws emr modify-instance-fleet \
--cluster-id j-2AL4XXXXXX5T9 \
--region us-west-2 \
--instance-fleet instanceFleet.json
```

using the Java SDK

> **Note**
>
> Negli esempi seguenti, sostituisci j-2 AL4 XXXXXX5 T9 con il tuo ID del cluster e sostituisci if-1xxxxxxx9 con l'ID del tuo parco istanze.

Il seguente frammento di codice fornisce una nuova configurazione per un parco di istanze utilizzando l' 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);
```

Risolvi i problemi di riconfigurazione del parco istanze

Se il processo di riconfigurazione per qualsiasi tipo di istanza all'interno di un parco istanze fallisce, Amazon EMR ripristina la riconfigurazione in corso e registra un messaggio di errore utilizzando un evento Events. AAmazon CloudWatch L'evento fornisce un breve riepilogo dell'errore di

riconfigurazione. Elenca le istanze per le quali la riconfigurazione non è riuscita e i messaggi di errore corrispondenti. Di seguito è riportato un esempio di messaggio di errore.

```
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"...
```

Per accedere ai log di provisioning dei nodi

Utilizza SSH per connetterti al nodo sul quale la riconfigurazione non è riuscita. Per istruzioni, consulta Connect alla tua istanza Linux in Amazon Elastic Compute Cloud.

Accessing logs by connecting to a node

1. Passa alla directory seguente, che contiene i file di log di provisioning dei nodi.

   ```
   /mnt/var/log/provision-node/
   ```

2. Apri la sottodirectory dei report e cerca il report di provisioning dei nodi per la riconfigurazione. La directory dei report organizza i log in base al numero di versione di riconfigurazione, all'identificatore univoco universale (UUID), all'indirizzo IP dell'istanza Amazon EC2 e al timestamp. Ogni report è un file YAML compresso che contiene informazioni dettagliate sul processo di riconfigurazione. Di seguito è riportato un esempio di nome del file di report e relativo percorso.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
   xxx.ec2.internal/202104061715.yaml.gz
   ```

3. Puoi esaminare un report utilizzando un visualizzatore di file come zless, come nell'esempio seguente.

   ```
   zless 202104061715.yaml.gz
   ```

Accessing logs using Amazon S3

Accedi a AWS Management Console e apri la console Amazon S3 all'indirizzo. https://console.aws.amazon.com/s3/ Apri il bucket Amazon S3 specificato quando il cluster è stato configurato per archiviare i file di log.

1. Passa alla cartella seguente, che contiene i file di log di provisioning dei nodi:

```
amzn-s3-demo-bucket/elasticmapreduce/cluster id/node/instance id/provision-node/
```

2. Apri la cartella dei report e cerca il report di provisioning dei nodi per la riconfigurazione. La cartella dei report organizza i log in base al numero di versione di riconfigurazione, all'identificatore univoco universale (UUID), all'indirizzo IP dell'istanza Amazon EC2 e al timestamp. Ogni report è un file YAML compresso che contiene informazioni dettagliate sul processo di riconfigurazione. Di seguito è riportato un esempio di nome del file di report e relativo percorso.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
xxx.ec2.internal/202104061715.yaml.gz
```

Per visualizzare un file di log, puoi scaricarlo da Amazon S3 sul tuo computer come file di testo. Per istruzioni, consulta Download di un oggetto.

Ogni file di log contiene un report di provisioning dettagliato per la riconfigurazione associata. Per trovare informazioni sui messaggi di errore, cerca il livello di log `err` di un report. Il formato del report dipende dalla versione di Amazon EMR sul cluster. L'esempio seguente mostra le informazioni sugli errori per le versioni 5.32.0 e 6.2.0 di Amazon EMR e le versioni successive utilizzano il seguente formato:

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR

Per avviare parchi istanze on demand con opzioni di prenotazione della capacità, allega autorizzazioni aggiuntive per il ruolo di servizio necessarie per utilizzare le opzioni di prenotazione della capacità. Poiché le opzioni di prenotazione della capacità devono essere utilizzate insieme alla strategia di allocazione on demand, è necessario includere anche le autorizzazioni necessarie per la strategia di allocazione nel ruolo di servizio e nella policy gestita. Per ulteriori informazioni, consulta Autorizzazioni della strategia di allocazione.

Amazon EMR supporta prenotazioni della capacità aperte e mirate. Negli argomenti seguenti vengono illustrate le configurazioni dei gruppi di istanze che è possibile utilizzare con l'operazione `RunJobFlow` o il comando `create-cluster` per avviare parchi istanze mediante prenotazioni della capacità on demand.

Utilizzo delle prenotazioni della capacità aperte in base al miglior tentativo

Se le istanze on demand del cluster corrispondono agli attributi delle prenotazioni di capacità aperta (tipo di istanza, piattaforma, tenancy e zona di disponibilità) disponibili nell'account, le prenotazioni di capacità vengono applicate automaticamente. Tuttavia, non è garantito che le prenotazioni di capacità verranno utilizzate. Per il provisioning del cluster, Amazon EMR valuta tutti i pool di istanze specificati nella richiesta di avvio e utilizza quello con il prezzo più basso che dispone di capacità sufficienti per avviare tutti i nodi principali richiesti. Le prenotazioni della capacità aperte disponibili che corrispondono al pool di istanze vengono applicate automaticamente. Se le prenotazioni della capacità aperte disponibili non corrispondono al pool di istanze, rimangono inutilizzate.

Una volta eseguito il provisioning dei nodi principali, la zona di disponibilità viene selezionata e risolta. Amazon EMR effettua il provisioning dei nodi attività in pool di istanze, a partire da quelli con il prezzo più basso, nella zona di disponibilità selezionata fino a quando non viene eseguito il provisioning di tutti i nodi attività. Le prenotazioni della capacità aperte disponibili che corrispondono al pool di istanze vengono applicate automaticamente.

Di seguito sono riportati i casi d'uso della logica di allocazione della capacità di Amazon EMR per l'utilizzo delle prenotazioni della capacità aperte in base al miglior tentativo.

Esempio 1: il pool di istanze a prezzo più basso nella richiesta di avvio dispone di prenotazioni della capacità aperte disponibili

In questo caso, Amazon EMR avvia la capacità nel pool di istanze a prezzo più basso con istanze on demand. Le prenotazioni della capacità aperte disponibili in quel pool di istanze vengono utilizzate automaticamente.

Strategia on-demand	lowest-price		
Capacità richiesta	100		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
Prenotazioni della capacità aperte disponibili	150	100	100
Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	100	-	-
Prenotazione della capacità aperta utilizzata	100	-	-
Prenotazioni della capacità aperte disponibili	50	100	100

Dopo l'avvio del parco istanze, puoi eseguire `describe-capacity-reservations` per vedere quante prenotazioni di capacità sono rimaste inutilizzate.

Esempio 2: il pool di istanze a prezzo più basso nella richiesta di avvio non dispone di prenotazioni della capacità aperte disponibili

In questo caso, Amazon EMR avvia la capacità nel pool di istanze a prezzo più basso con istanze on demand. Tuttavia, le prenotazioni della capacità aperte rimangono inutilizzate.

Strategia su richiesta	lowest-price		
Capacità richiesta	100		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
	-	-	100

Prenotazioni della capacità aperte disponibili			
Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	100	-	-
Prenotazione della capacità aperta utilizzata	-	-	-
Prenotazioni della capacità aperte disponibili	-	-	100

Configurazione dei parchi istanze per utilizzare le prenotazioni della capacità aperte in base al miglior tentativo

Quando utilizzi l'operazione `RunJobFlow` per creare un cluster basato sul parco istanze, imposta la strategia di allocazione on demand `lowest-price` e `CapacityReservationPreference` per le opzioni di prenotazione di capacità su `open`. In alternativa, se lasci vuoto questo campo, Amazon EMR configura come impostazione predefinita la preferenza di prenotazione della capacità dell'istanza on demand su `open`.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "CapacityReservationPreference": "open"
         }
        }
    }
```

È inoltre possibile utilizzare l'interfaccia CLI di Amazon EMR per creare un cluster basato sul parco istanze utilizzando prenotazioni della capacità aperte.

```
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}}'}
```

Dove,

- `open-ODCR-cluster` viene sostituito con il nome del cluster che utilizza le prenotazioni della capacità aperte.
- `subnet-22XXXX01` viene sostituito con l'ID della sottorete.

Usa prima le prenotazioni della capacità aperte

Puoi scegliere di sovrascrivere la strategia di allocazione del prezzo più basso e assegnare priorità utilizzando le prenotazioni della capacità aperte disponibili prima durante il provisioning di un cluster Amazon EMR. In questo caso, Amazon EMR valuta tutti i pool di istanze con prenotazioni della capacità specificati nella richiesta di avvio e utilizza quello con il prezzo più basso che dispone di capacità sufficienti per avviare tutti i nodi principali richiesti. Se nessuno dei pool di istanze con prenotazioni della capacità dispone di capacità sufficiente per i nodi principali richiesti, Amazon EMR ritorna al caso sulla base del miglior tentativo descritto nell'argomento precedente. In altre parole, Amazon EMR valuta nuovamente tutti i pool di istanze specificati nella richiesta di avvio e utilizza quello con il prezzo più basso che dispone di capacità sufficiente per avviare tutti i nodi principali richiesti. Le prenotazioni della capacità aperte disponibili che corrispondono al pool di istanze vengono applicate automaticamente. Se le prenotazioni della capacità aperte disponibili non corrispondono al pool di istanze, rimangono inutilizzate.

Una volta eseguito il provisioning dei nodi principali, la zona di disponibilità viene selezionata e risolta. Amazon EMR effettua il provisioning dei nodi attività in pool di istanze con prenotazioni della capacità, a partire da quelli con il prezzo più basso, nella zona di disponibilità selezionata fino a quando non viene eseguito il provisioning di tutti i nodi attività. Amazon EMR utilizza prima

le prenotazioni della capacità aperte disponibili per ogni pool di istanze nella zona di disponibilità selezionata e, solo se necessario, utilizza la strategia di prezzo più basso per eseguire il provisioning di eventuali nodi attività rimanenti.

Di seguito sono riportati i casi d'uso della logica di allocazione della capacità di Amazon EMR per l'utilizzo delle prenotazioni della capacità aperte per prime.

Esempio 1: il pool di istanze con prenotazioni della capacità aperte disponibili nella richiesta di avvio ha una capacità sufficiente per i nodi principali

In questo caso, Amazon EMR avvia la capacità nel pool di istanze con prenotazioni della capacità aperte disponibili indipendentemente dal prezzo del pool di istanze. Di conseguenza, le prenotazioni della capacità aperte vengono utilizzate quando possibile, fino a quando non viene eseguito il provisioning di tutti i nodi principali.

Strategia su richiesta	lowest-price		
Capacità richiesta	100		
Strategia di utilizzo	use-capacity-reservations-first		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
Prenotazioni della capacità aperte disponibili	-	-	150
Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	-	-	100
Prenotazione della capacità aperta utilizzata	-	-	100
Prenotazioni della capacità aperte disponibili	-	-	50

Esempio 2: il pool di istanze con prenotazioni della capacità aperte disponibili nella richiesta di avvio non dispone di capacità sufficiente per i nodi principali

In questo caso, Amazon EMR torna ad avviare nodi principali utilizzando una strategia di prezzo più basso in base al miglior tentativo per utilizzare le prenotazioni di capacità.

Strategia su richiesta	lowest-price		
Capacità richiesta	100		
Strategia di utilizzo	use-capacity-reservations-first		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
Prenotazioni della capacità aperte disponibili	10	50	50
Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	100	-	-
Prenotazione della capacità aperta utilizzata	10	-	-
Prenotazioni della capacità aperte disponibili	-	50	50

Dopo l'avvio del parco istanze, puoi eseguire `describe-capacity-reservations` per vedere quante prenotazioni di capacità sono rimaste inutilizzate.

Configurazione dei parchi istanze per utilizzare prima le prenotazioni della capacità aperte

Quando utilizzi l'operazione `RunJobFlow` per creare un cluster basato sul parco istanze, imposta la strategia di allocazione on-demand su `lowest-price` e `UsageStrategy` per `CapacityReservationOptions` su `use-capacity-reservations-first`.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "UsageStrategy": "use-capacity-reservations-first"
         }
       }
    }
```

È inoltre possibile utilizzare l'interfaccia CLI Amazon EMR per creare un cluster basato sul parco istanze utilizzando prima prenotazioni della capacità.

```
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}}'}
```

Dove,

- `use-CR-first-cluster` viene sostituito con il nome del cluster che utilizza le prenotazioni della capacità aperte.
- `subnet-22XXXX01` viene sostituito con l'ID della sottorete.

Usa prima le prenotazioni della capacità mirate

Quando si esegue il provisioning di un cluster Amazon EMR, è possibile scegliere di sostituire la strategia di allocazione a prezzo più basso e assegnare priorità utilizzando innanzitutto le prenotazioni della capacità mirate disponibili. In questo caso, Amazon EMR valuta tutti i pool di istanze con prenotazioni della capacità mirate specificate nella richiesta di avvio e seleziona

quello con il prezzo più basso che dispone di capacità sufficiente per avviare tutti i nodi principali richiesti. Se nessuno dei pool di istanze con prenotazioni della capacità mirate dispone di capacità sufficiente per i nodi principali, Amazon EMR ritorna al caso sulla base del miglior tentativo descritto in precedenza. In altre parole, Amazon EMR valuta nuovamente tutti i pool di istanze specificati nella richiesta di avvio e seleziona quello con il prezzo più basso che dispone di capacità sufficiente per avviare tutti i nodi principali richiesti. Le prenotazioni della capacità aperte disponibili che corrispondono al pool di istanze vengono applicate automaticamente. Tuttavia, le prenotazioni della capacità mirate rimangono inutilizzate.

Una volta eseguito il provisioning dei nodi principali, la zona di disponibilità viene selezionata e risolta. Amazon EMR effettua il provisioning dei nodi attività in pool di istanze con prenotazioni della capacità mirate, a partire da quelli con il prezzo più basso, nella zona di disponibilità selezionata fino a quando non viene eseguito il provisioning di tutti i nodi attività. Amazon EMR tenta di utilizzare prima le prenotazioni della capacità mirate disponibili in ogni pool di istanze nella zona di disponibilità selezionata. Quindi, solo se necessario, Amazon EMR utilizza la strategia di prezzo più basso per effettuare il provisioning di eventuali nodi attività rimanenti.

Di seguito sono riportati i casi d'uso della logica di allocazione della capacità di Amazon EMR per l'utilizzo delle prenotazioni della capacità mirate per prime.

Esempio 1: il pool di istanze con prenotazioni della capacità mirate disponibili nella richiesta di avvio ha una capacità sufficiente per i nodi principali

In questo caso, Amazon EMR avvia la capacità nel pool di istanze con prenotazioni della capacità mirate disponibili indipendentemente dal prezzo del pool di istanze. Di conseguenza, le prenotazioni della capacità mirate vengono utilizzate quando possibile, fino a quando non viene eseguito il provisioning di tutti i nodi principali.

Strategia su richiesta	lowest-price		
Strategia di utilizzo	use-capacity-reservations-first		
Capacità richiesta	100		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
Prenotazioni della capacità mirate disponibili	-	-	150

Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	-	-	100
Prenotazione della capacità mirata utilizzata	-	-	100
Prenotazioni della capacità mirate disponibili	-	-	50

Example Esempio 2: il pool di istanze con prenotazioni della capacità mirate disponibili nella richiesta di avvio non dispone di capacità sufficiente per i nodi principali

Strategia su richiesta	lowest-price		
Capacità richiesta	100		
Strategia di utilizzo	use-capacity-reservations-first		
Tipo di istanza	c5.xlarge	m5.xlarge	r5.xlarge
Prenotazioni della capacità mirate disponibili	10	50	50
Prezzo su richiesta	$	$$	$$$
Istanze con provisioning	100	-	-
Prenotazioni della capacità mirate utilizzate	10	-	-

Prenotazioni della capacità mirate disponibili	-	50	50

Dopo l'avvio del parco istanze, puoi eseguire `describe-capacity-reservations` per vedere quante prenotazioni di capacità sono rimaste inutilizzate.

Configurazione dei parchi istanze per utilizzare prima le prenotazioni della capacità mirate

Quando utilizzi l'operazione `RunJobFlow` per creare un cluster basato sul parco istanze, imposta la strategia di allocazione on demand su `lowest-price`, `UsageStrategy` per `CapacityReservationOptions` su `use-capacity-reservations-first` e `CapacityReservationResourceGroupArn` per `CapacityReservationOptions` su `<your resource group ARN>`. Per ulteriori informazioni, consulta Lavora con le prenotazioni di capacità nella Amazon EC2 User Guide.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "UsageStrategy": "use-capacity-reservations-first",
            "CapacityReservationResourceGroupArn": "arn:aws:resource-groups:sa-
east-1:123456789012:group/MyCRGroup"
         }
       }
    }
```

Dove `arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup` viene sostituito con l'ARN del gruppo di risorse.

È inoltre possibile utilizzare la CLI Amazon EMR per creare un cluster basato sul parco istanze utilizzando prima prenotazioni della capacità mirate.

```
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}}'}
```

Dove,

- `targeted-CR-cluster` viene sostituito con il nome del cluster che utilizza le prenotazioni della capacità mirate.
- `subnet-22XXXX01` viene sostituito con l'ID della sottorete.
- `arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup` viene sostituito con l'ARN del gruppo di risorse.

Come evitare l'utilizzo delle prenotazioni della capacità aperte disponibili

Example

Se desideri evitare di utilizzare una qualsiasi delle prenotazioni della capacità aperte durante l'avvio di un cluster Amazon EMR, imposta la strategia di allocazione on demand su `lowest-price` e `CapacityReservationPreference` per `CapacityReservationOptions` su `none`. In caso contrario, Amazon EMR imposta come impostazione predefinita la preferenza di prenotazione della capacità dell'istanza on demand su `open` e prova a utilizzare le prenotazioni della capacità aperte disponibili sulla base della migliore ipotesi.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "CapacityReservationPreference": "none"
         }
       }
    }
```

È inoltre possibile utilizzare l'interfaccia CLI di Amazon EMR per creare un cluster basato sul parco istanze senza utilizzare prenotazioni della capacità aperte.

```
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}}'}
```

Dove,

- `none-CR-cluster` viene sostituito con il nome del cluster che non utilizza le prenotazioni della capacità aperte.
- `subnet-22XXXX01` viene sostituito con l'ID della sottorete.

Scenari per l'utilizzo delle prenotazioni della capacità

È possibile utilizzare le prenotazioni della capacità nei seguenti scenari.

Scenario 1: Rotazione di un cluster a esecuzione prolungata utilizzando le prenotazioni della capacità

Quando si ruota un cluster a esecuzione prolungata, è possibile che siano richiesti requisiti rigorosi per i tipi di istanza e le zone di disponibilità per le nuove istanze di cui si esegue il provisioning. Con le prenotazioni della capacità, è possibile utilizzare la garanzia della capacità per completare la rotazione del cluster senza interruzioni.

Scenario 2: Provisioning di cluster di breve durata successivi utilizzando le prenotazioni della capacità

È possibile utilizzare le prenotazioni della capacità anche per eseguire il provisioning di un gruppo di cluster successivi e di breve durata per singoli carichi di lavoro in modo che quando si chiude un

cluster, il cluster successivo possa utilizzare le prenotazioni della capacità. È possibile utilizzare le prenotazioni della capacità mirate per assicurarsi che solo i cluster previsti utilizzino le prenotazioni della capacità.

Configura gruppi di istanze uniformi per il tuo cluster Amazon EMR

Con la configurazione dei gruppi di istanze, ogni tipo di nodo (master, principale o di task) è costituito dallo stesso tipo di istanza e dalla stessa opzione di acquisto per istanze: on demand o Spot. Queste impostazioni vengono specificate al momento della creazione di un gruppo di istanze. Le impostazioni non possono essere modificate in seguito. Tuttavia, puoi aggiungere istanze dello stesso tipo e opzioni di acquisto a gruppi di istanze principali e dell'attività Puoi anche rimuovere istanze.

Se le istanze on demand del cluster corrispondono agli attributi delle prenotazioni di capacità aperta (tipo di istanza, piattaforma, tenancy e zona di disponibilità) disponibili nell'account, le prenotazioni di capacità vengono applicate automaticamente. È possibile utilizzare prenotazioni di capacità aperte per nodi primari, core e attività. Tuttavia, non è possibile utilizzare prenotazioni della capacità mirate o impedire l'avvio di istanze in prenotazioni della capacità aperte con attributi corrispondenti quando si esegue il provisioning di cluster utilizzando gruppi di istanze. Se si desidera utilizzare prenotazioni della capacità mirate o impedire l'avvio di istanze in prenotazioni della capacità aperte, utilizza piuttosto parchi istanze. Per ulteriori informazioni, consulta Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR.

Per aggiungere tipi di istanze differenti dopo che un cluster è stato creato, puoi aggiungere gruppi di istanze di attività aggiuntivi. Puoi scegliere tipi di istanze e opzioni di acquisto differenti per ogni gruppo di istanze. Per ulteriori informazioni, consulta Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione.

Quando si avviano delle istanze, la preferenza della prenotazione di capacità dell'istanza on demand è impostata in modo predefinito su `open`, che consente di eseguirla in qualsiasi prenotazione di capacità aperta che abbia gli attributi corrispondenti (tipo di istanza, piattaforma, zona di disponibilità). Per informazioni sulla condivisione delle prenotazioni di capacità on demand, consulta Usa le prenotazioni di capacità con flotte di istanze in Amazon EMR.

In questa sezione viene descritta la creazione di un cluster con gruppi di istanze uniformi. Per ulteriori informazioni sulla modifica di un gruppo di istanze esistente aggiungendo o rimuovendo istanze manualmente o con scalabilità automatica, consulta Gestione dei cluster Amazon EMR.

Utilizzo della console per configurare gruppi di istanze uniformi

Console

Creazione di un cluster con gruppi di istanze con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e scegli Crea cluster.
3. In Cluster configuration (Configurazione del cluster), scegli Instance groups (Gruppi di istanze).
4. In Node groups (Gruppi di nodi) è presente una sezione per ogni tipo di gruppo di nodi. Per il gruppo di nodi primari, seleziona la casella di controllo Use multiple primary nodes (Usa più nodi primari) se desideri avere 3 nodi primari. Seleziona la casella di controllo Use Spot purchasing option (Utilizza l'opzione di acquisto spot) se desideri utilizzare gli acquisti spot.
5. Per i gruppi di nodi primari e core, seleziona Add instance type (Aggiungi tipo di istanza) e scegli fino a 5 tipi di istanza. Per il gruppo attività, seleziona Add instance type (Aggiungi tipo di istanza) e scegli fino a 15 tipi di istanza. Amazon EMR può effettuare il provisioning di qualsiasi combinazione di questi tipi di istanze quando avvia il cluster.
6. In ogni tipo di gruppo di nodi, scegli il menu a discesa Actions (Operazioni) accanto a ciascuna istanza per modificare queste impostazioni:

   Aggiunta di volumi EBS

   Specifica i volumi EBS da collegare al tipo di istanza dopo che Amazon EMR ne effettua il provisioning.

   Modifica del prezzo spot massimo

   Per ogni tipo di istanza all'interno di un parco, specifica un prezzo spot massimo. Puoi impostare questo prezzo come una percentuale del prezzo on demand o come un importo in dollari specifico. Se il prezzo spot corrente in una zona di disponibilità è inferiore al prezzo spot massimo, Amazon EMR effettua il provisioning delle istanze spot. L'utente paga il prezzo Spot, non necessariamente il prezzo Spot massimo.
7. Facoltativamente, espandi Configurazione nodo per inserire una configurazione JSON o per caricare JSON da Amazon S3.
8. Scegli qualsiasi altra opzione applicabile al cluster.
9. Per avviare il cluster, scegli Create cluster (Crea cluster).

Usa AWS CLI per creare un cluster con gruppi di istanze uniformi

Per specificare la configurazione di gruppi di istanze per un cluster utilizzando AWS CLI, utilizza il comando `create-cluster` insieme al parametro `--instance-groups`. Amazon EMR presuppone l'opzione On-Demand Instance (Istanza on demand) a meno che non venga specificato l'argomento `BidPrice` per un gruppo di istanze. Per esempi di comandi `create-cluster` che avviano gruppi di istanze uniformi con istanze on demand e opzioni di cluster diverse, digita `aws emr create-cluster help` dalla riga di comando oppure consulta create-cluster nella Guida di riferimento ai comandi della AWS CLI .

È possibile utilizzare il AWS CLI per creare gruppi di istanze uniformi in un cluster che utilizza istanze Spot. Il prezzo Spot offerto dipende dalla zona di disponibilità. Quando si utilizza la CLI o l'API, è possibile specificare la zona di disponibilità con l'`AvailabilityZone`argomento (se si utilizza una rete EC2 -classic) o l'`SubnetID` argomento del parametro. `--ec2-attributes` La zona di disponibilità o la sottorete selezionata vale per il cluster, pertanto viene utilizzata per tutti i gruppi di istanze. Se non specifichi una zona di disponibilità o sottorete in maniera esplicita, Amazon EMR seleziona la zona di disponibilità con il prezzo Spot più basso quando avvia il cluster.

Nel seguente esempio viene illustrato un comando `create-cluster` che crea un gruppo di istanze primarie, uno di core e due di attività, tutti i quali utilizzano istanze spot. Sostituiscilo *myKey* con il nome della tua coppia di EC2 chiavi Amazon.

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuovili o sostituiscili con un accento circonflesso (^).

```
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
```

Utilizzando la CLI, è possibile creare cluster di gruppi di istanze uniformi che specificano un'AMI personalizzata univoca per ogni tipo di istanza nel gruppo di istanze. Ciò consente di utilizzare architetture di istanze diverse nello stesso gruppo di istanze. Ogni tipo di istanza deve utilizzare un'AMI personalizzata con un'architettura corrispondente. Ad esempio, si dovrebbe configurare un tipo di istanza m5.xlarge con un'AMI personalizzata dell'architettura x86_64 e un tipo di istanza m6g.xlarge con una corrispondente (ARM) architettura AMI personalizzata `AWS AARCH64`.

L'esempio seguente mostra un cluster di gruppi di istanze uniforme creato con due tipi di istanza, ognuno con una propria AMI personalizzata. Nota che le personalizzazioni AMIs sono specificate solo a livello di tipo di istanza, non a livello di cluster. Questo serve a evitare conflitti tra il tipo di istanza AMIs e un'AMI a livello di cluster, che potrebbero causare il fallimento dell'avvio del cluster.

```
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
```

È possibile aggiungere più elementi personalizzati AMIs a un gruppo di istanze da aggiungere a un cluster in esecuzione. L'argomento `CustomAmiId` può essere utilizzato con il comando `add-instance-groups` come mostrato nell'esempio seguente.

```
aws emr add-instance-groups --cluster-id j-123456 \
  --instance-groups \

 InstanceGroupType=Task,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
```

Utilizzo di Java SDK per creare un gruppo di istanze

Crea un'istanza di un oggetto `InstanceGroupConfig` che specifica la configurazione di un gruppo di istanze per un cluster. Per utilizzare istanze Spot, imposta le proprietà `withBidPrice` e `withMarket` sull'oggetto `InstanceGroupConfig`. Nel codice seguente viene mostrato come definire gruppi di istanze primarie, core e attività che eseguono istanze 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");
```

Flessibilità della zona di disponibilità per un cluster Amazon EMR

Ciascuna Regione AWS ha più sedi isolate note come zone di disponibilità. Quando avvii un'istanza, puoi decidere di specificare una zona di disponibilità (AZ) nella Regione AWS utilizzata. La flessibilità delle zone di disponibilità è la distribuzione delle istanze su più AZs istanze. Se un'istanza fallisce, è possibile progettare l'applicazione in modo che un'istanza in un'altra zona dAZ possa gestire le richieste. Per ulteriori informazioni sulle zone di disponibilità, consulta la documentazione relativa alla regione e alle zone nella Amazon EC2 User Guide.

La flessibilità delle istanze consiste nell'utilizzare più tipi di istanze per soddisfare i requisiti di capacità. Quando esprimi la flessibilità con le istanze, puoi utilizzare la capacità aggregata attingendo da dimensioni, famiglie e generazioni di istanze diverse. Una maggiore flessibilità consente di migliorare la possibilità di trovare e allocare la quantità di capacità di elaborazione richiesta rispetto a un cluster che utilizza un singolo tipo di istanza.

La flessibilità delle istanze e delle zone di disponibilità riduce gli errori di capacità insufficiente (ICE) e le interruzioni Spot rispetto a un cluster con un singolo tipo di istanza o AZ. Utilizza le best practice qui descritte per determinare quali istanze diversificare dopo avere stabilito la famiglia e le dimensioni iniziali delle istanze. Questo approccio massimizza la disponibilità dei pool di EC2 capacità di Amazon con variazioni minime di prestazioni e costi.

Flessibilità nella scelta delle zone di disponibilità

Ti consigliamo di configurare tutte le zone di disponibilità per l'uso nel tuo cloud privato virtuale (VPC) e di selezionarle per il cluster EMR. I cluster devono esistere in una sola zona di disponibilità, ma con i parchi istanze Amazon EMR è possibile selezionare più sottoreti per diverse zone di disponibilità. Quando Amazon EMR avvia il cluster, viene eseguita una ricerca in tali sottoreti per individuare le istanze e le opzioni di acquisto che hai specificato. Quando effettui il provisioning di un cluster EMR per più sottoreti, il cluster può accedere a un pool di EC2 capacità Amazon più profondo rispetto ai cluster in una singola sottorete.

Se devi dare priorità a un certo numero di zone di disponibilità da utilizzare nel tuo cloud privato virtuale (VPC) per il tuo cluster EMR, puoi sfruttare la funzionalità di punteggio di posizionamento Spot con Amazon. EC2 Con il punteggio di posizionamento Spot, specifichi i requisiti di calcolo per le tue istanze Spot, quindi EC2 restituisci le prime dieci zone di disponibilità con un punteggio su una Regioni AWS scala da 1 a 10. Un punteggio pari a 10 indica che è molto probabile che la richiesta Spot abbia successo; un punteggio pari a 1 indica che è improbabile che la richiesta Spot abbia successo. Per ulteriori informazioni su come utilizzare il punteggio di posizionamento Spot, consulta il punteggio di posizionamento Spot nella Amazon EC2 User Guide.

Flessibilità nella scelta dei tipi di istanza

La flessibilità delle istanze consiste nell'utilizzare più tipi di istanze per soddisfare i requisiti di capacità. La flessibilità delle istanze avvantaggia sia l'utilizzo di Amazon EC2 Spot che delle istanze On-Demand. Con Spot Instances, la flessibilità delle istanze consente ad Amazon di EC2 avviare istanze da pool di capacità più profondi utilizzando dati sulla capacità in tempo reale. Inoltre, prevede quali sono le istanze più disponibili. Ciò garantisce un numero minore di interruzioni e può ridurre il costo complessivo di un carico di lavoro. Con le istanze on demand, la flessibilità delle istanze riduce gli errori di capacità insufficiente (ICE) quando la capacità totale viene fornita da un numero maggiore di pool di istanze.

Per i cluster di gruppi di istanze, puoi specificare fino a 50 EC2 tipi di istanze. Per le flotte di istanze con strategia di allocazione, puoi specificare fino a 30 tipi di EC2 istanze per ogni gruppo di nodi primario, principale e di task. Una gamma più ampia di istanze migliora i vantaggi della flessibilità delle istanze.

Espressione della flessibilità delle istanze

Tieni in considerazione le seguenti best practice per usufruire della flessibilità delle istanze per la tua applicazione.

Argomenti

- Determinazione della famiglia e delle dimensioni delle istanze
- Inclusione di istanze aggiuntive

Determinazione della famiglia e delle dimensioni delle istanze

Amazon EMR supporta diversi tipi di istanza per casi d'uso differenti. Questi tipi di istanze sono elencati nella documentazione di Tipi di istanze supportati con Amazon EMR. Ogni tipo di istanza appartiene a una famiglia di istanze che descrive l'applicazione per cui il tipo è ottimizzato.

Per i nuovi carichi di lavoro, è necessario eseguire un benchmark con i tipi di istanze della famiglia per uso generico, ad esempio `m5` o `c5`. Quindi, osserva i parametri del sistema operativo e YARN di Ganglia e Amazon CloudWatch per determinare i colli di bottiglia del sistema al massimo carico. I colli di bottiglia includono CPU, memoria, storage e operazioni. I/O Dopo avere identificato i colli di bottiglia, scegli una famiglia di istanze ottimizzate per l'elaborazione, per la memoria, per l'archiviazione o un'altra famiglia appropriata per i tuoi tipi di istanze. Per ulteriori dettagli, consulta la pagina Determina l'infrastruttura giusta per i carichi di lavoro Spark nella guida alle best practice di Amazon EMR su. GitHub

Quindi, identifica il container YARN o l'esecutore Spark più piccolo richiesto dall'applicazione. Questa è la dimensione dell'istanza più piccola adatta al container e la dimensione minima dell'istanza per il cluster. Utilizza questo parametro per determinare le istanze con cui puoi diversificare ulteriormente. Un'istanza più piccola consentirà una maggiore flessibilità delle istanze.

Per la massima flessibilità delle istanze, dovresti sfruttare il maggior numero possibile di istanze. Ti consigliamo di diversificare con istanze dotate di specifiche hardware simili. Ciò massimizza l'accesso ai pool di EC2 capacità con variazioni minime in termini di costi e prestazioni. Diversifica in base alle dimensioni. Per farlo, dai prima priorità ad AWS Graviton e alle generazioni precedenti. Una buona regola è quella di dimostrare flessibilità su almeno 15 tipi di istanza per ogni carico di lavoro. Ti consigliamo di iniziare con istanze per uso generico, ottimizzate per il calcolo oppure ottimizzate per la memoria. Questi tipi di istanze offriranno la massima flessibilità.

Inclusione di istanze aggiuntive

Per garantire la massima diversità, includi tipi di istanze aggiuntivi. Dai priorità alla flessibilità di dimensioni delle istanze, Graviton e di generazione. Ciò consente l'accesso a pool di EC2 capacità aggiuntivi con profili di costi e prestazioni simili. Se hai bisogno di ulteriore flessibilità a causa degli ICE o delle interruzioni spot, prendi in considerazione la flessibilità delle varianti e delle famiglie. Ogni approccio presenta dei compromessi che dipendono dal caso d'uso e dai requisiti.

- Flessibilità di dimensione: in primo luogo, diversifica scegliendo istanze di dimensioni diverse all'interno della stessa famiglia. Le istanze della medesima famiglia offrono lo stesso costo e le stesse prestazioni, ma possono avviare un numero diverso di container su ogni host. Ad esempio, se la dimensione minima dell'esecutore necessaria è 2 vCPU con 8 Gb di memoria, la dimensione minima dell'istanza è `m5.xlarge`. Per garantire la flessibilità delle dimensioni, includi `m5.xlarge`, `m5.2xlarge`, `m5.4xlarge`, `m5.8xlarge`, `m5.12xlarge`, `m5.16xlarge` e `m5.24xlarge`.
- Flessibilità di Graviton: oltre alle dimensioni, puoi diversificare con le istanze Graviton. Le istanze Graviton sono alimentate da processori AWS Graviton2 che offrono il miglior rapporto prezzo/prestazioni per i carichi di lavoro cloud su Amazon. EC2 Ad esempio, con la dimensione minima dell'istanza di `m5.xlarge`, è possibile includere `m6g.xlarge`, `m6g.2xlarge`, `m6g.4xlarge`, `m6g.8xlarge` e `m6g.16xlarge` per la flessibilità di Graviton.
- Flessibilità di generazione: analogamente alla flessibilità di dimensione e di Graviton, le istanze delle famiglie della generazione precedente condividono le stesse specifiche hardware. Ciò si traduce in un profilo di costi e prestazioni simile con un aumento del EC2 pool Amazon totale accessibile. Per garantire la flessibilità di generazione, includi `m4.xlarge`, `m4.2xlarge`, `m4.10xlarge` e `m4.16xlarge`.
- Flessibilità per famiglie e varianti
  - Capacità: per ottimizzare la capacità, consigliamo la flessibilità delle istanze scegliendo tra famiglie di istanze diverse. Le istanze comuni di famiglie di istanze diverse dispongono di pool di istanze più profondi che possono contribuire a soddisfare i requisiti di capacità. Tuttavia, istanze di famiglie diverse hanno rapporti tra vCPU e memoria differenti. Ciò comporta un sottoutilizzo se il container dell'applicazione previsto è dimensionato per un'istanza diversa. Ad esempio, con `m5.xlarge`, includi istanze ottimizzate per il calcolo come `c5` oppure istanze ottimizzate per la memoria come `r5` per la flessibilità della famiglia di istanze.
  - Costo: per ottimizzare i costi, consigliamo la flessibilità delle istanze scegliendo tra le varianti. Queste istanze hanno lo stesso rapporto tra memoria e vCPU dell'istanza iniziale. Il compromesso con la flessibilità delle varianti è che queste istanze hanno pool di capacità più piccoli, il che potrebbe comportare una capacità aggiuntiva limitata o interruzioni spot più numerose. Ad esempio, con `m5.xlarge`, includi istanze basate su AMD (`m5a`), istanze basate su SSD (`m5d`) o istanze ottimizzate per la rete (`m5n`) per garantire la flessibilità delle varianti.

Configurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot

Usa le linee guida in questa sezione per determinare i tipi di istanza, le opzioni di acquisto e la quantità di storage di cui effettuare il provisioning per ogni tipo di nodo in un cluster EMR.

Che tipo di istanza utilizzare?

Esistono diversi modi per aggiungere EC2 istanze Amazon a un cluster. Il metodo da scegliere dipende dal fatto che si utilizzi la configurazione dei gruppi di istanze o la configurazione dei parchi istanze per il cluster.

- Gruppi di istanze
  - Aggiungi manualmente istanze dello stesso tipo a gruppi di istanze principali e dell'attività esistenti.
  - Aggiungi manualmente un gruppo di istanze dell'attività, che può utilizzare un tipo di istanza diverso.
  - Configura la scalabilità automatica in Amazon EMR per un gruppo di istanze, aggiungendo e rimuovendo istanze automaticamente in base al valore di un parametro CloudWatch Amazon specificato. Per ulteriori informazioni, consulta Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione.
- Parchi istanze
  - Aggiungi un singolo parco istanze attività.
  - Cambia la capacità target per istanze on demand e Spot per parchi istanze principale e dell'attività. Per ulteriori informazioni, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR.

Un modo per pianificare le istanze del cluster è eseguire un cluster di test con un set di dati di esempio rappresentativo e monitorare l'utilizzo dei nodi del cluster. Per ulteriori informazioni, consulta Visualizza e monitora un cluster Amazon EMR mentre esegue il lavoro. Un altro modo è calcolare la capacità delle istanze che si stanno valutando e confrontare tale valore rispetto alle dimensioni dei dati.

In generale, il tipo di nodo principale, che assegna le attività, non richiede un'EC2 istanza con molta potenza di elaborazione; EC2 le istanze Amazon per il tipo di nodo core, che elaborano attività e archiviano dati in HDFS, richiedono sia potenza di elaborazione che capacità di archiviazione; EC2 le istanze Amazon per il tipo di nodo task, che non archiviano dati, richiedono solo potenza di elaborazione. Per linee guida sulle EC2 istanze Amazon disponibili e sulla loro configurazione, consultaConfigurazione dei tipi di EC2 istanze Amazon da utilizzare con Amazon EMR.

Le seguenti linee guida si applicano alla maggior parte dei cluster Amazon EMR.

- Esiste un limite di vCPU per il numero totale di istanze EC2 Amazon on-demand eseguite su AWS un account per account. Regione AWS Per ulteriori informazioni sul limite di vCPU e su come richiedere un aumento del limite per il tuo account, consulta On-Demand Instances nella Amazon EC2 User Guide for Linux Instances.
- Generalmente, il nodo primario non ha requisiti di calcolo elevati. Per i cluster con un numero elevato di nodi o per i cluster con applicazioni distribuite specificamente sul nodo primario (JupyterHub, Hue, ecc.), potrebbe essere necessario un nodo primario più grande che può contribuire a migliorare le prestazioni del cluster. Ad esempio, potresti prendere in considerazione l'utilizzo di un'istanza m5.xlarge per cluster piccoli (50 nodi o meno) e l'aumento a un tipo di istanza più grande per cluster più grandi.
- Le esigenze di calcolo dei nodi principali e di task dipendono dal tipo di elaborazione eseguita dall'applicazione. Molti processi possono essere eseguiti su tipi di istanze per uso generico, che offrono prestazioni bilanciate in termini di CPU, spazio su disco e input/output. Cluster che richiedono molti calcoli possono trarre beneficio dall'esecuzione su istanze con elevata CPU, che hanno proporzionalmente più CPU che RAM. Applicazioni di database e di caching nella memoria possono trarre vantaggio dall'esecuzione su istanze a memoria elevata. Applicazioni che prevedono un uso intensivo della rete e della CPU, ad esempio operazioni di analisi, procedimenti NLP e Machine Learning, possono trarre vantaggio dall'esecuzione su istanze Cluster Compute, che forniscono risorse di CPU proporzionalmente elevate e prestazioni di rete avanzate.
- Se fasi diverse del cluster hanno diverse esigenze di capacità, puoi iniziare con un numero ridotto di nodi principali e aumentare o diminuire il numero di nodi di task per soddisfare i diversi requisiti di capacità del flusso di elaborazione.
- La quantità di dati che puoi elaborare dipende dalla capacità dei nodi principali e dalla dimensione dei dati come input, durante l'elaborazione, e come output. I dataset di input, intermedio e di output risiedono tutti sul cluster durante l'elaborazione.

Quando occorre utilizzare le istanze Spot?

Quando avvii un cluster in Amazon EMR, puoi scegliere di avviare istanze primarie, core o attività su istanze spot. Poiché ogni tipo di gruppo di istanze gioca un ruolo diverso nel cluster, esistono implicazioni legate all'avvio di ciascun tipo di nodo sulle istanze Spot. Non è possibile modificare un'opzione di acquisto di un'istanza quando un cluster è in esecuzione. Per passare dalle istanze on demand alle istanze spot o viceversa per i nodi primari e core è necessario terminare il cluster e avviarne uno nuovo. Per i nodi di task, puoi avviare una nuovo gruppo di istanze di task o un parco istanze e rimuovere quella precedente.

Argomenti

- Impostazioni di Amazon EMR per impedire gli errori nei processi a causa dell'interruzione delle istanze Spot nei nodi attività
- Nodo primario su un'istanza spot
- Nodi principali sulle istanze Spot
- Nodi attività sulle istanze Spot
- Configurazioni di istanze per scenari applicativi

Impostazioni di Amazon EMR per impedire gli errori nei processi a causa dell'interruzione delle istanze Spot nei nodi attività

Poiché le istanze Spot vengono spesso utilizzate per eseguire nodi attività, Amazon EMR dispone delle funzionalità predefinite per la pianificazione dei processi YARN in modo che i processi in esecuzione non abbiano esito negativo quando i nodi attività in esecuzione su istanze Spot vengono terminati. Amazon EMR esegue questa operazione consentendo ai processi master delle applicazioni di funzionare solo sui nodi principali. Il processo master dell'applicazione controlla i processi in esecuzione e deve rimanere attivo per tutta la durata del processo.

Amazon EMR rilascio 5.19.0 e successivi utilizzano la caratteristica integrata etichette nodo YARN per questo scopo. (Le versioni precedenti utilizzavano una patch di codice). Le proprietà nelle classificazioni di configurazione `yarn-site` e `capacity-scheduler` sono configurate per impostazione predefinita in modo che capacity-scheduler e fair-scheduler YARN sfruttino le etichette dei nodi. Amazon EMR etichetta in automatico i nodi principali con l'etichetta `CORE` e imposta le proprietà in modo che i master dell'applicazione siano pianificati solo sui nodi con l'etichetta CORE. La modifica manuale delle proprietà correlate nelle classificazioni di configurazione del sito di YARN e del pianificatore di capacità o direttamente nei file XML associati potrebbe interrompere o alterare questa funzionalità.

Per impostazione predefinita, Amazon EMR configura le proprietà e i valori riportati di seguito. Fai attenzione quando configuri queste proprietà.

> **Note**
>
> A partire dalla serie di rilascio Amazon EMR 6.x, la funzione etichette nodo YARN è disabilitata per impostazione predefinita. Per impostazione predefinita, i processi primari dell'applicazione possono essere eseguiti sia sui nodi core sia su quelli attività. È possibile abilitare la caratteristica etichette nodo YARN configurando le seguenti proprietà:

    - `yarn.node-labels.enabled: true`
    - `yarn.node-labels.am.default-node-label-expression: 'CORE'`

- yarn-site (yarn-site.xml) Su tutti i nodi
  - `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) su nodi primari e core
  - `yarn.nodemanager.node-labels.provider: 'config'`
  - `yarn.nodemanager.node-labels.provider.configured-node-partition: 'CORE'`
- capacity-scheduler (capacity-scheduler.xml) Su tutti i nodi
  - `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 primario su un'istanza spot

Il nodo primario controlla e gestisce il cluster. Quando termina, termina anche il cluster, pertanto devi avviare il nodo primario come istanza spot solo se stai eseguendo un cluster in cui la terminazione improvvisa è accettabile. Ciò è possibile se stai eseguendo il test di una nuova applicazione, disponi di un cluster che conserva periodicamente i dati in uno store esterno come Amazon S3 o stai eseguendo un cluster in cui il costo è più importante che garantire il completamento del cluster.

Quando avvii il gruppo di istanze primarie come istanza spot, il cluster non si avvia finché tale richiesta di istanza spot non viene soddisfatta. È importare tenere presente ciò quando si seleziona il prezzo Spot massimo.

Puoi aggiungere un nodo primario di istanza spot solo quando avvii il cluster. I nodi primari non possono essere aggiunti o rimossi da un cluster in esecuzione.

In genere, devi eseguire il nodo primario come istanza spot solo se stai eseguendo l'intero cluster (tutti i gruppi di istanze) come istanze spot.

Nodi principali sulle istanze Spot

I nodi principali elaborano dati e memorizzano informazioni utilizzando HDFS. La terminazione di un'istanza principale comporta la perdita dei dati. Per questo motivo, è consigliabile eseguire solo i nodi principali sulle istanze Spot quando la perdita dei dati HDFS parziale è tollerabile.

Quando avvii il gruppo di istanze principale come istanze Spot, Amazon EMR attende finché non puoi effettuare il provisioning di tutte le istanze principali richieste prima di avviare il gruppo di istanze. In altre parole, se richiedi sei EC2 istanze Amazon e solo cinque sono disponibili al prezzo Spot massimo o inferiore, il gruppo di istanze non verrà avviato. Amazon EMR continua ad attendere che tutte e sei le EC2 istanze Amazon siano disponibili o fino alla chiusura del cluster. È possibile modificare il numero di istanze Spot in un gruppo di istanze principali per aggiungere capacità a un cluster in esecuzione. Per ulteriori informazioni sulle operazioni con i gruppi di istanze e sul funzionamento delle istanze Spot con il parco istanze, consulta the section called “Configurazione di parchi istanze o gruppi di istanze”.

Nodi attività sulle istanze Spot

I nodi di task elaborano i dati, ma non mantengono dati persistenti in HDFS. Se terminano perché il prezzo Spot ha superato il prezzo Spot massimo, nessun dato viene perso e l'effetto sul cluster è minimo.

Quando avvii uno o più gruppi di istanze attività come istanze Spot, Amazon EMR esegue il provisioning di quanti più nodi attività possibili, utilizzando il prezzo Spot massimo. Questo significa che se richiedi un gruppo di istanze attività con sei nodi e solo cinque istanze Spot sono disponibili al o sotto il prezzo Spot massimo, Amazon EMR avvia il gruppo di istanze con cinque nodi, aggiungendo il sesto in seguito, se possibile.

Avviare gruppi di istanze dell'attività come istanze Spot è un modo strategico per espandere la capacità del cluster riducendo i costi. Se avvii i gruppi di istanze primarie e core come istanze on demand, la loro capacità è garantita per l'esecuzione del cluster. Puoi aggiungere istanze dell'attività ai gruppi di istanze dell'attività in base alle esigenze, per gestire picchi di traffico o velocizzare l'elaborazione dei dati.

Puoi aggiungere o rimuovere nodi di attività utilizzando la console o l'API AWS CLI. Puoi anche aggiungere altri gruppi di attività, ma non puoi rimuovere un gruppo di attività dopo che è stato creato.

Configurazioni di istanze per scenari applicativi

La tabella seguente è un riferimento rapido per le opzioni di acquisto e le configurazioni dei tipi di nodo in genere appropriate per vari scenari applicativi. Scegli il link per visualizzare ulteriori informazioni su ciascun tipo di scenario.

Scenario applicativo	Opzione di acquisto per i nodi primari	Opzione di acquisto per i nodi principali	Opzione di acquisto per i nodi attività
Cluster di lunga durata e data warehouse	On demand	Combinazione di on demand o parco istanze	Combinazione di Spot o parco istanze
Carichi di lavoro basati sui costi	Spot	Spot	Spot
Carichi di lavoro critici	On demand	On demand	Combinazione di Spot o parco istanze
Test dell'applicazione	Spot	Spot	Spot

Esistono diversi scenari in cui le istanze Spot sono utili per l'esecuzione di un cluster Amazon EMR.

Cluster di lunga durata e data warehouse

Se stai eseguendo un cluster Amazon EMR persistente caratterizzato da una variazione prevedibile della capacità di elaborazione, ad esempio un data warehouse, puoi gestire picchi di domanda a costi inferiori con istanze Spot. Puoi avviare gruppi di istanze primarie e core come istanze on demand per gestire la normale capacità e avviare il gruppo di istanze attività come istanze spot per gestire i requisiti di carico di picco.

Carichi di lavoro basati sui costi

Se stai eseguendo cluster transitori per i quali un costo inferiore è più importante del tempo necessario per il completamento e la perdita parziale del lavoro è accettabile, puoi eseguire l'intero cluster (gruppi di istanze primarie, core e attività) come istanze spot per trarre vantaggio dal maggiore risparmio sui costi.

Carichi di lavoro critici

Se stai eseguendo un cluster per i quale un costo inferiore è più importante del tempo necessario per il completamento, ma la perdita parziale del lavoro non è accettabile, avvia i gruppi di istanze primarie e core come istanze on demand e integra con uno o più gruppi di istanze attività di istanze spot. L'esecuzione di gruppi di istanze primarie e core come istanze on demand garantisce che i dati vengono mantenuti in HDFS e che il cluster sia protetto dalla terminazione dovuta a fluttuazioni del mercato spot, fornendo al contempo risparmi sui costi derivanti dall'esecuzione di gruppi di istanze attività come istanze spot.

Test dell'applicazione

Quando esegui il test di una nuova applicazione come preparazione all'avvio in un ambiente di produzione, puoi eseguire l'intero cluster (gruppi di istanze primarie, core e attività) come istanze spot per ridurre i costi di test.

Calcolo della capacità HDFS richiesta di un cluster

La quantità di storage HDFS disponibile per il cluster dipende dai seguenti fattori:

- Il numero di EC2 istanze Amazon utilizzate per i nodi principali.
- La capacità dell' EC2 instance store di Amazon per il tipo di istanza utilizzato. Per ulteriori informazioni sui volumi degli instance store, consulta Amazon Amazon EC2 Instance Store nella Amazon EC2 User Guide.
- Il numero e la dimensione dei volumi Amazon EBS collegati ai nodi principali.
- Un fattore di replica, che tiene conto del modo in cui ogni blocco di dati viene archiviato in HDFS per ridondanza di tipo RAID. Per impostazione predefinita, il fattore di replica è tre per un cluster di almeno 10 nodi principali, due per un cluster di 4-9 nodi principali e uno per un cluster di massimo tre nodi.

Per calcolare la capacità HDFS di un cluster, per ogni nodo principale aggiungi la capacità del volume instance store alla capacità di archiviazione Amazon EBS (se utilizzata). Moltiplica il risultato per il numero di nodi principali, quindi dividi il totale per il fattore di replica in base al numero di nodi principali. Ad esempio, un cluster con 10 nodi principali di tipo i2.xlarge, che dispone di 800 GB di storage dell'istanza senza alcun volume Amazon EBS collegato, dispone di un totale di circa 2.666 GB disponibili per HDFS (10 nodi x 800 GB ÷ fattore di replica 3).

Se il valore di capacità HDFS calcolato è inferiore ai dati, puoi aumentare la quantità di storage HDFS nei seguenti modi:

- Creando un cluster con volumi Amazon EBS aggiuntivi o aggiungendo gruppi di istanze con volumi Amazon EBS collegati a un cluster esistente
- Aggiungendo più nodi principali
- Scelta di un tipo di EC2 istanza Amazon con maggiore capacità di storage
- Utilizzando la compressione dati
- Modificando le impostazioni di configurazione Hadoop per ridurre il fattore di replica

La riduzione del fattore di replica deve essere utilizzata con attenzione poiché riduce la ridondanza dei dati HDFS e la capacità del cluster di recuperare da blocchi HDFS persi o danneggiati.

Configurazione del logging e del debug dei cluster Amazon EMR

Quando pianifichi il cluster, devi determinare la quantità di supporto di debug che vuoi rendere disponibile. Durante lo sviluppo iniziale dell'applicazione di elaborazione dati, ti consigliamo di testare l'applicazione su un cluster che elabora un piccolo, ma rappresentativo, sottoinsieme dei tuoi dati. È probabile che per eseguire questa operazione tu intenda utilizzare tutti gli strumenti di debug forniti da Amazon EMR, come l'archiviazione dei file di log in Amazon S3.

Una volta terminata la fase di sviluppo dell'applicazione di elaborazione dati e avviata quella di produzione, puoi scegliere di ridurre il debug. In questo modo, azzeri il costo relativo all'archiviazione degli archivi di file di log in Amazon S3 e riduci il carico di elaborazione sul cluster in quanto non è più necessario scrivere lo stato su Amazon S3. L'inconveniente di questa scelta è che in caso di problemi avrai a disposizione un minor numero di strumenti per gestirli.

File di log di default

Per impostazione predefinita, ogni cluster scrive file di log sul nodo primario. Questi file sono scritti nella directory `/mnt/var/log/`. Puoi accedervi utilizzando SSH per la connessione al nodo primario come descritto in Connect al nodo primario del cluster Amazon EMR tramite SSH. Amazon EMR raccoglie determinati log di sistema e applicazioni generati dai daemon Amazon EMR e da altri processi Amazon EMR per garantire operazioni di servizio efficaci.

> **Note**
>
> Se utilizzi Amazon EMR versione 6.8.0 o precedenti, i file di log vengono salvati in Amazon S3 durante la terminazione del cluster, quindi non puoi accedere ai file di log una volta terminato il nodo primario. Amazon EMR versione 6.9.0 e successive archiviano i log

in Amazon S3 durante la riduzione del cluster, cosicché i file di log generati sul cluster persistono anche dopo la terminazione del nodo.

Non è necessario attivare alcuna opzione affinché i file di log siano scritti sul nodo primario. In effetti, questo è il comportamento predefinito di Amazon EMR e Hadoop.

Un cluster genera vari tipi di file di log, tra cui:

- Step logs (Log di fase): questi log sono generati dal servizio Amazon EMR e contengono informazioni sul cluster e i risultati di ogni fase. I file di log sono archiviati nella directory `/mnt/var/log/hadoop/steps/` nel nodo primario. Ogni fase registra i risultati a essa relativi in una sottodirectory numerata distinta: `/mnt/var/log/hadoop/steps/s-`*`stepId1`*`/` per la prima fase, `/mnt/var/log/hadoop/steps/s-`*`stepId2`*`/`, per la seconda e così via. Gli identificatori di fase a 13 caratteri (ad esempio stepId1, stepId2) sono esclusivi di un cluster.
- Registri dei componenti Hadoop e YARN: i log dei componenti associati sia ad Apache YARN che, ad esempio, sono contenuti in cartelle separate in. MapReduce `/mnt/var/log` Le posizioni dei file di log per i componenti Hadoop in `/mnt/var/log` sono le seguenti: hadoop-hdfs, hadoop-mapreduce, hadoop-httpfs e hadoop-yarn. La hadoop-state-pusher directory serve per l'output del processo Hadoop state pusher.
- Bootstrap action logs (Log delle operazioni di bootstrap): se il processo utilizza operazioni di bootstrap, i risultati di queste operazioni sono registrati. I file di registro sono memorizzati in/mnt/var/log/bootstrap-actions/ sul nodo primario. Ogni operazione di bootstrap registra i risultati a essa relativi in una sottodirectory numerata distinta: `/mnt/var/log/bootstrap-actions/1/` per la prima operazione di bootstrap, `/mnt/var/log/bootstrap-actions/2/` per la seconda e così via.
- Instance state logs (Log di stato dell'istanza): forniscono informazioni su CPU, stato della memoria e thread del garbage collector del nodo. I file di log sono archiviati in `/mnt/var/log/instance-state/` nel nodo primario.

Archiviazione di file di log in Amazon S3

> **Note**
>
> Non è al momento possibile utilizzare l'aggregazione dei log in Amazon S3 con l'utility `yarn logs`.

Amazon EMR rilascio 6.9.0 e successivi archiviano i log in Amazon S3 durante la riduzione del cluster, cosicché i file di log generati sul cluster persistono anche dopo la terminazione del nodo. Questo comportamento è abilitato automaticamente, quindi non devi fare nulla per attivarlo.
Per Amazon EMR rilascio 6.8.0 e precedenti, è possibile configurare un cluster per archiviare periodicamente i file di log presenti nel nodo primario in Amazon S3. In questo modo, i file di log saranno disponibili anche dopo la terminazione del cluster, indipendentemente dalla causa della stessa (arresto normale, errore, ecc.). Amazon EMR archivia i file di log in Amazon S3 ogni 5 minuti.

Per Amazon EMR rilascio 6.8.0 e successivi, per archiviare i file di log in Amazon S3 devi abilitare tale caratteristica all'avvio del cluster. Puoi eseguire tale operazione mediante la console, la CLI o l'API. Per impostazione predefinita, l'archiviazione dei log è abilitata per i cluster avviati utilizzando la console. Per i cluster avviati mediante la CLI o l'API, la registrazione in Amazon S3 deve essere abilitata manualmente.

Console

Archiviazione dei file di log in Amazon S3 con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Cluster logs (Log del cluster), seleziona la casella di controllo Publish cluster-specific logs to Amazon S3 (Pubblica log specifici del cluster su Amazon S3).
4. Nel campo Amazon S3 location (Posizione Amazon S3), digita o cerca il percorso Amazon S3 in cui archiviare i log. Se digiti il nome di una cartella che non esiste nel bucket, Amazon S3 crea tale cartella.

   Quando imposti questo valore, Amazon EMR copia i file di log dalle EC2 istanze del cluster su Amazon S3. In questo modo si evita la perdita dei file di log alla chiusura del cluster e si EC2 interrompono le istanze che ospitano il cluster. Questi log sono utili per la risoluzione dei problemi. Per ulteriori informazioni, consulta Visualizzazione dei file di log.
5. Facoltativamente, seleziona la casella di controllo Encrypt cluster-specific logs (Crittografa log specifici del cluster). Quindi, seleziona una AWS KMS chiave dall'elenco, inserisci un ARN della chiave o crea una nuova chiave. Questa opzione è disponibile solo con Amazon EMR versione 5.30.0 e successive, esclusa la versione 6.0.0. Per utilizzare questa opzione, aggiungi l'autorizzazione al AWS KMS profilo dell' EC2istanza e al ruolo Amazon EMR. Per

ulteriori informazioni, consulta Crittografia dei file di log archiviati in Amazon S3 con una chiave gestita dal cliente di AWS KMS.

6. Scegli qualsiasi altra opzione applicabile al cluster.
7. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per archiviare i file di registro su Amazon S3 con AWS CLI

Per archiviare i file di log su Amazon S3 utilizzando AWS CLI, digita il `create-cluster` comando e specifica il percorso di log di Amazon S3 utilizzando il parametro. `--log-uri`

1. Per registrare i file su Amazon S3, digita il seguente comando e sostituiscilo *myKey* con il nome della tua coppia di EC2 chiavi.

```
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. Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

> **Note**
>
> Se in precedenza non hai creato il ruolo e il profilo di EC2 istanza del servizio Amazon EMR predefiniti, inserisci `aws emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Crittografia dei file di log archiviati in Amazon S3 con una chiave gestita dal cliente di AWS KMS

Con Amazon EMR versione 5.30.0 e successive (eccetto Amazon EMR 6.0.0), puoi crittografare i file di log archiviati in Amazon S3 con una chiave gestita dal cliente KMS. AWS Per abilitare questa opzione nella console, segui le fasi in Archiviazione di file di log in Amazon S3. Il tuo profilo di EC2 istanza Amazon e il tuo ruolo Amazon EMR devono soddisfare i seguenti prerequisiti:

- Il profilo di EC2 istanza Amazon utilizzato per il tuo cluster deve essere autorizzato all'usokms:GenerateDataKey.
- Il ruolo Amazon EMR utilizzato per il cluster deve disporre dell'autorizzazione per utilizzare kms:DescribeKey.
- Il profilo dell' EC2 istanza Amazon e il ruolo Amazon EMR devono essere aggiunti all'elenco degli utenti chiave per la chiave gestita dal cliente AWS KMS specificata, come dimostrano i seguenti passaggi:
  1. Apri la console AWS Key Management Service (AWS KMS) in /kms. https://console.aws.amazon.com
  2. Per cambiare la AWS regione, usa il selettore della regione nell'angolo in alto a destra della pagina.
  3. Seleziona l'alias della chiave KMS da modificare.
  4. Nella pagina dei dettagli della chiave, in Key Users (Utenti di chiavi), scegli Add (Aggiungi).
  5. Nella finestra di dialogo Aggiungi utenti chiave, seleziona il tuo profilo di EC2 istanza Amazon e il ruolo Amazon EMR.
  6. Scegli Aggiungi.
- È inoltre necessario configurare la chiave KMS per consentire a persistentappui.elasticmapreduce.amazonaws.com and elasticmapreduce.amazonaws.com Service Principal di, e. kms:GenerateDataKey kms:GenerateDataKeyWithoutPlaintext kms:Decrypt Ciò consente a EMR di leggere e scrivere registri crittografati con la chiave KMS per lo storage gestito. L'utente IAM Role deve disporre dell'autorizzazione per utilizzare e. kms:GenerateDataKey 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"
        }
     }
```

```
}
```

Come best practice di sicurezza, ti consigliamo di aggiungere le `aws:SourceArn` condizioni `kms:EncryptionContext` and. Queste condizioni aiutano a garantire che la chiave venga utilizzata solo da Amazon EMR su EC2 e solo per i log generati dai lavori in esecuzione in un cluster specifico.

Per ulteriori informazioni, consulta i ruoli del servizio IAM utilizzati da Amazon EMR e Utilizzo delle politiche chiave nella guida per sviluppatori di AWS Key Management Service.

Aggregazione di log in Amazon S3 mediante la AWS CLI

> **Note**
>
> Non è al momento possibile utilizzare l'aggregazione dei log con l'utility `yarn logs`. È possibile utilizzare soltanto l'aggregazione supportata da questa procedura.

L'aggregazione di log (Hadoop 2.x) compila i log di tutti i container di una singola applicazione in un unico file. Per abilitare l'aggregazione dei log su Amazon S3 utilizzando AWS CLI il, è necessario utilizzare un'azione bootstrap all'avvio del cluster per abilitare l'aggregazione dei log e specificare il bucket in cui archiviare i log.

- Per abilitare l'aggregazione di log, crea il file di configurazione denominato `myConfig.json`, che contiene quanto segue:

```
[
  {
    "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"
    }
  }
]
```

  Digita il comando seguente e *myKey* sostituiscilo con il nome della tua EC2 key pair. Puoi inoltre sostituire il testo in rosso con le configurazioni desiderate.

```
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
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo primario e le istanze rimanenti vengono avviate come nodi core. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

> **Note**
>
> Se in precedenza non sono stati creati il ruolo e il profilo di EC2 istanza del servizio EMR predefiniti, esegui `aws emr create-default-roles` per crearli prima di eseguire il `create-cluster` sottocomando.

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta AWS CLI Command Reference.

Posizione dei log

L'elenco seguente include tutti i tipi di log e le relative posizioni in Amazon S3. Puoi utilizzarli per risolvere i problemi di Amazon EMR.

Log delle fasi

`s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/steps/<step-id>/`

Log di applicazioni

`s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/containers/`

Questa posizione include i log del container `stderr` e `stdout`, `directory.info`, `prelaunch.out` e `launch_container.sh`.

Log del gestore delle risorse

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
applications/hadoop-yarn/
```

Hadoop HDFS

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/
applications/hadoop-hdfs/
```

Questa posizione include NameNode DataNode, e i log YARN. TimelineServer

Log del gestore dei nodi

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/
applications/hadoop-yarn/
```

Log degli stati istanza

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/daemons/
instance-state/
```

Log di provisioning di Amazon EMR

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
provision-node/*
```

Log Hive

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
applications/hive/*
```

- Per trovare i log di Hive sul tuo cluster, rimuovi l'asterisco (`*`) e aggiungi `/var/log/hive/` al link precedente.
- Per trovare HiveServer 2 log, rimuovi l'asterisco (`*`) e `var/log/hive/hiveserver2.log` aggiungilo al link precedente.
- Per trovare i log HiveCLI, rimuovi l'asterisco (`*`) e aggiungi `/var/log/hive/user/hadoop/hive.log` al link precedente.
- Per trovare i log di Hive Metastore Server, rimuovi l'asterisco (`*`) e aggiungi `/var/log/hive/user/hive/hive.log` al link precedente.

Se l'errore si trova nel nodo primario o nel nodo attività dell'applicazione Tez, fornisci i log del container Hadoop appropriato.

Etichetta e classifica le risorse del cluster Amazon EMR

Può essere conveniente classificare le AWS risorse in diversi modi, ad esempio per scopo, proprietario o ambiente. Puoi eseguire questa operazione in Amazon EMR assegnando metadati personalizzati ai cluster Amazon EMR mediante tag. Un tag consiste di una chiave e di un valore che definisci. Per Amazon EMR, il cluster è il livello di risorse che puoi taggare. Ad esempio, puoi definire un set di tag per i cluster del tuo account per tenere traccia del proprietario di ogni cluster o per identificare un cluster di produzione rispetto a un cluster di testing. Ti consigliamo di creare un set di tag coerente per soddisfare i requisiti dell'organizzazione.

Quando aggiungi un tag a un cluster Amazon EMR, il tag viene anche propagato a ogni EC2 istanza Amazon attiva associata al cluster. Allo stesso modo, quando rimuovi un tag da un cluster Amazon EMR, tale tag viene rimosso da ogni istanza Amazon EC2 attiva associata.

> **Important**
>
> Utilizza la console o l'interfaccia a riga di comando di Amazon EMR per gestire i tag sulle EC2 istanze Amazon che fanno parte di un cluster anziché sulla console Amazon EC2 o l'interfaccia a riga di comando, poiché le modifiche apportate in Amazon EC2 non vengono sincronizzate con il sistema di tagging Amazon EMR.

Puoi identificare un' EC2 istanza Amazon che fa parte di un cluster Amazon EMR cercando i seguenti tag di sistema. In questo esempio, *CORE* è il valore per il ruolo del gruppo di istanze ed *j-12345678* è un esempio di valore identificativo del flusso di lavoro (cluster):

- aws:elasticmapreduce: = instance-group-role *CORE*
- aws: elasticmapreduce: = job-flow-id *j-12345678*

> **Note**
>
> Amazon EMR e Amazon EC2 interpretano i tag come una stringa di caratteri senza significato semantico.

Puoi lavorare con i tag utilizzando AWS Management Console, la CLI e l'API.

Puoi aggiungere tag durante la creazione di un nuovo cluster Amazon EMR nonché aggiungere, modificare o rimuovere tag da un cluster Amazon EMR in esecuzione. La modifica di un tag è un concetto che si applica alla console di Amazon EMR. Tuttavia, se utilizzi la CLI e l'API, la modifica di un tag consiste nel rimuovere il vecchio tag e aggiungerne un nuovo. Puoi modificare chiavi e valori di tag e rimuovere tag da una risorsa in qualsiasi momento durante l'esecuzione di un cluster. Tuttavia, non puoi aggiungere, modificare o rimuovere tag da un cluster o da istanze terminati precedentemente associati a un cluster che è ancora attivo. Puoi inoltre impostare il valore di un tag su una stringa vuota, ma non su null.

Se utilizzi AWS Identity and Access Management (IAM) con le tue EC2 istanze Amazon per le autorizzazioni basate sulle risorse per tag, le tue policy IAM vengono applicate ai tag che Amazon EMR propaga alle istanze Amazon di un cluster. EC2 Affinché i tag Amazon EMR si propaghino alle tue EC2 istanze Amazon, la tua policy IAM per Amazon EC2 deve consentire le autorizzazioni per chiamare Amazon e. EC2 CreateTags DeleteTags APIs Inoltre, i tag propagati possono influire sulle autorizzazioni basate sulle risorse EC2 di Amazon. I tag propagati su Amazon EC2 possono essere letti come condizioni nella tua policy IAM, proprio come gli altri EC2 tag Amazon. Fai riferimento alla policy IAM quando aggiungi tag ai cluster Amazon EMR, in modo da evitare che gli utenti dispongano di autorizzazioni non corrette per un cluster. Per evitare problemi, assicurati che le policy IAM non includano condizioni sui tag che intendi utilizzare anche sui cluster Amazon EMR. Per ulteriori informazioni, consulta Controllare l'accesso alle EC2 risorse Amazon.

Restrizioni che si applicano all'etichettatura delle risorse in Amazon EMR

Ai tag si applicano le seguenti limitazioni di base:

- Le restrizioni che si applicano alle EC2 risorse Amazon si applicano anche ad Amazon EMR. Per ulteriori informazioni, consulta https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Using_Tags.html#tag-restrictions.
- Non utilizzare il `aws:` prefisso nei nomi e nei valori dei tag perché è AWS riservato all'uso. Inoltre, non puoi modificare né eliminare nomi o valori di tag con tale prefisso.
- Non puoi cambiare o modificare i tag su un cluster terminato.
- Un valore di tag può essere una stringa vuota, ma non nullo. Inoltre, una chiave di tag non può essere una stringa vuota.
- Chiavi e valori possono contenere qualsiasi carattere alfabetico in qualsiasi lingua, qualsiasi carattere numerico, spazi bianchi, separatori sottointesi e i seguenti simboli: _ . : / = + - @

Per ulteriori informazioni sull'utilizzo dei tag AWS Management Console, consulta Lavorare con i tag nella console nella Amazon EC2 User Guide. Per ulteriori informazioni sull'etichettatura tramite l' EC2API o la riga di comando di Amazon, consulta la panoramica di API e CLI nella EC2 Amazon User Guide.

Contrassegna le risorse Amazon EMR per la fatturazione

È possibile utilizzare i tag per organizzare la AWS fattura in modo che rispecchino la propria struttura dei costi. A tale scopo, registrati per ricevere una fattura sul tuo AWS account con i valori chiave dell'etichetta inclusi. Puoi quindi organizzare le informazioni di fatturazione in base ai valori di chiave di tag per visualizzare il costo delle risorse combinate. Sebbene Amazon EMR e Amazon EC2 abbiano dichiarazioni di fatturazione diverse, i tag di ogni cluster vengono posizionati anche su ogni istanza associata, in modo da poter utilizzare i tag per collegare i costi correlati di Amazon EMR e Amazon. EC2

Puoi ad esempio applicare tag a numerose risorse con un nome di applicazione specifico, quindi organizzare le informazioni di fatturazione per visualizzare il costo totale dell'applicazione in più servizi. Per ulteriori informazioni, consulta Assegnazione di tag e allocazione dei costi nella Guida per l'utente di AWS Billing .

Aggiungere tag a un cluster Amazon EMR

Puoi aggiungere tag ai cluster al momento della loro creazione.

Console

Aggiunta di tag durante la creazione di un cluster con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Tags (Tag), seleziona Add new tag (Aggiungi nuovo tag). Specifica un tag nel campo Key (Chiave). Facoltativamente, puoi specificare un tag nel campo Value (Valore).
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per aggiungere tag quando crei un cluster con AWS CLI

L'esempio seguente illustra come aggiungere un tag a un nuovo cluster mediante l' AWS CLI. Per aggiungere dei tag quando si crea un cluster, digita il sottocomando `create-cluster` con il parametro `--tags`.

- Per aggiungere un tag denominato *`costCenter`* key value *`marketing`* quando crei un cluster, digita il seguente comando e sostituiscilo *`myKey`* con il nome della tua coppia di EC2 chiavi.

```
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
```

Quando si specifica il numero di istanze senza utilizzare il parametro `--instance-groups`, viene avviato un singolo nodo master e le istanze rimanenti vengono avviate come nodi principali. Tutti i nodi utilizzeranno il tipo di istanza specificato nel comando.

> **Note**
>
> Se in precedenza non hai creato il ruolo del servizio EMR e il profilo di EC2 istanza predefiniti, digita aws `emr create-default-roles` per crearli prima di digitare il sottocomando. `create-cluster`

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Puoi anche aggiungere tag a un cluster esistente.

Console

Aggiunta di tag a un cluster esistente con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.

2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Tags (Tag) nella pagina dei dettagli del cluster, seleziona Manage tags (Gestisci tag). Specifica un tag nel campo Key (Chiave). Facoltativamente, puoi specificare un tag nel campo Value (Valore).
4. Seleziona Salva modifiche. La scheda Tags (Tag) si aggiorna con il nuovo numero di tag presenti nel cluster. Ad esempio, se ora hai due tag, l'etichetta della tua scheda è Tags (2).

AWS CLI

Per aggiungere tag a un cluster in esecuzione con AWS CLI

- Inserisci il sottocomando `add-tags` con il parametro `--tag` per assegnare tag a un ID del cluster. Puoi recuperare l'ID del cluster tramite la console o il comando `list-clusters`. Il sottocomando `add-tags` accetta attualmente un solo ID di risorsa.

  Ad esempio, per aggiungere due tag a un cluster in esecuzione, uno con una chiave denominata *`costCenter`* con un valore di *`marketing`* e un altro denominato *`other`* con un valore di*`accounting`*, immetti il comando seguente e sostituiscilo *`j-KT4XXXXXXXX1NM`* con il tuo ID del cluster.

  ```
  aws emr add-tags --resource-id j-KT4XXXXXXXX1NM --tag "costCenter=marketing" --
  tag "other=accounting"
  ```

  Nota che quando i tag vengono aggiunti utilizzando la AWS CLI, non viene generato alcun output dal comando. Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Visualizza i tag su un cluster Amazon EMR

Se desideri visualizzare tutti i tag associati a un cluster, puoi farlo nella console o nella AWS CLI.

Console

Visualizzazione dei tag di un cluster con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.

2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Per visualizzare tutti i tag, seleziona la scheda Tags (Tag) nella pagina dei dettagli del cluster.

AWS CLI

Per visualizzare i tag su un cluster con AWS CLI

Per visualizzare i tag su un cluster utilizzando il AWS CLI, digitate il `describe-cluster` sottocomando con il `--query` parametro.

- Per visualizzare i tag di un cluster, digita il comando seguente e sostituiscilo *`j-KT4XXXXXXXX1NM`* con il tuo ID del cluster.

```
aws emr describe-cluster --cluster-id j-KT4XXXXXX1NM --query Cluster.Tags
```

L'output visualizza tutte le informazioni sui tag del cluster in un formato simile al seguente:

```
Value: accounting     Value: marketing
Key: other            Key: costCenter
```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Rimuovere i tag da un cluster Amazon EMR

Se un tag non è più necessario, puoi rimuoverlo dal cluster.

Console

Rimozione dei tag di un cluster con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Tags (Tag) nella pagina dei dettagli del cluster, seleziona Manage tags (Gestisci tag).

4. Scegli Remove (Rimuovi) per ogni coppia chiave-valore che desideri rimuovere.
5. Scegli Save changes (Salva modifiche).

AWS CLI

Per rimuovere i tag su un cluster con AWS CLI

Digita il sottocomando `remove-tags` con il parametro `--tag-keys`. Quando si rimuove un tag, è richiesto solo il nome di chiave.

- Per rimuovere un tag da un cluster, digita il comando seguente e sostituiscilo *`j-KT4XXXXXXXX1NM`* con il tuo ID del cluster.

```
aws emr remove-tags --resource-id j-KT4XXXXXX1NM --tag-keys "costCenter"
```

> **Note**
>
> Non è attualmente possibile rimuovere più tag utilizzando un unico comando.

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Integrazione di driver e applicazioni di terze parti su Amazon EMR

Puoi eseguire diverse popolari applicazioni per i Big Data su Amazon EMR con prezzi di tipo utility: ciò significa che pagherai una tariffa oraria aggiuntiva nominale per l'applicazione di terze parti mentre il cluster è in esecuzione. In questo modo puoi utilizzare l'applicazione senza dover acquistare una licenza annuale. Le seguenti sezioni descrivono alcuni degli strumenti che è possibile utilizzare con EMR.

Argomenti

- Utilizzo degli strumenti di Business Intelligence con Amazon EMR

Utilizzo degli strumenti di Business Intelligence con Amazon EMR

Puoi utilizzare i più diffusi strumenti di business intelligence come Microsoft Excel e Tableau con Amazon EMR per esplorare e visualizzare i tuoi dati. MicroStrategy QlikView Molti di questi strumenti richiedono un driver ODBC (Open Database Connectivity) o JDBC (Open Database Connectivity). Per scaricare e installare i driver più recenti, consulta la pagina http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/.

Per trovare le versioni precedenti dei driver, consulta la pagina http://awssupportdatasvcs.com/bootstrap-actions/Simba/.

Sicurezza in Amazon EMR

La sicurezza e la conformità sono una responsabilità che condividi. AWS Questo modello di responsabilità condivisa può contribuire ad alleggerire il carico operativo in quanto AWS gestisce, gestisce e controlla i componenti dal sistema operativo host e dal livello di virtualizzazione fino alla sicurezza fisica delle strutture in cui operano i cluster EMR. Ti assumi la responsabilità, la gestione e l'aggiornamento dei cluster Amazon EMR, oltre a configurare il software applicativo e AWS i controlli di sicurezza forniti. Questa differenziazione di responsabilità viene comunemente definita sicurezza del cloud rispetto alla sicurezza nel cloud.

- Sicurezza del cloud: AWS è responsabile della protezione dell'infrastruttura in esecuzione Servizi AWS . AWS AWS ti offre anche servizi che puoi utilizzare in modo sicuro. I revisori di terze parti testano e verificano regolarmente l'efficacia della sicurezza come parte dei programmi di conformitàAWS. Per maggiori informazioni sui programmi di conformità che si applicano ad Amazon EMR, consulta Servizi AWS la sezione «Scope by compliance program».
- Sicurezza nel cloud: sei anche responsabile dell'esecuzione di tutte le attività di configurazione e gestione della sicurezza necessarie per proteggere un cluster Amazon EMR. I clienti che implementano un cluster Amazon EMR sono responsabili della gestione del software applicativo installato sulle istanze e della configurazione delle funzionalità fornite come i gruppi di sicurezza, AWS la crittografia e il controllo degli accessi in base ai requisiti, alle leggi e ai regolamenti applicabili.

La presente documentazione aiuta a comprendere come applicare il modello di responsabilità condivisa quando si utilizza Amazon EMR. Gli argomenti di questo capitolo mostrano come configurare Amazon EMR e utilizzarne altri Servizi AWS per raggiungere i tuoi obiettivi di sicurezza e conformità.

Sicurezza della rete e dell'infrastruttura

In quanto servizio gestito, Amazon EMR è protetto dalle procedure di sicurezza di rete AWS globali descritte nel white paper Amazon Web Services: panoramica dei processi di sicurezza. AWS i servizi di protezione della rete e dell'infrastruttura offrono protezioni granulari sia a livello di host che a livello di rete. Supporti Servizi AWS e funzionalità applicative di Amazon EMR che soddisfano i requisiti di protezione e conformità della rete.

- I gruppi EC2 di sicurezza di Amazon fungono da firewall virtuale per le istanze di cluster Amazon EMR, limitando il traffico di rete in entrata e in uscita. Per ulteriori informazioni, consulta Controllare il traffico di rete con gruppi di sicurezza.
- Amazon EMR block public access (BPA) impedisce l'avvio di un cluster in una sottorete pubblica se il cluster ha una configurazione di sicurezza che consente il traffico in entrata da indirizzi IP pubblici su una porta. Per ulteriori informazioni, consulta Using Amazon EMR block public access.
- Secure Shell (SSH) aiuta a fornire agli utenti un modo sicuro per connettersi alla riga di comando sulle istanze del cluster. È inoltre possibile utilizzare SSH per visualizzare le interfacce Web ospitate dalle applicazioni sul nodo master di un cluster. Per ulteriori informazioni, consulta Use an EC2 key pair for SSH credenziali e Connect to a cluster.

Aggiornamenti per l'AMI predefinita di Amazon Linux per Amazon EMR

> **Important**
>
> I cluster EMR che eseguono Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizzano il comportamento predefinito di Amazon Linux e non scaricano e installano automaticamente aggiornamenti del kernel importanti e critici che richiedono un riavvio. Questo è lo stesso comportamento delle altre EC2 istanze Amazon che eseguono l'AMI Amazon Linux predefinita. Se nuovi aggiornamenti software Amazon Linux che richiedono un riavvio (ad esempio, aggiornamenti del kernel, NVIDIA e CUDA) risultano disponibili dopo il rilascio di una versione di Amazon EMR, le istanze del cluster EMR che eseguono l'AMI predefinita non scaricano e installano automaticamente tali aggiornamenti. Per ottenere gli aggiornamenti del kernel, puoi personalizzare l'AMI di Amazon EMR per utilizzare l'AMI di Amazon Linux più recente.

A seconda dell'assetto di sicurezza dell'applicazione e la durata di esecuzione di un cluster, è possibile scegliere di riavviare periodicamente il cluster per applicare gli aggiornamenti di sicurezza, oppure creare un'operazione di bootstrap per personalizzare l'installazione dei pacchetti e degli aggiornamenti. È anche possibile scegliere di provare e quindi installare determinati aggiornamenti di sicurezza sulle istanze del cluster in esecuzione. Per ulteriori informazioni, consulta Utilizzo dell'AMI Amazon Linux predefinita per Amazon EMR. Tieni presente che la configurazione di rete

deve consentire l'uscita di HTTP e HTTPS verso i repository Linux in Amazon S3, altrimenti gli aggiornamenti di sicurezza non avranno esito positivo.

AWS Identity and Access Management con Amazon EMR

AWS Identity and Access Management (IAM) è un AWS servizio che aiuta un amministratore a controllare in modo sicuro l'accesso alle AWS risorse. Gli amministratori IAM controllano chi è authenticated (autenticato) (accesso effettuato) e authorized (autorizzato) (dispone di autorizzazioni) a utilizzare risorse Amazon EMR. Le identità IAM includono utenti, gruppi e ruoli. Un ruolo IAM è simile a un utente IAM, ma non è associato a una persona specifica ed è pensato per essere assunto da qualsiasi utente che necessiti di autorizzazioni. Per ulteriori informazioni, consulta AWS Identity and Access Management Amazon EMR. Amazon EMR utilizza più ruoli IAM per aiutarti a implementare i controlli di accesso per i cluster Amazon EMR. IAM è un AWS servizio che puoi utilizzare senza costi aggiuntivi.

- Ruolo IAM per Amazon EMR (ruolo EMR): controlla in che modo il servizio Amazon EMR è in grado di accedere ad altri Servizi AWS per tuo conto, ad esempio il provisioning delle istanze Amazon all'avvio EC2 del cluster Amazon EMR. Per ulteriori informazioni, consulta Configurare i ruoli del servizio IAM per le autorizzazioni e le risorse di Amazon EMR. Servizi AWS
- Ruolo IAM per EC2 le istanze del cluster (profilo dell'EC2 istanza): un ruolo assegnato a ogni EC2 istanza del cluster Amazon EMR all'avvio dell'istanza. I processi applicativi eseguiti sul cluster utilizzano questo ruolo per interagire con altri Servizi AWS, come Amazon S3. Per ulteriori informazioni, consulta il ruolo IAM per le EC2 istanze del cluster.
- Ruolo IAM per le applicazioni (ruolo di runtime): un ruolo IAM che puoi specificare quando invii un lavoro o una query a un cluster Amazon EMR. Il job o la query che invii al tuo cluster Amazon EMR utilizza il ruolo di runtime per accedere alle AWS risorse, come gli oggetti in Amazon S3. Puoi specificare i ruoli di runtime con Amazon EMR per i processi Spark e Hive. Utilizzando i ruoli di runtime, puoi isolare i job in esecuzione sullo stesso cluster utilizzando ruoli IAM diversi. Per ulteriori informazioni, consulta Utilizzo del ruolo IAM come ruolo di runtime con Amazon EMR.

Le identità della forza lavoro si riferiscono agli utenti che creano o gestiscono carichi di lavoro. AWS Amazon EMR fornisce supporto per le identità della forza lavoro con quanto segue:

- AWS IAM identity center (Idc) è consigliato Servizio AWS per gestire l'accesso degli utenti alle risorse. AWS È un unico posto in cui è possibile assegnare le identità della forza lavoro e accedere in modo coerente a più AWS account e applicazioni. Amazon EMR supporta le identità della forza lavoro tramite una propagazione affidabile delle identità. Grazie alla funzionalità affidabile

di propagazione delle identità, un utente può accedere all'applicazione e tale applicazione può passare l'identità dell'utente ad altri Servizi AWS per autorizzare l'accesso a dati o risorse. Per ulteriori informazioni, consulta la sezione Abilitazione del supporto per AWS IAM Identity Center con Amazon EMR.

Lightweight Directory Access Protocol (LDAP) è un protocollo applicativo standard di settore aperto, indipendente dal fornitore e per l'accesso e la gestione di informazioni su utenti, sistemi, servizi e applicazioni sulla rete. LDAP è comunemente usato per l'autenticazione degli utenti su server di identità aziendali come Active Directory (AD) e OpenLDAP. Abilitando LDAP con cluster EMR, consenti agli utenti di utilizzare le credenziali esistenti per autenticare e accedere ai cluster. Per ulteriori informazioni, consulta Attivazione del supporto per LDAP con Amazon EMR.

Kerberos è un protocollo di autenticazione di rete progettato per fornire un'autenticazione avanzata per client/server le applicazioni utilizzando la crittografia a chiave segreta. Quando usi Kerberos, Amazon EMR configura Kerberos per le applicazioni, i componenti e i sottosistemi che installa nel cluster in modo che siano autenticati tra loro. Per accedere a un cluster con Kerberos configurato, un principale kerberos deve essere presente nel Kerberos Domain Controller (KDC). Per ulteriori informazioni, consulta Abilitazione del supporto per Kerberos con Amazon EMR.

Cluster single-tenant e multi-tenant

Per impostazione predefinita, un cluster è configurato per una singola tenancy con il profilo EC2 Instance come identità IAM. In un cluster single-tenant, ogni job ha accesso completo e completo al cluster e l'accesso a tutte le Servizi AWS risorse viene effettuato sulla base del profilo dell' EC2 istanza. In un cluster multi-tenant, i tenant sono isolati gli uni dagli altri e non hanno accesso completo ai cluster e alle istanze del cluster. EC2 L'identità nei cluster multi-tenant è rappresentata dai ruoli di runtime o dagli identificativi della forza lavoro. In un cluster multi-tenant, puoi anche abilitare il supporto per il controllo granulare degli accessi (FGAC) tramite o Apache Ranger. AWS Lake Formation In un cluster con ruoli di runtime o FGAC abilitati, l'accesso al profilo Instance viene disabilitato anche tramite iptables. EC2

> **Important**
>
> Tutti gli utenti che hanno accesso a un cluster single-tenant possono installare qualsiasi software sul sistema operativo Linux (OS), modificare o rimuovere i componenti software installati da Amazon EMR e influire sulle EC2 istanze che fanno parte del cluster. Se vuoi assicurarti che gli utenti non possano installare o modificare le configurazioni di un cluster

Amazon EMR, ti consigliamo di abilitare la multi-tenancy per il cluster. Puoi abilitare la multi-tenancy su un cluster abilitando il supporto per runtime role, AWS IAM Identity Center, Kerberos o LDAP.

Protezione dei dati

Con AWS, puoi controllare i tuoi dati utilizzando Servizi AWS strumenti per determinare in che modo i dati sono protetti e chi può accedervi. Servizi come AWS Identity and Access Management (IAM) consentono di gestire in modo sicuro l'accesso Servizi AWS e le risorse. AWS CloudTrail consente il rilevamento e il controllo. Amazon EMR semplifica la crittografia dei dati inattivi in Amazon S3 utilizzando chiavi gestite AWS o completamente gestite da te. Amazon EMR supporta anche l'abilitazione della crittografia per i dati in transito. Per ulteriori informazioni, consulta crittografare i dati a riposo e in transito.

Controllo dell'accesso ai dati

Con il controllo dell'accesso ai dati, puoi controllare a quali dati può accedere un'identità IAM o un'identità della forza lavoro. Amazon EMR supporta i seguenti controlli di accesso:

- Policy basate sull'identità IAM: gestisci le autorizzazioni per i ruoli IAM che usi con Amazon EMR. Le policy IAM possono essere combinate con l'etichettatura per controllare l'accesso su base individuale. cluster-by-cluster Per ulteriori informazioni, consulta AWS Identity and Access Management Amazon EMR.
- AWS Lake Formationcentralizza la gestione delle autorizzazioni dei dati e ne semplifica la condivisione all'interno dell'organizzazione e all'esterno. Puoi usare Lake Formation per abilitare un accesso granulare a livello di colonna a database e tabelle nel Glue Data Catalog. AWS Per ulteriori informazioni, consulta Using AWS Lake Formation with Amazon EMR.
- L'accesso ad Amazon S3 concede identità di mappe, identità di mappe in directory come Active Directory o AWS Identity and Access Management (IAM) principal, ai set di dati in S3. Inoltre, l'accesso S3 garantisce l'identità dell'utente finale del registro e l'applicazione utilizzata per accedere ai dati S3 in. AWS CloudTrail Per ulteriori informazioni, consulta Usare le concessioni di accesso di Amazon S3 con Amazon EMR.
- Apache Ranger è un framework per abilitare, monitorare e gestire una sicurezza completa dei dati su tutta la piattaforma Hadoop. Amazon EMR supporta il controllo granulare degli accessi basato su Apache Ranger per Apache Hive Metastore e Amazon S3. Per ulteriori informazioni, consulta Integrazione di Apache Ranger con Amazon EMR.

Usa le configurazioni di sicurezza per configurare la sicurezza dei cluster Amazon EMR

Puoi utilizzare le configurazioni di sicurezza Amazon EMR per configurare la crittografia dei dati, l'autenticazione Kerberos e l'autorizzazione Amazon S3 per EMRFS sui tuoi cluster. Per prima cosa, crea una configurazione di sicurezza. Dopodiché, la configurazione di sicurezza sarà disponibile per l'utilizzo e il riutilizzo durante la creazione dei cluster.

Puoi usare AWS Management Console, the AWS Command Line Interface (AWS CLI) o the AWS SDKs per creare configurazioni di sicurezza. È inoltre possibile utilizzare un AWS CloudFormation modello per creare una configurazione di sicurezza. Per ulteriori informazioni, consulta la Guida per AWS CloudFormation l'utente e il modello di riferimento per AWS::EMR::SecurityConfiguration.

Argomenti

- Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI
- Specificare una configurazione di sicurezza per un cluster Amazon EMR

Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI

Questo argomento descrive le procedure generali per creare una configurazione di sicurezza con la console Amazon EMR e AWS CLI, seguite da un riferimento per i parametri che comprendono crittografia, autenticazione e ruoli IAM per EMRFS. Per ulteriori informazioni su queste caratteristiche, consulta i seguenti argomenti:

- Crittografa i dati inattivi e in transito con Amazon EMR
- Utilizzo di Kerberos per l'autenticazione con Amazon EMR
- Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3

Per creare una configurazione di sicurezza mediante la console

1. Apri la console Amazon EMR in /emr. https://console.aws.amazon.com
2. Nel riquadro di navigazione, scegliere Security Configurations (Configurazioni di sicurezza), quindi Create security configuration (Crea configurazione di sicurezza).
3. Digitare un nome in Name (Nome) per la configurazione di sicurezza.

4. Scegli le opzioni per Encryption (Crittografia) e Authentication (Autenticazione) come descritto nelle sezioni successive, quindi scegli Create (Crea).

Per creare una configurazione di sicurezza utilizzando AWS CLI

- Usa il comando `create-security-configuration` come mostrato nell'esempio seguente.
  - Per*SecConfigName*, specificare il nome della configurazione di sicurezza. Si tratta del nome che hai specificato alla creazione di un cluster che utilizza questa configurazione di sicurezza.
  - Per *SecConfigDef*, specifica una struttura JSON inline o il percorso di un file JSON locale *file://MySecConfig.json*. I parametri JSON definiscono le opzioni per Encryption (Crittografia), IAM Roles for EMRFS access to Amazon S3 (Ruoli IAM per l'accesso EMRFS ad Amazon S3) e Authentication (Autenticazione) come descritto nelle sezioni successive.

```
aws emr create-security-configuration --name "SecConfigName" --security-
configuration SecConfigDef
```

Configurazione della crittografia di dati

Prima di configurare la crittografia in una configurazione di sicurezza, è necessario creare le chiavi e i certificati utilizzati per la crittografia. Per ulteriori informazioni, consultare Fornitura di chiavi per crittografare i dati inattivi e Fornitura di certificati per la crittografia di dati in transito con Amazon EMR.

Quando crei una configurazione di sicurezza, specifichi due set di opzioni di crittografia: la crittografia di dati inattivi e la crittografia di dati in transito. Le opzioni per la crittografia di dati a riposo includono Amazon S3 con EMRFS e la crittografia per dischi locali. Le opzioni per la crittografia In transito abilitano le funzionalità di crittografia open source per determinate applicazioni che supportano il protocollo Transport Layer Security (TLS). Le opzioni per la crittografia inattiva e per quella in transito possono essere attivate insieme o separatamente. Per ulteriori informazioni, consulta Crittografa i dati inattivi e in transito con Amazon EMR.

> **Note**
>
> Quando si utilizza AWS KMS, vengono addebitati costi per l'archiviazione e l'uso delle chiavi di crittografia. Per ulteriori informazioni, consultare AWS KMS Prezzi.

Impostazione delle opzioni di crittografia mediante la console

Scegli le opzioni in Encryption (Crittografia) in base alle linee guida riportate di seguito.

- Scegli le opzioni in At rest encryption (Crittografia inattiva) per crittografare i dati archiviati nel file system.

  Puoi scegliere di crittografare i dati in Amazon S3, nei dischi locali o in entrambi.
- In S3 data encryption (Crittografia di dati S3), per Encryption mode (Modalità di crittografia), scegli un valore per determinare il modo in cui Amazon EMR crittografa i dati Amazon S3 con EMRFS.

  La fase successiva dipende dalla modalità di crittografia scelta:

  - SSE-S3

    Specifica la crittografia lato server con le chiavi di crittografia gestite da Amazon S3. Non è necessaria alcuna altra operazione in quanto Amazon S3 gestisce automaticamente le chiavi.
  - SSE-KMS o CSE-KMS

    Specifica la crittografia lato server con chiavi AWS KMS gestite (SSE-KMS) o la crittografia lato client con chiavi gestite (CSE-KMS). AWS KMS Per AWS KMS key, seleziona una chiave. La chiave deve esistere nella stessa regione del cluster EMR. Per i requisiti relativi alle chiavi, consulta Utilizzo AWS KMS keys per la crittografia.
  - CSE-Custom

    Specifica la crittografia lato client mediante una chiave root lato client personalizzata (CSE-Custom). Per S3 object (Oggetto S3), immetti il percorso in Amazon S3 oppure l'ARN Amazon S3 del file JAR del provider di chiavi personalizzato. Quindi, per la classe Key provider, inserisci il nome completo della classe dichiarata nell'applicazione che implementa l'interfaccia. EncryptionMaterialsProvider
- In Local disk encryption (Crittografia per dischi locali), scegli un valore per Key provider type (Tipo di provider di chiavi).

  - AWS KMS key

    Selezionare questa opzione per specificare una AWS KMS key. Per AWS KMS key, seleziona una chiave. La chiave deve esistere nella stessa regione del cluster EMR. Per ulteriori informazioni sui requisiti relativi alle chiavi, consulta Utilizzo AWS KMS keys per la crittografia.

    Crittografia EBS

Se lo specifichi AWS KMS come provider di chiavi, puoi abilitare la crittografia EBS per crittografare i dispositivi root e i volumi di archiviazione EBS. Per abilitare tale opzione, è necessario concedere al ruolo di servizio Amazon EMR `EMR_DefaultRole` le autorizzazioni per utilizzare la AWS KMS key specificata. Per ulteriori informazioni sui requisiti relativi alle chiavi, consulta Abilitazione della crittografia EBS fornendo autorizzazioni aggiuntive per le chiavi KMS.

- Personalizza

  Seleziona questa opzione per specificare un provider di chiavi personalizzato. Per S3 object (Oggetto S3), immetti il percorso in Amazon S3 oppure l'ARN Amazon S3 del file JAR del provider di chiavi personalizzato. Per la classe Key provider, inserisci il nome completo di una classe dichiarata nell'applicazione che implementa l'interfaccia. EncryptionMaterialsProvider Il nome di classe qui specificato deve essere diverso dal nome di classe fornito per CSE-Custom.

- Scegli In-transit encryption (Crittografia in transito) per abilitare le funzionalità della crittografia TLS open source per dati in transito. Scegli Certificate provider type (Tipo di provider di certificati) in base alle seguenti linee guida:

  - PEM

    Seleziona questa opzione per utilizzare i file PEM forniti in un file ZIP. Due artefatti devono essere inclusi nel file ZIP: privateKey.pem e certificateChain.pem. Un terzo file, trustedCertificates.pem, è facoltativo. Per informazioni dettagliate, consulta Fornitura di certificati per la crittografia di dati in transito con Amazon EMR. Per S3 object (Oggetto S3), specifica il percorso in Amazon S3 o l'ARN Amazon S3 del campo del file ZIP.

  - Personalizza

    Seleziona questa opzione per specificare un provider di certificati personalizzato, quindi, per S3 object (Oggetto S3), immetti il percorso in Amazon S3, o l'ARN Amazon S3, del file JAR del provider di certificati personalizzato. Per la classe Key provider, inserite il nome completo di una classe dichiarata nell'applicazione che implementa l'interfaccia TLSArtifacts Provider.

Specificare le opzioni di crittografia utilizzando AWS CLI

Le sezioni seguenti includono scenari di esempio per illustrare il JSON ben formato --security-configuration per differenti configurazioni e provider di chiavi nonché un riferimento per i parametri JSON e i valori appropriati.

Esempio di opzioni di crittografia di dati in transito

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è abilitata e la crittografia di dati inattivi è disabilitata.
- Un file ZIP contenente certificati in Amazon S3 è utilizzato come provider di chiavi (consulta Fornitura di certificati per la crittografia di dati in transito con Amazon EMR per i requisiti relativi ai certificati).

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": true,
  "EnableAtRestEncryption": false,
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "PEM",
    "S3Object": "s3://MyConfigStore/artifacts/MyCerts.zip"
   }
  }
 }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è abilitata e la crittografia di dati inattivi è disabilitata.
- È utilizzato un provider di chiavi personalizzato (vedi Fornitura di certificati per la crittografia di dati in transito con Amazon EMR per i requisiti relativi ai certificati).

```
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"
```

```
   }
  }
  }
}'
```

Esempio di opzioni di crittografia di dati a riposo

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è disabilitata e la crittografia di dati inattivi è abilitata.
- SSE-S3 è utilizzato per la crittografia di Amazon S3.
- La crittografia del disco locale viene utilizzata AWS KMS come fornitore di chiavi.

```
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"
   }
  }
  }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati in transito è abilitata e fa riferimento a un file ZIP con certificati PEM in Amazon S3 mediante l'ARN.
- SSE-KMS è utilizzato per la crittografia di Amazon S3.
- La crittografia del disco locale viene utilizzata AWS KMS come fornitore di chiavi.

```
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"
   }
  }
 }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati in transito è abilitata e fa riferimento a un file ZIP con certificati PEM in Amazon S3.
- CSE-KMS è utilizzato per la crittografia di Amazon S3.
- La crittografia per dischi locali utilizza un provider di chiavi personalizzato a cui si fa riferimento con il relativo 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"
   }
  }
 }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati in transito è abilitata con un provider di chiavi personalizzato.
- CSE-Custom è utilizzato per i dati Amazon S3.
- La crittografia per dischi locali utilizza un provider di chiavi personalizzato.

```
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"
   }
  }
 }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è disabilitata e la crittografia di dati inattivi è abilitata.
- La crittografia Amazon S3 è abilitata con SSE-KMS.
- Vengono utilizzate più AWS KMS chiavi, una per ogni bucket S3, e le eccezioni di crittografia vengono applicate a questi singoli bucket S3.
- La crittografia del disco locale è disabilitata.

```
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
  }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è disabilitata e la crittografia di dati inattivi è abilitata.
- La crittografia Amazon S3 è abilitata con SSE-S3 e la crittografia dei dischi locali è disabilitata.

```
aws emr create-security-configuration --name "MyS3EncryptionConfig" --security-
configuration '{
    "EncryptionConfiguration": {
        "EnableInTransitEncryption": false,
        "EnableAtRestEncryption": true,
        "AtRestEncryptionConfiguration": {
            "S3EncryptionConfiguration": {
                "EncryptionMode": "SSE-S3"
            }
        }
     }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è disabilitata e la crittografia di dati inattivi è abilitata.
- La crittografia del disco locale è abilitata AWS KMS come provider di chiavi e la crittografia Amazon S3 è disabilitata.

```
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"
            }
        }
     }
}'
```

L'esempio successivo illustra lo scenario seguente:

- La crittografia di dati In transito è disabilitata e la crittografia di dati inattivi è abilitata.
- La crittografia del disco locale è abilitata AWS KMS come provider di chiavi e la crittografia Amazon S3 è disabilitata.
- La crittografia EBS è abilitata.

```
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"
            }
        }
     }
}'
```

L'esempio successivo illustra lo scenario seguente:

SSE-EMR-WAL viene utilizzato per la crittografia EMR WAL

```
aws emr create-security-configuration --name "MySecConfig" \
    --security-configuration '{
        "EncryptionConfiguration": {
            "EMRWALEncryptionConfiguration":{ },
            "EnableInTransitEncryption":false, "EnableAtRestEncryption":false
        }
```

```
    }'
```

`EnableInTransitEncryption`e potrebbe `EnableAtRestEncryption` comunque essere vero, se si desidera abilitare la crittografia correlata.

L'esempio successivo illustra lo scenario seguente:

- SSE-KMS-WAL viene utilizzato per la crittografia EMR WAL
- La crittografia lato server viene utilizzata AWS Key Management Service come fornitore di chiavi

```
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`e potrebbe `EnableAtRestEncryption` comunque essere vero, se si desidera abilitare la crittografia correlata.

Riferimento JSON per impostazioni di crittografia

La tabella seguente elenca i parametri JSON per le impostazioni di crittografia e fornisce una descrizione dei valori accettabili per ogni parametro.

Parametro	Descrizione
`"EnableInTransitEncryption" : true | false`	Specificare `true` per abilitare la crittografia in transito e `false` per disabilitarla. Se omesso, per impostazione predefinita viene utilizzato `false` e la crittografia in transito viene disabilit ata.
`"EnableAtRestEncryption": true | false`	Specificare `true` per abilitare la crittografia inattiva e `false` per disabilitarla. Se omesso,

Parametro	Descrizione
	per impostazione predefinita viene utilizzato `false` e la crittografia inattiva viene disabilitata.
Parametri di crittografia in transito	
`"InTransitEncryptionConfiguration" :`	Specifica una raccolta di valori utilizzati per configurare la crittografia in transito quando `EnableInTransitEncryption` è `true`.
`"CertificateProviderType": "PEM" | "Custom"`	Specifica se utilizzare certificati `PEM` a cui si fa riferimento con un file zippato oppure un provider di certificati `Custom`. Se `PEM` specificato, `S3Object` deve essere un riferimento alla posizione in Amazon S3 di un file zip contenente i certificati. Se viene specificato Custom, `S3Object` deve essere un riferimento alla posizione in Amazon S3 di un file JAR, seguito da una `CertificateProviderClass` voce.
`"S3Object" : " ZipLocation "| "JarLocation "`	Fornisce la posizione in Amazon S3 a un file zip quando `PEM` specificato o a un file JAR quando `Custom` specificato. Il formato può essere un percorso (ad esempio, `s3://MyConfig/artifacts/CertFiles.zip`) oppure un ARN (ad esempio, `arn:aws:s3:::Code/MyCertProvider.jar`). Se si specifica un file ZIP, deve contenere file i cui nomi sono esattamente `privateKey.pem` e `certificateChain.pem`. Un file denominato `trustedCertificates.pem` è facoltativo.

Parametro	Descrizione
`"CertificateProviderClass" : "MyClassID "`	Obbligatorio solo `Custom` se specifica to per`CertificateProviderType` . *`MyClassID`* specifica un nome di classe completo dichiarato nel file JAR, che implementa l'interfaccia TLSArtifacts Provider. Ad esempio, `com.mycompany.MyCertProvider` .
Parametri di crittografia inattiva	
`"AtRestEncryptionConfiguration" :`	Specifica una raccolta di valori per la crittografia a riposo quando `EnableAtRestEncryption` è attiva`true`, inclusa la crittografia Amazon S3 e la crittografia del disco locale.
Parametri di crittografia Amazon S3	
`"S3EncryptionConfiguration" :`	Speciifica una raccolta di valori utilizzati per la crittografia di Amazon S3 con Amazon EMR File System (EMRFS).
`"EncryptionMode" : "SSE-S3" | "SSE-KMS" | "CSE-KMS" | "CSE-Custom"`	Speciifica il tipo di crittografia Amazon S3 da utilizzare. Se `SSE-S3` specificato, non sono richiesti altri valori di crittografia Amazon S3. Se `CSE-KMS` viene specificato uno dei due `SSE-KMS` o, è necessario specificare un AWS KMS key ARN come valore. `AwsKmsKey` Se `CSE-Custom` è specificato, i valori `S3Object` e `EncryptionKeyProviderClass` devono essere specificati.

Parametro	Descrizione
`"AwsKmsKey" : " MyKeyARN"`	Richiesto solo quando è specificato `SSE-KMS` o `CSE-KMS` per `EncryptionMode` . *`MyKeyARN`* deve essere un ARN completo per una chiave (ad esempio, `arn:aws:kms:us-east-1:123456789012:key/12345678-1234-1234-1234-123456789012`).
`"S3Object" : "JarLocation "`	Richiesto solo quando `CSE-Custom` è specificato per`CertificateProviderType` . *`JarLocation`* fornisce la posizione in Amazon S3 a un file JAR. Il formato può essere un percorso (ad esempio, `s3://MyConfig/artifacts/MyKeyProvider.jar`) oppure un ARN (ad esempio, `arn:aws:s3:::Code/MyKeyProvider.jar)` .
`"EncryptionKeyProviderClass" : "`*`MyS3KeyClassID`*` "`	Richiesto solo quando `CSE-Custom` è specificato per`EncryptionMode` . *`MyS3KeyClassID`* specifica il nome completo di una classe dichiarata nell'applicazione che implementa l' EncryptionMaterialsProvider interfaccia; ad esempio,. *`com.mycompany.MyS3KeyProvider`*
Parametri di crittografia per dischi locali	
`"LocalDiskEncryptionConfiguration"`	Specifica il provider di chiavi e i valori corrispondenti da utilizzare per la crittografia per dischi locali.

Parametro	Descrizione
`"EnableEbsEncryption": true | false`	Specificare `true` per abilitare la crittografia EBS. La crittografia EBS crittografa il volume del dispositivo root EBS e i volumi di archiviazione collegati. Per utilizzare la crittografia EBS, è necessario specificare come. `AwsKms` `EncryptionKeyProviderType`
`"EncryptionKeyProviderType": "AwsKms" | "Custom"`	Specifica il provider di chiavi. Se `AwsKms` è specificato, è necessario specificare un ARN della chiave KMS come `AwsKmsKey` valore. Se **`Custom`** è specificato, i valori `S3Object` e `EncryptionKeyProviderClass` devono essere specificati.
`"AwsKmsKey : " MyKeyARN"`	Richiesto solo quando `AwsKms` è specificato per. `Type` *`MyKeyARN`* deve essere un ARN completamente specificato per una chiave (ad esempio, `arn:aws:kms:us-east-1:123456789012:key/12345678-1234-1234-1234-456789012123`).
`"S3Object" : "JarLocation "`	Richiesto solo quando `CSE-Custom` è specificato per `CertificateProviderType` . *`JarLocation`* fornisce la posizione in Amazon S3 a un file JAR. Il formato può essere un percorso (ad esempio, `s3://MyConfig/artifacts/MyKeyProvider.jar`) oppure un ARN (ad esempio, `arn:aws:s3:::Code/MyKeyProvider.jar)` .

Parametro	Descrizione
`"EncryptionKeyProviderClass" : "MyLocalDiskKeyClassID "`	Richiesto solo quando `Custom` è specifica to perType. *`MyLocalDiskKeyClassID`* specifica il nome completo di una classe dichiarata nell'applicazione che implement a l' EncryptionMaterialsProviderinterfacc ia; ad esempio,. *`com.mycompany.MyLocalDiskKeyProvider`*
Parametri di crittografia EMR WAL	
`"EMRWALEncryptionConfiguration"`	Specifica il valore per la crittografia WAL EMR.
`"AwsKmsKey"`	Specifica l'ID della chiave CMK Arn.

Configurazione dell'autenticazione Kerberos

Una configurazione di sicurezza con impostazioni Kerberos può essere utilizzata da un cluster creato con attributi Kerberos, altrimenti si verifica un errore. Per ulteriori informazioni, consulta Utilizzo di Kerberos per l'autenticazione con Amazon EMR. Kerberos è disponibile solo in Amazon EMR versione 5.10.0 e versioni successive.

Configurazione delle impostazioni di Kerberos mediante la console

Scegliere le opzioni in Kerberos authentication (Autenticazione Kerberos) in base alle linee guida seguenti.

Parametro	Descrizione
Kerberos	Specifica che Kerberos è abilitato per i cluster che utilizzano questa configurazione di protezione. Se un cluster utilizza questa configurazione di protezione, deve avere anche le impostazioni Kerberos specificate o, in caso contrario, genererà un errore.

Parametro		Descrizione
Provider	KDC dedicato del cluster	Specifica che Amazon EMR crea un KDC sul nodo primario di qualsiasi cluster che utilizza questa configurazione di sicurezza. Quando crei il cluster, puoi specificare il nome del realm e la password di amministratore KDC. Se necessario, puoi fare riferimento a questo KDC da altri cluster. Crea tali cluster utilizzando una configura zione di sicurezza diversa, specifica un KDC esterno e utilizza il nome del realm e la password di amministr atore KDC specificati per il KDC dedicato al cluster.
	KDC esterno	Disponibili solo in Amazon EMR versione 5.20.0 e successive. Specifica che i cluster che utilizzan o questa configurazione di sicurezza autenticano le entità Kerberos principali utilizzando un server KDC esterno al cluster. Non viene creato un KDC nel cluster. Quando crei il cluster, specifichi il nome del realm e la password di amministratore KDC per il KDC esterno.
Durata del ticket		Facoltativo. Specifica il periodo di validità di un ticket Kerberos emesso dal KDC nei cluster che utilizzano questa configurazione di sicurezza. La durata dei ticket è limitata per motivi di sicurezza . Le applicazioni e i servizi del cluster rinnovano automaticamente i ticket dopo la loro scadenza. Gli utenti che si connettono al cluster tramite SSH utilizzan do le credenziali Kerberos devono eseguire `kinit` dalla linea di comando del nodo primario per rinnovare un ticket dopo la relativa scadenza.

Parametro		Descrizione
Fiducia tra realm		Specifica una relazione di fiducia tra realm tra un KDC dedicato al cluster in cluster che utilizzano questa configurazione di sicurezza e un KDC in un altro realm Kerberos. Le entità principali (in genere gli utenti) di un altro realm vengono autenticate nei cluster che utilizzano questa configurazione. È necessaria una configura zione aggiuntiva nell'altro realm Kerberos. Per ulteriori informazioni, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.
Proprietà di fiducia tra realm	Realm (Dominio)	Specifica il nome di realm Kerberos dell'altro realm della relazione di fiducia. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.
	Dominio	Specifica il nome di dominio dell'altro realm della relazione di fiducia.
	Server amministratore	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server di amministrazione nell'altro realm della relazione di fiducia. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :749`).

Parametro		Descrizione
	Server KDC	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server KDC nell'altro realm della relazione di fiducia. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :88`).
KDC esterno		Specifica che i cluster KDC esterni vengono utilizzati dal cluster.
Proprietà del KDC esterno	Server amministratore	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server amministratore esterno. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :749`).

Parametro			Descrizione
	Server KDC		Specifica il nome di dominio completo (FQDN) del server KDC esterno. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :88`).
	Integrazione con Active Directory		Specifica che l'autenticazione dell'entità Kerberos principale è integrata con un dominio Microsoft Active Directory.
	Proprietà di integrazione con Active Directory	Realm di Active Directory	Specifica il nome del realm Kerberos del dominio Active Directory. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.
		Dominio Active Directory	Specifica il nome del dominio di Active Directory.
		Server Active Directory	Specifica il nome di dominio completo (FQDN) del controller di dominio Microsoft Active Directory.

Specificare le impostazioni Kerberos utilizzando AWS CLI

La seguente tabella mostra i parametri JSON per le impostazioni di Kerberos in una configurazione di sicurezza. Per gli esempi di configurazione, consulta Esempi di configurazione.

Parametro	Descrizione
`"AuthenticationConfiguration": {`	Obbligatorio per Kerberos. Specifica che una configurazione di autentica zione fa parte di questa configura zione di sicurezza.
`"KerberosConfiguration": {`	Obbligatorio per Kerberos. Specifica le proprietà di configurazione Kerberos.

`"Provider": "ClusterD edicatedKdc",`
—oppure—
`"Provider: "ExternalKdc",`
	`ClusterDedicatedKdc` specifica che Amazon EMR crea un KDC sul nodo primario di qualsiasi cluster che utilizza questa configura zione di sicurezza. Quando crei il cluster, puoi specificare il nome del realm e la password di amministr atore KDC. Se necessario, puoi fare riferimento a questo KDC da altri cluster. Crea tali cluster utilizzan do una configurazione di sicurezza diversa, specifica un KDC esterno e utilizza il nome del realm e la password di amministratore KDC che hai specificato quando hai creato il cluster con il KDC dedicato al cluster.
	
	`ExternalKdc` specifica che il cluster utilizza un KDC esterno. Amazon EMR non crea un KDC sul nodo primario. Un cluster che utilizza questa configurazione di sicurezza deve specificare il nome del realm e la password di amministratore KDC del KDC esterno.

Parametro	Descrizione
`"ClusterDedicatedKdcConfiguration": {`	Opzione obbligatoria quando specifichi *`ClusterDedicatedKdc`* .
`"TicketLifetimeInHours": 24,`	Facoltativo. Specifica il periodo di validità di un ticket Kerberos emesso dal KDC nei cluster che utilizzano questa configurazione di sicurezza. La durata dei ticket è limitata per motivi di sicurezza. Le applicazioni e i servizi del cluster rinnovano automaticamente i ticket dopo la loro scadenza. Gli utenti che si connettono al cluster tramite SSH utilizzando le credenziali Kerberos devono eseguire `kinit` dalla linea di comando del nodo primario per rinnovare un ticket dopo la relativa scadenza.

Parametro				Descrizione
		`"CrossRealmTrustConfiguration": {`		Specifica una relazione di fiducia tra realm tra un KDC dedicato al cluster in cluster che utilizzano questa configurazione di sicurezza e un KDC in un altro realm Kerberos. Le entità principali (in genere gli utenti) di un altro realm vengono autenticate nei cluster che utilizzano questa configurazione. È necessari a una configurazione aggiuntiva nell'altro realm Kerberos. Per ulteriori informazioni, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.
			`"Realm": "KDC2.COM",`	Specifica il nome di realm Kerberos dell'altro realm della relazione di fiducia. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.
			`"Domain": "kdc2.com",`	Specifica il nome di dominio dell'altro realm della relazione di fiducia.

Parametro	Descrizione
"AdminServer": "*kdc.com:749* ",	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server di amministrazione nell'altro realm della relazione di fiducia. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com` *:749*).
"KdcServer": "*kdc.com:88* "	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server KDC nell'altro realm della relazione di fiducia. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com` *:88*).

Parametro	Descrizione
}	
}	
"ExternalKdcConfiguration": {	Opzione obbligatoria quando specifichi *ExternalKdc* .
"TicketLifetimeInHours": *24*,	Facoltativo. Specifica il periodo di validità di un ticket Kerberos emesso dal KDC nei cluster che utilizzano questa configurazione di sicurezza. La durata dei ticket è limitata per motivi di sicurezza. Le applicazioni e i servizi del cluster rinnovano automaticamente i ticket dopo la loro scadenza. Gli utenti che si connettono al cluster tramite SSH utilizzando le credenziali Kerberos devono eseguire `kinit` dalla linea di comando del nodo primario per rinnovare un ticket dopo la relativa scadenza.
"KdcServerType": "Single",	Specifica che viene fatto riferimento a un singolo server KDC. Attualmente, `Single` è l'unico valore supportato.

Parametro				Descrizione
			"AdminServer": "*kdc.com:749* ",	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server amministratore esterno. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com` *:749*).
			"KdcServer": "*kdc.com:88* ",	Specifica il nome di dominio completo (FQDN) del server KDC esterno. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com` *:88*).

Parametro	Descrizione
"AdIntegrationConfiguration": {	Specifica che l'autenticazione dell'entità Kerberos principale è integrata con un dominio Microsoft Active Directory.
"AdRealm": "*AD.DOMAIN.COM* ",	Specifica il nome del realm Kerberos del dominio Active Directory. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.
"AdDomain": "*ad.domain.com* "	Specifica il nome del dominio di Active Directory.
"AdServer": "*ad.domain.com* "	Specifica il nome di dominio completo (FQDN) del controller di dominio Microsoft Active Directory.
}	
}	
}	
}	

Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3

I ruoli IAM per EMRFS ti consentono di fornire differenti autorizzazioni per dati EMRFS in Amazon S3. A questo proposito, crei mappature che specificano un ruolo IAM utilizzato per le autorizzazioni quando una richiesta di accesso contiene un identificatore da te specificato. L'identificatore può essere un ruolo o un utente Hadoop oppure un prefisso Amazon S3.

Per ulteriori informazioni, consulta Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3.

Specificare i ruoli IAM per EMRFS utilizzando il AWS CLI

Di seguito è riportato un esempio di frammento JSON per specificare ruoli IAM personalizzati per EMRFS all'interno di una configurazione di sicurezza. Vengono illustrate le mappature dei ruoli per i tre diversi tipi di identificatori, seguite da un riferimento ai parametri.

```
{
  "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" ]
      }]
    }
  }
}
```

Parametro	Descrizione
`"AuthorizationConfiguration":`	Obbligatorio.
`"EmrFsConfiguration":`	Obbligatorio. Contiene mappature dei ruoli.
`"RoleMappings":`	Obbligatorio. Contiene una o più definizioni di mappatura dei ruoli. Le mappature dei ruoli vengono valutate dall'alto verso il basso nell'ordine in cui vengono visualizzate. Se una mappatura dei ruoli viene valutata come true (vera) per una chiamata EMRFS per i dati in Amazon S3, non vengono valutate altre mappature dei ruoli ed EMRFS utilizza il ruolo

Parametro	Descrizione
	IAM specificato per la richiesta. Le mappature dei ruoli sono costituite dai parametri obbligatori seguenti:
`"Role":`	Specifica l'identificatore ARN di un ruolo IAM nel formato `arn:aws:iam::` *`account-id`* `:role/`*`role-name`* . Questo è il ruolo IAM che Amazon EMR assume se la richiesta EMRFS ad Amazon S3 corrisponde a uno degli `Identifiers` specificato.
`"IdentifierType":`	Il valore può essere uno dei seguenti: • `"User"` specifica che gli identificatori sono uno o più utenti Hadoop, i quali possono essere utenti di account Linux o entità Kerberos. Quando la richiesta EMRFS viene originata dall'utente o dagli utenti specificati, viene assunto il ruolo IAM. • `"Prefix"` specifica che l'identificatore è un percorso Amazon S3. Il ruolo IAM viene assunto per le chiamate alla posizione o alle posizioni con i prefissi specificati. Ad esempio, il prefisso `s3://amzn-s3-demo-bucket/` corrisponde a `s3://amzn-s3-demo-bucket/mydir` e `s3://amzn-s3-demo-bucket/yetanotherdir` . • `"Group"` specifica che gli identificatori sono uno o più Gruppi Hadoop. Il ruolo IAM viene assunto se la richiesta proviene da un utente nel gruppo o nei gruppi specificati.

Parametro	Descrizione
`"Identifiers":`	Specifica uno o più identificatori del tipo di identificatore appropriato. Separa più identific atori con virgole senza spazi.

Configura le richieste di servizi di metadati per le istanze Amazon EC2

I metadati dell'istanza sono dati relativi all'istanza che puoi utilizzare per configurare o gestire un'istanza in esecuzione. Puoi accedere ai metadati dell'istanza da un'istanza in esecuzione utilizzando uno dei metodi seguenti:

- Instance Metadata Service versione 1 (IMDSv1): un metodo di richiesta/risposta
- Instance Metadata Service Version 2 (IMDSv2): un metodo orientato alla sessione

Sebbene Amazon EC2 supporti entrambi IMDSv1 e IMDSv2, Amazon EMR supporta IMDSv2 Amazon EMR 5.23.1, 5.27.1, 5.32 o versioni successive e 6.2 o versioni successive. In queste versioni, i componenti di Amazon EMR vengono utilizzati IMDSv2 per tutte le chiamate IMDS. Per le chiamate IMDS nel codice dell'applicazione, puoi utilizzare entrambi IMDSv1 e IMDSv2 oppure configurare l'IMDS in modo che venga utilizzato solo IMDSv2 per una maggiore sicurezza. Quando si specifica che IMDSv2 deve essere utilizzato, IMDSv1 non funziona più.

Per ulteriori informazioni, consulta Configurare il servizio di metadati dell'istanza nella Amazon EC2 User Guide.

> **Note**
>
> Nelle versioni precedenti di Amazon EMR 5.x o 6.x, la disattivazione IMDSv1 causa un errore di avvio del cluster poiché i componenti di Amazon EMR vengono utilizzati per tutte le chiamate IMDS. IMDSv1 Quando si spegne IMDSv1, assicurati che qualsiasi software personalizzato che utilizza sia aggiornato a. IMDSv1 IMDSv2

Specifica della configurazione del servizio di metadati dell'istanza utilizzando la AWS CLI

Di seguito è riportato un esempio di frammento JSON per specificare Amazon EC2 Instance Metadata Service (IMDS) all'interno di una configurazione di sicurezza. L'utilizzo di una configurazione di sicurezza personalizzata è facoltativo.

```
{
  "InstanceMetadataServiceConfiguration" : {
      "MinimumInstanceMetadataServiceVersion": integer,
      "HttpPutResponseHopLimit": integer
   }
}
```

Parametro	Descrizione
`"InstanceMetadataServiceConfiguration":`	Se non specifichi IMDS all'interno di una configurazione di sicurezza e utilizzi una release di Amazon EMR che lo richiede IMDSv1, per impostazione predefinita Amazon EMR IMDSv1 utilizza come istanza minima la versione del servizio di metadati. Se desideri utilizzare la tua configurazione, sono necessari entrambi i seguenti parametri.
`"MinimumInstanceMetadataServiceVersion":`	Obbligatorio. Specificare 1 o 2. Un valore di 1 allows IMDSv1 e IMDSv2. Un valore di 2 allows only IMDSv2.
`"HttpPutResponseHopLimit":`	Obbligatorio. Il limite di hop della risposta HTTP PUT per le richieste di metadati dell'istanza. Maggiore è il numero, più è lungo il tragitto che le richieste di metadati dell'istanza possono percorrere. Default: 1. Specifica un numero intero da 1 a 64.

Specifica della configurazione del servizio di metadati dell'istanza utilizzando la console

Puoi configurare l'utilizzo di IMDS per un cluster quando lo avvii dalla console di Amazon EMR.

Per configurare l'utilizzo di IMDS tramite la console:

1. Quando crei una nuova configurazione di sicurezza nella pagina Configurazioni di sicurezza, seleziona Configura il servizio di metadati dell' EC2EC2istanza nell'impostazione Instance Metadata Service. Questa configurazione è supportata solo in Amazon EMR 5.23.1, 5.27.1, 5.32 o versioni successive e 6.2 o versioni successive.
2. Per l'opzione Minimum Instance Metadata Service Version (Versione del servizio di metadati dell'istanza minima), seleziona una delle seguenti opzioni:
   - Disattiva IMDSv1 e consenti IMDSv2 solo se desideri consentire solo IMDSv2 su questo cluster. Consulta la sezione Transizione all'utilizzo del servizio di metadati delle istanze versione 2 nella Amazon EC2 User Guide.
   - Consenti entrambi IMDSv1 e IMDSv2 sul cluster, se lo desideri, IMDSv1 ed è orientato alla sessione IMDSv2 su questo cluster.
3. Infatti IMDSv2, puoi anche configurare il numero consentito di hop di rete per il token di metadati impostando il limite HTTP put response hop su un numero intero compreso tra e. 1 64

Per ulteriori informazioni, consulta Configurare il servizio di metadati dell'istanza nella Amazon EC2 User Guide.

Consulta Configurare i dettagli dell'istanza e Configurare il servizio di metadati dell'istanza nella Amazon EC2 User Guide.

Specificare una configurazione di sicurezza per un cluster Amazon EMR

Puoi definire impostazioni di crittografia quando crei un cluster specificando la configurazione di sicurezza. Puoi usare il AWS Management Console o il AWS CLI.

Console

Per specificare una configurazione di sicurezza con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Security configuration and permissions (Configurazione e autorizzazioni di sicurezza), cerca il campo Security configuration (Configurazione di sicurezza). Seleziona il menu a

discesa o scegli Browse (Sfoglia) per selezionare il nome di una configurazione di sicurezza che hai creato in precedenza. In alternativa, scegli Create security configuration (Crea configurazione di sicurezza) per creare una configurazione da utilizzare per il cluster.

4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

CLI

Per specificare una configurazione di sicurezza con AWS CLI

- Utilizza `aws emr create-cluster` per applicare facoltativamente una configurazione di sicurezza con `--security-configuration` *`MySecConfig`*, dove *`MySecConfig`* è il nome della configurazione di sicurezza, come mostrato nell'esempio seguente. Il valore `--release-label` specificato deve essere 4.8.0 o successivo e il parametro `--instance-type` può essere qualsiasi tipo disponibile.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-5.0.0 --
security-configuration mySecConfig
```

Protezione dei dati in Amazon EMR

Il modello di responsabilità AWS condivisa si applica alla protezione dei dati in Amazon EMR. Come descritto in questo modello, AWS è responsabile della protezione dell'infrastruttura globale che gestisce tutto il AWS cloud. L'utente è responsabile del controllo dei contenuti ospitati su questa infrastruttura. Questo contenuto include le attività di configurazione e gestione della sicurezza per AWS ciò che utilizzi. Per ulteriori informazioni sulla privacy dei dati, consulta Domande frequenti sulla privacy dei dati. Per informazioni sulla protezione dei dati in Europa, consulta il modello di responsabilità condivisa di Amazon e il post sul GDPR sul AWS Security Blog.

Ai fini della protezione dei dati, ti consigliamo di proteggere le credenziali AWS dell'account e di configurare account individuali con AWS Identity and Access Management. In questo modo, a ogni utente verranno assegnate solo le autorizzazioni necessarie per svolgere il proprio lavoro. Ti suggeriamo, inoltre, di proteggere i dati nei seguenti modi:

- Utilizza l'autenticazione a più fattori (MFA) con ogni account.
- Utilizza TLS per comunicare con AWS le risorse. È richiesto TLS 1.2.

- Configura l'API e la registrazione delle attività degli utenti con. AWS CloudTrail
- Utilizza soluzioni di AWS crittografia, insieme a tutti i controlli di sicurezza predefiniti all'interno AWS dei servizi.
- Utilizza i servizi di sicurezza gestiti avanzati, ad esempio Amazon Macie, che aiutano a individuare e proteggere i dati personali archiviati in Amazon S3.
- Se si richiedono moduli crittografici convalidati FIPS 140-2 quando si accede ad AWS tramite una CLI o un'API, utilizzare un endpoint FIPS. Per ulteriori informazioni sugli endpoint FIPS disponibili, consulta il Federal Information Processing Standard (FIPS) 140-2.

Consigliamo di non inserire mai informazioni identificative sensibili, ad esempio i numeri di account dei clienti, in campi a formato libero come un campo Nome. Ciò include quando lavori con Amazon EMR o altri AWS servizi utilizzando la console, l'API o. AWS CLI AWS SDKs Gli eventuali dati immessi in Amazon EMR o altri servizi potrebbero essere prelevati per l'inserimento nei registri di diagnostica. Quando fornisci un URL a un server esterno, non includere informazioni sulle credenziali nell'URL per convalidare la tua richiesta a tale server.

Crittografa i dati inattivi e in transito con Amazon EMR

La crittografia dei dati consente di impedire agli utenti non autorizzati di leggere dati su un cluster e sui sistemi di archiviazione di dati associati. Sono inclusi i dati salvati su supporti persistenti, noti come dati inattivi, e i dati che possono essere intercettati quando circolano in rete, noti come dati in transito.

A partire dalla versione Amazon EMR 4.8.0, puoi utilizzare le configurazioni di sicurezza di Amazon EMR per configurare più facilmente impostazioni di crittografia di dati per cluster. Le configurazioni di sicurezza offrono impostazioni che assicurano la sicurezza dei dati in transito e a riposo nei volumi di Amazon Elastic Block Store (Amazon EBS) e in EMRFS su Amazon S3.

Facoltativamente, a partire da Amazon EMR versione 4.1.0 e successive, puoi scegliere di configurare la crittografia trasparente in HDFS, che non è configurata utilizzando le configurazioni di sicurezza. Per ulteriori informazioni, consulta Crittografia trasparente in HDFS su Amazon EMR nella Guida ai rilasci di Amazon EMR.

Argomenti

- Opzioni di crittografia per Amazon EMR
- Crittografia inattiva utilizzando una chiave KMS del cliente per il servizio EMR WAL

- Crea chiavi e certificati per la crittografia dei dati con Amazon EMR
- Comprendere la crittografia in transito

Opzioni di crittografia per Amazon EMR

Con le versioni 4.8.0 e successive di Amazon EMR, puoi utilizzare una configurazione di sicurezza per specificare le impostazioni per crittografare i dati inattivi, i dati in transito o entrambi. Quando abiliti la crittografia dei dati a riposo, puoi scegliere di crittografare i dati EMRFS in Amazon S3, i dati in dischi locali o entrambi. Ogni configurazione di sicurezza creata viene archiviata in Amazon EMR anziché nella configurazione del cluster. Di conseguenza, puoi facilmente riutilizzare una configurazione per specificare impostazioni di crittografia dei dati ogni volta che crei un cluster. Per ulteriori informazioni, consulta Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI.

Il diagramma seguente mostra le differenti opzioni di crittografia dei dati disponibili con le configurazioni di sicurezza.

Anche le seguenti opzioni di crittografia sono disponibili e non sono configurate utilizzando una configurazione di sicurezza:

- Facoltativamente, con Amazon EMR versione 4.1.0 e versioni successive, puoi scegliere di configurare la crittografia trasparente in HDFS. Per ulteriori informazioni, consulta Crittografia trasparente in HDFS su Amazon EMR nella Guida ai rilasci di Amazon EMR.
- Se utilizzi una versione di Amazon EMR che non supporta configurazioni di sicurezza, puoi configurare manualmente la crittografia per i dati EMRFS in Amazon S3. Per ulteriori informazioni, consulta Specifica della crittografia Amazon S3 utilizzando le proprietà EMRFS.
- Se utilizzi una versione Amazon EMR precedente alla 5.24.0, un volume del dispositivo di root EBS crittografato è supportato solo quando utilizzi un'AMI personalizzata. Per ulteriori informazioni, consulta la sezione Creazione di un'AMI personalizzata con un volume del dispositivo di root Amazon EBS crittografato nella Guida alla gestione di Amazon EMR.

> **Note**
>
> A partire dalla versione 5.24.0 di Amazon EMR, puoi utilizzare un'opzione di configurazione di sicurezza per crittografare il dispositivo root e i volumi di storage EBS quando lo specifichi

come provider di chiavi. AWS KMS Per ulteriori informazioni, consulta Crittografia del disco locale.

La crittografia dei dati richiede chiavi e certificati. Una configurazione di sicurezza ti offre la flessibilità di scegliere tra diverse opzioni, tra cui chiavi gestite da AWS Key Management Service, chiavi gestite da Amazon S3 e chiavi e certificati di provider personalizzati da te forniti. Quando lo utilizzi AWS KMS come fornitore di chiavi, vengono addebitati costi per l'archiviazione e l'uso delle chiavi di crittografia. Per ulteriori informazioni, consulta Prezzi di AWS KMS.

Prima di specificare le opzioni di crittografia, scegli i sistemi di gestione di chiavi e certificati che intendi utilizzare, in modo da cominciare a creare chiavi e certificati o i provider personalizzati che specifichi durante l'impostazione dei parametri di crittografia.

Crittografia dei dati a riposo per dati EMRFS in Amazon S3

La crittografia Amazon S3 funziona con gli oggetti Amazon EMR File System (EMRFS) letti e scritti su Amazon S3. Puoi specificare la crittografia lato server (SSE) o la crittografia lato client (CSE) di Amazon S3 come modalità di crittografia predefinita quando abiliti la crittografia dei dati a riposo. Facoltativamente, è possibile specificare diversi metodi di crittografia per singoli bucket utilizzando override di crittografia per bucket. Indipendentemente dall'abilitazione o meno della crittografia di Amazon S3, il protocollo Transport Layer Security (TLS) esegue la crittografia degli oggetti EMRFS in transito tra i nodi cluster EMR e Amazon S3. Per ulteriori informazioni sulla crittografia di Amazon S3, consulta Protezione dei dati mediante crittografia nella Guida per l'utente di Amazon Simple Storage Service.

> **Note**
>
> Quando si utilizza AWS KMS, vengono addebitati costi per l'archiviazione e l'uso delle chiavi di crittografia. Per ulteriori informazioni, consultare AWS KMS Prezzi.

Crittografia lato server di Amazon S3

Tutti i bucket Amazon S3 hanno la crittografia configurata di default e tutti i nuovi oggetti caricati su un bucket S3 vengono crittografati automaticamente quando sono inattivi. Amazon S3 crittografa i dati a livello di oggetto mentre scrive i dati su disco e decrittografa i dati quando vi si accede. Per ulteriori informazioni sulla SEE, consulta Protezione dei dati con la crittografia lato server nella Guida per l'utente di Amazon Simple Storage Service.

Puoi scegliere tra due diversi sistemi di gestione delle chiavi quando specifichi la SSE in Amazon EMR:

- SSE-S3: Amazon S3 gestisce le chiavi per te.
- SSE-KMS: si utilizza un file AWS KMS key di configurazione con politiche adatte per Amazon EMR. Per ulteriori informazioni sui requisiti chiave per Amazon EMR, consulta Using AWS KMS keys for encryption.

La SSE con chiavi fornite dal cliente (SSE-C) non è disponibile per l'utilizzo con Amazon EMR.

File crittografato lato client Amazon S3

Con la crittografia lato client di Amazon S3, la crittografia e la decrittografia Amazon S3 avvengono nel client EMRFS nel tuo cluster. Gli oggetti vengono crittografati prima di essere caricati su Amazon S3 e decrittati dopo il loro download. Il provider specificato fornisce la chiave di crittografia utilizzata dal client. Il client può utilizzare le chiavi fornite da AWS KMS (CSE-KMS) o una classe Java personalizzata che fornisce la chiave principale lato client (CSE-C). Le specifiche di crittografia sono leggermente diverse tra CSE-KMS e CSE-C, a seconda del provider specificato e dei metadati dell'oggetto da decrittare o crittografare. Per ulteriori informazioni su queste differenze, consulta Protezione dei dati tramite la crittografia lato client nella Guida per l'utente di Amazon Simple Storage Service.

> **Note**
>
> Amazon S3 CSE garantisce solo che i dati EMRFS scambiati con Amazon S3 siano crittografati; non tutti i dati sui volumi delle istanze del cluster sono crittografati. Inoltre, poiché Hue non utilizza EMRFS, gli oggetti che il browser di file Hue S3 scrive su Amazon S3 non vengono crittografati.

Crittografia inattiva per i dati in Amazon EMR WAL

Quando configuri la crittografia lato server (SSE) per il write-ahead logging (WAL), Amazon EMR crittografa i dati inattivi. Puoi scegliere tra due diversi sistemi di gestione delle chiavi quando specifichi SSE in Amazon EMR:

SSE-EMR-WAL

Amazon EMR gestisce le chiavi per te. Per impostazione predefinita, Amazon EMR crittografa i dati archiviati in Amazon EMR WAL. SSE-EMR-WAL

SSE-KMS-WAL

Usa una AWS KMS chiave per configurare le politiche che si applicano a Amazon EMR WAL. Per ulteriori informazioni sulla configurazione della crittografia a riposo per EMR WAL utilizzando una chiave KMS del cliente, vedere Crittografia a riposo utilizzando una chiave KMS del cliente per il servizio EMR WAL.

> **Note**
>
> Non puoi usare la tua chiave con SSE quando abiliti WAL con Amazon EMR. Per ulteriori informazioni, consulta WRITE-ahead logs (WAL) per Amazon EMR.

Crittografia del disco locale

I seguenti meccanismi interagiscono per crittografare i dischi locali quando abiliti la crittografia dei dischi locali utilizzando una configurazione Amazon EMR di sicurezza.

Crittografia HDFS open source

HDFS scambia i dati tra le istanze del cluster durante l'elaborazione distribuita. Inoltre, legge e scrive i dati nei volumi instance store e nei volumi EBS collegati alle istanze. Le seguenti opzioni di crittografia Hadoop open source sono attivate quando abiliti la crittografia per dischi locali:

- Secure Hadoop RPC (RPC Hadoop protetto) è impostata su `Privacy`, che utilizza Simple Authentication Security Layer (SASL).
- Data encryption on HDFS block data transfer (Crittografia di dati su trasferimento di dati di blocco HDFS) è impostata su `true` ed è configurata per utilizzare la crittografia AES 256.

> **Note**
>
> Puoi attivare una crittografia Apache Hadoop supplementare abilitando la crittografia in transito. Per ulteriori informazioni, consulta Crittografia in transito. Queste impostazioni di

crittografia non attivano la crittografia trasparente HDFS, che puoi configurare manualmente. Per ulteriori informazioni, consulta Crittografia trasparente in HDFS su Amazon EMR nella Guida ai rilasci di Amazon EMR.

Crittografia dell'archivio istanza

Per i tipi di EC2 istanze che utilizzano NVMe based SSDs come volume di archiviazione dell'istanza, NVMe la crittografia viene utilizzata indipendentemente dalle impostazioni di crittografia di Amazon EMR. Per ulteriori informazioni, consulta i volumi NVMe SSD nella Amazon EC2 User Guide. Per altri volumi di archivio istanza, Amazon EMR utilizza LUKS per crittografare il volume di archivio istanza quando la crittografia su disco locale è abilitata, indipendentemente dal fatto che i volumi EBS siano crittografati utilizzando la crittografia EBS o LUKS.

Crittografia dei volumi EBS

Se crei un cluster in una regione in cui la EC2 crittografia Amazon dei volumi EBS è abilitata per impostazione predefinita per il tuo account, i volumi EBS vengono crittografati anche se la crittografia del disco locale non è abilitata. Per ulteriori informazioni, consulta Encryption by default nella Amazon EC2 User Guide. Con la crittografia locale del disco abilitata in una configurazione di sicurezza, le impostazioni di Amazon EMR hanno la precedenza sulle impostazioni di Amazon per EC2 encryption-by-default le istanze di cluster. EC2

Sono disponibili le seguenti opzioni per crittografare i volumi EBS utilizzando una configurazione di sicurezza:

- Crittografia EBS: a partire da Amazon EMR versione 5.24.0, puoi scegliere di abilitare la crittografia EBS. L'opzione di crittografia EBS crittografa il volume dispositivo root EBS e i volumi di storage collegati. L'opzione di crittografia EBS è disponibile solo se specifichi AWS Key Management Service come provider di chiavi. Ti consigliamo di utilizzare la crittografia EBS.
- Crittografia LUKS: se scegli di utilizzare la crittografia LUKS per i volumi Amazon EBS, la crittografia LUKS si applica solo ai volumi di archiviazione collegati, non al volume del dispositivo di root. Per ulteriori informazioni sulla crittografia LUKS, vedi la specifica di LUKS su disco.

  Per il tuo fornitore di chiavi, puoi configurare una classe AWS KMS key con policy adatte ad Amazon EMR o una classe Java personalizzata che fornisca gli artefatti di crittografia. Quando si utilizza AWS KMS, vengono addebitati costi per l'archiviazione e l'uso delle chiavi di crittografia. Per ulteriori informazioni, consulta Prezzi di AWS KMS.

> **Note**
>
> Per verificare se la crittografia EBS è abilitata sul cluster, è consigliabile utilizzare la chiamata API `DescribeVolumes`. Per ulteriori informazioni, consulta DescribeVolumes. L'esecuzione di `lsblk` sul cluster verificherà solo lo stato della crittografia LUKS anziché la crittografia EBS.

Crittografia in transito

Diversi meccanismi di crittografia sono abilitati con la crittografia in transito. Si tratta di funzionalità open source, sono specifiche dell'applicazione e possono variare in base alla versione di Amazon EMR. Per abilitare la crittografia in transito, usala the section called “Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI” in Amazon EMR. Per i cluster EMR con crittografia in transito abilitata, Amazon EMR configura automaticamente le configurazioni delle applicazioni open source per abilitare la crittografia in transito. Per i casi d'uso avanzati, puoi configurare direttamente le configurazioni delle applicazioni open source per sovrascrivere il comportamento predefinito in Amazon EMR. Per ulteriori informazioni, consulta la matrice di supporto per la crittografia in transito e la sezione Configurazione delle applicazioni.

Consulta quanto segue per ulteriori dettagli specifici sulle applicazioni open source relative alla crittografia in transito:

- Quando abiliti la crittografia in transito con una configurazione di sicurezza, Amazon EMR abilita la crittografia in transito per tutti gli endpoint di applicazioni open source che supportano la crittografia in transito. Il supporto per la crittografia in transito per diversi endpoint applicativi varia a seconda della versione di rilascio di Amazon EMR. Per ulteriori informazioni, consulta la matrice di supporto per la crittografia in transito.
- È possibile sovrascrivere le configurazioni open source, il che consente di effettuare le seguenti operazioni:
  - Disattiva la verifica del nome host TLS se i certificati TLS forniti dall'utente non soddisfano i requisiti
  - Disattiva la crittografia in transito per determinati endpoint in base ai requisiti di prestazioni e compatibilità
  - Controlla quali versioni TLS e suite di crittografia utilizzare.

Puoi trovare maggiori dettagli sulle configurazioni specifiche dell'applicazione nella matrice di supporto per la crittografia in transito

- Oltre ad abilitare la crittografia in transito con una configurazione di sicurezza, alcuni canali di comunicazione richiedono anche configurazioni di sicurezza aggiuntive per abilitare la crittografia in transito. Ad esempio, alcuni endpoint applicativi open source utilizzano Simple Authentication and Security Layer (SASL) per la crittografia in transito, che richiede che l'autenticazione Kerberos sia abilitata nella configurazione di sicurezza del cluster EMR. Per ulteriori informazioni su questi endpoint, consulta la matrice di supporto per la crittografia in transito.
- Ti consigliamo di utilizzare software che supporti TLS v1.2 o versioni successive. Amazon EMR offre la distribuzione JDK Corretto predefinita, che determina quali versioni TLS, suite di crittografia e dimensioni delle chiavi sono consentite dalle reti open source eseguite su Java. EC2 Al momento, la maggior parte dei framework open source applica TLS v1.2 o versioni successive per Amazon EMR 7.0.0 e versioni successive. Questo perché la maggior parte dei framework open source funziona su Java 17 per Amazon EMR 7.0.0 e versioni successive. Le versioni precedenti di Amazon EMR potrebbero supportare TLS v1.0 e v1.1 perché utilizzano versioni Java precedenti, ma Corretto JDK potrebbe modificare le versioni TLS supportate da Java, il che potrebbe influire sulle versioni esistenti di Amazon EMR.

Puoi specificare gli artifact di crittografia utilizzati per la crittografia di dati in transito in due modi: fornendo un file zippato di certificati che hai caricato in Amazon S3 oppure facendo riferimento a una classe Java personalizzata che fornisce artifact di crittografia. Per ulteriori informazioni, consulta Fornitura di certificati per la crittografia di dati in transito con Amazon EMR.

Crittografia inattiva utilizzando una chiave KMS del cliente per il servizio EMR WAL

I registri WAL (write-ahead logs) EMR forniscono supporto chiave KMS ai clienti. encryption-at-rest Di seguito sono riportati alcuni dettagli di alto livello su come Amazon EMR WAL è integrato con: AWS KMS

I registri di scrittura anticipata EMR (WAL) interagiscono AWS durante le seguenti operazioni:`CreateWAL`,,,,, `AppendEdit` `ArchiveWALCheckPoint` `CompleteWALFlush` `DeleteWAL` `GetCurrentWALTimeReplayEdits`, `TrimWAL` tramite l'impostazione `EMR_EC2_DefaultRole` predefinita Quando viene richiamata una delle operazioni precedenti elencate, il WAL EMR crea e contro la chiave KMS. `Decrypt` `GenerateDataKey`

Considerazioni

Considerate quanto segue quando utilizzate la crittografia AWS KMS basata per EMR WAL:

- La configurazione di crittografia non può essere modificata dopo la creazione di un WAL EMR.
- Quando utilizzi la crittografia KMS con la tua chiave KMS, la chiave deve esistere nella stessa regione del cluster Amazon EMR.
- Sei responsabile del mantenimento di tutte le autorizzazioni IAM richieste e si consiglia di non revocare le autorizzazioni necessarie durante il ciclo di vita del WAL. In caso contrario, causerà scenari di errore imprevisti, come l'impossibilità di eliminare EMR WAL, poiché la chiave di crittografia associata non esiste.
- L'utilizzo delle chiavi comporta un costo. AWS KMS Per ulteriori informazioni, consulta Prezzi di AWS Key Management Service.

Autorizzazioni IAM richieste

Per utilizzare la chiave KMS del cliente per crittografare EMR WAL a riposo, è necessario assicurarsi di impostare l'autorizzazione appropriata per il ruolo client EMR WAL e il principale del servizio EMR WAL. `emrwal.amazonaws.com`

Autorizzazioni per il ruolo client EMR WAL

Di seguito è riportata la politica IAM necessaria per il ruolo del client EMR WAL:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey*"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowKMSDecrypt"
```

```
    }
  ]
}
```

Il client EMR WAL sul cluster EMR verrà utilizzato per impostazione predefinita. `EMR_EC2_DefaultRole` Se utilizzi un ruolo diverso per il profilo dell'istanza nel cluster EMR, assicurati che ogni ruolo disponga delle autorizzazioni appropriate.

Per ulteriori informazioni sulla gestione della politica dei ruoli, consulta Aggiungere e rimuovere i permessi di identità IAM.

Autorizzazioni per la policy chiave KMS

È necessario fornire il ruolo del cliente EMR WAL e il servizio `Decrypt` e `GenerateDataKey*` l'autorizzazione EMR WAL nella politica KMS. Per ulteriori informazioni sulla gestione delle politiche chiave, consulta la politica chiave di KMS.

JSON

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

Il ruolo specificato nello snippet può cambiare se si modifica il ruolo predefinito.

Monitoraggio dell'interazione WAL di Amazon EMR con AWS KMS

Contesto di crittografia WAL di Amazon EMR

Un contesto di crittografia è un insieme di coppie chiave-valore che contiene dati arbitrari non segreti. Quando includi un contesto di crittografia in una richiesta di crittografia dei dati, associa AWS KMS crittograficamente il contesto di crittografia ai dati crittografati. lo stesso contesto di crittografia sia necessario per decrittografare i dati.

Nelle sue richieste GenerateDataKeye Decrypt a, AWS KMS Amazon EMR WAL utilizza un contesto di crittografia con coppie nome-valore che identificano il nome WAL EMR.

```
"encryptionContext": {
    "aws:emrwal:walname": "111222333444555-testworkspace-emrwalclustertest-
emrwaltestwalname"
}
```

Puoi utilizzare il contesto di crittografia per identificare queste operazioni crittografiche nei record e nei log di controllo, come AWS CloudTrail Amazon CloudWatch Logs, e come condizione per l'autorizzazione nelle politiche e nelle concessioni.

Crea chiavi e certificati per la crittografia dei dati con Amazon EMR

Prima di specificare le opzioni di crittografia utilizzando una configurazione di sicurezza, scegli il provider che desideri utilizzare per le chiavi e gli artefatti di crittografia. Ad esempio, puoi utilizzare AWS KMS o un provider personalizzato da te creato. Quindi, crea le chiavi o il provider di chiavi come descritto in questa sezione.

Fornitura di chiavi per crittografare i dati inattivi

Puoi utilizzare AWS Key Management Service (AWS KMS) o un provider di chiavi personalizzate per la crittografia dei dati a riposo in Amazon EMR. Quando si utilizza AWS KMS, vengono addebitati costi per l'archiviazione e l'uso delle chiavi di crittografia. Per ulteriori informazioni, consulta Prezzi di AWS KMS.

Questo argomento fornisce dettagli sulla policy delle chiavi per una chiave KMS da utilizzare con Amazon EMR, nonché linee guida ed esempi di codice per scrivere una classe di provider di chiavi personalizzato per la crittografia di Amazon S3. Per ulteriori informazioni sulla creazione di chiavi, consulta Creazione di chiavi nella Guida per gli sviluppatori di AWS Key Management Service .

Utilizzo AWS KMS keys per la crittografia

La chiave di AWS KMS crittografia deve essere creata nella stessa regione dell'istanza del cluster Amazon EMR e dei bucket Amazon S3 utilizzati con EMRFS. Se la chiave specificata si trova in un account diverso da quello utilizzato per configurare un cluster, è necessario specificare la chiave utilizzando il relativo ARN.

Il ruolo per il profilo dell' EC2 istanza Amazon deve disporre delle autorizzazioni per utilizzare la chiave KMS specificata. Il ruolo predefinito per il profilo dell'istanza in Amazon EMR è `EMR_EC2_DefaultRole`. Se utilizzi un ruolo diverso per il profilo dell'istanza o utilizzi ruoli IAM per le richieste EMRFS ad Amazon S3, assicurati che ogni ruolo venga aggiunto come utente chiave a seconda dei casi. Ciò concede al ruolo l'autorizzazione a utilizzare la chiave KMS. Per ulteriori informazioni, consulta Utilizzo di policy delle chiavi nella Guida per gli sviluppatori di AWS Key Management Service e Configurazione dei ruoli IAM per richieste EMRFS ad Amazon S3.

Puoi usare il AWS Management Console per aggiungere il tuo profilo di istanza o il tuo profilo di EC2 istanza all'elenco degli utenti chiave per la chiave KMS specificata, oppure puoi utilizzare il AWS CLI o un AWS SDK per allegare una policy chiave appropriata.

Nota che Amazon EMR supporta solo Chiavi KMS simmetriche. Non è possibile utilizzare una chiave KMS asimmetrica per crittografare i dati a riposo in un cluster Amazon EMR. Per informazioni su come determinare se una chiave KMS è simmetrica o asimmetrica, consulta Identificazione di chiavi KMS simmetriche e asimmetriche.

La procedura seguente descrive come aggiungere il profilo dell'istanza Amazon EMR predefinito, `EMR_EC2_DefaultRole` come utente di chiavi utilizzando la AWS Management Console. Presuppone che tu abbia già creato una chiave KMS. Per creare una nuova chiave KMS, consulta Creazione di chiavi nella Guida per gli sviluppatori di AWS Key Management Service .

Per aggiungere il profilo di EC2 istanza per Amazon EMR all'elenco degli utenti delle chiavi di crittografia

1. Accedi a AWS Management Console e apri la console AWS Key Management Service (AWS KMS) su https://console.aws.amazon.com/kms.
2. Per modificare il Regione AWS, usa il selettore della regione nell'angolo in alto a destra della pagina.
3. Seleziona l'alias della chiave KMS da modificare.
4. Nella pagina dei dettagli della chiave, in Key Users (Utenti di chiavi), scegli Add (Aggiungi).

5. Nella finestra di dialogo Add key users (Aggiungi utenti chiave) selezionare il ruolo appropriato. Il nome del ruolo di default è `EMR_EC2_DefaultRole`.
6. Scegliere Add (Aggiungi).

Abilitazione della crittografia EBS fornendo autorizzazioni aggiuntive per le chiavi KMS

A partire da Amazon EMR versione 5.24.0, puoi crittografare il dispositivo di root EBS e i volumi di archiviazione utilizzando un'opzione di configurazione di sicurezza. Per abilitare tale opzione, è necessario specificare AWS KMS come fornitore di chiavi. Inoltre, è necessario concedere al ruolo di servizio `EMR_DefaultRole` le autorizzazioni per utilizzare AWS KMS key quelle specificate.

È possibile utilizzare il AWS Management Console per aggiungere il ruolo di servizio all'elenco degli utenti chiave per la chiave KMS specificata oppure utilizzare il AWS CLI o un AWS SDK per allegare una politica di chiave appropriata.

La procedura seguente descrive come utilizzare per AWS Management Console aggiungere il ruolo di servizio Amazon EMR predefinito `EMR_DefaultRole` come utente chiave. Presuppone che tu abbia già creato una chiave KMS. Per creare una nuova chiave KMS, consulta Creazione di chiavi nella Guida per gli sviluppatori di AWS Key Management Service .

Per aggiungere il ruolo del servizio Amazon EMR all'elenco degli utenti delle chiavi di crittografia

1. Accedi a AWS Management Console e apri la console AWS Key Management Service (AWS KMS) su https://console.aws.amazon.com/kms.
2. Per modificare il Regione AWS, usa il selettore della regione nell'angolo in alto a destra della pagina.
3. Nella barra laterale sinistra, scegli Customer managed keys (Chiavi gestite dal cliente).
4. Seleziona l'alias della chiave KMS da modificare.
5. Nella pagina dei dettagli della chiave, in Key Users (Utenti di chiavi), scegli Add (Aggiungi).
6. Nella sezione Aggiungi utenti chiave, seleziona il ruolo appropriato. Il nome del ruolo di servizio predefinito per Amazon EMR è. `EMR_DefaultRole`
7. Scegli Aggiungi.

Creazione di un provider di chiavi personalizzato

Quando utilizzi una configurazione di sicurezza, devi specificare un nome di classe di provider differente per la crittografia per dischi locali e la crittografia di Amazon S3. I requisiti per il fornitore

di chiavi personalizzate dipendono dall'utilizzo della crittografia locale del disco e della crittografia Amazon S3, nonché della versione di rilascio di Amazon EMR.

A seconda del tipo di crittografia utilizzato per la creazione di un provider di chiavi personalizzate, l'applicazione deve inoltre implementare interfacce diverse EncryptionMaterialsProvider . Entrambe le interfacce sono disponibili nella versione AWS SDK for Java 1.11.0 e successive.

- Per implementare la crittografia Amazon S3, usa com.amazonaws.services.s3.model. EncryptionMaterialsProvider interfaccia.
- Per implementare la crittografia locale del disco, usa com.amazonaws.services.elasticmapreduce.spi.security. EncryptionMaterialsProvider interfaccia.

È possibile utilizzare qualsiasi strategia per fornire materiali di crittografia per l'implementazione. Ad esempio, è possibile scegliere di fornire materiali di crittografia statici o integrarli con un sistema di gestione delle chiavi più complesso.

Se utilizzi la crittografia Amazon S3, devi utilizzare gli algoritmi di crittografia AES/GCM/ NoPaddingper materiali di crittografia personalizzati.

Se utilizzi la crittografia del disco locale, l'algoritmo di crittografia da utilizzare per i materiali di crittografia personalizzati varia in base alla versione EMR. Per Amazon EMR 7.0.0 e versioni precedenti, è necessario utilizzare. AES/GCM/NoPadding Per Amazon EMR 7.1.0 e versioni successive, devi utilizzare AES.

La EncryptionMaterialsProvider classe ottiene materiali di crittografia in base al contesto di crittografia. Amazon EMR popola le informazioni sul contesto di crittografia al runtime per aiutare il chiamante a determinare i materiali di crittografia corretti da restituire.

Example Esempio: utilizzo di un provider di chiavi personalizzato per la crittografia Amazon S3 con EMRFS

Quando Amazon EMR recupera i materiali di crittografia dalla EncryptionMaterialsProvider classe per eseguire la crittografia, EMRFS compila facoltativamente l'argomento MaterialsDescription con due campi: l'URI Amazon S3 per l'oggetto JobFlowId e il cluster, che possono essere utilizzati dalla classe per restituire i materiali di crittografia in modo selettivo. EncryptionMaterialsProvider

Ad esempio, il provider potrebbe restituire chiavi diverse per diversi prefissi URI di Amazon S3. Sarà la descrizione dei materiali di crittografia restituiti a essere memorizzata con l'oggetto Amazon S3 anziché il valore materialsDescription generato da EMRFS e passato al provider. Durante la

decrittografia di un oggetto Amazon S3, la descrizione dei materiali di crittografia viene passata alla EncryptionMaterialsProvider classe, in modo che possa, ancora una volta, restituire selettivamente la chiave corrispondente per decrittografare l'oggetto.

Di seguito viene fornita un'implementazione di riferimento EncryptionMaterialsProvider . Un altro provider personalizzato è disponibile presso GitHub. EMRFSRSAEncryptionMaterialsProvider

```
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;
  }
}
```

Fornitura di certificati per la crittografia di dati in transito con Amazon EMR

Con Amazon EMR versione 4.8.0 o successive, sono disponibili due opzioni per specificare artifact di crittografia dei dati in transito utilizzando una configurazione di sicurezza:

- Puoi creare manualmente certificati PEM, includerli in un file .zip e quindi fare riferimento al file .zip in Amazon S3.
- Puoi implementare un provider di certificati personalizzato come classe Java, specificare il file JAR dell'applicazione in Amazon S3 e infine fornire il nome di classe completo del provider come dichiarato nell'applicazione. La classe deve implementare l'interfaccia TLSArtifactsProvider disponibile a partire dalla AWS SDK per Java versione 1.11.0.

Amazon EMR scarica automaticamente artifact in ogni nodo nel cluster e successivamente li utilizza per implementare le funzionalità di crittografia in transito open source. Per ulteriori informazioni sulle opzioni disponibili, consulta Crittografia in transito.

Utilizzo di certificati PEM

Quando specifichi un file ZIP per la crittografia in transito, per la configurazione di sicurezza i file PEM nel file ZIP devono essere denominati esattamente come illustrato di seguito:

Certificati di crittografia in transito

Nome file	Obbligatorio/facoltativo	Informazioni
privateKey.pem	Richiesto	Chiave privata

Nome file	Obbligatorio/facoltativo	Informazioni
certificateChain.pem	Richiesto	Catena di certificati
trustedCertificates.pem	Facoltativo	Si consiglia di fornire un certificato non firmato dalla Trusted Root Certification Authority (CA) predefinita di Java o da una CA intermedia che possa collegarsi alla CA root affidabile predefinita di Java. Non è consigliabile utilizzare il formato public CAs quando si utilizzano certificati wildcard o quando si disabilita la verifica del nome host.

È probabile che tu intenda configurare il file PEM della chiave privata affinché sia un certificato generico che consente l'accesso al dominio Amazon VPC in cui si trovano le istanze di cluster. Ad esempio, se il cluster risiede in us-east-1 (Virginia settentrionale), puoi scegliere di specificare un nome comune nella configurazione del certificato che autorizza l'accesso al cluster specificando `CN=*.ec2.internal` nella definizione di oggetto del certificato. Se il cluster risiede in us-west-2 (Oregon), puoi specificare `CN=*.us-west-2.compute.internal`.

Se il file PEM fornito nell'elemento di crittografia non contiene un carattere jolly per il dominio del nome comune, devi modificare il valore di to. `hadoop.ssl.hostname.verifier` `ALLOW_ALL` Per farlo nelle versioni 7.3.0 e successive di Amazon EMR, aggiungi la `core-site` classificazione quando invii le configurazioni a un cluster. Nelle versioni precedenti alla 7.3.0, aggiungi la configurazione `"hadoop.ssl.hostname.verifier": "ALLOW_ALL"` direttamente nel file. `core-site.xml` Questa modifica è necessaria perché il verificatore del nome host predefinito richiede un nome host senza il carattere jolly perché tutti gli host del cluster lo utilizzano. Per ulteriori informazioni sulla configurazione di cluster EMR in Amazon VPC, consulta Configurazione della rete in un VPC per Amazon EMR.

L'esempio seguente dimostra come utilizzare OpenSSL per generare un certificato X.509 autofirmato con una chiave privata RSA a 2048 bit. La chiave consente di accedere alle istanze del cluster

Amazon EMR dell'emittente nella Regione `us-west-2` (Oregon) come specificato dal nome di dominio *`*.us-west-2.compute.internal`* come nome comune.

Sono specificati altri elementi di oggetto facoltativi, come paese (C), stato (S) e lingua (L). Poiché viene generato un certificato autofirmato, il secondo comando dell'esempio copia il file `certificateChain.pem` nel file `trustedCertificates.pem`. Il terzo comando usa `zip` per creare il file `my-certs.zip` che contiene i certificati.

> **Important**
>
> Questo esempio è solo una proof-of-concept dimostrazione. L'utilizzo di certificati autofirmati non è consigliato e presenta un potenziale rischio per la sicurezza. Per i sistemi di produzione, utilizza un'autorità di certificazione attendibile per emettere certificati.

```
$ 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
```

Comprendere la crittografia in transito

È possibile configurare un cluster EMR per eseguire framework open source come Apache Spark, Apache Hive e Presto. Ciascuno di questi framework open source ha una serie di processi in esecuzione sulle istanze di un cluster. EC2 Ciascuno di questi processi può ospitare endpoint di rete per le comunicazioni di rete.

Se la crittografia in transito è abilitata su un cluster EMR, diversi endpoint di rete utilizzano meccanismi di crittografia diversi. Consulta le sezioni seguenti per saperne di più sugli endpoint di rete con framework open source specifici supportati dalla crittografia in transito, sui relativi meccanismi di crittografia e sulla versione di Amazon EMR che ha aggiunto il supporto. Ogni applicazione open source può inoltre avere best practice e configurazioni di framework open source diverse che puoi modificare.

Per una copertura ottimale della crittografia in transito, consigliamo di abilitare sia la crittografia in transito che Kerberos. Se abiliti solo la crittografia in transito, la crittografia in transito sarà disponibile solo per gli endpoint di rete che supportano TLS. Kerberos è necessario perché alcuni endpoint di

rete con framework open source utilizzano Simple Authentication and Security Layer (SASL) per la crittografia in transito.

Tieni presente che tutti i framework open source non supportati nelle versioni di Amazon EMR 7.x.x non sono inclusi.

Spark

Quando abiliti la crittografia in transito nelle configurazioni di sicurezza, `spark.authenticate` viene automaticamente impostata e utilizza la crittografia basata su AES per `true` le connessioni RPC.

A partire da Amazon EMR 7.3.0, se utilizzi la crittografia in transito e l'autenticazione Kerberos, non puoi utilizzare le applicazioni Spark che dipendono dal metastore Hive. Hive 3 risolve questo problema in HIVE-16340. HIVE-44114 risolve completamente questo problema quando Spark open source può essere aggiornato a Hive 3. Nel frattempo, puoi decidere di risolvere il problema. `hive.metastore.use.SSL` `false` Per ulteriori informazioni, consulta la sezione Configurazione delle applicazioni.

Per ulteriori informazioni, consulta la sicurezza di Spark nella documentazione di Apache Spark.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Spark History Server	spark.ssl.history.port	18480	TLS	emr-5,3,0+, emr-6,0+, emr-7,0+
Interfaccia utente di Spark	spark.ui.port	4440	TLS	emr-5,3,0+, emr-6,0+, emr-7,0+
Driver Spark	spark.driver.port	Dinamico	Crittografia basata su Spark AES	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Esecutore Spark	Executor Port (nessuna configurazione con nome)	Dinamico	Crittografia basata su Spark AES	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0+
FILATO NodeManager	spark.shuffle.service.porta 1	7337	Crittografia basata su Spark AES	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0+

1 è [ospitato su] `spark.shuffle.service.port` YARN ma è utilizzato solo da Apache NodeManager Spark.

Hadoop YARN

Secure Hadoop RPC è impostato e utilizza la crittografia in transito basata su SASL. `privacy` Ciò richiede che l'autenticazione Kerberos sia abilitata nella configurazione di sicurezza. Se non desideri la crittografia in transito per Hadoop RPC, configura. `hadoop.rpc.protection = authentication` Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

Se i tuoi certificati TLS non sono in grado di soddisfare i requisiti di verifica del nome host TLS, puoi configurarli. `hadoop.ssl.hostname.verifier = ALLOW_ALL` Ti consigliamo di utilizzare la configurazione predefinita di`hadoop.ssl.hostname.verifier = DEFAULT`, che impone la verifica del nome host TLS.

Per disabilitare HTTPS per gli endpoint dell'applicazione web YARN, configura. `yarn.http.policy = HTTP_ONLY` In questo modo il traffico verso questi endpoint rimane non crittografato. Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

Per ulteriori informazioni, consulta Hadoop in secure mode (Hadoop in modalità protetta) nella documentazione di Apache Hadoop.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
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+, emr-6.0,0+, emr-7.0+
ResourceManager	yarn.resourcemanager.scheduler.address	8030	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
ResourceManager	yarn.resourcemanager.address	8032	SASL + Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
ResourceManager	yarn.resourcemanager.admin.address	8033	SASL + Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
TimelineServer	yarn.timeline-service.indirizzo	10200	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
TimelineServer	yarn.timeline-service.webapp.address	8188	TLS	emr-7.3.0+
WebApplicationProxy	yarn.web-proxy.address	2088	SASL + Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
NodeManager	yarn.nodemanager.address	8041	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
NodeManager	yarn.nodemanager.localizer.address	8040	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
NodeManager	yarn.nodemanager.webapp.address	8044	TLS	emr-7.3.0+
NodeManager	mapreduce.shuffle.porta 1	13562	TLS	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
NodeManager	spark.shuffle.service.porta [2]	7337	Crittografia basata su Spark AES	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0+

1 è ospitato su `mapreduce.shuffle.port` YARN ma viene utilizzato solo da NodeManager Hadoop. MapReduce

[2] `spark.shuffle.service.port` è ospitato su YARN NodeManager ma viene utilizzato solo da Apache Spark.

Problema noto

La `yarn.log.server.url` configurazione in utilizza attualmente HTTP con porta 19888, che impedisce l'accesso ai log delle applicazioni dall'interfaccia utente di Resource Manager. Influisce sulla versione: da EMR - 7.3.0 a EMR - 7.8.0.

Utilizza la seguente soluzione alternativa:

1. Modifica la `yarn.log.server.url` configurazione `yarn-site.xml` per utilizzare il `HTTPS` protocollo e il numero di `19890` porta.
2. Riavvia YARN Resource Manager:`sudo systemctl restart hadoop-yarn-resourcemanager.service`.

Hadoop HDFS

Il nodo nome Hadoop, il nodo dati e il nodo journal supportano tutti TLS per impostazione predefinita se la crittografia in transito è abilitata nei cluster EMR.

Secure Hadoop RPC è impostato e utilizza la crittografia in transito basata su SASL. `privacy` Ciò richiede che l'autenticazione Kerberos sia abilitata nella configurazione di sicurezza.

Ti consigliamo di non modificare le porte predefinite utilizzate per gli endpoint HTTPS.

La crittografia dei dati sul trasferimento a blocchi HDFS utilizza AES 256 e richiede che la crittografia a riposo sia abilitata nella configurazione di sicurezza.

Per ulteriori informazioni, consulta Hadoop in secure mode (Hadoop in modalità protetta) nella documentazione di Apache Hadoop.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
NomeNode	indirizzo dfs.namenode.https-address	9871	TLS	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
Nodo del nome	dfs.namenode.rpc-indirizzo-pc	8020	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
Nodo dati	dfs.datanode.https.address	9865	TLS	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
Nodo dati	dfs.datanode.address	9866	SASL + Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
Nodo Journal	indirizzo dfs.journalnode.https-address	8481	TLS	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Nodo Journal	indirizzo dfs.journalnode.rpc	8485	SASL+ Kerberos	emr-4.8.0+, emr-5.0+, emr-6.0,0+, emr-7.0+
DFSZKFailoverControllore	dfs.ha.zkfc.port	8019	Nessuno	TLS per ZKFC è supportato solo in Hadoop 3.4.0. Per ulteriori informazioni, vedere HADOOP-18919. La versione 7.1.0 di Amazon EMR è attualmente su Hadoop 3.3.6. Le versioni successive di Amazon EMR saranno disponibili su Hadoop 3.4.0 in futuro

Hadoop MapReduce

Hadoop MapReduce, Job History Server e MapReduce shuffle supportano tutti TLS per impostazione predefinita quando la crittografia in transito è abilitata nei cluster EMR.

Lo shuffle crittografato Hadoop MapReduce utilizza TLS.

Ti consigliamo di non modificare le porte predefinite per gli endpoint HTTPS.

Per ulteriori informazioni, consulta Hadoop in secure mode (Hadoop in modalità protetta) nella documentazione di Apache Hadoop.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
JobHistoryServer	mapreduce .jobhisto ry.webapp .https.address	19890	TLS	emr-7.3.0+
FILATO NodeManager	mapreduce.shuffle. porta 1	13562	TLS	emr-4.8.0 +, emr-5.0+, emr-6.0,0+, emr-7.0+

1 è [1] ospitato su `mapreduce.shuffle.port` YARN ma viene utilizzato solo da NodeManager Hadoop. MapReduce

Presto

Nelle versioni 5.6.0 e successive di Amazon EMR, la comunicazione interna tra il coordinatore Presto e i lavoratori utilizza TLS Amazon EMR configura tutte le configurazioni richieste per consentire una comunicazione interna sicura in Presto.

Se il connettore utilizza il metastore Hive come archivio di metadati, anche la comunicazione tra il comunicatore e il metastore Hive viene crittografata con TLS.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Coordinatore Presto	http-server.https. port	8446	TLS	emr-5,6,0 +, emr-6,0+, emr-7,0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Lavoratore Presto	http-server.https.port	8446	TLS	emr-5,6,0+, emr-6,0+, emr-7,0+

Trino

Nelle versioni 6.1.0 e successive di Amazon EMR, la comunicazione interna tra il coordinatore di Presto e i lavoratori utilizza TLS Amazon EMR configura tutte le configurazioni richieste per consentire una comunicazione interna sicura in Trino.

Se il connettore utilizza il metastore Hive come archivio di metadati, anche la comunicazione tra il comunicatore e il metastore Hive viene crittografata con TLS.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Coordinatore di Trino	http-server.https.port	8446	TLS	emr-6,1+, emr-7,0+
Operaio di Trino	http-server.https.port	8446	TLS	emr-6,1+, emr-7,0+

Hive e Tez

Per impostazione predefinita, Hive server 2, Hive metastore server, Hive LLAP Daemon Web UI e Hive LLAP shuffle supportano tutti TLS quando la crittografia in transito è abilitata nei cluster EMR. Per ulteriori informazioni sulle configurazioni Hive, vedere Proprietà di configurazione.

L'interfaccia utente Tez ospitata sul server Tomcat è anche abilitata per HTTPS quando la crittografia in transito è abilitata nel cluster EMR. Tuttavia, HTTPS è disabilitato per il servizio

di interfaccia utente web Tez AM, quindi gli utenti AM non hanno accesso al file keystore per il listener SSL di apertura. È inoltre possibile abilitare questo comportamento con le configurazioni booleane e. `tez.am.tez-ui.webservice.enable.ssl` `tez.am.tez-ui.webservice.enable.client.auth`

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
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+
HiveServer2	hive.server2.webui.port	10002	TLS	emr-7.3.0+
HiveMetastoreServer	hive.metastore.port	9083	TLS	emr-7.3.0+
Demone LLAP	hive.llap.daemon.yarn.shuffle.port	1551	TLS	emr-7.3.0+
Demone LLAP	hive.llap.daemon.web.port	15002	TLS	emr-7.3.0+
Demone LLAP	hive.llap.daemon.output.service.port	15003	Nessuno	Hive non supporta la crittografia in transito per questo endpoint

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Demone LLAP	hive.llap.management.rpc.port	15004	Nessuno	Hive non supporta la crittografia in transito per questo endpoint
Demone LLAP	hive.llap.plugin.rpc.port	Dinamico	Nessuno	Hive non supporta la crittografia in transito per questo endpoint
Demone LLAP	hive.llap.daemon.rpc.port	Dinamico	Nessuno	Hive non supporta la crittografia in transito per questo endpoint
Web HCat	templeton.port	5011	TLS	emr-7.3.0+
Tez Application Master	tez.am.client.am.port-range tez.am.task.am.port-range	Dinamico	Nessuno	Tez non supporta la crittografia in transito per questo endpoint

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Tez Application Master	tez.am.tez-ui.webservice.port-range	Dinamico	Nessuno	Disabilitato per impostazione predefinita. Può essere abilitato utilizzando le configurazioni Tez in emr-7.3.0+
Tez Task	N/A: non configurabile	Dinamico	Nessuno	Tez non supporta la crittografia in transito per questo endpoint
Interfaccia utente Tez	Configurabile tramite il server Tomcat su cui è ospitata l'interfaccia utente Tez	8080	TLS	emr-7.3.0+

Flink

Gli endpoint REST Apache Flink e la comunicazione interna tra i processi flink supportano TLS per impostazione predefinita quando si abilita la crittografia in transito nei cluster EMR.

`security.ssl.internal.enabled`è impostato `true` e utilizza la crittografia in transito per la comunicazione interna tra i processi Flink. Se non desideri la crittografia in transito per le comunicazioni interne, disabilita tale configurazione. Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

Amazon EMR imposta `true` e utilizza `security.ssl.rest.enabled`la crittografia in transito per gli endpoint REST. Inoltre, Amazon EMR imposta su true `historyserver.web.ssl.enabled`per utilizzare la comunicazione TLS con il server di cronologia Flink. Se non desideri la crittografia in transito per i punti REST, disabilita queste configurazioni. Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

Amazon EMR utilizza `security.ssl.algorithms`. per specificare l'elenco di cifrari che utilizzano la crittografia basata su AES. Sostituisci questa configurazione per utilizzare i codici che desideri.

Per ulteriori informazioni, consulta SSL Setup nella documentazione di Flink.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Flink History Server	historyserver.web.port	8082	TLS	emr-7.3.0+
Server Rest di Job Manager	rest.bind-port porta di riposo	Dinamico	TLS	emr-7.3.0+

HBase

Amazon EMR imposta Secure Hadoop RPC su. `privacy` HMaster e RegionServer utilizza la crittografia in transito basata su SASL. Ciò richiede che l'autenticazione Kerberos sia abilitata nella configurazione di sicurezza.

Amazon EMR imposta su true e utilizza TLS `hbase.ssl.enabled` per gli endpoint dell'interfaccia utente. Se non desideri utilizzare TLS per gli endpoint dell'interfaccia utente, disabilita questa configurazione. Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

Amazon EMR imposta `hbase.rest.ssl.enabled` `hbase.thrift.ssl.enabled` e utilizza TLS per gli endpoint dei server REST e Thirft, rispettivamente. Se non desideri utilizzare TLS per questi endpoint, disabilita questa configurazione. Ti consigliamo di utilizzare la configurazione predefinita per la massima sicurezza.

A partire da EMR 7.6.0, TLS è supportato sugli endpoint. HMaster RegionServer Amazon EMR imposta `hbase.server.netty.tls.enabled` anche e. `hbase.client.netty.tls.enabled` Se non desideri utilizzare TLS per questi endpoint, disabilita questa configurazione. Ti consigliamo di utilizzare la configurazione predefinita, che fornisce la crittografia e quindi una maggiore sicurezza. Per ulteriori informazioni, consulta Transport Level Security (TLS) nella comunicazione HBase RPC nella Apache HBase Reference Guide.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
HMaster	HMaster	16000	SASL + Kerberos TLS	SASL + Kerberos in emr-4.8.0+, emr-5.0.0+, emr-6.0.0+ ed emr-7.0.0+ TLS in emr-7.6.0 +
HMaster	HMaster INTERFACCIA UTENTE	16010	TLS	emr-7.3.0+
RegionServer	RegionServer	16020	SASL+ Kerberos TLS	SASL + Kerberos in emr-4.8.0+, emr-5.0.0+, emr-6.0.0+ ed emr-7.0.0+ TLS in emr-7.6.0 +
RegionServer		16030	TLS	emr-7.3.0+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
	RegionServer Informazioni			
HBase Server REST	Server di riposo	8070	TLS	emr-7.3.0+
HBase Server REST	Interfaccia utente Rest	8085	TLS	emr-7.3.0+
Hbase Thrift Server	Server Thrift	9090	TLS	emr-7.3.0+
Hbase Thrift Server	Interfaccia utente di Thrift Server	9095	TLS	emr-7.3.0+

Phoenix

Se hai abilitato la crittografia in transito nel tuo cluster EMR, Phoenix Query Server supporta la `phoenix.queryserver.tls.enabled` proprietà TLS, che è impostata su di default. `true`

Per ulteriori informazioni, consulta Configurazioni relative a HTTPS nella documentazione di Phoenix Query Server.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Server di interrogazione	phoenix.queryserver.http.port	8765	TLS	emr-7.3.0+

Oozie

OOZIE-3673 è disponibile su Amazon EMR se esegui Oozie su Amazon EMR 7.3.0 e versioni successive. Se devi configurare protocolli SSL o TLS personalizzati quando esegui un'azione e-mail, puoi impostare la proprietà nel file. `oozie.email.smtp.ssl.protocols` `oozie-site.xml` Per impostazione predefinita, se hai abilitato la crittografia in transito, Amazon EMR utilizza il protocollo TLS v1.3.

OOZIE-3677 e OOZIE-3674 sono disponibili anche su Amazon EMR se esegui Oozie su Amazon EMR 7.3.0 e versioni successive. Ciò consente di specificare `keyStoreType` `trustStoreType` le `oozie-site.xml` proprietà e le impostazioni. OOZIE-3674 aggiunge il parametro `--insecure` al client Oozie in modo che possa ignorare gli errori dei certificati.

Oozie applica la verifica del nome host TLS, il che significa che qualsiasi certificato utilizzato per la crittografia in transito deve soddisfare i requisiti di verifica del nome host. Se il certificato non soddisfa i criteri, il cluster potrebbe rimanere bloccato nella `oozie share lib update` fase in cui Amazon EMR effettua il provisioning del cluster. Ti consigliamo di aggiornare i certificati per assicurarti che siano conformi alla verifica del nome host. Tuttavia, se non riesci ad aggiornare i certificati, puoi disabilitare SSL per Oozie impostando la `oozie.https.enabled` proprietà su nella configurazione del cluster. `false`

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
EmbeddedOozieServer	oozie.https.port	11443	TLS	emr-7.3.0+
EmbeddedOozieServer	oozie.email.smtp.port	25	TLS	emr-7.3.0+

Hue

Per impostazione predefinita, Hue supporta TLS quando la crittografia in transito è abilitata nei cluster Amazon EMR. Per ulteriori informazioni sulle configurazioni Hue, consulta Configurare Hue con HTTPS/SSL.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Hue	http_port	8888	TLS	emr-7,4+

Livy

Per impostazione predefinita, Livy supporta TLS quando la crittografia in transito è abilitata nei cluster Amazon EMR. Per ulteriori informazioni sulle configurazioni Livy, consulta Abilitazione di HTTPS con Apache Livy.

A partire da Amazon EMR 7.3.0, se utilizzi la crittografia in transito e l'autenticazione Kerberos, non puoi utilizzare il server Livy per le applicazioni Spark che dipendono dal metastore Hive. Questo problema è stato risolto in HIVE-16340 e completamente risolto in SPARK-44114 quando l'applicazione open source Spark può essere aggiornata a Hive 3. Nel frattempo, puoi aggirare il problema se lo imposti. `hive.metastore.use.SSL` `false` Per ulteriori informazioni, consulta la sezione Configurazione delle applicazioni.

Per ulteriori informazioni, consulta Abilitazione di HTTPS con Apache Livy.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
livy-server	livy.server.port	8998	TLS	emr-7.4.0+

JupyterEnterpriseGateway

Per impostazione predefinita, Jupyter Enterprise Gateway supporta TLS quando la crittografia in transito è abilitata nei cluster Amazon EMR. Per ulteriori informazioni sulle configurazioni di Jupyter Enterprise Gateway, consulta Securing Enterprise Gateway Server.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
jupyter_enterprise_gateway	c. porta EnterpriseGatewayApp	9547	TLS	emr-7,4+

JupyterHub

Per impostazione predefinita, JupyterHub supporta TLS quando la crittografia in transito è abilitata nei cluster Amazon EMR. Per ulteriori informazioni, consulta Attivazione della crittografia SSL nella documentazione. JupyterHub Non è consigliabile disabilitare la crittografia.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
jupyter_hub	c. porta JupyterHub	9443	TLS	emr-5.14.0+, emr-6.0+, emr-7.0+

Zeppelin

Per impostazione predefinita, Zeppelin supporta TLS quando abiliti la crittografia in transito nel tuo cluster EMR. Per ulteriori informazioni sulle configurazioni Zeppelin, consulta Configurazione SSL nella documentazione di Zeppelin.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
zeppelin	zeppelin.server.ssl.porta	8890	TLS	7,3,0+

Zookeeper

Amazon EMR si configura `serverCnxnFactory` per abilitare il TLS `org.apache.zookeeper.server.NettyServerCnxnFactory` per il quorum di Zookeeper e la comunicazione con i clienti.

`secureClientPort`specifica la porta che ascolta le connessioni TLS. Se il client non supporta le connessioni TLS a Zookeeper, i client possono connettersi alla porta non sicura 2181 specificata in. `clientPort` Puoi sovrascrivere o disabilitare queste due porte.

Amazon EMR imposta entrambi `sslQuorum` e `admin.forceHttps` per abilitare la comunicazione TLS `true` per il quorum e il server di amministrazione. Se non desideri la crittografia in transito per il quorum e il server di amministrazione, puoi disabilitare tali configurazioni. Ti consigliamo di utilizzare le configurazioni predefinite per la massima sicurezza.

Per ulteriori informazioni, consulta Crittografia, Autenticazione, Opzioni di autorizzazione nella documentazione di Zookeeper.

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Server Zookeeper	secureClientPort	2281	TLS	emr-7,4+
Server Zookeeper	Porte Quorum	Ce ne sono 2: I follower usano 2888 per connettersi al leader. L'elezione del leader utilizza 3888	TLS	emr-7.4.0+
		8341	TLS	emr-7,4+

Componente	Endpoint	Porta	Meccanismo di crittografia in transito	Supportato dalla versione
Server Zookeeper	Porta admin.server			

AWS Identity and Access Management per Amazon EMR

AWS Identity and Access Management (IAM) è un software Servizio AWS che aiuta un amministratore a controllare in modo sicuro l'accesso alle AWS risorse. Gli amministratori IAM controllano chi è authenticated (autenticato) (accesso effettuato) e authorized (autorizzato) (dispone di autorizzazioni) a utilizzare risorse Amazon EMR. IAM è un software Servizio AWS che puoi utilizzare senza costi aggiuntivi.

Argomenti

- Destinatari
- Autenticazione con identità
- Gestione dell'accesso con policy
- Funzionamento di Amazon EMR con IAM
- Ruoli di runtime per le fasi di Amazon EMR
- Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS
- Esempi di policy basate su identità per Amazon EMR

Destinatari

Il modo in cui utilizzi AWS Identity and Access Management (IAM) varia a seconda del lavoro svolto in Amazon EMR.

Utente del servizio: se utilizzi il servizio Amazon EMR per eseguire il tuo lavoro, l'amministratore fornisce le credenziali e le autorizzazioni necessarie. All'aumentare del numero di caratteristiche Amazon EMR utilizzate per il lavoro, potrebbero essere necessarie ulteriori autorizzazioni. La comprensione della gestione dell'accesso ti consente di richiedere le autorizzazioni corrette

all'amministratore. Se non riesci ad accedere a una funzionalità in Amazon EMR, consulta Risoluzione dei problemi relativi all'identità e all'accesso di Amazon EMR.

Amministratore del servizio: se sei il responsabile delle risorse Amazon EMR presso la tua azienda, probabilmente disponi dell'accesso completo ad Amazon EMR. Il tuo compito è determinare le funzionalità e le risorse di Amazon EMR a cui gli utenti del servizio devono accedere. Devi inviare le richieste all'amministratore IAM per cambiare le autorizzazioni degli utenti del servizio. Esamina le informazioni contenute in questa pagina per comprendere i concetti di base relativi a IAM. Per ulteriori informazioni su come la tua azienda può utilizzare IAM con Amazon EMR, consulta Funzionamento di Amazon EMR con IAM.

Amministratore IAM: se sei un amministratore IAM, potresti essere interessato a ottenere informazioni su come scrivere policy per gestire l'accesso ad Amazon EMR. Per visualizzare policy basate su identità Amazon EMR di esempio che possono essere utilizzate in IAM, consulta Esempi di policy basate su identità per Amazon EMR.

Autenticazione con identità

L'autenticazione è il modo in cui accedi AWS utilizzando le tue credenziali di identità. Devi essere autenticato (aver effettuato l' Utente root dell'account AWS accesso AWS) come utente IAM o assumendo un ruolo IAM.

Puoi accedere AWS come identità federata utilizzando le credenziali fornite tramite una fonte di identità. AWS IAM Identity Center Gli utenti (IAM Identity Center), l'autenticazione Single Sign-On della tua azienda e le tue credenziali di Google o Facebook sono esempi di identità federate. Se accedi come identità federata, l'amministratore ha configurato in precedenza la federazione delle identità utilizzando i ruoli IAM. Quando accedi AWS utilizzando la federazione, assumi indirettamente un ruolo.

A seconda del tipo di utente, puoi accedere al AWS Management Console o al portale di AWS accesso. Per ulteriori informazioni sull'accesso a AWS, vedi Come accedere al tuo Account AWS nella Guida per l'Accedi ad AWS utente.

Se accedi a AWS livello di codice, AWS fornisce un kit di sviluppo software (SDK) e un'interfaccia a riga di comando (CLI) per firmare crittograficamente le tue richieste utilizzando le tue credenziali. Se non utilizzi AWS strumenti, devi firmare tu stesso le richieste. Per ulteriori informazioni sul metodo consigliato per la firma delle richieste, consulta Signature Version 4 AWS per le richieste API nella Guida per l'utente IAM.

A prescindere dal metodo di autenticazione utilizzato, potrebbe essere necessario specificare ulteriori informazioni sulla sicurezza. Ad esempio, ti AWS consiglia di utilizzare l'autenticazione a più fattori (MFA) per aumentare la sicurezza del tuo account. Per ulteriori informazioni, consulta Autenticazione a più fattori nella Guida per l'utente di AWS IAM Identity Center e Utilizzo dell'autenticazione a più fattori (MFA)AWS in IAM nella Guida per l'utente IAM.

Account AWS utente root

Quando si crea un account Account AWS, si inizia con un'identità di accesso che ha accesso completo a tutte Servizi AWS le risorse dell'account. Questa identità è denominata utente Account AWS root ed è accessibile effettuando l'accesso con l'indirizzo e-mail e la password utilizzati per creare l'account. Si consiglia vivamente di non utilizzare l'utente root per le attività quotidiane. Conserva le credenziali dell'utente root e utilizzale per eseguire le operazioni che solo l'utente root può eseguire. Per un elenco completo delle attività che richiedono l'accesso come utente root, consulta la sezione Attività che richiedono le credenziali dell'utente root nella Guida per l'utente IAM.

Identità federata

Come procedura consigliata, richiedi agli utenti umani, compresi gli utenti che richiedono l'accesso come amministratore, di utilizzare la federazione con un provider di identità per accedere Servizi AWS utilizzando credenziali temporanee.

Un'identità federata è un utente dell'elenco utenti aziendale, di un provider di identità Web AWS Directory Service, della directory Identity Center o di qualsiasi utente che accede utilizzando le Servizi AWS credenziali fornite tramite un'origine di identità. Quando le identità federate accedono Account AWS, assumono ruoli e i ruoli forniscono credenziali temporanee.

Per la gestione centralizzata degli accessi, consigliamo di utilizzare AWS IAM Identity Center. Puoi creare utenti e gruppi in IAM Identity Center oppure puoi connetterti e sincronizzarti con un set di utenti e gruppi nella tua fonte di identità per utilizzarli su tutte le tue applicazioni. Account AWS Per ulteriori informazioni su IAM Identity Center, consulta Cos'è IAM Identity Center? nella Guida per l'utente di AWS IAM Identity Center .

Utenti e gruppi IAM

Un utente IAM è un'identità interna Account AWS che dispone di autorizzazioni specifiche per una singola persona o applicazione. Ove possibile, consigliamo di fare affidamento a credenziali temporanee invece di creare utenti IAM con credenziali a lungo termine come le password e le chiavi di accesso. Tuttavia, se si hanno casi d'uso specifici che richiedono credenziali a lungo termine con

utenti IAM, si consiglia di ruotare le chiavi di accesso. Per ulteriori informazioni, consulta la pagina Rotazione periodica delle chiavi di accesso per casi d'uso che richiedono credenziali a lungo termine nella Guida per l'utente IAM.

Un gruppo IAM è un'identità che specifica un insieme di utenti IAM. Non è possibile eseguire l'accesso come gruppo. È possibile utilizzare gruppi per specificare le autorizzazioni per più utenti alla volta. I gruppi semplificano la gestione delle autorizzazioni per set di utenti di grandi dimensioni. Ad esempio, potresti avere un gruppo denominato IAMAdminse concedere a quel gruppo le autorizzazioni per amministrare le risorse IAM.

Gli utenti sono diversi dai ruoli. Un utente è associato in modo univoco a una persona o un'applicazione, mentre un ruolo è destinato a essere assunto da chiunque ne abbia bisogno. Gli utenti dispongono di credenziali a lungo termine permanenti, mentre i ruoli forniscono credenziali temporanee. Per ulteriori informazioni, consulta Casi d'uso per utenti IAM nella Guida per l'utente IAM.

Ruoli IAM

Un ruolo IAM è un'identità interna all'utente Account AWS che dispone di autorizzazioni specifiche. È simile a un utente IAM, ma non è associato a una persona specifica. Per assumere temporaneamente un ruolo IAM in AWS Management Console, puoi passare da un ruolo utente a un ruolo IAM (console). Puoi assumere un ruolo chiamando un'operazione AWS CLI o AWS API o utilizzando un URL personalizzato. Per ulteriori informazioni sui metodi per l'utilizzo dei ruoli, consulta Utilizzo di ruoli IAM nella Guida per l'utente IAM.

I ruoli IAM con credenziali temporanee sono utili nelle seguenti situazioni:

- Accesso utente federato: per assegnare le autorizzazioni a una identità federata, è possibile creare un ruolo e definire le autorizzazioni per il ruolo. Quando un'identità federata viene autenticata, l'identità viene associata al ruolo e ottiene le autorizzazioni da esso definite. Per ulteriori informazioni sulla federazione dei ruoli, consulta Create a role for a third-party identity provider (federation) nella Guida per l'utente IAM. Se utilizzi IAM Identity Center, configura un set di autorizzazioni. IAM Identity Center mette in correlazione il set di autorizzazioni con un ruolo in IAM per controllare a cosa possono accedere le identità dopo l'autenticazione. Per informazioni sui set di autorizzazioni, consulta Set di autorizzazioni nella Guida per l'utente di AWS IAM Identity Center .
- Autorizzazioni utente IAM temporanee: un utente IAM o un ruolo può assumere un ruolo IAM per ottenere temporaneamente autorizzazioni diverse per un'attività specifica.

- Accesso multi-account: è possibile utilizzare un ruolo IAM per permettere a un utente (un principale affidabile) con un account diverso di accedere alle risorse nell'account. I ruoli sono lo strumento principale per concedere l'accesso multi-account. Tuttavia, con alcuni Servizi AWS, è possibile allegare una policy direttamente a una risorsa (anziché utilizzare un ruolo come proxy). Per informazioni sulle differenze tra ruoli e policy basate su risorse per l'accesso multi-account, consulta Accesso a risorse multi-account in IAM nella Guida per l'utente IAM.
- Accesso a più servizi: alcuni Servizi AWS utilizzano le funzionalità di altri Servizi AWS. Ad esempio, quando effettui una chiamata in un servizio, è normale che quel servizio esegua applicazioni in Amazon EC2 o archivi oggetti in Amazon S3. Un servizio può eseguire questa operazione utilizzando le autorizzazioni dell'entità chiamante, utilizzando un ruolo di servizio o utilizzando un ruolo collegato al servizio.
  - Sessioni di accesso inoltrato (FAS): quando utilizzi un utente o un ruolo IAM per eseguire azioni AWS, sei considerato un principale. Quando si utilizzano alcuni servizi, è possibile eseguire un'operazione che attiva un'altra operazione in un servizio diverso. FAS utilizza le autorizzazioni del principale che chiama an Servizio AWS, combinate con la richiesta Servizio AWS per effettuare richieste ai servizi downstream. Le richieste FAS vengono effettuate solo quando un servizio riceve una richiesta che richiede interazioni con altri Servizi AWS o risorse per essere completata. In questo caso è necessario disporre delle autorizzazioni per eseguire entrambe le azioni. Per i dettagli delle policy relative alle richieste FAS, consulta Forward access sessions.
  - Ruolo di servizio: un ruolo di servizio è un ruolo IAM che un servizio assume per eseguire operazioni per tuo conto. Un amministratore IAM può creare, modificare ed eliminare un ruolo di servizio dall'interno di IAM. Per ulteriori informazioni, consulta la sezione Create a role to delegate permissions to an Servizio AWS nella Guida per l'utente IAM.
  - Ruolo collegato al servizio: un ruolo collegato al servizio è un tipo di ruolo di servizio collegato a un. Servizio AWS Il servizio può assumere il ruolo per eseguire un'azione per tuo conto. I ruoli collegati al servizio vengono visualizzati nel tuo account Account AWS e sono di proprietà del servizio. Un amministratore IAM può visualizzare le autorizzazioni per i ruoli collegati ai servizi, ma non modificarle.
- Applicazioni in esecuzione su Amazon EC2: puoi utilizzare un ruolo IAM per gestire le credenziali temporanee per le applicazioni in esecuzione su un' EC2 istanza e che AWS CLI effettuano richieste AWS API. È preferibile archiviare le chiavi di accesso all'interno dell' EC2 istanza. Per assegnare un AWS ruolo a un' EC2 istanza e renderlo disponibile per tutte le sue applicazioni, create un profilo di istanza collegato all'istanza. Un profilo di istanza contiene il ruolo e consente ai programmi in esecuzione sull' EC2 istanza di ottenere credenziali temporanee. Per ulteriori

informazioni, consulta Utilizzare un ruolo IAM per concedere le autorizzazioni alle applicazioni in esecuzione su EC2 istanze Amazon nella IAM User Guide.

Gestione dell'accesso con policy

Puoi controllare l'accesso AWS creando policy e collegandole a AWS identità o risorse. Una policy è un oggetto AWS che, se associato a un'identità o a una risorsa, ne definisce le autorizzazioni. AWS valuta queste politiche quando un principale (utente, utente root o sessione di ruolo) effettua una richiesta. Le autorizzazioni nelle policy determinano l'approvazione o il rifiuto della richiesta. La maggior parte delle politiche viene archiviata AWS come documenti JSON. Per ulteriori informazioni sulla struttura e sui contenuti dei documenti delle policy JSON, consulta Panoramica delle policy JSON nella Guida per l'utente IAM.

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale principale può eseguire operazioni su quali risorse e in quali condizioni.

Per impostazione predefinita, utenti e ruoli non dispongono di autorizzazioni. Per concedere agli utenti l'autorizzazione a eseguire operazioni sulle risorse di cui hanno bisogno, un amministratore IAM può creare policy IAM. L'amministratore può quindi aggiungere le policy IAM ai ruoli e gli utenti possono assumere i ruoli.

Le policy IAM definiscono le autorizzazioni relative a un'operazione, a prescindere dal metodo utilizzato per eseguirla. Ad esempio, supponiamo di disporre di una policy che consente l'operazione `iam:GetRole`. Un utente con tale policy può ottenere informazioni sul ruolo dall' AWS Management Console AWS CLI, dall'o dall' AWS API.

Policy basate sull'identità

Le policy basate su identità sono documenti di policy di autorizzazione JSON che è possibile allegare a un'identità (utente, gruppo di utenti o ruolo IAM). Tali policy definiscono le operazioni che utenti e ruoli possono eseguire, su quali risorse e in quali condizioni. Per informazioni su come creare una policy basata su identità, consulta Definizione di autorizzazioni personalizzate IAM con policy gestite dal cliente nella Guida per l'utente IAM.

Le policy basate su identità possono essere ulteriormente classificate come policy inline o policy gestite. Le policy inline sono integrate direttamente in un singolo utente, gruppo o ruolo. Le politiche gestite sono politiche autonome che puoi allegare a più utenti, gruppi e ruoli nel tuo Account AWS. Le politiche gestite includono politiche AWS gestite e politiche gestite dai clienti. Per informazioni su

come scegliere tra una policy gestita o una policy inline, consulta Scelta fra policy gestite e policy inline nella Guida per l'utente IAM.

Policy basate sulle risorse

Le policy basate su risorse sono documenti di policy JSON che è possibile collegare a una risorsa. Esempi di policy basate sulle risorse sono le policy di attendibilità dei ruoli IAM e le policy dei bucket Amazon S3. Nei servizi che supportano policy basate sulle risorse, gli amministratori dei servizi possono utilizzarli per controllare l'accesso a una risorsa specifica. Quando è collegata a una risorsa, una policy definisce le operazioni che un principale può eseguire su tale risorsa e a quali condizioni. È necessario specificare un principale in una policy basata sulle risorse. I principali possono includere account, utenti, ruoli, utenti federati o. Servizi AWS

Le policy basate sulle risorse sono policy inline che si trovano in tale servizio. Non puoi utilizzare le policy AWS gestite di IAM in una policy basata sulle risorse.

Elenchi di controllo degli accessi () ACLs

Le liste di controllo degli accessi (ACLs) controllano quali principali (membri dell'account, utenti o ruoli) dispongono delle autorizzazioni per accedere a una risorsa. ACLs sono simili alle politiche basate sulle risorse, sebbene non utilizzino il formato del documento di policy JSON.

Amazon S3 e Amazon VPC sono esempi di servizi che supportano. AWS WAF ACLs Per ulteriori informazioni ACLs, consulta la panoramica della lista di controllo degli accessi (ACL) nella Amazon Simple Storage Service Developer Guide.

Altri tipi di policy

AWS supporta tipi di policy aggiuntivi e meno comuni. Questi tipi di policy possono impostare il numero massimo di autorizzazioni concesse dai tipi di policy più comuni.

- Limiti delle autorizzazioni: un limite delle autorizzazioni è una funzionalità avanzata nella quale si imposta il numero massimo di autorizzazioni che una policy basata su identità può concedere a un'entità IAM (utente o ruolo IAM). È possibile impostare un limite delle autorizzazioni per un'entità. Le autorizzazioni risultanti sono l'intersezione delle policy basate su identità dell'entità e i relativi limiti delle autorizzazioni. Le policy basate su risorse che specificano l'utente o il ruolo nel campo `Principal`sono condizionate dal limite delle autorizzazioni. Un rifiuto esplicito in una qualsiasi di queste policy sostituisce l'autorizzazione. Per ulteriori informazioni sui limiti delle autorizzazioni, consulta Limiti delle autorizzazioni per le entità IAM nella Guida per l'utente IAM.

- Politiche di controllo del servizio (SCPs): SCPs sono politiche JSON che specificano le autorizzazioni massime per un'organizzazione o un'unità organizzativa (OU) in. AWS Organizations AWS Organizations è un servizio per il raggruppamento e la gestione centralizzata di più di proprietà dell' Account AWS azienda. Se abiliti tutte le funzionalità di un'organizzazione, puoi applicare le politiche di controllo del servizio (SCPs) a uno o tutti i tuoi account. L'SCP limita le autorizzazioni per le entità presenti negli account dei membri, inclusa ciascuna di esse. Utente root dell'account AWS Per ulteriori informazioni su Organizations and SCPs, consulta le politiche di controllo dei servizi nella Guida AWS Organizations per l'utente.
- Politiche di controllo delle risorse (RCPs): RCPs sono politiche JSON che puoi utilizzare per impostare le autorizzazioni massime disponibili per le risorse nei tuoi account senza aggiornare le politiche IAM allegate a ciascuna risorsa di tua proprietà. L'RCP limita le autorizzazioni per le risorse negli account dei membri e può influire sulle autorizzazioni effettive per le identità, incluse le Utente root dell'account AWS, indipendentemente dal fatto che appartengano o meno all'organizzazione. Per ulteriori informazioni su Organizations e RCPs, incluso un elenco di Servizi AWS tale supporto RCPs, vedere Resource control policies (RCPs) nella Guida per l'AWS Organizations utente.
- Policy di sessione: le policy di sessione sono policy avanzate che vengono trasmesse come parametro quando si crea in modo programmatico una sessione temporanea per un ruolo o un utente federato. Le autorizzazioni della sessione risultante sono l'intersezione delle policy basate su identità del ruolo o dell'utente e le policy di sessione. Le autorizzazioni possono anche provenire da una policy basata su risorse. Un rifiuto esplicito in una qualsiasi di queste policy sostituisce l'autorizzazione. Per ulteriori informazioni, consulta Policy di sessione nella Guida per l'utente IAM.

Più tipi di policy

Quando più tipi di policy si applicano a una richiesta, le autorizzazioni risultanti sono più complicate da comprendere. Per scoprire come si AWS determina se consentire o meno una richiesta quando sono coinvolti più tipi di policy, consulta la logica di valutazione delle policy nella IAM User Guide.

Funzionamento di Amazon EMR con IAM

Prima di utilizzare IAM per gestire l'accesso ad Amazon EMR, scopri quali funzionalità IAM sono disponibili per l'uso con Amazon EMR.

Funzionalità IAM utilizzabili con Amazon EMR

Funzionalità IAM	Supporto per Amazon EMR
Policy basate su identità	Sì
Policy basate su risorse	Sì
Azioni di policy	Sì
Risorse relative alle policy	Sì
Chiavi di condizione delle policy	Sì
ACLs	No
ABAC (tag nelle policy)	Sì
Credenziali temporanee	Sì
Autorizzazioni del principale	Sì
Ruoli di servizio	No
Ruoli collegati al servizio	Sì

Per avere una visione di alto livello di come Amazon EMR e AWS altri servizi funzionano con la maggior parte delle funzionalità IAM, AWS consulta i servizi che funzionano con IAM nella IAM User Guide.

Policy basate su identità per Amazon EMR

Supporta le policy basate su identità: sì

Le policy basate su identità sono documenti di policy di autorizzazione JSON che è possibile allegare a un'identità (utente, gruppo di utenti o ruolo IAM). Tali policy definiscono le operazioni che utenti e ruoli possono eseguire, su quali risorse e in quali condizioni. Per informazioni su come creare una policy basata su identità, consulta Definizione di autorizzazioni personalizzate IAM con policy gestite dal cliente nella Guida per l'utente IAM.

Con le policy basate su identità di IAM, è possibile specificare quali operazioni e risorse sono consentite o respinte, nonché le condizioni in base alle quali le operazioni sono consentite o respinte. Non è possibile specificare l'entità principale in una policy basata sull'identità perché si applica all'utente o al ruolo a cui è associato. Per informazioni su tutti gli elementi utilizzabili in una policy JSON, consulta Guida di riferimento agli elementi delle policy JSON IAM nella Guida per l'utente IAM.

Esempi di policy basate su identità per Amazon EMR

Per visualizzare esempi di policy basate su identità Amazon EMR, consulta Esempi di policy basate su identità per Amazon EMR.

Policy basate su risorse all'interno di Amazon EMR

Supporta le policy basate sulle risorse: sì

Le policy basate su risorse sono documenti di policy JSON che è possibile collegare a una risorsa. Esempi di policy basate sulle risorse sono le policy di attendibilità dei ruoli IAM e le policy dei bucket Amazon S3. Nei servizi che supportano policy basate sulle risorse, gli amministratori dei servizi possono utilizzarli per controllare l'accesso a una risorsa specifica. Quando è collegata a una risorsa, una policy definisce le operazioni che un principale può eseguire su tale risorsa e a quali condizioni. È necessario specificare un principale in una policy basata sulle risorse. I principali possono includere account, utenti, ruoli, utenti federati o. Servizi AWS

Per consentire l'accesso multi-account, puoi specificare un intero account o entità IAM in un altro account come principale in una policy basata sulle risorse. L'aggiunta di un principale multi-account a una policy basata sulle risorse rappresenta solo una parte della relazione di trust. Quando il principale e la risorsa sono diversi Account AWS, un amministratore IAM dell'account affidabile deve inoltre concedere all'entità principale (utente o ruolo) l'autorizzazione ad accedere alla risorsa. L'autorizzazione viene concessa collegando all'entità una policy basata sull'identità. Tuttavia, se una policy basata su risorse concede l'accesso a un principale nello stesso account, non sono richieste ulteriori policy basate su identità. Per ulteriori informazioni, consulta Accesso a risorse multi-account in IAM nella Guida per l'utente IAM.

Operazioni di policy per Amazon EMR

Supporta le operazioni di policy: si

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale principale può eseguire operazioni su quali risorse, e in quali condizioni.

L'elemento `Action`di una policy JSON descrive le operazioni che è possibile utilizzare per consentire o negare l'accesso a un criterio. Le azioni politiche in genere hanno lo stesso nome dell'operazione AWS API associata. Ci sono alcune eccezioni, ad esempio le operazioni di sola autorizzazione che non hanno un'operazione API corrispondente. Esistono anche alcune operazioni che richiedono più operazioni in una policy. Queste operazioni aggiuntive sono denominate operazioni dipendenti.

Includi le operazioni in una policy per concedere le autorizzazioni a eseguire l'operazione associata.

Per visualizzare un elenco di operazioni Amazon EMR, consulta Operazioni, risorse e chiavi di condizione per Amazon EMR in Guida di riferimento all'autorizzazione del servizio.

Le operazioni delle policy in Amazon EMR utilizzano il seguente prefisso prima dell'operazione:

```
EMR
```

Per specificare più operazioni in una sola istruzione, occorre separarle con la virgola.

```
"Action": [
      "EMR:action1",
      "EMR:action2"
         ]
```

Per visualizzare esempi di policy basate su identità Amazon EMR, consulta Esempi di policy basate su identità per Amazon EMR.

Risorse delle policy per Amazon EMR

Supporta le risorse di policy: sì

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale principale può eseguire operazioni su quali risorse, e in quali condizioni.

L'elemento JSON `Resource`della policy specifica l'oggetto o gli oggetti ai quali si applica l'operazione. Le istruzioni devono includere un elemento `Resource`o un elemento `NotResource`. Come best practice, specifica una risorsa utilizzando il suo nome della risorsa Amazon (ARN). È possibile eseguire questa operazione per operazioni che supportano un tipo di risorsa specifico, note come autorizzazioni a livello di risorsa.

Per le azioni che non supportano le autorizzazioni a livello di risorsa, ad esempio le operazioni di elenco, utilizza un carattere jolly (*) per indicare che l'istruzione si applica a tutte le risorse.

```
"Resource": "*"
```

Per visualizzare un elenco dei tipi di risorse Amazon EMR e relativi ARNs, consulta Resources Defined by Amazon EMR nel Service Authorization Reference. Per informazioni sulle operazioni per cui è possibile specificare l'ARN di ciascuna risorsa, consulta Operazioni, risorse e chiavi di condizione per Amazon EMR.

Per visualizzare esempi di policy basate su identità Amazon EMR, consulta Esempi di policy basate su identità per Amazon EMR.

Chiavi di condizione delle policy per Amazon EMR

Supporta le chiavi di condizione delle policy specifiche del servizio: sì

Gli amministratori possono utilizzare le policy AWS JSON per specificare chi ha accesso a cosa. In altre parole, quale principale può eseguire operazioni su quali risorse, e in quali condizioni.

L'elemento `Condition`(o blocco `Condition`) consente di specificare le condizioni in cui un'istruzione è in vigore. L'elemento `Condition`è facoltativo. È possibile compilare espressioni condizionali che utilizzano operatori di condizione, ad esempio uguale a o minore di, per soddisfare la condizione nella policy con i valori nella richiesta.

Se specifichi più elementi `Condition`in un'istruzione o più chiavi in un singolo elemento `Condition`, questi vengono valutati da AWS utilizzando un'operazione `AND`logica. Se si specificano più valori per una singola chiave di condizione, AWS valuta la condizione utilizzando un'operazione logica. `OR` Tutte le condizioni devono essere soddisfatte prima che le autorizzazioni dell'istruzione vengano concesse.

È possibile anche utilizzare variabili segnaposto quando specifichi le condizioni. Ad esempio, è possibile autorizzare un utente IAM ad accedere a una risorsa solo se è stata taggata con il relativo nome utente IAM. Per ulteriori informazioni, consulta Elementi delle policy IAM: variabili e tag nella Guida per l'utente di IAM.

AWS supporta chiavi di condizione globali e chiavi di condizione specifiche del servizio. Per visualizzare tutte le chiavi di condizione AWS globali, consulta le chiavi di contesto delle condizioni AWS globali nella Guida per l'utente IAM.

Per visualizzare un elenco di chiavi di condizione Amazon EMR e scoprire quali operazioni e risorse è possibile utilizzare con una chiave di condizione, consulta Operazioni, risorse e chiavi di condizione per Amazon EMR in Guida di riferimento all'autorizzazione del servizio.

Per visualizzare esempi di policy basate su identità Amazon EMR, consulta Esempi di policy basate su identità per Amazon EMR.

Elenchi di controllo degli accessi (ACLs) in Amazon EMR

Supporti ACLs: no

Le liste di controllo degli accessi (ACLs) controllano quali principali (membri dell'account, utenti o ruoli) dispongono delle autorizzazioni per accedere a una risorsa. ACLs sono simili alle politiche basate sulle risorse, sebbene non utilizzino il formato del documento di policy JSON.

Controllo degli accessi basato su attributi (ABAC) con Amazon EMR

Supporta ABAC (tag nelle policy)	Sì

Il controllo dell'accesso basato su attributi (ABAC) è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. In AWS, questi attributi sono chiamati tag. Puoi allegare tag a entità IAM (utenti o ruoli) e a molte AWS risorse. L'assegnazione di tag alle entità e alle risorse è il primo passaggio di ABAC. In seguito, vengono progettate policy ABAC per consentire operazioni quando il tag dell'entità principale corrisponde al tag sulla risorsa a cui si sta provando ad accedere.

La strategia ABAC è utile in ambienti soggetti a una rapida crescita e aiuta in situazioni in cui la gestione delle policy diventa impegnativa.

Per controllare l'accesso basato su tag, fornisci informazioni sui tag nell'elemento condizione di una policy utilizzando le chiavi di condizione `aws:ResourceTag/`*`key-name`*, `aws:RequestTag/`*`key-name`* o `aws:TagKeys`.

Se un servizio supporta tutte e tre le chiavi di condizione per ogni tipo di risorsa, il valore per il servizio è Yes (Sì). Se un servizio supporta tutte e tre le chiavi di condizione solo per alcuni tipi di risorsa, allora il valore sarà Parziale.

Per ulteriori informazioni su ABAC, consulta Definizione delle autorizzazioni con autorizzazione ABAC nella Guida per l'utente IAM. Per visualizzare un tutorial con i passaggi per l'impostazione di ABAC, consulta Utilizzo del controllo degli accessi basato su attributi (ABAC) nella Guida per l'utente IAM.

Utilizzo di credenziali temporanee con Amazon EMR

Supporta le credenziali temporanee: sì

Alcuni Servizi AWS non funzionano quando accedi utilizzando credenziali temporanee. Per ulteriori informazioni, incluse quelle che Servizi AWS funzionano con credenziali temporanee, consulta la sezione relativa alla Servizi AWS compatibilità con IAM nella IAM User Guide.

Stai utilizzando credenziali temporanee se accedi AWS Management Console utilizzando qualsiasi metodo tranne nome utente e password. Ad esempio, quando accedi AWS utilizzando il link Single Sign-On (SSO) della tua azienda, tale processo crea automaticamente credenziali temporanee. Le credenziali temporanee vengono create in automatico anche quando accedi alla console come utente e poi cambi ruolo. Per ulteriori informazioni sullo scambio dei ruoli, consulta Passaggio da un ruolo utente a un ruolo IAM (console) nella Guida per l'utente IAM.

È possibile creare manualmente credenziali temporanee utilizzando l'API or. AWS CLI AWS È quindi possibile utilizzare tali credenziali temporanee per accedere. AWS AWS consiglia di generare dinamicamente credenziali temporanee anziché utilizzare chiavi di accesso a lungo termine. Per ulteriori informazioni, consulta Credenziali di sicurezza provvisorie in IAM.

Autorizzazioni delle entità principali tra servizi per Amazon EMR

Supporta l'inoltro delle sessioni di accesso (FAS): sì

Quando utilizzi un utente o un ruolo IAM per eseguire azioni AWS, sei considerato un principale. Quando si utilizzano alcuni servizi, è possibile eseguire un'operazione che attiva un'altra operazione in un servizio diverso. FAS utilizza le autorizzazioni del principale che chiama an Servizio AWS, in combinazione con la richiesta Servizio AWS per effettuare richieste ai servizi downstream. Le richieste FAS vengono effettuate solo quando un servizio riceve una richiesta che richiede interazioni con altri Servizi AWS o risorse per essere completata. In questo caso è necessario disporre delle autorizzazioni per eseguire entrambe le azioni. Per i dettagli delle policy relative alle richieste FAS, consulta Forward access sessions.

Ruoli di servizio per Amazon EMR

Supporta i ruoli di servizio	No

Ruoli collegati ai servizi per Amazon EMR

Supporta i ruoli collegati ai servizi	Sì

Per ulteriori informazioni su come creare e gestire i ruoli collegati ai servizi, consulta Servizi AWS supportati da IAM. Trova un servizio nella tabella che include un `Yes` nella colonna Service-linked role (Ruolo collegato ai servizi). Scegli il collegamento Sì per visualizzare la documentazione relativa al ruolo collegato ai servizi per tale servizio.

Utilizzo di tag di cluster e notebook con policy IAM per il controllo degli accessi

Le autorizzazioni per operazioni Amazon EMR associate a EMR Notebooks e cluster EMR possono essere ottimizzate utilizzando il controllo degli accessi basato su tag con policy IAM basate su identità. Puoi utilizzare chiavi di condizione all'interno di un elemento `Condition` (chiamato anche blocco `Condition`) per consentire determinate operazioni solo quando un notebook, un cluster o entrambi dispongono di una determinata combinazione di chiave di tag o di chiave-valore. Puoi anche limitare l'operazione `CreateEditor` (che crea un notebook EMR) e l'operazione `RunJobFlow` (che crea un cluster) in modo che una richiesta per un tag deve essere inviata quando la risorsa viene creata.

In Amazon EMR, le chiavi di condizioni che possono essere utilizzate in un elemento `Condition` si applicano solo a tali operazioni API Amazon EMR in cui `ClusterID` o `NotebookID` è un parametro di richiesta obbligatorio. Ad esempio, l'ModifyInstanceGroupsazione non supporta le chiavi di contesto perché `ClusterID` è un parametro opzionale.

Quando crei un notebook EMR, viene applicato un tag predefinito con una stringa di chiavi `creatorUserId` impostata sul valore dell'ID utente IAM che ha creato il notebook. Ciò è utile per limitare le operazioni consentite per il notebook al solo creatore.

Le seguenti chiavi di condizione sono disponibili in Amazon EMR:

- Utilizza la chiave di contesto di condizione `elasticmapreduce:ResourceTag/`*`TagKeyString`* per concedere o negare agli utenti operazioni su cluster p notebook con tag che hanno *`TagKeyString`* specificato. Se un'operazione trasferisce `ClusterID` e `NotebookID`, la condizione si applica al cluster e al notebook. Questo significa che entrambe le risorse devono avere la stringa della chiave di tag o la combinazione chiave-valore specificata. Puoi utilizzare l'elemento `Resource` per limitare l'istruzione in modo che si applichi solo ai cluster o ai notebook

in base alle esigenze. Per ulteriori informazioni, consulta Esempi di policy basate su identità per Amazon EMR.

- Utilizzate la chiave di contesto della `elasticmapreduce:RequestTag/`*`TagKeyString`* condizione per richiedere un tag specifico con actions/API le chiamate. Ad esempio, puoi utilizzare questa chiave di contesto di condizione insieme all'operazione `CreateEditor` per richiedere che una chiave con *`TagKeyString`* sia applicata a un notebook quando viene creato.

Esempi

Per visualizzare un elenco delle operazioni Amazon EMR, consulta Actions Defined by Amazon EMR (Operazioni definite da Amazon EMR) nella IAM User Guide (Guida per l'utente IAM).

Ruoli di runtime per le fasi di Amazon EMR

Un ruolo di runtime è un ruolo AWS Identity and Access Management (IAM) che puoi specificare quando invii un lavoro o una query a un cluster Amazon EMR. Il job o la query che invii al tuo cluster Amazon EMR utilizza il ruolo di runtime per accedere alle AWS risorse, come gli oggetti in Amazon S3. Puoi specificare i ruoli di runtime con Amazon EMR per i processi Spark e Hive.

Puoi anche specificare i ruoli di runtime quando ti connetti a cluster Amazon EMR in Amazon SageMaker AI e quando colleghi un WorkSpace Amazon EMR Studio a un cluster EMR. Per ulteriori informazioni, consulta Connect a un cluster Amazon EMR di SageMaker AI Studio e. Esecuzione di un WorkSpace EMR Studio con un ruolo di runtime

In precedenza, i cluster Amazon EMR eseguivano processi o query Amazon EMR con autorizzazioni basate sulla policy IAM associata al profilo dell'istanza utilizzato per avviare il cluster. Ciò significava che le policy dovevano contenere l'unione di tutte le autorizzazioni per tutti i processi e le query eseguiti su un cluster Amazon EMR. Con i ruoli di runtime, ora puoi gestire il controllo degli accessi per ogni singolo processo o query, invece di condividere il profilo dell'istanza Amazon EMR del cluster.

Sui cluster Amazon EMR con ruoli di runtime, puoi anche applicare il controllo di accesso AWS Lake Formation basato ai job e alle query di Spark, Hive e Presto sui tuoi data lake. Per ulteriori informazioni su come effettuare l'integrazione con, consulta. AWS Lake FormationIntegra Amazon EMR con AWS Lake Formation

> **Note**
>
> Quando specifichi un ruolo di runtime per una fase di Amazon EMR, i job o le query che invii possono accedere solo alle AWS risorse consentite dalle policy associate al ruolo di runtime. Questi job e queste query non possono accedere all'Instance Metadata Service sulle EC2 istanze del cluster o utilizzare il profilo dell' EC2 istanza del cluster per accedere a qualsiasi risorsa. AWS

Prerequisiti per l'avvio di un cluster Amazon EMR con un ruolo di runtime

Argomenti

- Fase 1: impostazione delle configurazioni di sicurezza in Amazon EMR
- Fase 2: configurare un profilo di EC2 istanza per il cluster Amazon EMR
- Passaggio 3: Configurazione di una policy di attendibilità

Fase 1: impostazione delle configurazioni di sicurezza in Amazon EMR

Utilizza la seguente struttura JSON per creare una configurazione di sicurezza su AWS Command Line Interface (AWS CLI) e imposta su. `EnableApplicationScopedIAMRole` `true` Per ulteriori informazioni sulle configurazioni della sicurezza, consulta Usa le configurazioni di sicurezza per configurare la sicurezza dei cluster Amazon EMR.

```
{
    "AuthorizationConfiguration":{
        "IAMConfiguration":{
            "EnableApplicationScopedIAMRole":true
        }
    }
}
```

Ti consigliamo di abilitare sempre le opzioni di crittografia in transito nella configurazione di sicurezza, in modo che i dati trasferiti su Internet siano crittografati anziché in testo semplice. Puoi ignorare queste opzioni se non desideri connetterti ai cluster Amazon EMR con ruoli SageMaker di runtime di Runtime Studio o EMR Studio. Per configurare la crittografia dei dati, consulta Configurazione della crittografia di dati.

In alternativa, puoi creare una configurazione di sicurezza con impostazioni personalizzate con la AWS Management Console.

Fase 2: configurare un profilo di EC2 istanza per il cluster Amazon EMR

I cluster Amazon EMR utilizzano il ruolo del profilo di EC2 istanza Amazon per assumere i ruoli di runtime. Per utilizzare i ruoli di runtime con le fasi di Amazon EMR, aggiungi le seguenti policy al ruolo IAM che intendi utilizzare come ruolo del profilo dell'istanza. Per aggiungere le policy a un ruolo IAM o modificare una policy inline o gestita esistente, consulta Aggiunta e rimozione di autorizzazioni per identità IAM.

JSON

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

Passaggio 3: Configurazione di una policy di attendibilità

Per ogni ruolo IAM che intendi utilizzare come ruolo di runtime, imposta la seguente policy di attendibilità, sostituendo `EMR_EC2_DefaultRole` con il ruolo del profilo dell'istanza. Per modificare la policy di attendibilità di un ruolo IAM, consulta Modifica di una policy di attendibilità del ruolo.

```
{
    "Sid":"AllowAssumeRole",
    "Effect":"Allow",
    "Principal":{
```

```
        "AWS":"arn:aws:iam::<AWS_ACCOUNT_ID>:role/EMR_EC2_DefaultRole"
    },
    "Action":[
             "sts:AssumeRole",
             "sts:TagSession"
            ]
}
```

Avvio di un cluster Amazon EMR con controllo degli accessi basato su ruoli

Dopo aver impostato le configurazioni, puoi avviare un cluster Amazon EMR con la configurazione di sicurezza da Fase 1: impostazione delle configurazioni di sicurezza in Amazon EMR. Per utilizzare i ruoli di runtime con le fasi di Amazon EMR, usa release label `emr-6.7.0` o versioni successive e seleziona Hive, Spark o entrambi come applicazione cluster. CloudWatchAgent è supportato su Runtime Role Clusters per EMR 7.6 e versioni successive. Per connetterti da SageMaker AI Studio, usa la release `emr-6.9.0` o una versione successiva e seleziona Livy, Spark, Hive o Presto come applicazione cluster. Per istruzioni sull'avvio del cluster, consulta la sezione Specificare una configurazione di sicurezza per un cluster Amazon EMR.

Invio di processi Spark tramite la procedura di Amazon EMR

Di seguito è riportato un esempio di come eseguire l' HdfsTest esempio incluso in Apache Spark. Questa chiamata API ha esito positivo solo se il ruolo di runtime per Amazon EMR fornito dispone dell'accesso 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**
>
> Ti consigliamo di disattivare l'accesso al cluster Amazon EMR di SSH e di consentirne l'accesso solo all'API `AddJobFlowSteps` di Amazon EMR.

Invio di processi Hive tramite la procedura di Amazon EMR

L'esempio seguente utilizza Apache Hive con fasi Amazon EMR per inviare un processo per l'esecuzione del file `QUERY_FILE.hql`. Questa query ha esito positivo solo se il ruolo di runtime fornito può accedere al percorso Amazon S3 del file di query.

```
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
```

Connect ai cluster Amazon EMR con ruoli di runtime da un SageMaker notebook AI Studio

Puoi applicare i ruoli di runtime di Amazon EMR alle query eseguite nei cluster Amazon EMR da AI Studio. SageMaker A tale scopo, segui la procedura seguente.

1. Segui le istruzioni in Launch Amazon SageMaker AI Studio per creare un SageMaker AI Studio.
2. Nell'interfaccia utente di SageMaker AI Studio, avvia un notebook con kernel supportati. Ad esempio, avvia un' SparkMagic immagine con un PySpark kernel.
3. Scegli un cluster Amazon EMR in SageMaker AI Studio, quindi scegli Connect.
4. Scegli un ruolo di runtime, quindi seleziona Connect (Connetti).

Questo creerà una cella notebook SageMaker AI con comandi magici per connettersi al tuo cluster Amazon EMR con il ruolo di runtime Amazon EMR scelto. Nella cella del notebook, puoi inserire ed eseguire query con il ruolo di runtime e il controllo degli accessi basato su Lake Formation. Per un esempio più dettagliato, consulta Applica controlli granulari di accesso ai dati con AWS Lake Formation Amazon EMR di Amazon AI Studio. SageMaker

Controllo dell'accesso al ruolo di runtime di Amazon EMR

Puoi controllare l'accesso al ruolo di runtime con la chiave di condizione `elasticmapreduce:ExecutionRoleArn`. La seguente policy consente a un principale IAM

di utilizzare un ruolo IAM denominato `Caller` o qualsiasi ruolo IAM che inizia con la stringa `CallerTeamRole`, come il ruolo di runtime.

> **Important**
>
> È necessario creare una condizione basata sulla chiave di `elasticmapreduce:ExecutionRoleArn` contesto quando si concede a un chiamante l'accesso per chiamare `AddJobFlowSteps` o `GetClusterSessionCredentials` APIs, come illustrato nell'esempio seguente.

```
{
    "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*"
            ]
        }
    }
}
```

Configurazione di un rapporto di attendibilità tra i ruoli di runtime e i cluster Amazon EMR

Amazon EMR genera un identificatore univoco `ExternalId` per ogni configurazione di sicurezza con autorizzazione del ruolo di runtime attivata. Questa autorizzazione consente a ogni utente di possedere una serie di ruoli di runtime da utilizzare sui cluster di sua proprietà. Ad esempio, ogni reparto di un'azienda può utilizzare il proprio ID esterno per aggiornare la policy di attendibilità sulla propria serie di ruoli di runtime.

Puoi trovare l'ID esterno con l'API `DescribeSecurityConfiguration` di Amazon EMR, come mostrato nell'esempio seguente.

```
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"
}
```

Per informazioni su come utilizzare un ID esterno, vedi Come utilizzare un ID esterno per concedere l'accesso alle tue AWS risorse a terzi.

Audit

Per monitorare e controllare le azioni che gli utenti finali compiono con i ruoli IAM, è possibile attivare la funzione di identità di origine. Per ulteriori informazioni sull'origine dell'identità, consulta la sezione Monitoraggio e controllo delle operazioni intraprese con i ruoli assunti.

Per tenere traccia dell'identità dell'origine, imposta `ApplicationScopedIAMRoleConfiguration/PropagateSourceIdentity` su `true` nella configurazione di sicurezza, come indicato di seguito.

```
{
    "AuthorizationConfiguration":{
        "IAMConfiguration":{
            "EnableApplicationScopedIAMRole":true,
            "ApplicationScopedIAMRoleConfiguration":{
                "PropagateSourceIdentity":true
            }
        }
    }
}
```

Quando imposti `PropagateSourceIdentity` su `true`, Amazon EMR applica l'identità dell'origine dalle credenziali di chiamata a una sessione di processo o query creata con il ruolo di runtime. Se

nelle credenziali di chiamata non è presente alcuna identità di origine, Amazon EMR non imposta l'identità di origine.

Per utilizzare questa proprietà, fornisci le autorizzazioni `sts:SetSourceIdentity` al tuo profilo dell'istanza, come indicato di seguito.

```
{ // PropagateSourceIdentity statement
    "Sid":"PropagateSourceIdentity",
    "Effect":"Allow",
    "Action":"sts:SetSourceIdentity",
    "Resource":[
        <runtime-role-ARN>
    ],
    "Condition":{
        "StringEquals":{
            "sts:SourceIdentity":<source-identity>
        }
    }
}
```

È necessario anche aggiungere l'istruzione `AllowSetSourceIdentity` alla policy di attendibilità dei ruoli di runtime.

```
{ // 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>
        }
    }
}
```

Ulteriori considerazioni

> **Note**
>
> Con la versione di Amazon EMR`emr-6.9.0`, potresti riscontrare errori intermittenti quando ti connetti ai cluster Amazon EMR da AI Studio. SageMaker Per risolvere questo problema, puoi installare la patch con un'operazione di bootstrap all'avvio del cluster. Per informazioni dettagliate sulle patch, consulta la sezione Amazon EMR release 6.9.0 known issues (Problemi noti di Amazon EMR versione 6.9.0).

Inoltre, quando configuri i ruoli di runtime per Amazon EMR, tieni presente quanto segue.

- Amazon EMR supporta i ruoli di runtime in tutte le Regioni AWS commerciali.
- Le fasi di Amazon EMR supportano i processi Apache Spark e Apache Hive con i ruoli di runtime quando utilizzi la versione `emr-6.7.0` o successive.
- SageMaker AI Studio supporta le query Spark, Hive e Presto con ruoli di runtime quando usi release o versioni successive. `emr-6.9.0`
- I seguenti kernel per notebook in SageMaker AI supportano i ruoli di runtime:
  - DataScience — Kernel Python 3
  - DataScience 2.0 — Kernel Python 3
  - DataScience 3.0 — Kernel Python 3
  - SparkAnalytics 1.0 — SparkMagic e kernel PySpark
  - SparkAnalytics 2.0 — SparkMagic e PySpark kernel
  - SparkMagic — PySpark kernel
- Amazon EMR supporta le fasi che utilizzano `RunJobFlow` solo al momento della creazione del cluster. Questa API non supporta i ruoli di runtime.
- Amazon EMR non supporta i ruoli di runtime su cluster configurati per la disponibilità elevata.
- A partire dalla versione 7.5.0 e successive di Amazon EMR, i ruoli di runtime supportano la visualizzazione delle interfacce utente Spark e YARN (UIs), come le seguenti: Spark Live UI, Spark History Server, YARN e YARN. NodeManager ResourceManager Quando accedi a questi UIs, viene richiesto un nome utente e una password. I nomi utente e le password possono essere generati tramite l'uso dell'API EMR. GetClusterSessionCredentials Per ulteriori informazioni sui dettagli di utilizzo dell'API, consulta. GetClusterSessionCredentials

Un esempio di utilizzo dell' GetClusterSessionCredentials API EMR è il seguente:

```
aws emr  get-cluster-session-credentials --cluster-id <cluster_ID> --execution-role-
arn <IAM_role_arn>
```

- È necessario evitare gli argomenti del comando Bash quando si eseguono comandi con il file `command-runner.jar` 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>
```

  Inoltre, è necessario evitare gli argomenti del comando Bash quando si eseguono comandi con lo script runner. Di seguito è riportato un esempio che mostra l'impostazione delle proprietà di Spark, con caratteri di escape inclusi:

```
"\"--conf spark.sql.autoBroadcastJoinThreshold=-1\n--conf
 spark.cradle.RSv2Mode.enabled=true\""
```

- I ruoli di runtime non forniscono supporto per il controllo dell'accesso alle risorse del cluster, come HDFS e HMS.
- I ruoli di runtime non forniscono supporto per docker/container.

Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS

Amazon EMR e applicazioni come Hadoop e Spark hanno bisogno di autorizzazioni per accedere ad altre risorse AWS ed eseguire operazioni quando sono in esecuzione. Ogni cluster in Amazon EMR deve avere un ruolo di servizio e un ruolo per il profilo dell' EC2 istanza Amazon. Per ulteriori informazioni, consulta Ruoli IAM e Utilizzo dei profili dell'istanza nella Guida per l'utente IAM. Le policy IAM allegate a questi ruoli forniscono al cluster le autorizzazioni per interagire con altri servizi AWS per conto di un utente.

Un ruolo aggiuntivo, il ruolo Auto Scaling, è necessario se il cluster utilizza la scalabilità automatica in Amazon EMR. Il ruolo AWS di servizio per EMR Notebooks è richiesto se si utilizzano EMR Notebooks.

Amazon EMR fornisce ruoli e policy gestite predefiniti che determinano le autorizzazioni per ciascun ruolo. Le policy gestite vengono create e gestite da AWS, quindi vengono aggiornate automaticamente se i requisiti del servizio cambiano. Per ulteriori informazioni, consulta Policy gestite da AWS nella Guida per l'utente IAM.

Se si sta creando un cluster o un notebook per la prima volta in un account, i ruoli per Amazon EMR non esistono ancora. Dopo averle create, puoi visualizzare i ruoli, le policy ad esse associate e le autorizzazioni consentite o negate dalle politiche nella console IAM (https://console.aws.amazon.com/iam/). È possibile specificare ruoli predefiniti per Amazon EMR da creare e utilizzare, è possibile creare ruoli propri e specificarli individualmente al momento della creazione di un cluster per personalizzare le autorizzazioni, infine è possibile specificare ruoli predefiniti da utilizzare al momento della creazione di un cluster utilizzando la AWS CLI. Per ulteriori informazioni, consulta Personalizza i ruoli IAM con Amazon EMR.

Modifica di policy basate sull'identità per le autorizzazioni a passare i ruoli del servizio per Amazon EMR

Le policy gestite predefinite con autorizzazioni complete di Amazon EMR incorporano configurazioni di sicurezza `iam:PassRole`, tra cui:

- Autorizzazioni `iam:PassRole` solo per specifici ruoli Amazon EMR predefiniti.
- `iam:PassedToService`condizioni che consentono di utilizzare la policy solo con AWS servizi specifici, come `elasticmapreduce.amazonaws.com` e `ec2.amazonaws.com`.

Puoi visualizzare la versione JSON delle policy Amazon _v2 EMRFullAccessPolicye Amazon EMRService Policy_v2 nella console IAM. Ti consigliamo di creare i nuovi cluster con le policy gestite v2.

Riepilogo del ruolo di servizio

La tabella seguente elenca i ruoli del servizio IAM associati ad Amazon EMR per riferimento rapido.

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
Ruolo di servizio per Amazon EMR (ruolo EMR)	`EMR_DefaultRole_V2`	Consente ad Amazon EMR di chiamare altri AWS servizi per	`AmazonEMRServicePolicy_v2`

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
		tuo conto durante il provisioning di risorse e l'esecuzione di azioni a livello di servizio. Questo ruolo è obbligatorio per tutti i cluster.	**Important** Per richiedere le Istanze spot è necessari o un ruolo collegato al servizio. Se questo ruolo non esiste, il ruolo di servizio Amazon EMR deve avere l'autoriz zazione per crearlo, altrimenti si verifica un errore di autorizza zione. Se si prevede di richiedere istanze Spot, è necessari o aggiornare questa policy per includere un'istruzione che consenta la creazione di questo ruolo collegato

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
			al servizio. Per ulteriori informazioni, consulta Ruolo di servizio per Amazon EMR (ruolo EMR) il ruolo collegato ai servizi per le richieste di istanze Spot nella Amazon EC2 User Guide.

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2	`EMR_EC2_DefaultRole`	I processi applicativi eseguiti sull'ecosistema Hadoop su istanze cluster utilizzano questo ruolo quando chiamano altri servizi. AWS Per accedere ai dati in Amazon S3 utilizzando EMRFS, è possibile specificare diversi ruoli da assumere in base alla posizione dei dati in Amazon S3. Ad esempio, più team possono accedere a un singolo "account di archiviazione" dei dati di Amazon S3. Per ulteriori informazioni, consulta Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3. Questo ruolo è obbligatorio per tutti i cluster.	`AmazonElasticMapReduceforEC2Role`. Per ulteriori informazioni, consulta Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2.

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
Ruolo di servizio per il dimensionamento automatico in Amazon EMR (ruolo Auto Scaling)	`EMR_AutoScaling_DefaultRole`	Consente di eseguire azioni aggiuntive per ambienti con dimensionamento dinamico. Necessario solo per i cluster che utilizzano la scalabilità automatica in Amazon EMR. Per ulteriori informazioni, consulta Utilizzo del ridimensionamento automatico con una politica personalizzata, ad esempio gruppi in Amazon EMR.	`AmazonElasticMapReduceforAutoScalingRole` . Per ulteriori informazioni, consulta Ruolo di servizio per il dimensionamento automatico in Amazon EMR (ruolo Auto Scaling).

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
Ruolo di servizio per EMR Notebooks	`EMR_Notebooks_DefaultRole`	Fornisce le autorizzazioni necessarie a un notebook EMR per accedere ad AWS altre risorse ed eseguire azioni. Richiesto solo se si utilizza EMR Notebooks.	`AmazonElasticMapReduceEditorsRole` . Per ulteriori informazioni, consulta Ruolo di servizio per EMR Notebooks. `S3FullAccessPolicy` è inoltre collegato per impostazione predefinita. Il contenuto di questa policy è mostrato di seguito. JSON `{ "Version": "2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "s3:*" ],`

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
			"Resource": ["*"], "Sid": "AllowS3" }] }

Funzione	Ruolo predefinito	Descrizione	Policy gestita predefinita
Ruolo collegato ai servizi	`AWSServiceRoleForEMRCleanup`	Amazon EMR crea automaticamente un ruolo collegato ai servizi. Se il servizio per Amazon EMR non è più in grado di pulire EC2 le risorse Amazon, Amazon EMR può utilizzare questo ruolo per eseguire la pulizia. Se un cluster usa le istanze Spot, le policy di autorizzazione associate a Ruolo di servizio per Amazon EMR (ruolo EMR) devono consentire la creazione di un ruolo collegato al servizio. Per ulteriori informazioni, consulta Utilizzo di ruoli collegati ai servizi per Amazon EMR.	`AmazonEMRCleanupPolicy`

Argomenti

- Ruoli di servizio IAM utilizzati da Amazon EMR
- Personalizza i ruoli IAM con Amazon EMR
- Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3
- Utilizzo di policy basate su risorse per l'accesso Amazon EMR ad AWS Glue Data Catalog
- I ruoli IAM possono essere utilizzati con le applicazioni che richiamano direttamente i servizi AWS

- Consentire a utenti e gruppi di creare e modificare i ruoli

Ruoli di servizio IAM utilizzati da Amazon EMR

Amazon EMR impiega i ruoli di servizio IAM per eseguire operazioni per conto dell'utente durante il provisioning delle risorse cluster, l'esecuzione di applicazioni, il dimensionamento dinamico delle risorse e la creazione ed esecuzione di EMR Notebooks. Amazon EMR utilizza i seguenti ruoli durante l'interazione con altri servizi AWS . Ogni ruolo dispone di una funzione univoca all'interno di Amazon EMR. Gli argomenti in questa sezione descrivono la funzione del ruolo e forniscono i ruoli e le policy di autorizzazioni predefiniti per ciascun ruolo.

Se nel cluster è presente un codice applicativo che chiama direttamente AWS i servizi, potrebbe essere necessario utilizzare l'SDK per specificare i ruoli. Per ulteriori informazioni, consulta I ruoli IAM possono essere utilizzati con le applicazioni che richiamano direttamente i servizi AWS.

Argomenti

- Ruolo di servizio per Amazon EMR (ruolo EMR)
- Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2
- Ruolo di servizio per il dimensionamento automatico in Amazon EMR (ruolo Auto Scaling)
- Ruolo di servizio per EMR Notebooks
- Utilizzo di ruoli collegati ai servizi per Amazon EMR

Ruolo di servizio per Amazon EMR (ruolo EMR)

Il ruolo Amazon EMR definisce le azioni consentite per Amazon EMR quando fornisce risorse ed esegue attività a livello di servizio che non vengono eseguite nel contesto di un'istanza Amazon in esecuzione all'interno di un cluster. EC2 Ad esempio, il ruolo di servizio viene utilizzato per fornire EC2 istanze all'avvio di un cluster.

- Il nome del ruolo predefinito è `EMR_DefaultRole_V2`.
- La policy gestita predefinita necessaria per Amazon EMR e associata a `EMR_DefaultRole_V2` è `AmazonEMRServicePolicy_v2`. Questa policy v2 sostituisce la policy gestita predefinita e obsoleta `AmazonElasticMapReduceRole`.

`AmazonEMRServicePolicy_v2` dipende dall'accesso ridotto alle risorse che Amazon EMR fornisce o utilizza. Quando si utilizza questa policy, è necessario passare il tag utente `for-use-with-`

`amazon-emr-managed-policies = true` durante il provisioning del cluster. Amazon EMR propaga automaticamente tali tag. Inoltre, potrebbe essere necessario aggiungere manualmente un tag utente a tipi specifici di risorse, ad esempio gruppi di EC2 sicurezza che non sono stati creati da Amazon EMR. Consultare Assegnazione di tag alle risorse per l'utilizzo delle policy gestite.

> **Important**
>
> Amazon EMR utilizza questo ruolo del servizio Amazon EMR e il `AWSServiceRoleForEMRCleanup` ruolo per ripulire le risorse del cluster nel tuo account che non usi più, come le istanze Amazon. EC2 È necessario includere azioni relative alle policy del ruolo per eliminare o terminare le risorse. Altrimenti, Amazon EMR non può eseguire queste azioni di pulizia e potresti incorrere in costi per le risorse inutilizzate che rimangono nel cluster.

Nell'esempio seguente viene mostrato il contenuto della policy `AmazonEMRServicePolicy_v2` corrente. È anche possibile visualizzare il contenuto corrente della policy gestita `AmazonEMRServicePolicy_v2` sulla console 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"
        }
      }
    }
  ]
}
```

Il tuo ruolo di servizio dovrebbe utilizzare la seguente policy di attendibilità:

> **Important**
>
> La policy di attendibilità seguente include le chiavi di condizione globale `aws:SourceArn` e `aws:SourceAccount` per limitare le autorizzazioni concesse ad Amazon EMR per determinate risorse dell'account. Il loro utilizzo può proteggerti dal problema del "confused deputy".

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:*"
        }
      }
    }
  ]
}
```

Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2

Il ruolo di servizio per EC2 le istanze di cluster (chiamato anche profilo di EC2 istanza per Amazon EMR) è un tipo speciale di ruolo di servizio assegnato a EC2 ogni istanza di un cluster Amazon EMR all'avvio dell'istanza. I processi delle applicazioni eseguite sull'ecosistema Hadoop presuppongono che questo ruolo per le autorizzazioni interagisca con altri servizi AWS .

Per ulteriori informazioni sui ruoli di servizio per EC2 le istanze, consulta Usare un ruolo IAM per concedere le autorizzazioni alle applicazioni in esecuzione su EC2 istanze Amazon nella IAM User Guide.

> **Important**
>
> Il ruolo di servizio predefinito per EC2 le istanze di cluster e la politica gestita AWS predefinita associata `AmazonElasticMapReduceforEC2Role` sono sulla via dell'obsolescenza e non sono previste politiche gestite sostitutive. AWS Sarà necessario creare e specificare un profilo di istanza per sostituire il ruolo obsoleto e la policy predefinita.

Ruolo predefinito e policy gestita

- Il nome del ruolo predefinito è `EMR_EC2_DefaultRole`.
- Il supporto per la policy gestita da `EMR_EC2_DefaultRole` predefinita (`AmazonElasticMapReduceforEC2Role`) è quasi al termine. Invece di utilizzare una policy gestita predefinita per il profilo dell' EC2 istanza, applica policy basate sulle risorse ai bucket S3 e ad altre risorse di cui Amazon EMR ha bisogno, oppure usa la tua policy gestita dal cliente con un ruolo IAM come profilo di istanza. Per ulteriori informazioni, consulta Creazione di un ruolo di servizio per le istanze di cluster con autorizzazioni con privilegi minimi EC2 .

Di seguito viene mostrato il contenuto della versione 3 di `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"
```

```
    }
  ]
}
```

Il tuo ruolo di servizio dovrebbe utilizzare la seguente policy di attendibilità:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowSTSAssumerole",
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMR_EC2_DefaultRole"
    }
  ]
}
```

Creazione di un ruolo di servizio per le istanze di cluster con autorizzazioni con privilegi minimi EC2

Come procedura consigliata, consigliamo vivamente di creare un ruolo di servizio per le EC2 istanze del cluster e una politica di autorizzazioni che disponga delle autorizzazioni minime per gli altri servizi richiesti dall'applicazione. AWS

La policy gestita predefinita `AmazonElasticMapReduceforEC2Role` offre le autorizzazioni che facilitano l'avvio del primo cluster. Tuttavia, `AmazonElasticMapReduceforEC2Role` è sulla via dell'obsolescenza e Amazon EMR non fornirà una policy predefinita AWS gestita sostitutiva per il ruolo obsoleto. Per avviare un cluster iniziale, è necessario fornire una policy basata sulle risorse gestita dal cliente o basata su ID.

Le seguenti dichiarazioni di policy forniscono esempi di autorizzazioni richieste per differenti caratteristiche di Amazon EMR. È consigliabile utilizzare queste autorizzazioni per la creazione di una policy di autorizzazioni che limita l'accesso alle sole caratteristiche e risorse che richiede il cluster. Tutte le dichiarazioni politiche di esempio utilizzano la regione e l'ID dell'account fittizio*us-west-2*. AWS *123456789012* Sostituirli nel modo appropriato per il cluster.

Per ulteriori informazioni sulla creazione e sula specifica di ruoli personalizzati, consulta Personalizza i ruoli IAM con Amazon EMR.

> **Note**
>
> Se crei un ruolo EMR personalizzato per EC2, segui il flusso di lavoro di base, che crea automaticamente un profilo di istanza con lo stesso nome. Amazon EC2 consente di creare profili e ruoli di istanza con nomi diversi, ma Amazon EMR non supporta questa configurazione e genera un errore di «profilo di istanza non valido» durante la creazione del cluster.

Lettura e scrittura di dati su Amazon S3 utilizzando EMRFS

Quando un'applicazione in esecuzione su un cluster Amazon EMR fa riferimento ai dati utilizzando il `s3://mydata` formato, Amazon EMR utilizza il profilo dell' EC2 istanza per effettuare la richiesta. I cluster in genere leggono e scrivono dati su Amazon S3 in questo modo e Amazon EMR utilizza per impostazione predefinita le autorizzazioni associate al ruolo di servizio per le istanze di cluster. EC2 Per ulteriori informazioni, consulta Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3.

Poiché i ruoli IAM per EMRFS ricorreranno alle autorizzazioni associate al ruolo di servizio per le EC2 istanze cluster, come best practice, consigliamo di utilizzare i ruoli IAM per EMRFS e di limitare le autorizzazioni EMRFS e Amazon S3 associate al ruolo di servizio per le istanze cluster. EC2

L'istruzione di esempio di seguito mostra le autorizzazioni che EMRFS richiede per effettuare le richieste ad Amazon S3.

- `my-data-bucket-in-s3-for-emrfs-reads-and-writes` specifica il bucket in Amazon S3 dove il cluster legge e scrive i dati e tutte le sottocartelle utilizzando `/*`. Aggiungere solo i bucket e le cartelle necessari per la propria applicazione.
- La dichiarazione di policy che consente le azioni `dynamodb` è necessaria solo se è abilitata la visualizzazione coerente di EMRFS. `EmrFSMetadata` specifica la cartella predefinita per la visualizzazione coerente di 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"
    }
  ]
}
```

Archiviazione di file di log in Amazon S3

La seguente istruzione di policy consente al cluster Amazon EMR di archiviare i file di log nel percorso Amazon S3 specificato. Nell'esempio seguente, quando il cluster è stato creato, *s3://MyLoggingBucket/MyEMRClusterLogs* è stato specificato utilizzando la posizione della cartella Log S3 nella console, utilizzando l'`--log-uri` AWS CLI opzione from o utilizzando il `LogUri` parametro nel `RunJobFlow` comando. Per ulteriori informazioni, consulta Archiviazione di file di log in Amazon S3.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::MyLoggingBucket/MyEMRClusterLogs/*"
      ],
      "Sid": "AllowS3Putobject"
    }
  ]
}
```

Utilizzo del AWS Glue Data Catalog

La seguente dichiarazione sulla politica consente le azioni necessarie se si utilizza il AWS Glue Data Catalog come metastore per le applicazioni. Per ulteriori informazioni, consulta Using the AWS Glue Data Catalog come metastore per Spark SQL, Using the AWS Glue Data Catalog come metastore per Hive e Using Presto with the Glue AWS Data Catalog nella Amazon EMR Release Guide.

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"
    }
  ]
}
```

Ruolo di servizio per il dimensionamento automatico in Amazon EMR (ruolo Auto Scaling)

Il ruolo Auto Scaling per Amazon EMR svolge una funzione simile a quella del ruolo di servizio, ma consente azioni aggiuntive per ambienti di scalabilità dinamica.

- Il nome del ruolo predefinito è `EMR_AutoScaling_DefaultRole`.
- La policy gestita predefinita associata a `EMR_AutoScaling_DefaultRole` è `AmazonElasticMapReduceforAutoScalingRole`.

I contenuti della versione 1 di `AmazonElasticMapReduceforAutoScalingRole` sono elencati di seguito.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "cloudwatch:DescribeAlarms",
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowCLOUDWATCHDescribealarms"
    }
  ]
}
```

Il tuo ruolo di servizio dovrebbe utilizzare la seguente policy di attendibilità:

> **Important**
>
> La policy di attendibilità seguente include le chiavi di condizione globale `aws:SourceArn` e `aws:SourceAccount` per limitare le autorizzazioni concesse ad Amazon EMR per determinate risorse dell'account. Il loro utilizzo può proteggerti dal problema del "confused deputy".

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"
    }
  ]
}
```

Ruolo di servizio per EMR Notebooks

Ogni notebook EMR necessita delle autorizzazioni per accedere ad altre AWS risorse ed eseguire azioni. Le policy IAM associate a questo ruolo di servizio forniscono le autorizzazioni per consentire al notebook di interagire con altri servizi. AWS Quando si crea un notebook utilizzando AWS Management Console, si specifica un ruolo di AWS servizio. È possibile utilizzare il ruolo predefinito, `EMR_Notebooks_DefaultRole`, oppure specificare un ruolo creato. Se un notebook non è stato creato in precedenza, è possibile scegliere di creare il ruolo predefinito.

- Il nome del ruolo predefinito è `EMR_Notebooks_DefaultRole`.
- Le policy gestite di default allegate a `EMR_Notebooks_DefaultRole` sono `AmazonElasticMapReduceEditorsRole` e `S3FullAccessPolicy`.

Il tuo ruolo di servizio dovrebbe utilizzare la seguente policy di attendibilità:

> **Important**
>
> La policy di attendibilità seguente include le chiavi di condizione globale `aws:SourceArn` e `aws:SourceAccount` per limitare le autorizzazioni concesse ad Amazon EMR per determinate risorse dell'account. Il loro utilizzo può proteggerti dal problema del "confused deputy".

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"
    }
  ]
}
```

Il contenuto della versione 1 di `AmazonElasticMapReduceEditorsRole` è il seguente.

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"
    }
  ]
}
```

Di seguito sono riportati il contenuto di `S3FullAccessPolicy`. La `S3FullAccessPolicy` consente al tuo ruolo di servizio per i EMR Notebooks di eseguire tutte le operazioni di Amazon S3 sugli oggetti nel Account AWS. Quando crei un ruolo di servizio personalizzato per i EMR Notebooks, devi assegnare al tuo ruolo di servizio le autorizzazioni Amazon S3.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:*"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowS3"
    }
  ]
}
```

È possibile individuare l'accesso in lettura e scrittura per il ruolo di servizio nella posizione Amazon S3 in cui si desidera salvare i file del notebook. Utilizza il seguente set minimo di autorizzazioni Amazon S3.

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

Se il bucket Amazon S3 è crittografato, devi includere le seguenti autorizzazioni per AWS Key Management Service.

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

Quando si collegano i repository Git al notebook e si deve creare un segreto per il repository, è necessario aggiungere l'autorizzazione `secretsmanager:GetSecretValue` nella policy IAM

collegata al ruolo di servizio per i notebooks Amazon EMR. Di seguito è illustrato un esempio di policy:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor0",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Autorizzazioni del ruolo di servizio EMR Notebooks

In questa tabella sono elencate le azioni eseguite da EMR Notebooks utilizzando il ruolo di servizio, assieme alle autorizzazioni necessarie per ogni azione.

Azione	Autorizzazioni
Stabilisci un canale di rete protetto tra un notebook e un cluster Amazon EMR ed esegui le azioni di pulizia necessarie.	`"ec2:CreateNetworkInterface", "ec2:CreateNetworkInterfacePermission", "ec2:DeleteNetworkInterface", "ec2:DeleteNetworkInterfacePermission", "ec2:DescribeNetworkInterfaces", "ec2:ModifyNetworkInterfaceAttribute", "ec2:AuthorizeSecurityGroupEgress", "ec2:AuthorizeSecurityGroupIngress", "ec2:CreateSecurityGroup", "ec2:DescribeSecurityGroups", "ec2:RevokeSecurityGroupEgress", "ec2:DescribeTags",`

Azione	Autorizzazioni
	"ec2:DescribeInstances", "ec2:DescribeSubnets", "ec2:DescribeVpcs", "elasticmapreduce:ListInstances", "elasticmapreduce:DescribeCluster", "elasticmapreduce:ListSteps"
Utilizza le credenziali Git memorizzate in AWS Secrets Manager per collegare i repository Git a un notebook.	"secretsmanager:GetSecretValue"
Applica i AWS tag all'interfaccia di rete e ai gruppi di sicurezza predefiniti creati da EMR Notebooks durante la configurazione del canale di rete sicuro. Per ulteriori informazioni, consulta Assegnazione di tag alle risorse AWS.	"ec2:CreateTags"
Accedi o carica i file notebook e i metadati in Amazon S3.	"s3:PutObject", "s3:GetObject", "s3:GetEncryptionConfiguration", "s3:ListBucket", "s3:DeleteObject" Le seguenti autorizzazioni sono necessarie solo se utilizzi un bucket Amazon S3 crittografato. "kms:Decrypt", "kms:GenerateDataKey", "kms:ReEncryptFrom", "kms:ReEncryptTo", "kms:DescribeKey"

EMR Notebooks: aggiornamenti alle policy gestite AWS

Visualizza i dettagli sugli aggiornamenti delle policy AWS gestite per EMR Notebooks dal 1° marzo 2021.

Modifica	Descrizione	Data
`AmazonElasticMapReduceEditorsRole - Added permissions`	EMR Notebooks ha aggiunto le autorizzazioni `ec2:describeVPCs` e `elastmicmapreduce:ListSteps` a `AmazonElasticMapReduceEditorsRole` .	8 febbraio 2023
EMR Notebooks ha avviato il tracciamento delle modifiche	EMR Notebooks ha iniziato a tenere traccia delle modifiche per le sue policy gestite. AWS	8 febbraio 2023

Utilizzo di ruoli collegati ai servizi per Amazon EMR

Amazon EMR utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo di ruolo IAM univoco collegato direttamente ad Amazon EMR. I ruoli collegati ai servizi sono predefiniti da Amazon EMR e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

Argomenti

- Utilizzo di ruoli collegati ai servizi per Amazon EMR for cleanup
- Utilizzo di ruoli collegati ai servizi con Amazon EMR per la registrazione write-ahead

Per informazioni su altri servizi che supportano i ruoli collegati ai servizi, consulta i AWS servizi che funzionano con IAM e cerca i servizi con Sì nella colonna Ruoli collegati ai servizi. Scegli Sì in corrispondenza di un link per visualizzare la documentazione relativa al ruolo collegato al servizio per tale servizio.

Utilizzo di ruoli collegati ai servizi per Amazon EMR for cleanup

Amazon EMR utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo di ruolo IAM univoco collegato direttamente ad Amazon EMR. I ruoli

collegati ai servizi sono predefiniti da Amazon EMR e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

I ruoli collegati ai servizi interagiscono con il ruolo di servizio Amazon EMR e il EC2 profilo di istanza Amazon per Amazon EMR. Per ulteriori informazioni sul ruolo del servizio e il profilo dell'istanza, consulta Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS.

Un ruolo collegato al servizio semplifica la configurazione di Amazon EMR perché non è necessario aggiungere manualmente le autorizzazioni necessarie. Amazon EMR definisce le autorizzazioni dei suoi ruoli collegati ai servizi e, se non diversamente definito, solo Amazon EMR può assumerne i ruoli. Le autorizzazioni definite includono la policy di attendibilità e la policy delle autorizzazioni che non può essere collegata a nessun'altra entità IAM.

Puoi eliminare questo ruolo collegato al servizio per Amazon EMR solo dopo aver eliminato tutte le risorse correlate e terminato tutti i cluster EMR nell'account. Ciò protegge le tue risorse Amazon EMR in modo da non poter rimuovere inavvertitamente l'autorizzazione ad accedere alle risorse.

Utilizzo di ruoli collegati ai servizi per la pulizia

Amazon EMR utilizza il ruolo basato sui servizi per concedere AWSServiceRoleForEMRCleanupad Amazon EMR l'autorizzazione a terminare ed eliminare le risorse Amazon per tuo conto se il ruolo EC2 collegato al servizio Amazon EMR perde tale capacità. Amazon EMR crea automaticamente il ruolo collegato al servizio durante la creazione del cluster, se non esiste già.

Il ruolo AWSService RoleFor EMRCleanup collegato ai servizi si affida ai seguenti servizi per l'assunzione del ruolo:

- `elasticmapreduce.amazonaws.com`

La politica AWSService RoleFor EMRCleanup di autorizzazione dei ruoli collegati al servizio consente ad Amazon EMR di completare le seguenti azioni sulle risorse specificate:

- Operazione: `DescribeInstances` su `ec2`
- Operazione: `DescribeLaunchTemplates` su `ec2`
- Operazione: `DeleteLaunchTemplate` su `ec2`
- Operazione: `DescribeSpotInstanceRequests` su `ec2`
- Operazione: `ModifyInstanceAttribute` su `ec2`

- Operazione: `TerminateInstances` su `ec2`
- Operazione: `CancelSpotInstanceRequests` su `ec2`
- Operazione: `DeleteNetworkInterface` su `ec2`
- Operazione: `DescribeInstanceAttribute` su `ec2`
- Operazione: `DescribeVolumeStatus` su `ec2`
- Operazione: `DescribeVolumes` su `ec2`
- Operazione: `DetachVolume` su `ec2`
- Operazione: `DeleteVolume` su `ec2`
- Operazione: `DescribePlacementGroups` su `ec2`
- Operazione: `DeletePlacementGroup` su `ec2`

Per consentire a un'entità IAM (come un utente, un gruppo o un ruolo) di creare, modificare o eliminare un ruolo collegato ai servizi devi configurare le relative autorizzazioni.

Creazione di un ruolo collegato ai servizi per Amazon EMR

Non è necessario creare manualmente il ruolo. AWSService RoleFor EMRCleanup Quando avvii un cluster, per la prima volta o quando il ruolo AWSService RoleFor EMRCleanup collegato al servizio non è presente, Amazon EMR crea il ruolo collegato al AWSService RoleFor EMRCleanup servizio per te. È necessario disporre delle autorizzazioni per creare un ruolo collegato al servizio. Per un'istruzione di esempio che aggiunge questa funzionalità alla politica di autorizzazione di un'entità IAM (come un utente, un gruppo o un ruolo):

Aggiungi la seguente dichiarazione alla politica di autorizzazione per l'entità IAM che deve creare il ruolo collegato al servizio.

```
{
   "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**
>
> Se hai utilizzato Amazon EMR prima del 24 ottobre 2017, quando i ruoli collegati ai servizi non erano supportati, Amazon EMR ha creato il ruolo collegato al servizio nel tuo account. AWSService RoleFor EMRCleanup Per ulteriori informazioni, consulta Comparsa di un nuovo ruolo nell'account IAM.

Modifica di un ruolo collegato ai servizi per Amazon EMR

Amazon EMR non consente di modificare il ruolo collegato al AWSService RoleFor EMRCleanup servizio. Dopo aver creato un ruolo collegato al servizio, non puoi modificare il nome del ruolo collegato al servizio perché varie entità potrebbero fare riferimento al ruolo collegato al servizio. Tuttavia, puoi modificare la descrizione del ruolo collegato al servizio utilizzando IAM.

Modifica della descrizione di un ruolo collegato ai servizi (console IAM)

È possibile utilizzare la console IAM per modificare la descrizione di un ruolo collegato ai servizi.

Per modificare la descrizione di un ruolo collegato ai servizi (console)

1. Nel pannello di navigazione della console IAM seleziona Roles (Ruoli).
2. Scegliere il nome del ruolo da modificare.
3. Nella parte destra di Role description (Descrizione ruolo), scegli Edit (Modifica).
4. Digita una nuova descrizione nella casella e scegli Save changes (Salva modifiche).

Modifica della descrizione di un ruolo collegato ai servizi (CLI IAM)

Puoi utilizzare i comandi IAM di AWS Command Line Interface per modificare la descrizione di un ruolo collegato al servizio.

Per modificare la descrizione di un ruolo collegato ai servizi (CLI)

1. (Facoltativo) Per visualizzare la descrizione attuale di un ruolo, utilizza i seguenti comandi:

```
$ aws iam get-role --role-name role-name
```

Per fare riferimento ai ruoli con i comandi della CLI utilizza il nome del ruolo, non l'ARN. Ad esempio, per fare riferimento a un ruolo il cui ARN è `arn:aws:iam::123456789012:role/myrole`, puoi usare **`myrole`**.

2. Per aggiornare la descrizione di un ruolo collegato ai servizi, utilizza uno dei seguenti comandi:

```
$ aws iam update-role-description --role-name role-name --description description
```

Modifica della descrizione di un ruolo collegato ai servizi (API IAM)

È possibile utilizzare l'API IAM per modificare la descrizione di un ruolo collegato ai servizi.

Per modificare la descrizione di un ruolo collegato ai servizi (API)

1. (Facoltativo) Per visualizzare la descrizione attuale di un ruolo, utilizza il seguente comando:

   API IAM: GetRole
2. Per aggiornare la descrizione di un ruolo, utilizza il seguente comando:

   API IAM: UpdateRoleDescription

Eliminazione di un ruolo collegato ai servizi per Amazon EMR

Se non hai più bisogno di utilizzare una funzionalità o un servizio che richiede un ruolo collegato al servizio, ti consigliamo di eliminare quel ruolo collegato al servizio. In questo modo non sarà più presente un'entità non utilizzata che non viene monitorata e gestita attivamente. Tuttavia, è necessario effettuare la pulizia delle risorse associate al ruolo collegato ai servizi prima di poterlo eliminare.

Pulizia di un ruolo collegato ai servizi

Prima di poter utilizzare IAM per eliminare un ruolo collegato al servizio, devi prima confermare che il ruolo collegato al servizio non abbia sessioni attive e rimuovere tutte le risorse utilizzate dal ruolo collegato al servizio.

Per verificare se il ruolo collegato ai servizi dispone di una sessione attiva nella console IAM

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.

2. Nel riquadro di navigazione, seleziona Ruoli. Seleziona il nome (non la casella di controllo) del ruolo collegato al servizio. AWSService RoleFor EMRCleanup
3. Nella pagina di riepilogo per il ruolo collegato al servizio selezionato, scegli Access Advisor.
4. Nella scheda Access Advisor (Consulente accessi) esaminare l'attività recente per il ruolo collegato ai servizi.

> **Note**
>
> Se non sei sicuro che Amazon EMR stia utilizzando AWSService RoleFor EMRCleanup il ruolo collegato al servizio, puoi provare a eliminare il ruolo collegato al servizio. Se il servizio utilizza il ruolo collegato al servizio, l'eliminazione non riesce e puoi visualizzare le regioni in cui viene utilizzato il ruolo collegato al servizio. Se viene utilizzato il ruolo collegato al servizio, è necessario attendere il termine della sessione prima di poter eliminare il ruolo collegato al servizio. Non puoi revocare la sessione per un ruolo collegato al servizio.

Per rimuovere le risorse Amazon EMR utilizzate da AWSService RoleFor EMRCleanup

- Chiudi tutti i cluster del tuo account. Per ulteriori informazioni, consulta Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

Eliminazione di un ruolo collegato ai servizi (console IAM)

È possibile utilizzare la console IAM per eliminare un ruolo collegato ai servizi.

Per eliminare un ruolo collegato ai servizi (console)

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Ruoli. Seleziona la casella di controllo accanto a AWSService RoleForEMRCleanup, non il nome o la riga stessa.
3. In operazioni Role (Ruolo) nella parte superiore della pagina, seleziona Delete (Elimina) ruolo.
4. Nella finestra di dialogo di conferma, esamina i dati dell'ultimo accesso al servizio, che mostrano l'ultima volta che ciascuno dei ruoli selezionati ha effettuato l'ultimo accesso a un AWS servizio. In questo modo potrai verificare se il ruolo è attualmente attivo. Per procedere, seleziona Yes, Delete (Sì, elimina).

5. Controlla le notifiche della console IAM per monitorare lo stato dell'eliminazione del ruolo collegato ai servizi. Poiché l'eliminazione del ruolo collegato al servizio IAM è asincrona, dopo aver inviato il ruolo collegato al servizio per l'eliminazione, l'operazione di eliminazione può avere esito positivo o negativo. Se il task non viene eseguito correttamente, puoi scegliere View details (Visualizza dettagli) o View Resources (Visualizza risorse) dalle notifiche per capire perché l'eliminazione non è stata effettuata. Se l'eliminazione non viene eseguita perché vi sono risorse nel servizio che sono usate dal ruolo, il motivo dell'errore include un elenco di risorse.

Eliminazione di un ruolo collegato ai servizi (CLI IAM)

Puoi utilizzare i comandi IAM di per eliminare un ruolo collegato al servizio. AWS Command Line Interface Poiché un ruolo collegato ai servizi non può essere eliminato se è in uso o se a esso sono associate delle risorse, occorre inviare una richiesta di eliminazione. Se queste condizioni non sono soddisfatte, la richiesta può essere rifiutata.

Per eliminare un ruolo collegato ai servizi (CLI)

1. Per controllare lo stato dell'attività di eliminazione, devi acquisire il `deletion-task-id` dalla risposta. Per inviare una richiesta di eliminazione per un ruolo collegato ai servizi, digita il seguente comando:

```
$ aws iam delete-service-linked-role --role-name AWSServiceRoleForEMRCleanup
```

2. Digita il seguente comando per verificare lo stato del task di eliminazione:

```
$ aws iam get-service-linked-role-deletion-status --deletion-task-id deletion-task-id
```

   Lo stato di un task di eliminazione può essere `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED`o `FAILED`. Se l'eliminazione non viene eseguita correttamente, la chiamata restituisce il motivo dell'errore per consentire all'utente di risolvere il problema.

Eliminazione di un ruolo collegato ai servizi (API IAM)

È possibile utilizzare l'API IAM per eliminare un ruolo collegato ai servizi. Poiché un ruolo collegato ai servizi non può essere eliminato se è in uso o se a esso sono associate delle risorse, occorre inviare una richiesta di eliminazione. Se queste condizioni non sono soddisfatte, la richiesta può essere rifiutata.

Per eliminare un ruolo collegato ai servizi (API)

1. Per inviare una richiesta di eliminazione per un ruolo collegato al servizio, chiama. DeleteServiceLinkedRole Nella richiesta, specifica il nome del AWSService RoleFor EMRCleanup ruolo.

   Per controllare lo stato dell'attività di eliminazione, devi acquisire il `DeletionTaskId` dalla risposta.

2. Chiamare GetServiceLinkedRoleDeletionStatus per controllare lo stato dell'eliminazione. Nella richiesta, specificare il `DeletionTaskId`.

   Lo stato di un task di eliminazione può essere `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED`o `FAILED`. Se l'eliminazione non viene eseguita correttamente, la chiamata restituisce il motivo dell'errore per consentire all'utente di risolvere il problema.

Regioni supportate per AWSService RoleFor EMRCleanup

Amazon EMR supporta l'utilizzo del ruolo AWSService RoleFor EMRCleanup collegato ai servizi nelle seguenti regioni.

Nome della Regione	Identità della regione	Supporto in Amazon EMR
US East (N. Virginia)	us-east-1	Sì
Stati Uniti orientali (Ohio)	us-east-2	Sì
US West (N. California)	us-west-1	Sì
US West (Oregon)	us-west-2	Sì
Asia Pacific (Mumbai)	ap-south-1	Sì
Asia Pacifico (Osaka-Locale)	ap-northeast-3	Sì
Asia Pacifico (Seul)	ap-northeast-2	Sì
Asia Pacifico (Singapore)	ap-southeast-1	Sì
Asia Pacifico (Sydney)	ap-southeast-2	Sì

Nome della Regione	Identità della regione	Supporto in Amazon EMR
Asia Pacifico (Tokyo)	ap-northeast-1	Sì
Canada (Central)	ca-central-1	Sì
Europe (Frankfurt)	eu-central-1	Sì
Europa (Irlanda)	eu-west-1	Sì
Europe (London)	eu-west-2	Sì
Europe (Paris)	eu-west-3	Sì
Sud America (São Paulo)	sa-east-1	Sì

Utilizzo di ruoli collegati ai servizi con Amazon EMR per la registrazione write-ahead

Amazon EMR utilizza ruoli collegati ai servizi AWS Identity and Access Management (IAM). Un ruolo collegato ai servizi è un tipo di ruolo IAM univoco collegato direttamente ad Amazon EMR. I ruoli collegati ai servizi sono predefiniti da Amazon EMR e includono tutte le autorizzazioni richieste dal servizio per chiamare altri servizi per tuo conto. AWS

I ruoli collegati ai servizi interagiscono con il ruolo di servizio Amazon EMR e il EC2 profilo di istanza Amazon per Amazon EMR. Per ulteriori informazioni sul ruolo del servizio e il profilo dell'istanza, consulta Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS.

Un ruolo collegato al servizio semplifica la configurazione di Amazon EMR perché non è necessario aggiungere manualmente le autorizzazioni necessarie. Amazon EMR definisce le autorizzazioni dei suoi ruoli collegati ai servizi e, se non diversamente definito, solo Amazon EMR può assumerne i ruoli. Le autorizzazioni definite includono la policy di attendibilità e la policy delle autorizzazioni che non può essere collegata a nessun'altra entità IAM.

Puoi eliminare questo ruolo collegato al servizio per Amazon EMR solo dopo aver eliminato le relative risorse e aver terminato tutti i cluster EMR nell'account. Ciò protegge le tue risorse Amazon EMR in modo da non poter rimuovere inavvertitamente l'autorizzazione ad accedere alle risorse.

Autorizzazioni di ruolo collegate al servizio per il write-ahead logging (WAL)

Amazon EMR utilizza il ruolo collegato al servizio AWSServiceRoleForEMRWAL per recuperare lo stato di un cluster.

Il ruolo collegato al servizio AWSService RoleFor EMRWAL si affida ai seguenti servizi per assumere il ruolo:

- `emrwal.amazonaws.com`

La politica di `EMRDescribeClusterPolicyForEMRWAL`autorizzazione per il ruolo collegato al servizio consente ad Amazon EMR di completare le seguenti azioni sulle risorse specificate:

- Operazione: `DescribeCluster` su `*`

È necessario configurare le autorizzazioni per consentire a un'entità IAM (in questo caso, Amazon EMR WAL) di creare, modificare o eliminare un ruolo collegato al servizio. Aggiungi le seguenti istruzioni, se necessario, alla politica di autorizzazione per il tuo profilo di istanza:

CreateServiceLinkedRole

Per consentire a un'entità IAM di creare il ruolo collegato al servizio AWSService RoleFor EMRWAL

Aggiungi la seguente istruzione alla policy delle autorizzazioni per l'entità IAM che deve creare il ruolo collegato ai servizi:

```
{
    "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

Per consentire a un'entità IAM di modificare la descrizione del ruolo collegato al servizio EMRWAL AWSService RoleFor

Aggiungi la seguente istruzione alla policy delle autorizzazioni per l'entità IAM che deve modificare la descrizione di un ruolo collegato ai servizi:

```
{
    "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

Per consentire a un'entità IAM di eliminare il ruolo collegato al servizio EMRWAL AWSService RoleFor

Aggiungi la seguente istruzione alla policy delle autorizzazioni per l'entità IAM che deve eliminare un ruolo collegato ai servizi:

```
{
    "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"
            ]
        }
    }
}
```

Creazione di un ruolo collegato ai servizi per Amazon EMR

Non è necessario creare manualmente il ruolo EMRWAL. AWSService RoleFor Amazon EMR crea automaticamente questo ruolo collegato al servizio quando crei uno spazio di lavoro WAL con l'EMRWAL CLI o AWS CloudFormation da, HBase oppure creerà il ruolo collegato al servizio quando configuri uno spazio di lavoro per Amazon EMR WAL e il ruolo collegato al servizio non esiste ancora. È necessario disporre delle autorizzazioni per creare un ruolo collegato al servizio. Per esempio le istruzioni che aggiungono questa funzionalità alla politica di autorizzazione di un'entità IAM (come un utente, un gruppo o un ruolo), consulta la sezione precedente,. Autorizzazioni di ruolo collegate al servizio per il write-ahead logging (WAL)

Modifica di un ruolo collegato ai servizi per Amazon EMR

Amazon EMR non consente di modificare il ruolo collegato al servizio AWSService RoleFor EMRWAL. Dopo aver creato un ruolo collegato al servizio, non puoi modificare il nome del ruolo collegato al servizio perché varie entità potrebbero fare riferimento al ruolo collegato al servizio. Tuttavia, puoi modificare la descrizione del ruolo collegato al servizio utilizzando IAM.

Modifica della descrizione di un ruolo collegato ai servizi (console IAM)

È possibile utilizzare la console IAM per modificare la descrizione di un ruolo collegato ai servizi.

Per modificare la descrizione di un ruolo collegato ai servizi (console)

1. Nel pannello di navigazione della console IAM seleziona Roles (Ruoli).
2. Scegliere il nome del ruolo da modificare.
3. Nella parte destra di Role description (Descrizione ruolo), scegli Edit (Modifica).
4. Digita una nuova descrizione nella casella e scegli Save changes (Salva modifiche).

Modifica della descrizione di un ruolo collegato ai servizi (CLI IAM)

Puoi utilizzare i comandi IAM di AWS Command Line Interface per modificare la descrizione di un ruolo collegato al servizio.

Per modificare la descrizione di un ruolo collegato ai servizi (CLI)

1. (Facoltativo) Per visualizzare la descrizione attuale di un ruolo, utilizza i seguenti comandi:

   ```
   $ aws iam get-role --role-name role-name
   ```

   Per fare riferimento ai ruoli con i comandi della CLI utilizza il nome del ruolo, non l'ARN. Ad esempio, per fare riferimento a un ruolo il cui ARN è `arn:aws:iam::123456789012:role/myrole`, puoi usare **`myrole`**.
2. Per aggiornare la descrizione di un ruolo collegato ai servizi, utilizza uno dei seguenti comandi:

   ```
   $ aws iam update-role-description --role-name role-name --description description
   ```

Modifica della descrizione di un ruolo collegato ai servizi (API IAM)

È possibile utilizzare l'API IAM per modificare la descrizione di un ruolo collegato ai servizi.

Per modificare la descrizione di un ruolo collegato ai servizi (API)

1. (Facoltativo) Per visualizzare la descrizione attuale di un ruolo, utilizza il seguente comando:

   API IAM: GetRole
2. Per aggiornare la descrizione di un ruolo, utilizza il seguente comando:

   API IAM: UpdateRoleDescription

Eliminazione di un ruolo collegato ai servizi per Amazon EMR

Se non hai più bisogno di utilizzare una funzionalità o un servizio che richiede un ruolo collegato al servizio, ti consigliamo di eliminare quel ruolo collegato al servizio. In questo modo non sarà più presente un'entità non utilizzata che non viene monitorata e gestita attivamente. Tuttavia, è necessario effettuare la pulizia delle risorse associate al ruolo collegato ai servizi prima di poterlo eliminare.

> **Note**
>
> L'operazione di registrazione write-ahead non viene influenzata dall'eliminazione del AWSService RoleFor ruolo EMRWAL, ma Amazon EMR non eliminerà automaticamente i log che ha creato una volta terminato il cluster EMR. Pertanto, dovrai eliminare manualmente i log WAL di Amazon EMR se elimini il ruolo collegato al servizio.

Pulizia di un ruolo collegato ai servizi

Prima di utilizzare IAM per eliminare un ruolo collegato ai servizi, devi innanzitutto verificare che il ruolo non abbia sessioni attive ed eliminare tutte le risorse utilizzate dal ruolo.

Per verificare se il ruolo collegato ai servizi dispone di una sessione attiva nella console IAM

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Ruoli. Seleziona il nome (non la casella di controllo) del ruolo EMRWAL. AWSService RoleFor
3. Nella pagina Summary (Riepilogo) per il ruolo selezionato, scegli Access Advisor (Consulente accessi).
4. Nella scheda Access Advisor (Consulente accessi) esaminare l'attività recente per il ruolo collegato ai servizi.

> **Note**
>
> Se non sei sicuro che Amazon EMR stia utilizzando AWSService RoleFor il ruolo EMRWAL, puoi provare a eliminare il ruolo collegato al servizio. Se il servizio utilizza il ruolo, l'eliminazione non riesce e puoi visualizzare le regioni in cui viene utilizzato il ruolo collegato al servizio. Se viene utilizzato il ruolo collegato al servizio, è necessario attendere il termine della sessione prima di poter eliminare il ruolo collegato al servizio. Non puoi revocare la sessione per un ruolo collegato al servizio.

Per rimuovere le risorse Amazon EMR utilizzate dall'EMRWAL AWSService RoleFor

- Chiudi tutti i cluster del tuo account. Per ulteriori informazioni, consulta Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

Eliminazione di un ruolo collegato ai servizi (console IAM)

È possibile utilizzare la console IAM per eliminare un ruolo collegato ai servizi.

Per eliminare un ruolo collegato ai servizi (console)

1. Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.
2. Nel riquadro di navigazione, seleziona Ruoli. Seleziona la casella di controllo accanto a AWSService RoleFor EMRWAL, non il nome o la riga stessa.
3. In operazioni Role (Ruolo) nella parte superiore della pagina, seleziona Delete (Elimina) ruolo.
4. Nella finestra di dialogo di conferma, esamina i dati dell'ultimo accesso al servizio, che mostrano l'ultima volta che ciascuno dei ruoli selezionati ha effettuato l'ultimo accesso a un AWS servizio. In questo modo potrai verificare se il ruolo è attualmente attivo. Per procedere, seleziona Yes, Delete (Sì, elimina).
5. Controlla le notifiche della console IAM per monitorare lo stato dell'eliminazione del ruolo collegato ai servizi. Poiché l'eliminazione del ruolo collegato ai servizi IAM è asincrona, una volta richiesta l'eliminazione del ruolo, il task di eliminazione può essere eseguito correttamente o meno. Se il task non viene eseguito correttamente, puoi scegliere View details (Visualizza dettagli) o View Resources (Visualizza risorse) dalle notifiche per capire perché l'eliminazione non è stata effettuata. Se l'eliminazione non viene eseguita perché vi sono risorse nel servizio che sono usate dal ruolo, il motivo dell'errore include un elenco di risorse.

Eliminazione di un ruolo collegato ai servizi (CLI IAM)

Puoi utilizzare i comandi IAM di AWS Command Line Interface per eliminare un ruolo collegato al servizio. Poiché un ruolo collegato ai servizi non può essere eliminato se è in uso o se a esso sono associate delle risorse, occorre inviare una richiesta di eliminazione. Se queste condizioni non sono soddisfatte, la richiesta può essere rifiutata.

Per eliminare un ruolo collegato ai servizi (CLI)

1. Per controllare lo stato dell'attività di eliminazione, devi acquisire il `deletion-task-id` dalla risposta. Per inviare una richiesta di eliminazione per un ruolo collegato ai servizi, digita il seguente comando:

```
$ aws iam delete-service-linked-role --role-name AWSServiceRoleForEMRWAL
```

2. Digita il seguente comando per verificare lo stato del task di eliminazione:

```
$ aws iam get-service-linked-role-deletion-status --deletion-task-id deletion-task-id
```

Lo stato di un task di eliminazione può essere `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED`o `FAILED`. Se l'eliminazione non viene eseguita correttamente, la chiamata restituisce il motivo dell'errore per consentire all'utente di risolvere il problema.

Eliminazione di un ruolo collegato ai servizi (API IAM)

È possibile utilizzare l'API IAM per eliminare un ruolo collegato ai servizi. Poiché un ruolo collegato ai servizi non può essere eliminato se è in uso o se a esso sono associate delle risorse, occorre inviare una richiesta di eliminazione. Se queste condizioni non sono soddisfatte, la richiesta può essere rifiutata.

Per eliminare un ruolo collegato ai servizi (API)

1. Per inviare una richiesta di eliminazione per un ruolo collegato al servizio, chiama. DeleteServiceLinkedRole Nella richiesta, specificare il nome del ruolo AWSService RoleFor EMRWAL.

   Per controllare lo stato dell'attività di eliminazione, devi acquisire il `DeletionTaskId` dalla risposta.
2. Chiamare GetServiceLinkedRoleDeletionStatus per controllare lo stato dell'eliminazione. Nella richiesta, specificare il `DeletionTaskId`.

   Lo stato di un task di eliminazione può essere `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED`o `FAILED`. Se l'eliminazione non viene eseguita correttamente, la chiamata restituisce il motivo dell'errore per consentire all'utente di risolvere il problema.

Regioni supportate per EMRWAL AWSService RoleFor

Amazon EMR supporta l'utilizzo del ruolo collegato AWSService RoleFor al servizio EMRWAL nelle seguenti regioni.

Nome della Regione	Identità della regione	Supporto in Amazon EMR
US East (N. Virginia)	us-east-1	Sì

Nome della Regione	Identità della regione	Supporto in Amazon EMR
Stati Uniti orientali (Ohio)	us-east-2	Sì
US West (N. California)	us-west-1	Sì
US West (Oregon)	us-west-2	Sì
Asia Pacific (Mumbai)	ap-south-1	Sì
Asia Pacifico (Singapore)	ap-southeast-1	Sì
Asia Pacifico (Sydney)	ap-southeast-2	Sì
Asia Pacifico (Tokyo)	ap-northeast-1	Sì
Europe (Frankfurt)	eu-central-1	Sì
Europa (Irlanda)	eu-west-1	Sì

Personalizza i ruoli IAM con Amazon EMR

È possibile personalizzare il ruolo del servizio IAM e le autorizzazioni per limitare i privilegi in base ai requisiti di sicurezza. Per personalizzare le autorizzazioni, si consiglia di creare nuovi ruoli e policy. Iniziare con le autorizzazioni nelle policy gestite per i ruoli predefiniti (ad esempio, `AmazonElasticMapReduceforEC2Role` e `AmazonElasticMapReduceRole`). Quindi, copiare e incollare il contenuto in nuove istruzione della policy, modificare le autorizzazioni come appropriato e collegare le policy di autorizzazione modificate ai ruoli creati. È necessario disporre delle autorizzazioni IAM appropriate per lavorare con i ruoli e le policy. Per ulteriori informazioni, consulta Consentire a utenti e gruppi di creare e modificare i ruoli.

Se crei un ruolo EMR personalizzato per EC2, segui il flusso di lavoro di base, che crea automaticamente un profilo di istanza con lo stesso nome. Amazon EC2 consente di creare profili e ruoli di istanza con nomi diversi, ma Amazon EMR non supporta questa configurazione e genera un errore di «profilo di istanza non valido» durante la creazione del cluster.

> **Important**
>
> Le policy in linea non vengono aggiornate automaticamente quando i requisiti di servizio cambiano. Se si creano e si collegano policy in linea, occorre tenere presente che potrebbero verificarsi aggiornamenti del servizio responsabili di errori di autorizzazione imprevisti. Per ulteriori informazioni, consulta la sezione relativa a Policy gestite e policy inline nella Guida per l'utente IAM e Specifica dei ruoli IAM personalizzati durante la creazione di un cluster.

Per ulteriori informazioni sull'utilizzo dei ruoli IAM, consulta i seguenti argomenti nella Guida per l'utente IAM:

- Creazione di un ruolo per delegare le autorizzazioni a un servizio AWS
- Modifica di un ruolo
- Eliminazione di un ruolo

Specifica dei ruoli IAM personalizzati durante la creazione di un cluster

Quando crei un cluster, specifichi il ruolo del servizio per Amazon EMR e il ruolo per il profilo dell' EC2 istanza Amazon. L'utente che crea i cluster necessita delle autorizzazioni per recuperare e assegnare ruoli ad Amazon EMR e alle istanze. EC2 In caso contrario, si verifica un errore di chiamata di un account non autorizzato. EC2 Per ulteriori informazioni, consulta Consentire a utenti e gruppi di creare e modificare i ruoli.

Utilizzo della console per specificare i ruoli personalizzati

Quando crei un cluster, puoi specificare un ruolo di servizio personalizzato per Amazon EMR, un ruolo personalizzato per il profilo dell' EC2 istanza e un ruolo Auto Scaling personalizzato utilizzando le opzioni avanzate. Quando utilizzi le opzioni Quick, vengono specificati il ruolo di servizio e il ruolo predefinito per il profilo dell' EC2 istanza. Per ulteriori informazioni, consulta Ruoli di servizio IAM utilizzati da Amazon EMR.

Console

Per specificare ruoli IAM personalizzati con la console

Quando crei un cluster con la console, devi specificare un ruolo di servizio personalizzato per Amazon EMR e un ruolo personalizzato per il profilo dell' EC2 istanza. Per ulteriori informazioni, consulta Ruoli di servizio IAM utilizzati da Amazon EMR.

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Security configuration and permissions (Configurazione e autorizzazioni di sicurezza), cerca i campi IAM role for instance profile (Ruolo IAM per il profilo dell'istanza) e Service role for Amazon EMR (Ruolo di servizio per i campi Amazon EMR). Per ogni tipo di ruolo, selezionare un ruolo dall'elenco. Vengono elencati solo i ruoli all'interno dell'account che dispongono della policy di affidabilità appropriata per quel tipo di ruolo.
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

Usa il AWS CLI per specificare ruoli personalizzati

Puoi specificare un ruolo di servizio per Amazon EMR e un ruolo di servizio per EC2 le istanze di cluster in modo esplicito utilizzando le opzioni con il `create-cluster` comando di. AWS CLI L'opzione `--service-role` consente di specificare il ruolo di servizio. Utilizza l'`InstanceProfile`argomento dell'`--ec2-attributes`opzione per specificare il ruolo per il profilo dell'istanza. EC2

Il ruolo Auto Scaling viene specificato mediante un'opzione distinta, `--auto-scaling-role`. Per ulteriori informazioni, consulta Utilizzo del ridimensionamento automatico con una politica personalizzata, ad esempio gruppi in Amazon EMR.

Per specificare ruoli IAM personalizzati utilizzando il AWS CLI

- Il comando seguente specifica il ruolo di servizio personalizzato e un ruolo personalizzato per il profilo dell' EC2 istanza all'avvio di un cluster. *`MyCustomServiceRoleForEMR`* *`MyCustomServiceRoleForClusterEC2Instances`* Questo esempio utilizza il ruolo Amazon EMR predefinito.

  > **Note**
  >
  > I caratteri di continuazione della riga Linux (\) sono inclusi per la leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

```
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
```

È possibile utilizzare queste opzioni per specificare esplicitamente i ruoli predefiniti anziché utilizzare l'opzione `--use-default-roles`. L'`--use-default-roles`opzione specifica il ruolo di servizio e il ruolo per il profilo di EC2 istanza definito nel `config` file per. AWS CLI

L'esempio seguente mostra il contenuto di un `config` file per AWS CLI i ruoli personalizzati specificati per Amazon EMR. Con questo file di configurazione, quando viene specificata l'opzione `--use-default-roles`, il cluster viene creato utilizzando *`MyCustomServiceRoleForEMR`* e *`MyCustomServiceRoleForClusterEC2Instances`*. Per impostazione predefinita, il file `config` specifica l'impostazione predefinita `service_role` come `AmazonElasticMapReduceRole` e l'impostazione predefinita `instance_profile` come `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
```

Configurazione di ruoli IAM per le richieste EMRFS ad Amazon S3

> **Note**
>
> La funzionalità di mappatura dei ruoli EMRFS descritta in questa pagina è stata migliorata con l'introduzione di Amazon S3 Access Grants in Amazon EMR 6.15.0. Per una soluzione scalabile di controllo degli accessi per i tuoi dati in Amazon S3, ti consigliamo di utilizzare S3 Access Grants con Amazon EMR.

Quando un'applicazione in esecuzione su un cluster fa riferimento ai dati utilizzando il formato `s3://mydata`, Amazon EMR utilizza EMRFS per effettuare la richiesta. Per interagire con Amazon S3, EMRFS presuppone le politiche di autorizzazione allegate al tuo profilo di istanza Amazon. EC2 Lo stesso profilo di EC2 istanza Amazon viene utilizzato indipendentemente dall'utente o dal gruppo che esegue l'applicazione o dalla posizione dei dati in Amazon S3.

Se si dispone di un cluster con più utenti che necessitano di diversi livelli di accesso ai dati in Amazon S3 tramite EMRFS, è possibile impostare una configurazione della sicurezza con i ruoli IAM per EMRFS. EMRFS può assumere un ruolo di servizio diverso per EC2 le istanze del cluster in base all'utente o al gruppo che effettua la richiesta o in base alla posizione dei dati in Amazon S3. Ogni ruolo IAM può avere differenti autorizzazioni per l'accesso ai dati in Amazon S3. Per ulteriori informazioni sul ruolo di servizio per le istanze di cluster EC2 , consulta. Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2

L'utilizzo di ruoli IAM personalizzati per EMRFS è supportato in Amazon EMR versioni 5.10.0 e successive. Se si utilizza una versione precedente o si dispone di requisiti che i ruoli IAM per EMRFS non possono soddisfare, è possibile creare un provider di credenziali personalizzato. Per ulteriori informazioni, consulta Autorizzazione dell'accesso ai dati EMRFS in Amazon S3.

Quando utilizzi una configurazione di sicurezza per specificare ruoli IAM per EMRFS, configuri mappature di ruoli. Ogni mappatura di ruoli specifica un ruolo IAM che corrisponde a identificatori. Questi identificatori determinano la base per l'accesso ad Amazon S3 tramite EMRFS. Gli identificatori possono essere utenti, gruppi o prefissi Amazon S3 che indicano un percorso di dati. Quando EMRFS effettua una richiesta ad Amazon S3, se la richiesta corrisponde alla base per l'accesso, EMRFS fa sì che le istanze del EC2 cluster assumano il ruolo IAM corrispondente per la richiesta. Le autorizzazioni IAM associate a quel ruolo si applicano invece delle autorizzazioni IAM associate al ruolo di servizio per le istanze del cluster. EC2

Gli utenti e i gruppi in una mappatura di ruoli sono utenti e gruppi Hadoop definiti nel cluster. Gli utenti e i gruppi sono passati a EMRFS nel contesto dell'applicazione che lo utilizza (ad esempio, la rappresentazione di utente YARN). Il prefisso Amazon S3 può essere un identificatore bucket di qualsiasi profondità (ad esempio, `s3://amzn-s3-demo-bucket` o `s3://amzn-s3-demo-bucket/myproject/mydata`). Puoi specificare più identificatori in una singola mappatura di ruoli, ma devono essere tutti dello stesso tipo.

> **Important**
>
> I ruoli IAM per EMRFS forniscono l'isolamento a livello di applicazione tra gli utenti dell'applicazione. Non offrono l'isolamento a livello di host tra gli utenti sull'host. Qualsiasi utente con accesso al cluster può ignorare l'isolamento per assumere uno qualsiasi dei ruoli.

Quando un'applicazione cluster invia una richiesta ad Amazon S3 tramite EMRFS, EMRFS valuta le mappature di ruoli nell'ordine decrescente in cui sono visualizzate nella configurazione di sicurezza. Se una richiesta effettuata tramite EMRFS non corrisponde a nessun identificatore, EMRFS torna a utilizzare il ruolo di servizio per le istanze del cluster. EC2 Per questo motivo, è consigliabile che le policy associate a questo ruolo limitino le autorizzazioni per Amazon S3. Per ulteriori informazioni, consulta Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2.

Configurazione dei ruoli

Prima di impostare una configurazione di sicurezza con ruoli IAM per EMRFS, pianifica e crea i ruoli e le policy di autorizzazione da associare ai ruoli. Per ulteriori informazioni, vedi Come funzionano i ruoli per le istanze? EC2 nella Guida per l'utente di IAM. Quando crei policy di autorizzazione, ti consigliamo di iniziare con la policy gestita associata al ruolo Amazon EMR predefinito EC2 per, quindi di modificarla in base alle tue esigenze. Il nome del ruolo predefinito è `EMR_EC2_DefaultRole` e la policy gestita predefinita da modificare è `AmazonElasticMapReduceforEC2Role`. Per ulteriori informazioni, consulta Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2.

Aggiornamento delle policy di affidabilità per le autorizzazioni del ruolo Assume

Ogni ruolo utilizzato da EMRFS deve avere una policy di fiducia che consenta al ruolo Amazon EMR del cluster di assumerlo EC2 . Analogamente, il ruolo Amazon EMR for del cluster EC2 deve avere una policy di fiducia che consenta ai ruoli EMRFS di assumerlo.

L'esempio di policy di trust riportata di seguito è collegata ai ruoli per EMRFS. L'istruzione consente al ruolo predefinito di Amazon EMR di EC2 assumere il ruolo. Se ad esempio si dispone di due ruoli fittizi EMRFS, `EMRFSRole_First` ed `EMRFSRole_Second`, questa istruzione di policy viene aggiunta alla policy di trust per ognuno di essi.

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMR_EC2_DefaultRole",
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

Inoltre, la seguente istruzione per policy di trust di esempio viene aggiunta al ruolo `EMR_EC2_DefaultRole` per consentire ai due ruoli EMRFS fittizi di assumerlo.

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"
    }
  ]
}
```

Aggiornamento della policy di affidabilità di un ruolo IAM

Aprire la console IAM all'indirizzo https://console.aws.amazon.com/iam/.

1. Scegliere Roles (Ruoli), immettere il nome del ruolo in Search (Cerca), quindi selezionare il relativo Role name (Nome ruolo).
2. Scegliere Trust relationships (Relazioni di trust), quindi Edit trust relationship (Modifica relazione di trust).
3. Aggiungi un'istruzione di attendibilità in base a quanto riportato nel Policy Document (Documento di policy) secondo le linee guida in alto, quindi scegli Update Trust Policy (Aggiorna policy di attendibilità).

Indicazione di un ruolo come utente chiave

Se un ruolo autorizza l'accesso a una posizione in Amazon S3 crittografata mediante una AWS KMS key, assicurati che il ruolo sia specificato come utente di chiavi. Ciò concede al ruolo l'autorizzazione a utilizzare la chiave KMS. Per ulteriori informazioni, consulta Policy delle chiavi in AWS KMS nella Guida per gli sviluppatori di AWS Key Management Service .

Impostazione di una configurazione di sicurezza con ruoli IAM per EMRFS

> **Important**
>
> Se non si applica nessuno dei ruoli IAM per EMRFS specificati, EMRFS torna al ruolo di Amazon EMR per. EC2 Valuta la possibilità di personalizzare questo ruolo per limitare le autorizzazioni per Amazon S3 come appropriato per la tua applicazione e di specificare questo ruolo personalizzato anziché `EMR_EC2_DefaultRole` quando crei un cluster. Per ulteriori informazioni, consultare Personalizza i ruoli IAM con Amazon EMR e Specifica dei ruoli IAM personalizzati durante la creazione di un cluster.

Specifica dei ruoli IAM per le richieste EMRFS ad Amazon S3 mediante la console

1. Creare una configurazione di sicurezza che specifichi mappature di ruoli:
   a. Nella console di Amazon EMR, seleziona Security configurations (Configurazioni di sicurezza) e Create (Crea).
   b. Digitare un nome in Name (Nome) per la configurazione di sicurezza. Questo nome è utilizzato per specificare la configurazione di sicurezza al momento della creazione di un cluster.
   c. Scegli Use IAM roles for EMRFS requests to Amazon S3 (Utilizza i ruoli IAM per le richieste EMRFS ad Amazon S3).

d. Seleziona un IAM role (Ruolo IAM) da applicare e in Basis for access (Base per accesso) seleziona un tipo di identificatore (Users [Utenti], Groups [Gruppi) o S3 prefixes [Prefissi S3]) dall'elenco e immetti gli identificatori corrispondenti. Se si utilizzano più identificatori, separarli con una virgola e senza inserire spazi. Per ulteriori informazioni su ogni tipo di identificatore, vedi JSON configuration reference qui sotto.

e. Scegliere Add role (Aggiungi ruolo) per configurare ulteriori mappature di ruoli come descritto nella fase precedente.

f. Configurare altre opzioni per la configurazione di sicurezza come appropriato e scegliere Create (Crea). Per ulteriori informazioni, consulta Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI.

2. Specificare la configurazione di sicurezza creata precedentemente alla creazione di un cluster. Per ulteriori informazioni, consulta Specificare una configurazione di sicurezza per un cluster Amazon EMR.

Per specificare i ruoli IAM per le richieste EMRFS ad Amazon S3 utilizzando AWS CLI

1. Utilizzare il comando `aws emr create-security-configuration`, specificando un nome per la configurazione di sicurezza, e i dettagli relativi a tale configurazione in formato JSON.

   L'esempio di comando riportato di seguito crea una configurazione di sicurezza con il nome `EMRFS_Roles_Security_Configuration`. Questa è basata su una struttura JSON nel file `MyEmrfsSecConfig.json`, che viene salvato nella stessa directory in cui viene eseguito il comando.

   ```
   aws emr create-security-configuration --name EMRFS_Roles_Security_Configuration --
   security-configuration file://MyEmrFsSecConfig.json.
   ```

   Utilizza le linee guida seguenti per la struttura del file `MyEmrFsSecConfig.json`. È possibile specificare questa struttura insieme alle strutture per altre opzioni della configurazione di sicurezza. Per ulteriori informazioni, consulta Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI.

   Di seguito è riportato un esempio di frammento JSON per specificare ruoli IAM personalizzati per EMRFS all'interno di una configurazione di sicurezza. Vengono illustrate le mappature dei ruoli per i tre diversi tipi di identificatori, seguite da un riferimento ai parametri.

   ```
   {
   ```

```
  "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" ]
      }]
    }
  }
}
```

Parametro	Descrizione
`"AuthorizationConfiguration":`	Obbligatorio.
`"EmrFsConfiguration":`	Obbligatorio. Contiene mappature dei ruoli.
`"RoleMappings":`	Obbligatorio. Contiene una o più definizioni di mappatura dei ruoli. Le mappature dei ruoli vengono valutate dall'alto verso il basso nell'ordine in cui vengono visualizzate. Se una mappatura dei ruoli viene valutata come true (vera) per una chiamata EMRFS per i dati in Amazon S3, non vengono valutate altre mappature dei ruoli ed EMRFS utilizza il ruolo IAM specificato per la richiesta. Le mappature dei ruoli sono costituite dai parametri obbligatori seguenti:

Parametro	Descrizione
`"Role":`	Specifica l'identificatore ARN di un ruolo IAM nel formato `arn:aws:iam::` *`account-id`* `:role/`*`role-name`* . Questo è il ruolo IAM che Amazon EMR assume se la richiesta EMRFS ad Amazon S3 corrisponde a uno degli `Identifiers` specificato.
`"IdentifierType":`	Il valore può essere uno dei seguenti: • `"User"` specifica che gli identificatori sono uno o più utenti Hadoop, i quali possono essere utenti di account Linux o entità Kerberos. Quando la richiesta EMRFS viene originata dall'utente o dagli utenti specificati, viene assunto il ruolo IAM. • `"Prefix"` specifica che l'identificatore è un percorso Amazon S3. Il ruolo IAM viene assunto per le chiamate alla posizione o alle posizioni con i prefissi specificati. Ad esempio, il prefisso `s3://amzn-s3-demo-bucket/` corrisponde a `s3://amzn-s3-demo-bucket/mydir` e `s3://amzn-s3-demo-bucket/yetanotherdir` . • `"Group"` specifica che gli identificatori sono uno o più Gruppi Hadoop. Il ruolo IAM viene assunto se la richiesta proviene da un utente nel gruppo o nei gruppi specificati.
`"Identifiers":`	Specifica uno o più identificatori del tipo di identificatore appropriato. Separa più identificatori con virgole senza spazi.

2. Utilizzare il comando `aws emr create-cluster` per creare un cluster e specificare la configurazione di sicurezza creata nella fase precedente.

   L'esempio seguente crea un cluster con le principali applicazioni Hadoop di default installate. Il cluster utilizza la configurazione di sicurezza creata sopra `EMRFS_Roles_Security_Configuration` e utilizza anche un ruolo Amazon EMR personalizzato per EC2`EC2_Role_EMR_Restrict_S3`, che viene specificato utilizzando l'`InstanceProfile`argomento del `--ec2-attributes` parametro.

   > **Note**
   >
   > I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

```
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
```

Utilizzo di policy basate su risorse per l'accesso Amazon EMR ad AWS Glue Data Catalog

Se utilizzi AWS Glue insieme a Hive, Spark o Presto in Amazon EMR, AWS Glue supporta politiche basate sulle risorse per controllare l'accesso alle risorse del Data Catalog. Queste risorse includono tabelle, database, connessioni e funzioni definite dall'utente. Per ulteriori informazioni, consulta Policy sulle risorse AWS Glue nella Guida per gli sviluppatori di AWS Glue.

Quando si utilizzano politiche basate sulle risorse per limitare l'accesso a AWS Glue dall'interno di Amazon EMR, il principio specificato nella politica delle autorizzazioni deve essere il ruolo ARN associato al profilo dell' EC2 istanza specificato al momento della creazione di un cluster. Ad esempio, per una policy basata sulle risorse allegata a un catalogo, è possibile specificare il ruolo ARN per il ruolo di servizio predefinito per le EC2 istanze del cluster, *EMR_EC2_DefaultRole* come`Principal`, utilizzando il formato illustrato nell'esempio seguente:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*Può essere diverso dall'ID dell'account AWS Glue. Ciò consente l'accesso da cluster EMR in account diversi. Puoi specificare più entità principali, ognuna da un account diverso.

I ruoli IAM possono essere utilizzati con le applicazioni che richiamano direttamente i servizi AWS

Le applicazioni in esecuzione sulle EC2 istanze di un cluster possono utilizzare il profilo dell' EC2 istanza per ottenere credenziali di sicurezza temporanee quando AWS chiamano i servizi.

Le versioni di Hadoop disponibili con Amazon EMR versione 2.3.0 e successive sono già state aggiornate per l'utilizzo di ruoli IAM. Se l'applicazione viene eseguita esclusivamente sull'architettura Hadoop e non richiama direttamente alcun servizio AWS, dovrebbe funzionare con i ruoli IAM senza alcuna modifica.

Se l'applicazione richiama AWS direttamente i servizi, è necessario aggiornarla per sfruttare i ruoli IAM. Ciò significa che invece di ottenere le credenziali dell'account dalle `/etc/hadoop/conf/core-site.xml` EC2 istanze del cluster, l'applicazione utilizza un SDK per accedere alle risorse utilizzando i ruoli IAM o chiama i metadati dell' EC2 istanza per ottenere le credenziali temporanee.

Per accedere alle AWS risorse con ruoli IAM utilizzando un SDK

- I seguenti argomenti mostrano come utilizzare diversi di essi per accedere AWS SDKs a credenziali temporanee utilizzando i ruoli IAM. Ogni argomento inizia con una versione di un'applicazione che non utilizza ruoli IAM e poi guida attraverso il processo di conversione di quell'applicazione per utilizzare ruoli IAM.
  - Utilizzo dei ruoli IAM per EC2 le istanze Amazon con l'SDK for Java nella AWS SDK per Java Developer Guide
  - Utilizzo dei ruoli IAM per EC2 le istanze Amazon con l'SDK for .NET nella AWS SDK per .NET Developer Guide
  - Utilizzo dei ruoli IAM per EC2 le istanze Amazon con l'SDK for PHP nella Developer Guide AWS SDK per PHP
  - Utilizzo dei ruoli IAM per EC2 le istanze Amazon con l'SDK for Ruby nella Developer Guide AWS SDK per Ruby

Per ottenere credenziali temporanee dai metadati dell'istanza EC2

- Chiama il seguente URL da un' EC2 istanza in esecuzione con il ruolo IAM specificato, che restituisce le credenziali di sicurezza temporanee associate (AccessKeyId, SecretAccessKey SessionToken, e scadenza). L'esempio seguente utilizza il profilo istanza predefinito per Amazon EMR, `EMR_EC2_DefaultRole`.

```
GET http://169.254.169.254/latest/meta-data/iam/security-
credentials/EMR_EC2_DefaultRole
```

Per ulteriori informazioni sulla scrittura di applicazioni che utilizzano ruoli IAM, consulta Concedere alle applicazioni eseguite su EC2 istanze Amazon l'accesso alle AWS risorse.

Per ulteriori informazioni sulle credenziali di sicurezza temporanee, consulta Utilizzo delle credenziali di sicurezza temporanee nella Guida all'utilizzo di credenziali di sicurezza temporanee.

Consentire a utenti e gruppi di creare e modificare i ruoli

Le entità principali (utenti e gruppi) IAM che creano, modificano e specificano i ruoli per un cluster, inclusi i ruoli predefiniti, devono essere autorizzate a eseguire le seguenti operazioni. Per ulteriori informazioni su ciascuna operazione, consulta Actions (Operazioni) nella Guida di riferimento alle API di IAM.

- `iam:CreateRole`
- `iam:PutRolePolicy`
- `iam:CreateInstanceProfile`
- `iam:AddRoleToInstanceProfile`
- `iam:ListRoles`
- `iam:GetPolicy`
- `iam:GetInstanceProfile`
- `iam:GetPolicyVersion`
- `iam:AttachRolePolicy`
- `iam:PassRole`

L'autorizzazione `iam:PassRole` consente la creazione del cluster. Le restanti autorizzazioni consentono la creazione di ruoli predefiniti.

Per ulteriori informazioni sull'assegnazione di autorizzazioni a un utente IAM, consulta Modifica delle autorizzazioni per un utente nella Guida per l'utente IAM.

Esempi di policy basate su identità per Amazon EMR

Per impostazione predefinita, gli utenti e i ruoli non dispongono dell'autorizzazione per creare o modificare risorse Amazon EMR. Inoltre, non possono eseguire attività utilizzando l'API AWS Management Console, AWS CLI, o AWS . Un amministratore IAM deve creare policy IAM che concedono a utenti e ruoli l'autorizzazione per eseguire operazioni API specifiche sulle risorse specificate di cui hanno bisogno. L'amministratore deve quindi collegare queste policy a utenti o gruppi che richiedono tali autorizzazioni.

Per informazioni su come creare una policy basata su identità IAM utilizzando questi documenti di policy JSON di esempio, consultare Creazione di policy nella scheda JSON nella Guida per l'utente IAM.

Argomenti

- Best practice delle policy per Amazon EMR
- Consentire agli utenti di visualizzare le loro autorizzazioni
- Policy gestite da Amazon EMR
- Policy IAM per l'accesso basato su tag ai cluster e a EMR Notebooks
- Negare l' ModifyInstanceGroup azione in Amazon EMR
- Risoluzione dei problemi relativi all'identità e all'accesso di Amazon EMR

Best practice delle policy per Amazon EMR

Le policy basate su identità sono molto efficaci. Determinano se qualcuno può creare, accedere o eliminare risorse Amazon EMR nell'account. Queste azioni possono comportare costi per il tuo AWS account. Quando crei o modifichi policy basate su identità, segui queste linee guida e raccomandazioni:

- Inizia a usare le policy AWS gestite: per iniziare a usare Amazon EMR rapidamente, utilizza policy AWS gestite per concedere ai dipendenti le autorizzazioni di cui hanno bisogno. Queste policy sono già disponibili nell'account e sono gestite e aggiornate da AWS. Per ulteriori informazioni,

consulta Introduzione all'utilizzo delle autorizzazioni con policy AWS gestite nella Guida per l'utente IAM e. Policy gestite da Amazon EMR

- Assegnare il privilegio minimo: quando si creano policy personalizzate, concedere solo le autorizzazioni richieste per eseguire un'attività. Inizia con un set di autorizzazioni minimo e concedi autorizzazioni aggiuntive quando necessario. Questo è più sicuro che iniziare con autorizzazioni che siano troppo permissive e cercare di limitarle in un secondo momento. Per ulteriori informazioni, consulta Assegnare il privilegio minimo nella Guida per l'utente IAM.
- Abilitare MFA per operazioni sensibili: per una maggiore sicurezza, richiedere agli utenti di utilizzare l'autenticazione a più fattori (MFA) per accedere a risorse sensibili o operazioni API. Per ulteriori informazioni, consulta Utilizzo dell'autenticazione a più fattori (MFA) in AWS nella Guida per l'utente IAM.
- Utilizzare le condizioni della policy per ulteriore sicurezza – Per quanto possibile, definire le condizioni in cui le policy basate su identità consentono l'accesso a una risorsa. Ad esempio, è possibile scrivere condizioni per specificare un intervallo di indirizzi IP consentiti dai quali deve provenire una richiesta. È anche possibile scrivere condizioni per consentire solo le richieste all'interno di un intervallo di date o ore specificato oppure per richiedere l'utilizzo di SSL o MFA. Per ulteriori informazioni, consulta Elementi delle policy JSON di IAM: Condizioni nella Guida per l'utente IAM.

Consentire agli utenti di visualizzare le loro autorizzazioni

Questo esempio mostra in che modo è possibile creare una policy che consente agli utenti di visualizzare le policy inline e gestite che sono collegate alla relativa identità utente. Questa policy include le autorizzazioni per completare questa azione sulla console o utilizzando l'API or a livello di codice. AWS CLI 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": [
        "*"
      ]
    }
  ]
}
```

Policy gestite da Amazon EMR

Il modo più semplice di concedere accesso completo o in sola lettura alle operazioni di Amazon EMR necessarie è utilizzare le policy gestite da IAM per Amazon EMR. Le policy gestite offrono il vantaggio di essere aggiornate automaticamente nel momento in cui i requisiti di autorizzazione cambiano. Se utilizzi policy inline, è possibile che si verifichino degli errori di autorizzazione in caso di modifiche al servizio.

Amazon EMR renderà obsolete le policy gestite esistenti (policy v1), a favore di nuove policy gestite (policy v2). Le nuove politiche gestite sono state ridotte in modo da allinearle alle migliori pratiche. AWS Dopo che le policy gestite v1 renderà obsolete, non sarà possibile allegare queste policy a nuovi ruoli o utenti IAM. I ruoli e gli utenti esistenti che utilizzano policy obsolete possono continuare

a utilizzarli. Le policy gestite v2 limitano l'accesso utilizzando i tag. Consentono solo operazioni Amazon EMR specifiche e richiedono risorse cluster contrassegnate con una chiave specifica per EMR. Si consiglia di esaminare attentamente la documentazione prima di utilizzare le nuove policy v2.

Le policy v1 verranno contrassegnate come obsolete con un'icona di avviso accanto a esse nell'elenco Policies (Policy) della console IAM. Le policy obsolete avranno le seguenti caratteristiche:

- Continuano a funzionare per tutti gli utenti, gruppi e ruoli attualmente collegati. Nessuna interruzione.
- Non possono essere collegate a nuovi utenti, gruppi o ruoli. Se una delle policy viene scollegata da un'entità corrente, non è possibile collegarla nuovamente.
- Dopo aver scollegato una policy v1 da tutte le entità correnti, la policy non sarà più visibile e non potrà essere utilizzata.

Nella tabella seguente sono riepilogate le modifiche tra le policy correnti (v1) e le policy v2.

Modifiche alle policy gestite da Amazon EMR

Tipo di policy	Nomi delle policy	Scopo della policy	Modifiche alla policy v2
Ruolo di servizio EMR predefinito e policy gestita collegata	Nome del ruolo: EMR_ DefaultRole Criterio V1 (da rendere obsoleto) : (ruolo del servizio AmazonElasticMapReduceRoleEMR) Nome della policy v2 (con ambito): **AmazonEMRServicePolicy_v2**	Consente ad Amazon EMR di chiamare altri AWS servizi per tuo conto durante il provisioning di risorse e l'esecuzione di azioni a livello di servizio. Questo ruolo è obbligatorio per tutti i cluster.	La policy aggiunge la nuova autorizzazione. `"ec2:DescribeInstanceTypeOfferings"` Questa operazione API restituisce un elenco di tipi di istanze supportati da un elenco di determinate zone di disponibilità.
Policy gestita da IAM per l'accesso	Nome della policy V2 (con ambito):	Consente agli utenti autorizza	La policy aggiunge il prerequisito secondo

Tipo di policy	Nomi delle policy	Scopo della policy	Modifiche alla policy v2
completo ad Amazon EMR per utente, ruolo o gruppo collegato	**AmazonEMRServicePolicy_v2**	zioni complete per le operazioni EMR. Include iam: PassRole autorizzazioni per le risorse.	cui gli utenti devono aggiungere tag utente alle risorse prima di poter utilizzare questa policy. Consultare Assegnazione di tag alle risorse per l'utilizzo delle policy gestite. iam: PassRole l'azione richiede iam: PassedToService condizione impostata sul servizio specificato. L'accesso ad Amazon EC2, Amazon S3 e ad altri servizi non è consentito per impostazione predefinita. Consulta Policy IAM gestita per l'accesso completo (policy predefinita gestita v2).

Tipo di policy	Nomi delle policy	Scopo della policy	Modifiche alla policy v2
Policy IAM gestita per un accesso di sola lettura parte di utente, ruolo o gruppo collegato	Policy V1 (da essere resa obsoleta) : **AmazonElasticMapReduceReadOnlyAccess** Nome della policy V2 (con ambito): **AmazonEMRReadOnlyAccessPolicy_v2**	Consente agli utenti autorizzazioni di sola lettura per le operazioni di Amazon EMR.	Le autorizzazioni si riferiscono esclusivamente alle operazioni elasticmapreduce di sola lettura specificate. Per impostazione predefinita, l'accesso ad Amazon S3 non è consentito. Consulta Policy IAM gestita per l'accesso di sola lettura (policy predefinita gestita v2).
Ruolo di servizio EMR predefinito e policy gestita collegata	Nome del ruolo: EMR_ DefaultRole Criterio V1 (da rendere obsoleto) : (ruolo del servizio AmazonElasticMapReduceRoleEMR) Nome della policy v2 (con ambito): **AmazonEMRServicePolicy_v2**	Consente ad Amazon EMR di chiamare altri AWS servizi per tuo conto durante il provisioning di risorse e l'esecuzione di azioni a livello di servizio. Questo ruolo è obbligatorio per tutti i cluster.	Il ruolo di servizio v2 e la policy predefinita v2 sostituiscono il ruolo e la policy obsoleti. La policy aggiunge il prerequisito secondo cui gli utenti devono aggiungere tag utente alle risorse prima di poter utilizzare questa policy. Consultare Assegnazione di tag alle risorse per l'utilizzo delle policy gestite. Consultare Ruolo di servizio per Amazon EMR (ruolo EMR).

Tipo di policy	Nomi delle policy	Scopo della policy	Modifiche alla policy v2
Ruolo di servizio per le istanze del cluster (profilo dell' EC2 istanza) EC2	Nome del ruolo: EC2EMR_ _ DefaultRole Nome della politica obsoleto: Role AmazonElasticMapReducefor EC2	Consente alle applicazioni in esecuzione su un cluster EMR di accedere ad altre risorse AWS , ad esempio Amazon S3. Ad esempio, se esegui i processi Apache Spark che elaborano i dati da Amazon S3, la policy deve consentire l'accesso a tali risorse.	Sia il ruolo predefinito che la policy predefinita diventeranno presto obsoleti. Non esiste un ruolo o una politica gestiti AWS predefiniti sostitutivi. È necessario fornire una policy basata su risorse o su identità. Ciò significa che, per impostazione predefinita, le applicazioni in esecuzione su un cluster EMR non hanno accesso ad Amazon S3 o ad altre risorse a meno che non vengano aggiunte manualmente alla policy. Consultare Ruolo predefinito e policy gestita.

Tipo di policy	Nomi delle policy	Scopo della policy	Modifiche alla policy v2
Altre politiche relative EC2 ai ruoli di servizio	Nomi delle politiche attuali: AmazonElasticMapReduceforAutoScalingRole, AmazonElasticMapReduceEditorsRole, Amazon EMRCleanupPolicy	Fornisce le autorizzazioni necessarie ad Amazon EMR per accedere ad AWS altre risorse ed eseguire azioni se si utilizza la scalabilità automatica, i notebook o per ripulire le risorse. EC2	Nessuna modifica per v2.

Proteggere l'obiettivo: PassRole

Le policy gestite predefinite con autorizzazioni complete di Amazon EMR incorporano configurazioni di sicurezza `iam:PassRole`, tra cui:

- Autorizzazioni `iam:PassRole` solo per specifici ruoli Amazon EMR predefiniti.
- `iam:PassedToService`condizioni che consentono di utilizzare la politica solo con AWS servizi specifici, come `elasticmapreduce.amazonaws.com` `eec2.amazonaws.com`.

Puoi visualizzare la versione JSON delle policy Amazon _v2 EMRFullAccessPolicye Amazon EMRService Policy_v2 nella console IAM. Ti consigliamo di creare i nuovi cluster con le policy gestite v2.

Per creare policy personalizzate, ti consigliamo di iniziare con le policy gestite e di modificarle in base alle tue esigenze.

Per informazioni su come collegare policy a utenti (entità principali), consulta Utilizzo di policy gestite mediante la AWS Management Console nella Guida per l'utente IAM.

Assegnazione di tag alle risorse per l'utilizzo delle policy gestite

Amazon EMRService Policy_v2 e EMRFullAccessPolicyAmazon _v2 dipendono dall'accesso limitato alle risorse fornite o utilizzate da Amazon EMR. L'ambito inferiore si ottiene limitando l'accesso solo a quelle risorse a cui è associato un tag utente predefinito. Quando si utilizza una di queste due policy,

è necessario passare il tag utente predefinito `for-use-with-amazon-emr-managed-policies = true` durante il provisioning del cluster. Amazon EMR propaga automaticamente tale tag. Inoltre, è necessario aggiungere un tag utente alle risorse elencate nella sezione seguente. Se utilizzi la console Amazon EMR per avviare il cluster, consulta la sezione Considerazioni sull'utilizzo della console Amazon EMR per avviare cluster con policy gestite v2.

Per utilizzare le policy gestite, passa il tag utente `for-use-with-amazon-emr-managed-policies = true` durante il provisioning di un cluster con la CLI, l'SDK o un altro metodo.

Quando passi il tag, Amazon EMR lo propaga alla sottorete privata ENI, all' EC2istanza e ai volumi EBS che crea. Amazon EMR assegna automaticamente tag anche ai gruppi di sicurezza creati. Tuttavia, se desideri che Amazon EMR venga avviato con un determinato gruppo di sicurezza, devi taggarlo. Per le risorse che non sono state create da Amazon EMR, è necessario aggiungere i tag a tali risorse. Ad esempio, devi etichettare le EC2 sottoreti Amazon, i gruppi EC2 di sicurezza (se non creati da Amazon EMR) e VPCs (se desideri che Amazon EMR crei gruppi di sicurezza). Per avviare cluster con policy gestite v2 VPCs, devi etichettarli VPCs con il tag utente predefinito. Per informazioni, consultare Considerazioni sull'utilizzo della console Amazon EMR per avviare cluster con policy gestite v2.

Assegnazione di tag propagata specificata dall'utente

Amazon EMR contrassegna le risorse create utilizzando i tag Amazon EMR specificati durante la creazione di un cluster. Amazon EMR applica tag alle risorse create durante il ciclo di vita del cluster.

Amazon EMR propaga i tag utente per le seguenti risorse:

- ENI della sottorete privata (interfacce di rete elastiche di accesso ai servizi)
- EC2 Istanze
- Volumi EBS
- EC2 Modello di avvio

Gruppi di sicurezza con tag automatici

Amazon EMR contrassegna i gruppi di EC2 sicurezza che crea con il tag richiesto per le politiche gestite v2 per Amazon EMR`for-use-with-amazon-emr-managed-policies`, indipendentemente dai tag specificati nel comando create cluster. Per un gruppo di sicurezza creato prima dell'introduzione delle policy gestite v2, Amazon EMR non tagga automaticamente il gruppo di sicurezza. Se si desidera utilizzare policy gestite v2 con i gruppi di sicurezza predefiniti già esistenti

nell'account, è necessario taggare manualmente i gruppi di sicurezza con `for-use-with-amazon-emr-managed-policies = true`.

Risorse cluster con tag manuali

Occorre taggare manualmente alcune risorse del cluster in modo che possano essere accessibili dai ruoli predefiniti di Amazon EMR.

- È necessario etichettare manualmente i gruppi EC2 di sicurezza e le EC2 sottoreti con il tag di policy gestita di Amazon EMR. `for-use-with-amazon-emr-managed-policies`
- Occorre taggare manualmente un VPC se vuoi che Amazon EMR crei gruppi di sicurezza predefiniti. EMR tenterà di creare un gruppo di sicurezza con il tag specifico se il gruppo di sicurezza predefinito non esiste già.

Amazon EMR tagga automaticamente le seguenti risorse:

- Gruppi di sicurezza creati da EMR EC2

È necessario aggiungere i tag manualmente alle risorse seguenti:

- EC2 Sottorete
- EC2 Gruppi di sicurezza

È necessario taggare manualmente le risorse seguenti:

- VPC: solo quando si desidera che Amazon EMR crei gruppi di sicurezza

Considerazioni sull'utilizzo della console Amazon EMR per avviare cluster con policy gestite v2

Puoi eseguire il provisioning di cluster con policy gestite v2 utilizzando la console Amazon EMR. Di seguito sono riportate alcune considerazioni quando utilizzi la console per avviare cluster Amazon EMR.

- Non è necessario passare il tag predefinito. Amazon EMR aggiunge automaticamente il tag e lo propaga ai componenti appropriati.
- Per i componenti che devono essere taggati manualmente, la vecchia console Amazon EMR tenta di applicare automaticamente i tag se disponi delle autorizzazioni necessarie per taggare le risorse.

Se non disponi delle autorizzazioni per etichettare le risorse o se desideri utilizzare la console, chiedi all'amministratore di taggare tali risorse.

- Non è possibile avviare cluster con policy gestite v2 a meno che non siano soddisfatti tutti i prerequisiti.
- La vecchia console Amazon EMR mostra quali risorse (VPC/sottorete) devono essere taggate.

Policy gestita da IAM per l'accesso completo (policy predefinita gestita v2) per Amazon EMR

Le policy predefinite gestite da EMR con ambito v2 concedono privilegi di accesso specifici agli utenti. Richiedono un tag di risorsa Amazon EMR predefinito e chiavi di condizione `iam:PassRole` per le risorse utilizzate da Amazon EMR, quali `Subnet` e `SecurityGroup` utilizzate per avviare il cluster.

Per concedere le operazioni necessarie per Amazon EMR, collega la policy `AmazonEMRFullAccessPolicy_v2` gestita. Questa policy gestita predefinita aggiornata sostituisce la policy gestita `AmazonElasticMapReduceFullAccess`.

`AmazonEMRFullAccessPolicy_v2` dipende dall'accesso ridotto alle risorse che Amazon EMR fornisce o utilizza. Quando si utilizza questa policy, è necessario passare il tag utente `for-use-with-amazon-emr-managed-policies = true` durante il provisioning del cluster. Amazon EMR propaga automaticamente tale tag. Inoltre, potrebbe essere necessario aggiungere manualmente un tag utente a tipi specifici di risorse, ad esempio gruppi di EC2 sicurezza che non sono stati creati da Amazon EMR. Per ulteriori informazioni, consulta Assegnazione di tag alle risorse per l'utilizzo delle policy gestite.

La policy `AmazonEMRFullAccessPolicy_v2` protegge le risorse eseguendo le operazioni elencate di seguito:

- Richiede il tag delle risorse con il tag `for-use-with-amazon-emr-managed-policies` predefinito delle policy gestite da Amazon EMR per la creazione del cluster e l'accesso ad Amazon EMR.
- Limita l'operazione `iam:PassRole` a ruoli predefiniti specifici e l'accesso `iam:PassedToService` a servizi specifici.
- Per impostazione predefinita, non fornisce più l'accesso ad Amazon EC2, Amazon S3 e altri servizi.

Il contenuto di questa policy è mostrato di seguito.

> **Note**
>
> Puoi anche utilizzare il link alla console AmazonEMRFullAccessPolicy_v2 per visualizzare la policy.

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": [
        "*"
```

```
      ]
    }
  ]
}
```

Policy IAM gestite per l'accesso completo (presto obsolete)

Le policy gestite `AmazonElasticMapReduceFullAccess` e `AmazonEMRFullAccessPolicy_v2` AWS Identity and Access Management (IAM) garantiscono tutte le azioni richieste per Amazon EMR e altri servizi.

> **Important**
>
> La policy gestita `AmazonElasticMapReduceFullAccess` diventerà presto obsoleta, pertanto ti consigliamo di non utilizzarla con Amazon EMR. Utilizza invece `AmazonEMRFullAccessPolicy_v2`. Quando il servizio IAM imposterà come obsoleta la policy v1, non sarà possibile collegarla a un ruolo. Tuttavia, puoi collegare un ruolo esistente a un cluster anche se tale ruolo utilizza la policy obsoleta.

Le policy gestite predefinite con autorizzazioni complete di Amazon EMR incorporano configurazioni di sicurezza `iam:PassRole`, tra cui:

- Autorizzazioni `iam:PassRole` solo per specifici ruoli Amazon EMR predefiniti.
- `iam:PassedToService`condizioni che consentono di utilizzare la politica solo con AWS servizi specifici, come `elasticmapreduce.amazonaws.com` e`ec2.amazonaws.com`.

Puoi visualizzare la versione JSON delle policy Amazon _v2 EMRFullAccessPolicye Amazon EMRService Policy_v2 nella console IAM. Ti consigliamo di creare i nuovi cluster con le policy gestite v2.

Puoi visualizzare il contenuto della policy v1 obsoleta all'indirizzo. AWS Management Console `AmazonElasticMapReduceFullAccess` L'`ec2:TerminateInstances`azione contenuta nella policy concede l'autorizzazione all'utente o al ruolo di chiudere qualsiasi EC2 istanza Amazon associata all'account IAM. Sono incluse le istanze che non fanno parte di un cluster EMR.

Policy gestita da IAM per l'accesso in sola lettura (policy predefinita gestita v2) per Amazon EMR

Per concedere privilegi di sola lettura ad Amazon EMR, collega la policy gestita Amazon _v2. EMRRead OnlyAccessPolicy Questa policy gestita predefinita sostituisce la policy gestita `AmazonElasticMapReduceReadOnlyAccess`. Il contenuto di questa dichiarazione di policy è riportato nel frammento seguente. Rispetto alla policy `AmazonElasticMapReduceReadOnlyAccess`, la policy `AmazonEMRReadOnlyAccessPolicy_v2` non utilizza caratteri jolly per l'elemento `elasticmapreduce`. Al contrario, gli ambiti delle policy v2 predefiniti specificano le operazioni `elasticmapreduce` consentite.

> **Note**
>
> Puoi anche utilizzare il link per visualizzare la policy. AWS Management Console `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": [
        "*"
      ]
    }
  ]
}
```

Policy IAM gestite per l'accesso di sola lettura (presto obsolete)

La policy gestita `AmazonElasticMapReduceReadOnlyAccess` diventerà presto obsoleta. Non è possibile allegare questa policy durante l'avvio di nuovi cluster. `AmazonElasticMapReduceReadOnlyAccess` è stato sostituito con `AmazonEMRReadOnlyAccessPolicy_v2` come policy gestita predefinita di Amazon EMR. Il contenuto di questa dichiarazione di policy è riportato nel frammento seguente. I caratteri jolly per l'elemento `elasticmapreduce` specificano che solo le operazioni che iniziano con le stringhe specificate sono consentite. Tieni presente che poiché questa policy non nega esplicitamente le operazioni, un'altra dichiarazione di policy può essere utilizzata per concedere l'accesso alle operazioni specificate.

> **Note**
>
> Puoi anche utilizzare il AWS Management Console per visualizzare la politica.

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 politica gestita: EMRDescribe ClusterPolicyFor EMRWAL

Non è possibile collegare `EMRDescribeClusterPolicyForEMRWAL`alle entità IAM. Questa policy è associata a un ruolo collegato al servizio che consente ad Amazon EMR di eseguire azioni per tuo conto. Per ulteriori informazioni su questo ruolo collegato al servizio, consulta. Utilizzo di ruoli collegati ai servizi con Amazon EMR per la registrazione write-ahead

Questa politica concede autorizzazioni di sola lettura che consentono al servizio WAL per Amazon EMR di trovare e restituire lo stato di un cluster. Per ulteriori informazioni su Amazon EMR WAL, consulta Write-ahead logs (WAL) per Amazon EMR.

Dettagli dell'autorizzazione

Questa policy include le seguenti autorizzazioni:

- `emr`— Consente ai responsabili di descrivere lo stato del cluster da Amazon EMR. Ciò è necessario affinché Amazon EMR possa confermare la chiusura di un cluster e quindi, dopo trenta giorni, ripulire tutti i log WAL lasciati dal cluster.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:DescribeCluster"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEDescribecluster"
    }
  ]
}
```

AWS politiche gestite per Amazon EMR

Una politica AWS gestita è una politica autonoma creata e amministrata da. AWS AWS le politiche gestite sono progettate per fornire autorizzazioni per molti casi d'uso comuni, in modo da poter iniziare ad assegnare autorizzazioni a utenti, gruppi e ruoli.

Tieni presente che le policy AWS gestite potrebbero non concedere le autorizzazioni con il privilegio minimo per i tuoi casi d'uso specifici, poiché sono disponibili per tutti i clienti. AWS Ti consigliamo pertanto di ridurre ulteriormente le autorizzazioni definendo policy gestite dal cliente specifiche per i tuoi casi d'uso.

Non è possibile modificare le autorizzazioni definite nelle politiche gestite. AWS Se AWS aggiorna le autorizzazioni definite in una politica AWS gestita, l'aggiornamento ha effetto su tutte le identità principali (utenti, gruppi e ruoli) a cui è associata la politica. AWS è più probabile che aggiorni una policy AWS gestita quando ne Servizio AWS viene lanciata una nuova o quando diventano disponibili nuove operazioni API per i servizi esistenti.

Per ulteriori informazioni, consultare Policy gestite da AWSnella Guida per l'utente di IAM.

Aggiornamenti di Amazon EMR alle AWS politiche gestite

Visualizza i dettagli sugli aggiornamenti delle politiche AWS gestite per Amazon EMR da quando questo servizio ha iniziato a tracciare queste modifiche.

Modifica	Descrizione	Data
AmazonEMRServicePolicy_v2 : aggiornamento a una policy esistente	Aggiunto `ec2:CreateVpcEndpoint` e `ec2:CreateTags` richiesto per un'esperienza ottimale, a partire dalla release 7.5.0 di Amazon EMR. `ec2:ModifyVpcEndpoint`	4 marzo 2025
AmazonEMRServicePolicy_v2 : aggiornamento a una policy esistente	Aggiunto `elasticmapreduce:CreatePersistentAppUI` `eelasticmapreduce:DescribePersistentAppUI` . `elasticmapreduce:GetPersistentAppUIPresignedURL`	28 febbraio 2025
EMRDescribeClusterPolicyForEMRWAL : nuova policy	È stata aggiunta una nuova policy in modo che Amazon EMR possa determinare lo stato del cluster per la pulizia WAL trenta giorni dopo la chiusura del cluster.	10 agosto 2023
AmazonEMRFullAccessPolicy_v2 e **AmazonEMRReadOnlyAccessPolicy_v2** :	Aggiunto `elasticmapreduce:DescribeReleaseLabel` e `elasticmapreduce:G`	21 aprile 2022

Modifica	Descrizione	Data
aggiornamento a una policy esistente	`etAutoTerminationPolicy` .	
`AmazonEMRFullAccessPolicy_v2` : aggiornamento a una policy esistente	Aggiunta di `ec2:DescribeImages` per Utilizzo di un'AMI personalizzata per fornire maggiore flessibilità per la configurazione del cluster Amazon EMR.	15 febbraio 2022
Policy gestite da Amazon EMR	Aggiornato per chiarire l'uso di tag utente predefiniti. È stata aggiunta una sezione sull'utilizzo della AWS console per avviare cluster con politiche gestite v2.	29 settembre 2021
`AmazonEMRFullAccessPolicy_v2` : aggiornamento a una policy esistente	Modifica delle operazioni `PassRoleForAutoScaling` e `PassRoleForEC2` per utilizzare l'operatore di condizione `StringLike` in modo che corrispondano a `"iam:PassedToService":"application-autoscaling.amazonaws.com*"` e `"iam:PassedToService":"ec2.amazonaws.com*"` rispettivamente.	20 maggio 2021

Modifica	Descrizione	Data
AmazonEMRFullAccessPolicy_v2: aggiornamento a una policy esistente	Rimozione dell'operazione `s3:ListBuckets` non valida e sostituzione con l'operazione `s3:ListAllMyBuckets`. Aggiornamento della creazione del ruolo collegato al servizio (SLR) in modo esplicito fino all'unico SLR di Amazon EMR con principi di servizio espliciti. I file SLRs che possono essere creati sono esattamente gli stessi di prima di questa modifica.	23 marzo 2021
AmazonEMRFullAccessPolicy_v2: nuova policy	Amazon EMR ha aggiunto nuove autorizzazioni per l'ambito dell'accesso alle risorse e per aggiungere il prerequisito secondo cui gli utenti devono aggiungere un tag utente predefinito alle risorse prima di poter utilizzare le policy gestite da Amazon EMR. L'operazione `iam:PassRole` richiede che la condizione `iam:PassedToService` sia impostata sul servizio specificato. L'accesso ad Amazon EC2, Amazon S3 e ad altri servizi non è consentito per impostazione predefinita.	11 marzo 2021

Modifica	Descrizione	Data
AmazonEMRServicePolicy_v2 : nuova policy	Aggiunge il prerequisito secondo cui gli utenti devono aggiungere tag utente alle risorse prima di poter utilizzare questa policy.	11 marzo 2021
AmazonEMRReadOnlyAccessPolicy_v2 : nuova policy	Le autorizzazioni si riferiscono no esclusivamente alle operazioni elasticmapreduce di sola lettura specificate. Per impostazione predefinita, l'accesso ad Amazon S3 non è consentito.	11 marzo 2021
Amazon EMR ha iniziato a monitorare le modifiche	Amazon EMR ha iniziato a tracciare le modifiche per le sue politiche AWS gestite.	11 marzo 2021

Policy IAM per l'accesso basato su tag ai cluster e a EMR Notebooks

Puoi utilizzare le condizioni nella policy basata su identità per controllare l'accesso ai cluster e ai EMR Notebooks basati su tag.

Per ulteriori informazioni sull'aggiunta di tag ai cluster, consulta Tagging dei cluster EMR.

I seguenti esempi illustrano differenti scenari e modi per utilizzare gli operatori di condizione con chiavi di condizione Amazon EMR. Queste dichiarazioni di policy IAM sono concepite unicamente per scopi dimostrativi e non devono essere utilizzate negli ambienti di produzione. Esistono vari modi di combinare dichiarazioni di policy per concedere o negare autorizzazioni in base alle tue esigenze. Per ulteriori informazioni sulla pianificazione e sul test di policy IAM, consulta la Guida per l'utente IAM.

> **Important**
>
> Negare esplicitamente l'autorizzazione per operazioni di assegnazione di tag è una possibilità da tenere in debita considerazione. Ciò impedisce agli utenti di concedersi personalmente

autorizzazioni tramite tag di una risorsa che non avevi intenzione di accordare. Se non neghi le operazioni di assegnazione di tag per una risorsa, un utente può modificare i tag e aggirare l'intenzione delle policy basate su tag.

Dichiarazioni di policy basate su identità di esempio per cluster

Gli esempi di seguito dimostrano le policy di autorizzazioni basate su identità che vengono utilizzate per controllare le operazioni che sono consentite con cluster EMR.

Important

L'operazione `ModifyInstanceGroup` in Amazon EMR non richiede che si specifichi un ID cluster. Per questo motivo, negare questa azione basata su tag del cluster richiede un'ulteriore considerazione. Per ulteriori informazioni, consulta Negare l' ModifyInstanceGroup azione in Amazon EMR.

Argomenti

- Autorizzazione di operazioni solo su cluster con specifici valori di tag
- Richiesta dell'assegnazione di tag a un cluster appena creato
- Autorizzazione di operazioni su cluster con uno specifico tag indipendentemente dal valore di tag

Autorizzazione di operazioni solo su cluster con specifici valori di tag

Gli esempi che seguono mostrano una policy che autorizza un utente a eseguire operazioni in base al tag di cluster *`department`* con il valore *`dev`* e un altro utente a contrassegnare cluster con quello stesso tag. L'esempio di policy finale mostra come negare privilegi per taggare cluster EMR con qualsiasi tag tranne quel tag.

Nell'esempio di policy seguente, l'operatore di condizione `StringEquals` cerca di far corrispondere *`dev`* con il valore del tag *`department`*. Se il tag *`department`* non è stato aggiunto al cluster, oppure non contiene il valore *`dev`*, la policy non viene applicata e le operazioni non sono consentite da questa policy. Se nessun'altra dichiarazione di policy consente le operazioni, l'utente può utilizzare solo i cluster che hanno questo tag con tale valore.

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"
        }
      }
    }
  ]
}
```

Puoi anche specificare più valori di tag utilizzando un operatore di condizione. Ad esempio, per consentire tutte le operazioni su cluster in cui il tag *`department`* contiene il valore *`dev`* o *`test`*, puoi sostituire il blocco di condizione nell'esempio precedente con quanto segue.

```
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department":["dev", "test"]
        }
      }
```

Richiesta dell'assegnazione di tag a un cluster appena creato

Come nell'esempio precedente, l'esempio seguente di policy cerca lo stesso tag corrispondente: il valore *dev* per il tag *department*. In questo esempio, però, la chiave di condizione `RequestTag` specifica che la policy si applica durante la creazione del tag. Quindi è necessario creare un cluster con un tag che corrisponda al valore specificato.

Per creare un cluster con un tag, devi anche disporre dell'autorizzazione per l'azione `elasticmapredue:AddTags`. Per questa dichiarazione, la chiave di condizione `elasticmapreduce:ResourceTag` garantisce che IAM conceda l'accesso solo alle risorse dei tag con il valore *dev* sul tag *department*. L'elemento `Resource` viene utilizzato per limitare questa autorizzazione alle risorse del cluster.

Per le `PassRole` risorse, devi fornire l'ID o l'alias dell' AWS account, il nome del ruolo di servizio nell'`PassRoleForEMR`istruzione e il nome del profilo dell'istanza nell'`PassRoleForEC2`istruzione. Per ulteriori informazioni sul formato IAM ARN, consulta IAM ARNs nella IAM User Guide.

Per ulteriori informazioni sulla corrispondenza delle chiavi di tag, consulta `aws:RequestTag/tag-key` nella Guida per l'utente 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*"
        }
```

```
      }
    }
  ]
}
```

Autorizzazione di operazioni su cluster con uno specifico tag indipendentemente dal valore di tag

Puoi anche consentire operazioni solo su cluster con un determinato tag, indipendentemente dal valore di tag. A questo proposito, puoi utilizzare l'operatore `Null`. Per ulteriori informazioni, consulta Operatore di condizione per verificare la presenza di chiavi di condizione nella Guida per l'utente IAM. Ad esempio, per consentire operazioni solo su cluster EMR con il tag *`department`*, indipendentemente dal valore che lo stesso contiene, puoi sostituire i blocchi di condizione nell'esempio precedente con quello illustrato di seguito. L'operatore `Null` cerca il tag *`department`* su un cluster EMR. Se il tag esiste, l'istruzione `Null` restituisce il valore false, corrispondente alla condizione specificata nella dichiarazione di policy, e le operazioni appropriate sono consentite.

```
"Condition": {
  "Null": {
    "elasticmapreduce:ResourceTag/department":"false"
  }
}
```

La dichiarazione di policy seguente consente a un utente di creare un cluster EMR solo se il cluster avrà un tag *`department`*, indipendentemente dal valore dello stesso. Per la `PassRole` risorsa, devi fornire l'ID o l'alias AWS dell'account e il nome del ruolo del servizio. Per ulteriori informazioni sul formato IAM ARN, consulta IAM ARNs nella IAM User Guide.

Per ulteriori informazioni su come specificare l'operatore di condizione null ("falso"), consulta Operatore di condizione per verificare la presenza di chiavi di condizione nella Guida per l'utente 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*"
        }
      }
    }
  ]
}
```

Dichiarazioni di policy basate su identità di esempio per EMR Notebooks

Le dichiarazioni di policy IAM di esempio in questa sezione illustrano scenari comuni per l'utilizzo di chiavi per limitare le operazioni consentite mediante EMR Notebooks. Finché nessun'altra policy associata al principale (utente) consente le operazioni, le chiavi di contesto della condizione limitano le operazioni consentite come indicato.

Example : consente l'accesso solo ai EMR Notebooks creati da un utente in base all'assegnazione di tag

L'esempio seguente di istruzione di policy, quando collegata a un ruolo o a un utente, consente all'utente di utilizzare solo i notebook che ha creato. Questa istruzione di policy usa il tag predefinito applicato durante la creazione di un notebook.

In questo esempio, l'operatore di condizione `StringEquals` cerca di mettere in corrispondenza una variabile che rappresenta l'ID dell'utente attuale (`{aws:userId}`) con il valore del tag `creatorUserID`. Se il tag `creatorUserID` non è stato aggiunto al notebook, oppure non contiene il valore dell'ID dell'utente corrente, la policy non viene applicata e le operazioni non sono consentite da questa policy. Se nessun'altra istruzione di policy consente le operazioni, l'utente può utilizzare solo i notebook che hanno questo tag con tale valore.

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 -Richiedere l'assegnazione di tag al notebook durante la creazione

In questo esempio viene utilizzata la chiave di contesto `RequestTag`. L'operazione `CreateEditor` è consentita solo se l'utente non modifica o elimina il tag `creatorUserID` aggiunto per impostazione predefinita. La variabile ${aws:userId}, specifica l'ID dell'utente attualmente attivo, che è il valore predefinito del tag.

L'istruzione della policy può essere utilizzata per garantire che gli utenti non rimuovano il tag `createUserId` o ne modifichino il valore.

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:CreateEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:RequestTag/creatorUserId": "${aws:userid}"
        }
      },
      "Sid": "AllowELASTICMAPREDUCECreateeditor"
    }
  ]
}
```

Questo esempio richiede che l'utente crei il cluster con un tag con la stringa di chiave `dept` e un valore impostato per uno dei seguenti: `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 -Limitare la creazione di notebook ai cluster con tag e richiedere i tag del notebook

Questo esempio consente la creazione di notebook solo se il notebook viene creato con un tag che abbia la stringa di chiave `owner` impostata su uno dei valori specificati. Inoltre, è possibile creare il notebook solo se il cluster dispone di un tag con la stringa di chiave `department` impostata su uno dei valori specificati.

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 -Limitare la possibilità di avviare un notebook basato su tag

Questo esempio limita la possibilità di avviare solo i notebook che dispongono di un tag con la stringa di chiave `owner` impostata su uno dei valori specificati. Poiché l'elemento `Resource` è utilizzato per specificare solo `editor`, la condizione non è valida per il cluster e non è necessario aggiungere tag al cluster.

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"
    }
  ]
}
```

Questo esempio è simile a uno precedente. Tuttavia, il limite si applica solo ai cluster a cui sono applicati tag, non ai notebook.

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"
    }
  ]
}
```

In questo esempio viene utilizzato un altro set di tag di notebook e cluster. Consente di avviare un notebook solo se:

- Il notebook dispone di un tag con la stringa di chiave `owner` impostata su uno dei valori specificati

  -e-

- Il cluster dispone di un tag con la stringa di chiave `department` impostata su uno dei valori specificati

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 -Limitare la possibilità di aprire l'editor del notebook basato su tag

Questo esempio consente di aprire l'editor del notebook solo se:

- Il notebook dispone di un tag con la stringa di chiave `owner` impostata su uno dei valori specificati.

  -e-

- Il cluster dispone di un tag con la stringa di chiave `department` impostata su uno dei valori specificati.

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"
    }
  ]
}
```

Negare l' ModifyInstanceGroup azione in Amazon EMR

L'ModifyInstanceGroupsazione in Amazon EMR non richiede che tu fornisca un ID cluster con l'azione. È invece possibile specificare solo un ID del gruppo di istanze. Per questo motivo, una policy di negazione apparentemente semplice per questa operazione basata sull'ID cluster o su un tag cluster potrebbe non avere l'effetto desiderato. Esaminiamo l'esempio di policy seguente.

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"
    }
  ]
}
```

Se un utente con questa policy associata esegue un'operazione `ModifyInstanceGroup` e specifica solo l'ID del gruppo di istanze, la policy non viene applicata. Poiché l'operazione è consentita su tutte le altre risorse, avrà esito positivo.

Una soluzione a questo problema consiste nell'allegare una dichiarazione politica all'identità che utilizza un NotResourceelemento per negare qualsiasi `ModifyInstanceGroup` azione eseguita senza un ID cluster. La seguente policy di esempio aggiunge tale dichiarazione di negazione in modo che qualsiasi richiesta `ModifyInstanceGroups` abbia esito negativo a meno che non venga specificato un ID cluster. Poiché un'identità deve specificare un ID cluster con l'operazione, le istruzioni negate basate sull'ID cluster sono efficaci.

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"
    }
  ]
}
```

Un problema simile si verifica quando si desidera negare l'operazione `ModifyInstanceGroups` in base al valore associato a un tag cluster. La soluzione è simile. Oltre a un'istruzione di negazione che specifica il valore del tag, è possibile aggiungere un'istruzione della policy che neghi l'operazione `ModifyInstanceGroup` se il tag specificato non è presente, indipendentemente dal valore.

Nell'esempio seguente viene illustrata una policy che, se collegata a un'identità, nega all'identità l'operazione `ModifyInstanceGroups` per qualsiasi cluster con il tag `department` impostato su `dev`. Questa istruzione è efficace solo a causa dell'istruzione di negazione che utilizza la condizione `StringNotLike` per negare l'operazione a meno che il tag `department` non sia 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"
    }
  ]
}
```

Risoluzione dei problemi relativi all'identità e all'accesso di Amazon EMR

Utilizza le informazioni seguenti per eseguire la diagnosi e risolvere i problemi comuni che possono verificarsi durante l'utilizzo di Amazon EMR e IAM.

Argomenti

- Non dispongo dell'autorizzazione per eseguire un'operazione in Amazon EMR
- Non sono autorizzato a eseguire iam: PassRole
- Voglio consentire a persone esterne al mio AWS account di accedere alle mie risorse Amazon EMR

Non dispongo dell'autorizzazione per eseguire un'operazione in Amazon EMR

Se ti AWS Management Console dice che non sei autorizzato a eseguire un'azione, devi contattare l'amministratore per ricevere assistenza. L'amministratore è la persona da cui si sono ricevuti il nome utente e la password.

Il seguente esempio di errore si verifica quando l'utente `mateojackson` prova a utilizzare la console per visualizzare i dettagli relativi a una risorsa *`my-example-widget`* fittizia, ma non dispone di autorizzazioni `EMR:`*`GetWidget`* fittizie.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 EMR:GetWidget on resource: my-example-widget
```

In questo caso, Mateo richiede al suo amministratore di aggiornare le policy per poter accedere alla risorsa *`my-example-widget`* utilizzando l'operazione `EMR:`*`GetWidget`*.

Non sono autorizzato a eseguire iam: PassRole

Se ricevi un errore che indica che non disponi dell'autorizzazione per eseguire l'operazione `iam:PassRole`, per poter passare un ruolo ad Amazon EMR dovrai aggiornare le policy.

Alcuni Servizi AWS consentono di passare un ruolo esistente a quel servizio invece di creare un nuovo ruolo di servizio o un ruolo collegato al servizio. Per eseguire questa operazione, è necessario disporre delle autorizzazioni per trasmettere il ruolo al servizio.

Il seguente esempio di errore si verifica quando un utente IAM denominato `marymajor` cerca di utilizzare la console per eseguire un'operazione in Amazon EMR. Tuttavia, l'azione richiede che il servizio disponga delle autorizzazioni concesse da un ruolo di servizio. Mary non dispone delle autorizzazioni per passare il ruolo al servizio.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

In questo caso, le policy di Mary devono essere aggiornate per poter eseguire l'operazione `iam:PassRole`.

Se hai bisogno di aiuto, contatta il tuo AWS amministratore. L'amministratore è la persona che ti ha fornito le credenziali di accesso.

Voglio consentire a persone esterne al mio AWS account di accedere alle mie risorse Amazon EMR

È possibile creare un ruolo con il quale utenti in altri account o persone esterne all'organizzazione possono accedere alle tue risorse. È possibile specificare chi è attendibile per l'assunzione del ruolo. Per i servizi che supportano politiche basate sulle risorse o liste di controllo degli accessi (ACLs), puoi utilizzare tali politiche per consentire alle persone di accedere alle tue risorse.

Per ulteriori informazioni, consulta gli argomenti seguenti:

- Per sapere se Amazon EMR supporta queste funzionalità, consulta Funzionamento di Amazon EMR con IAM.
- Per scoprire come fornire l'accesso alle risorse di tua proprietà, consulta Fornire l'accesso a un utente IAM in Account AWS un altro Account AWS di tua proprietà nella IAM User Guide.
- Per scoprire come fornire l'accesso alle tue risorse a terze parti Account AWS, consulta Fornire l'accesso a soggetti Account AWS di proprietà di terze parti nella Guida per l'utente IAM.
- Per informazioni su come fornire l'accesso tramite la federazione delle identità, consulta Fornire l'accesso a utenti autenticati esternamente (Federazione delle identità) nella Guida per l'utente IAM.
- Per informazioni sulle differenze di utilizzo tra ruoli e policy basate su risorse per l'accesso multi-account, consulta Accesso a risorse multi-account in IAM nella Guida per l'utente IAM.

Utilizzo di Amazon S3 Access Grants con Amazon EMR

Panoramica di S3 Access Grants per Amazon EMR

Con le versioni 6.15.0 e successive di Amazon EMR, Amazon S3 Access Grants fornisce una soluzione di controllo degli accessi scalabile che puoi utilizzare per aumentare l'accesso ai tuoi dati Amazon S3 da Amazon EMR. Se hai una configurazione di autorizzazioni complessa o di grandi dimensioni per i dati S3, puoi utilizzare Access Grants per dimensionare le autorizzazioni relative ai dati S3 per utenti, ruoli e applicazioni sul cluster.

Utilizza S3 Access Grants per aumentare l'accesso ai dati di Amazon S3 oltre alle autorizzazioni concesse dal ruolo di runtime o dai ruoli IAM collegati alle identità con accesso al tuo cluster EMR. Per ulteriori informazioni, consulta Gestione degli accessi con S3 Access Grants nella Guida per l'utente di Amazon S3.

Per conoscere i passaggi per utilizzare S3 Access Grants con altre implementazioni Amazon EMR, consulta la seguente documentazione:

- Utilizzo di S3 Access Grants con Amazon EMR su EKS
- Utilizzo di S3 Access Grants con Amazon EMR Serverless

Come funziona Amazon EMR con S3 Access Grants

I rilasci di Amazon EMR versione 6.15.0 e successive forniscono un'integrazione nativa con S3 Access Grants. Puoi abilitare S3 Access Grants su Amazon EMR ed eseguire processi Spark. Quando il processo Spark effettua una richiesta di dati S3, Amazon S3 fornisce credenziali temporanee che rientrano nell'ambito del bucket, del prefisso o dell'oggetto specifico.

Di seguito è riportata una panoramica generale sul modo in cui Amazon EMR ottiene l'accesso ai dati protetti da S3 Access Grants.

1. Un utente invia un processo Spark di Amazon EMR che utilizza dati archiviati in Amazon S3.
2. Amazon EMR elabora una richiesta a S3 Access Grants per consentire l'accesso al bucket, al prefisso o all'oggetto per conto di quell'utente.
3. Amazon S3 restituisce credenziali temporanee sotto forma di token AWS Security Token Service (STS) per l'utente. Il token ha accesso al bucket, al prefisso o all'oggetto S3.
4. Amazon EMR utilizza il token STS per recuperare dati da S3.
5. Amazon EMR riceve i dati da S3 e restituisce i risultati all'utente.

Considerazioni su S3 Access Grants con Amazon EMR

Prendi nota dei seguenti comportamenti e limitazioni quando usi S3 Access Grants con Amazon EMR.

Supporto funzionalità

- S3 Access Grants è supportato con Amazon EMR versioni 6.15.0 e successive.
- Spark è l'unico motore di query supportato quando utilizzi S3 Access Grants con Amazon EMR.
- Delta Lake e Hudi sono gli unici formati a tabella aperta supportati quando utilizzi S3 Access Grants con Amazon EMR.
- Le seguenti funzionalità di Amazon EMR non sono supportate per l'utilizzo con S3 Access Grants:
  - Tabelle Apache Iceberg
  - Autenticazione LDAP nativa
  - Autenticazione nativa di Apache Ranger
  - AWS CLI richieste ad Amazon S3 che utilizzano ruoli IAM
  - Accesso a S3 tramite il protocollo S3A open source
- L'opzione `fallbackToIAM` non è supportata per i cluster EMR che utilizzano la propagazione affidabile delle identità con il Centro identità IAM.
- S3 Access Grants with AWS Lake Formation è supportato solo con i cluster Amazon EMR eseguiti su Amazon. EC2

Considerazioni comportamentali

- L'integrazione nativa di Apache Ranger con Amazon EMR include funzionalità congruenti con S3 Access Grants come parte del plug-in EMRFS S3 Apache Ranger. Se utilizzi Apache Ranger per il controllo granulare degli accessi (FGAC), ti consigliamo di utilizzare quel plug-in anziché S3 Access Grants.
- Amazon EMR fornisce una cache delle credenziali in EMRFS per garantire che un utente non debba effettuare richieste ripetute per le stesse credenziali all'interno di un processo Spark. Pertanto, Amazon EMR richiede sempre il privilegio di livello predefinito quando richiede le credenziali. Per ulteriori informazioni, consulta Richiedi l'accesso ai dati di S3 nella Guida per l'utente di Amazon S3.
- Nel caso in cui un utente esegua un'azione che S3 Access Grants non supporta, Amazon EMR è impostato per utilizzare il ruolo IAM specificato per l'esecuzione dei processi. Per ulteriori informazioni, consulta Torna ai ruoli IAM.

Avvio di un cluster Amazon EMR con S3 Access Grants

Questa sezione descrive come avviare un cluster EMR eseguito su Amazon EC2 e utilizza S3 Access Grants per gestire l'accesso ai dati in Amazon S3. Per conoscere i passaggi per utilizzare S3 Access Grants con altre implementazioni Amazon EMR, consulta la seguente documentazione:

- Utilizzo di S3 Access Grants con Amazon EMR su EKS
- Utilizzo di S3 Access Grants con EMR Serverless

Utilizza i seguenti passaggi per avviare un cluster EMR eseguito su Amazon EC2 e utilizza S3 Access Grants per gestire l'accesso ai dati in Amazon S3.

1. Configura un ruolo di esecuzione dei processi per il cluster EMR. Includi le autorizzazioni IAM necessarie per eseguire i processi Spark, `s3:GetDataAccess` e `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 le autorizzazioni impostate nei ruoli IAM. Se i ruoli IAM specificati per l'esecuzione dei processi dispongono di autorizzazioni per accedere direttamente a S3, gli utenti potrebbero essere in grado di accedere a più dati rispetto a quelli definiti in S3 Access Grants.

2. Quindi, usa AWS CLI per creare un cluster con Amazon EMR 6.15 o versione successiva e la `emrfs-site` classificazione per abilitare S3 Access Grants, in modo simile al seguente esempio:

```
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

Se utilizzi Amazon EMR con l'integrazione AWS Lake Formation, puoi utilizzare Amazon S3 Access Grants per l'accesso diretto o tabulare ai dati in Amazon S3.

> **Note**
>
> S3 Access Grants with AWS Lake Formation è supportato solo con i cluster Amazon EMR eseguiti su Amazon. EC2

Accesso diretto

L'accesso diretto include tutte le chiamate per accedere ai dati S3 che non richiamano l'API per il servizio AWS Glue che Lake Formation utilizza come metastore con Amazon EMR, ad esempio per chiamare: `spark.read`

```
spark.read.csv("s3://...")
```

Quando utilizzi S3 Access Grants con AWS Lake Formation Amazon EMR, tutti i modelli di accesso diretto passano attraverso S3 Access Grants per ottenere credenziali S3 temporanee.

Accesso tabulare

L'accesso tabulare si verifica quando Lake Formation richiama l'API del metastore per accedere alla posizione di S3, ad esempio per eseguire query sui dati della tabella:

```
spark.sql("select * from test_tbl")
```

Quando utilizzi S3 Access Grants con AWS Lake Formation Amazon EMR, tutti i modelli di accesso tabulari passano attraverso Lake Formation.

Torna ai ruoli IAM

Se un utente tenta di eseguire un'operazione che S3 Access Grants non supporta, Amazon EMR è impostato per utilizzare il ruolo IAM specificato per l'esecuzione dei processi quando la configurazione `fallbackToIAM` è `true`. Ciò consente agli utenti di tornare al proprio ruolo di esecuzione dei processi per fornire le credenziali per l'accesso a S3 in scenari non coperti da S3 Access Grants.

Se `fallbackToIAM` è abilitato, gli utenti possono accedere ai dati consentiti dall'Access Grant. Se non esiste un token S3 Access Grants per i dati di destinazione, Amazon EMR verifica l'autorizzazione per il ruolo di esecuzione dei processi.

> **Note**
>
> Ti consigliamo di testare le tue autorizzazioni di accesso con la configurazione `fallbackToIAM` abilitata anche se prevedi di disabilitare l'opzione per i carichi di lavoro di produzione. Con i processi Spark, ci sono altri modi in cui gli utenti potrebbero accedere a tutti i set di autorizzazioni con le proprie credenziali IAM. Se abilitate sui cluster EMR, le autorizzazioni di S3 consentono ai processi Spark di accedere alle posizioni S3. Assicurti di proteggere queste posizioni S3 dall'accesso esterno a EMRFS. Ad esempio, devi proteggere le posizioni S3 dall'accesso da parte dei client S3 utilizzati nei notebook o dalle applicazioni che non sono supportate da S3 Access Grants come Hive o Presto.

Autenticazione nei nodi cluster Amazon EMR

I client SSH possono utilizzare una coppia di EC2 chiavi Amazon per autenticarsi nelle istanze del cluster. In alternativa, con Amazon EMR rilasci 5.10.0 e successivi, puoi configurare Kerberos per autenticare utenti e connessioni SSH sul nodo primario. E con i rilasci 5.12.0 e successivi di Amazon EMR, puoi autenticarti con LDAP.

Argomenti

- Usa una coppia di EC2 chiavi per le credenziali SSH per Amazon EMR
- Utilizzo di Kerberos per l'autenticazione con Amazon EMR
- Utilizzo di server Active Directory o LDAP per l'autenticazione con Amazon EMR

Usa una coppia di EC2 chiavi per le credenziali SSH per Amazon EMR

I nodi del cluster Amazon EMR vengono eseguiti su istanze Amazon EC2 . Puoi connetterti ai nodi del cluster nello stesso modo in cui puoi connetterti alle EC2 istanze Amazon. Puoi usare Amazon EC2 per creare una coppia di chiavi o importare una coppia di chiavi. Quando crei un cluster, puoi specificare la coppia di EC2 chiavi Amazon che verrà utilizzata per le connessioni SSH a tutte le istanze del cluster. Puoi anche creare un cluster senza una coppia di chiavi. Questa operazione viene di solito effettuata con cluster transitori che si avviano, eseguono fasi e quindi terminano in automatico.

Il client SSH che utilizzi per connetterti al cluster deve utilizzare il file della chiave privata associato a tale coppia di chiavi. Questo è un file .pem per client SSH che utilizzano Linux, Unix e macOS. Devi impostare le autorizzazioni in modo che solo il proprietario della chiave disponga dell'autorizzazione per accedere al file. Si tratta di un file .ppk per i client SSH che utilizzano Windows ed è in genere creato a partire dal file .pem.

- Per ulteriori informazioni sulla creazione di una coppia di EC2 chiavi Amazon, consulta Amazon EC2 key pair nella Amazon EC2 User Guide.
- Per istruzioni sull'uso di Pu TTYgen per creare un file.ppk da un file.pem, consulta Convertire la tua chiave privata usando Pu nella TTYgen Amazon User Guide. EC2
- Per ulteriori informazioni sull'impostazione delle autorizzazioni per i file.pem e su come connettersi al nodo primario di un cluster EMR utilizzando metodi diversi, ad esempio da `ssh` Linux o macOS, PuTTY da Windows o da qualsiasi sistema operativo supportato, vedere. AWS CLI Connect al nodo primario del cluster Amazon EMR tramite SSH

Utilizzo di Kerberos per l'autenticazione con Amazon EMR

Amazon EMR rilascio 5.10.0 e successivi supporta Kerberos. Kerberos è un protocollo di autenticazione di rete che utilizza la crittografia con chiave segreta per fornire un'autenticazione efficace affinché le password o altre credenziali non siano inviate in rete in un formato non crittografato.

In Kerberos, i servizi e gli utenti che devono effettuare l'autenticazione sono denominati principali. I principali si trovano in un realm Kerberos. Nel realm, un server Kerberos, denominato Key Distribution Center (KDC), fornisce ai principali i mezzi per eseguire l'autenticazione. A questo proposito, il KDC emette dei ticket per l'autenticazione. Il KDC gestisce un database dei principali nel realm, le relative password e altre informazioni amministrative su ogni principale. Un KDC può anche

accettare credenziali di autenticazione da principali in altri realm; questa condizione è nota come trust tra realm. Inoltre, un cluster EMR può utilizzare un server KDC esterno per autenticare i principali.

La definizione di un trust tra realm o l'uso di un server KDC esterno è di uso comune nell'autenticazione degli utenti da un dominio di Active Directory. Questo consente agli utenti di accedere a un cluster EMR con il proprio account di dominio se si connettono a un cluster o lavorano con applicazioni Big Data utilizzando SSH.

Quando si utilizza l'autenticazione Kerberos, Amazon EMR configura Kerberos per le applicazioni, i componenti e i sottosistemi che installa sul cluster di modo che si autentichino a vicenda.

> **Important**
>
> Amazon EMR non supporta AWS Directory Service for Microsoft Active Directory un trust cross-realm o come KDC esterno.

Prima di configurare Kerberos utilizzando Amazon EMR, ti consigliamo di acquisire familiarità con i concetti di Kerberos, i servizi eseguiti su un KDC e gli strumenti di amministrazione dei servizi Kerberos. Per ulteriori informazioni, consulta la Documentazione di MIT Kerberos, pubblicata dal Kerberos Consortium.

Argomenti

- Applicazioni supportate con Amazon EMR
- Opzioni di architettura Kerberos con Amazon EMR
- Configurazione di Kerberos su Amazon EMR
- Utilizzo di SSH per connettersi a cluster kerberizzati con Amazon EMR
- Tutorial: configura un KDC dedicato al cluster con Amazon EMR
- Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory

Applicazioni supportate con Amazon EMR

In un cluster EMR, i principali Kerberos sono i sottosistemi e i servizi di applicazioni di Big Data eseguiti su tutti i nodi di cluster. Amazon EMR può configurare le applicazioni e i componenti elencati di seguito per utilizzare Kerberos. Ogni applicazione ha un principale utente Kerberos a cui è associato.

Amazon EMR non supporta relazioni di fiducia tra realm con AWS Directory Service for Microsoft Active Directory.

Amazon EMR configura solo le caratteristiche di autenticazione Kerberos open source per le applicazioni e i componenti elencati di seguito. Qualsiasi altra applicazione installata non utilizza Kerberos e ciò può comportare un'incapacità a comunicare con i componenti che utilizzano Kerberos e causare errori nelle applicazioni. Per le applicazioni e i componenti che non utilizzano Kerberos, l'autenticazione non è abilitata. Le applicazioni e i componenti supportati possono variare a seconda dei rilasci di Amazon EMR.

L'interfaccia utente di Livy è l'unica interfaccia utente Web ospitata sul cluster che è Kerberizzato.

- Hadoop MapReduce
- Hbase
- HCatalog
- HDFS
- Hive
  - Non abilitare Hive con l'autenticazione LDAP. Questa operazione può causare problemi di comunicazione con l'applicazione YARN che utilizza Kerberos.
- Hue
  - L'autenticazione utente Hue non è impostata automaticamente e può essere configurata utilizzando l'API di configurazione.
  - Il server Hue utilizza Kerberos. Il front-end Hue (interfaccia utente) non è configurato per l'autenticazione. L'autenticazione LDAP può essere configurata per l'interfaccia utente Hue.
- Livy
  - La rappresentazione di Livy con cluster che utilizzano Kerberos è supportata in Amazon EMR 5.22.0 e rilasci successivi.
- Oozie
- Phoenix
- Presto
  - Presto supporta l'autenticazione Kerberos nei rilasci di Amazon EMR 6.9.0 e successivi.
  - Per utilizzare l'autenticazione Kerberos per Presto, devi abilitare la crittografia in transito.
- Spark
- Tez

- Trino
  - Trino supporta l'autenticazione Kerberos in Amazon EMR rilasci 6.11.0 e successivi.
  - Per utilizzare l'autenticazione Kerberos per Trino, devi abilitare la crittografia in transito.
- YARN
- Zeppelin
  - Zeppelin è configurato solo per l'utilizzo di Kerberos con l'interprete Spark. Non è configurato per altri interpreti.
  - La rappresentazione dell'utente non è supportata per gli interpreti Zeppelin che utilizzano Kerberos diversi da Spark.
- Zookeeper
  - Il client Zookeeper non è supportato.

Opzioni di architettura Kerberos con Amazon EMR

Utilizzando Kerberos con Amazon EMR, è possibile scegliere tra le architetture elencate in questa sezione. Indipendentemente dall'architettura scelta, Kerberos può essere configurato con la stessa procedura. Si crea una configurazione di sicurezza, si specifica la configurazione di sicurezza e le opzioni compatibili di Kerberos specifiche per il cluster quando si crea il cluster stesso, infine si creano directory HDFS per utenti Linux sul cluster corrispondenti ai principali per l'utente nel server KDC. Per una spiegazione delle opzioni di configurazione ed esempi di configurazione per ogni architettura, consulta Configurazione di Kerberos su Amazon EMR.

KDC dedicato del cluster (KDC su nodo primario)

Questa configurazione è disponibile con Amazon EMR 5.10.0 e rilasci successivi.

Vantaggi

- Amazon EMR ha la piena proprietà del server KDC.
- Il server KDC nel cluster EMR è indipendente dalle implementazioni KDC centralizzate, ad esempio Microsoft Active Directory o AWS Managed Microsoft AD.
- L'impatto sulle prestazioni è minimo perché il server KDC gestisce l'autenticazione solo per i nodi locali all'interno del cluster.
- Altri cluster con Kerberos possono facoltativamente fare riferimento al server KDC come KDC esterno. Per ulteriori informazioni, consulta KDC esterno - Nodo primario su un altro cluster.

Considerazioni e limitazioni

- I cluster con Kerberos non possono eseguire l'autenticazione tra di loro, quindi le applicazioni non possono interagire. Se le applicazioni del cluster devono interagire, è necessario stabilire un trust tra realm tra i cluster o impostare un cluster come server KDC esterno per altri cluster. Se viene stabilita una relazione di fiducia tra più aree, KDCs deve avere aree Kerberos diverse.
- È necessario creare utenti Linux sull' EC2 istanza del nodo primario che corrispondono ai principali utenti KDC, insieme alle directory HDFS per ogni utente.
- I principali utenti devono utilizzare un file di chiave EC2 privata e `kinit` credenziali per connettersi al cluster tramite SSH.

Fiducia tra realm

In questa configurazione, i principali (di solito gli utenti) di un altro realm Kerberos eseguono l'autenticazione per i componenti applicativi su un cluster EMR con Kerberos, che ha il proprio server KDC. Il KDC sul nodo primario stabilisce una relazione di fiducia con un altro KDC utilizzando un principio cross-realm esistente in entrambi. KDCs Il nome e la password del principale corrispondono in modo preciso in ciascun KDC. I trust tra realm sono più comuni nelle implementazioni con Active Directory, come illustrato nel seguente diagramma. Sono supportati anche trust tra realm con un server KDC MIT esterno o un KDC su un altro cluster Amazon EMR.

Vantaggi

- Il cluster EMR su cui è installato il server KDC ne mantiene la piena proprietà.
- Con Active Directory, Amazon EMR crea automaticamente utenti Linux corrispondenti alle entità principali per l'utente dal server KDC. È comunque necessario creare directory HDFS per ogni utente. Inoltre, i principali utenti nel dominio Active Directory possono accedere ai cluster Kerberized utilizzando le credenziali, senza il file della chiave privata. `kinit` EC2 In questo modo viene meno la necessità di condividere il file di chiave privata tra gli utenti del cluster.
- Poiché ogni KDC del cluster gestisce l'autenticazione per i nodi del cluster, sono ridotti al minimo gli effetti della latenza di rete e il sovraccarico di elaborazione per un numero elevato di nodi tra i cluster.

Considerazioni e limitazioni

- Se si sta creando un trust con un realm Active Directory, è necessario fornire un nome utente e una password di Active Directory con le autorizzazioni per l'aggiunta di principali al dominio in cui si crea il cluster.
- Non è possibile stabilire trust tra realm Kerberos con lo stesso nome.
- I trust tra realm devono essere stabiliti in modo esplicito. Ad esempio, se il Cluster A e il Cluster B stabiliscono entrambi un trust tra realm con un server KDC, non hanno intrinsecamente una reciproca relazione di trust e le loro applicazioni non possono autenticarsi l'una con l'altra per interagire.
- KDCs devono essere gestiti in modo indipendente e coordinato in modo che le credenziali dei principali utenti corrispondano esattamente.

KDC esterno

Le configurazioni con un server KDC esterno sono supportate con Amazon EMR 5.20.0 e versioni successive.

- KDC esterno-KDC MIT
- KDC esterno - Nodo primario su un altro cluster
- KDC esterno: KDC del cluster su un cluster diverso con la relazione di fiducia tra realm Active Directory

KDC esterno-KDC MIT

Questa configurazione consente a uno o più cluster EMR di utilizzare principali definiti e gestiti in un server KDC MIT.

Vantaggi

- La gestione dei principali è consolidata in un unico KDC.
- Più cluster possono utilizzare lo stesso KDC nello stesso realm di Kerberos. Per ulteriori informazioni, consulta Requisiti per l'utilizzo di più cluster con lo stesso KDC.
- La gestione del KDC non comporta rallentamenti nelle prestazioni per un nodo primario su un cluster con Kerberos.

Considerazioni e limitazioni

- È necessario creare utenti Linux sull' EC2 istanza di ogni nodo primario del cluster Kerberized che corrispondono ai principali utenti KDC, insieme alle directory HDFS per ogni utente.
- I principali utenti devono utilizzare un file di chiave EC2 privata e `kinit` credenziali per connettersi ai cluster Kerberizzati tramite SSH.
- Ogni nodo nei cluster EMR con Kerberos deve avere un instradamento di rete al server KDC.
- L'autenticazione di ogni nodo nei cluster con Kerberos comporta un peso per il server KDC esterno, perciò la configurazione del KDC influisce sulle prestazioni del cluster. Quando si configura l'hardware del server KDC, è opportuno considerare il numero massimo di nodi Amazon EMR da supportare contemporaneamente.
- Le prestazioni del cluster dipendono dalla latenza di rete tra nodi nei cluster con Kerberos e il server KDC.
- La risoluzione dei problemi può essere più difficile a causa delle interdipendenze.

KDC esterno - Nodo primario su un altro cluster

Questa configurazione è quasi identica all'implementazione KDC MIT esterno qui sopra, l'unica differenza è che il KDC è attivo nel nodo primario di un cluster EMR. Per ulteriori informazioni, consultare KDC dedicato del cluster (KDC su nodo primario) e Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.

Vantaggi

- La gestione dei principali è consolidata in un unico KDC.
- Più cluster possono utilizzare lo stesso KDC nello stesso realm di Kerberos. Per ulteriori informazioni, consulta Requisiti per l'utilizzo di più cluster con lo stesso KDC.

Considerazioni e limitazioni

- È necessario creare utenti Linux sull' EC2 istanza del nodo primario di ogni cluster Kerberized che corrispondono ai principali utenti KDC, insieme alle directory HDFS per ogni utente.
- I principali utenti devono utilizzare un file di chiave EC2 privata e `kinit` credenziali per connettersi ai cluster Kerberizzati tramite SSH.
- Ogni nodo in ciascun cluster EMR deve avere un instradamento di rete al server KDC.

- L'autenticazione di ogni nodo Amazon EMR nei cluster con Kerberos comporta un peso per il server KDC esterno, perciò la configurazione del KDC influisce sulle prestazioni del cluster. Quando si configura l'hardware del server KDC, è opportuno considerare il numero massimo di nodi Amazon EMR da supportare contemporaneamente.
- Le prestazioni del cluster dipendono dalla latenza di rete tra nodi nei cluster e il server KDC.
- La risoluzione dei problemi può essere più difficile a causa delle interdipendenze.

KDC esterno: KDC del cluster su un cluster diverso con la relazione di fiducia tra realm Active Directory

In questa configurazione, si crea prima un cluster con un server KDC dedicato che dispone di un trust tra realm monodirezionale con Active Directory. Per un tutorial dettagliato, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory. È quindi possibile avviare ulteriori cluster facendo riferimento al KDC del cluster con il trust come KDC esterno. Per vedere un esempio, consulta KDC esterno del cluster con trust tra realm Active Directory. In questo modo ogni cluster Amazon EMR che utilizza il KDC esterno può autenticare le entità principali definite e gestite in un dominio Microsoft Active Directory.

Vantaggi

- La gestione dei principali è consolidata nel dominio Active Directory.
- Amazon EMR entra nel realm di Active Directory; si elimina così la necessità di creare utenti Linux corrispondenti agli utenti di Active Directory. È comunque necessario creare directory HDFS per ogni utente.
- Più cluster possono utilizzare lo stesso KDC nello stesso realm di Kerberos. Per ulteriori informazioni, consulta Requisiti per l'utilizzo di più cluster con lo stesso KDC.
- I principali utenti nel dominio Active Directory possono accedere ai cluster Kerberized utilizzando le credenziali, senza il file della chiave privata. `kinit` EC2 In questo modo viene meno la necessità di condividere il file di chiave privata tra gli utenti del cluster.
- Spetta a un solo un nodo primario Amazon EMR l'onere della manutenzione del KDC e solo quel cluster deve essere creato con credenziali di Active Directory per il trust tra realm tra il KDC e Active Directory.

Considerazioni e limitazioni

- Ogni nodo in ogni cluster EMR deve avere un instradamento di rete al server KDC e al controller di dominio di Active Directory.
- L'autenticazione di ogni nodo Amazon EMR comporta un peso per il server KDC esterno, perciò la configurazione del KDC influisce sulle prestazioni del cluster. Quando si configura l'hardware del server KDC, è opportuno considerare il numero massimo di nodi Amazon EMR da supportare contemporaneamente.
- Le prestazioni del cluster dipendono dalla latenza di rete tra nodi nei cluster e il server KDC.
- La risoluzione dei problemi può essere più difficile a causa delle interdipendenze.

Requisiti per l'utilizzo di più cluster con lo stesso KDC

Più cluster possono utilizzare lo stesso KDC nello stesso realm di Kerberos. Tuttavia, se i cluster vengono eseguiti contemporaneamente, potrebbero fallire se utilizzano nomi Kerberos in conflitto. ServicePrincipal

Se disponi di più cluster simultanei con lo stesso KDC esterno, assicurati che i cluster utilizzino realm Kerberos diversi. Se i cluster devono utilizzare lo stesso realm Kerberos, assicurati che si trovino in sottoreti diverse e che i relativi intervalli CIDR non si sovrappongano.

Configurazione di Kerberos su Amazon EMR

Questa sezione fornisce dettagli ed esempi di configurazione per configurare Kerberos con architetture comuni. Indipendentemente dall'architettura scelta, i passaggi di base delle configurazione sono gli stessi e vengono effettuati in tre fasi. Se si utilizza un KDC esterno o si imposta un trust tra realm, è necessario accertarsi che ogni nodo in un cluster disponga di un instradamento di rete al KDC esterno, compresa la configurazione dei gruppi di sicurezza applicabili per consentire il traffico Kerberos in entrata e in uscita.

Fase 1: creazione di una configurazione di sicurezza con le proprietà di Kerberos

La configurazione di sicurezza specifica i dettagli del server KDC Kerberos e consente il riutilizzo della configurazione di Kerberos a ogni creazione di un cluster. Puoi creare una configurazione di sicurezza utilizzando la console Amazon EMR AWS CLI, o l'API EMR. La configurazione di sicurezza può inoltre contenere altre opzioni di sicurezza, ad esempio la crittografia. Per ulteriori informazioni sulla creazione di configurazioni di sicurezza e la specifica di una configurazione di sicurezza al momento della creazione di un cluster, consulta Usa le configurazioni di sicurezza per configurare la

sicurezza dei cluster Amazon EMR. Per informazioni sulle proprietà di Kerberos in una configurazione di sicurezza, consulta Impostazioni Kerberos per configurazioni di sicurezza.

Fase 2: creazione di un cluster e specifica di attributi Kerberos appositi per il cluster

Quando crei un cluster, specifichi una configurazione di sicurezza Kerberos insieme alle opzioni di Kerberos specifiche del cluster stesso. Quando utilizzi la console di Amazon EMR, sono disponibili solo le opzioni di Kerberos compatibili con la configurazione di sicurezza specificata. Quando utilizzi l'API AWS CLI o Amazon EMR, assicurati di specificare opzioni Kerberos compatibili con la configurazione di sicurezza specificata. Ad esempio, se quando crei un cluster con la CLI specifichi una password principale per un trust tra più realm ma la configurazione di sicurezza specificata non include parametri per questa situazione, si verifica un errore. Per ulteriori informazioni, consulta Impostazioni Kerberos per i cluster.

Fase 3: configurazione del nodo primario del cluster

In base ai requisiti dell'architettura e dell'implementazione, può essere necessaria una configurazione aggiuntiva sul cluster. È possibile farlo dopo la sua creazione o utilizzando operazioni di bootstrap o fasi durante il processo di creazione.

Per ogni utente autenticato Kerberos che si connette al cluster tramite SSH, è necessario accertarsi che vengano creati account Linux corrispondenti agli utenti di Kerberos. Se le entità principali dell'utente vengono fornite da un controller di dominio Active Directory, sia come KDC esterno che attraverso un trust tra realm, Amazon EMR crea automaticamente account Linux. Se Active Directory non viene utilizzato, è necessario creare principali per ciascun utente corrispondenti al rispettivo utente Linux. Per ulteriori informazioni, consulta Configurazione di un cluster Amazon EMR per utenti HDFS autenticati da Kerberos e connessioni SSH.

Ogni utente deve inoltre disporre di una directory HDFS di proprietà, che è necessario creare. Inoltre, SSH deve essere configurato con GSSAPI abilitato per consentire le connessioni dagli utenti autenticati con Kerberos. GSSAPI deve essere abilitato sul nodo primario e l'applicazione client SSH deve essere configurata per l'utilizzo di GSSAPI. Per ulteriori informazioni, consulta Configurazione di un cluster Amazon EMR per utenti HDFS autenticati da Kerberos e connessioni SSH.

Configurazione di sicurezza e impostazioni del cluster per Kerberos su Amazon EMR

Quando crei un cluster che utilizza Kerberos, specifichi la configurazione di sicurezza con attributi Kerberos specifici del cluster. Non puoi specificare un set senza l'altro, altrimenti si verifica un errore.

Questo argomento fornisce una panoramica dei parametri di configurazione disponibili per Kerberos al momento della creazione di una configurazione di sicurezza e di un cluster. Inoltre, sono riportati

esempi di CLI per la creazione di configurazioni di sicurezza e cluster compatibili per architetture comuni.

Impostazioni Kerberos per configurazioni di sicurezza

Puoi creare una configurazione di sicurezza che specifichi gli attributi Kerberos utilizzando la console Amazon EMR, AWS CLI o l'API EMR. La configurazione di sicurezza può inoltre contenere altre opzioni di sicurezza, ad esempio la crittografia. Per ulteriori informazioni, consulta Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI.

Utilizza i seguenti riferimenti per comprendere le impostazioni di configurazione di sicurezza disponibili per l'architettura Kerberos scelta. Vengono visualizzate le impostazioni della console di Amazon EMR. Per le corrispondenti opzioni con la CLI, consulta Specificare le impostazioni Kerberos utilizzando AWS CLI o Esempi di configurazione.

Parametro		Descrizione
Kerberos		Specifica che Kerberos è abilitato per i cluster che utilizzano questa configurazione di protezione. Se un cluster utilizza questa configurazione di protezione, deve avere anche le impostazioni Kerberos specificate o, in caso contrario, genererà un errore.
Provider	KDC dedicato del cluster	Specifica che Amazon EMR crea un KDC sul nodo primario di qualsiasi cluster che utilizza questa configurazione di sicurezza. Quando crei il cluster, puoi specificare il nome del realm e la password di amministratore KDC. Se necessario, puoi fare riferimento a questo KDC da altri cluster. Crea tali cluster utilizzando una configura zione di sicurezza diversa, specifica un KDC esterno e utilizza il nome del realm e la password di amministr atore KDC specificati per il KDC dedicato al cluster.
	KDC esterno	Disponibili solo in Amazon EMR versione 5.20.0 e successive. Specifica che i cluster che utilizzan o questa configurazione di sicurezza autenticano le entità Kerberos principali utilizzando un server

Parametro		Descrizione
		KDC esterno al cluster. Non viene creato un KDC nel cluster. Quando crei il cluster, specifichi il nome del realm e la password di amministratore KDC per il KDC esterno.
Durata del ticket		Facoltativo. Specifica il periodo di validità di un ticket Kerberos emesso dal KDC nei cluster che utilizzano questa configurazione di sicurezza. La durata dei ticket è limitata per motivi di sicurezza . Le applicazioni e i servizi del cluster rinnovano automaticamente i ticket dopo la loro scadenza. Gli utenti che si connettono al cluster tramite SSH utilizzan do le credenziali Kerberos devono eseguire `kinit` dalla linea di comando del nodo primario per rinnovare un ticket dopo la relativa scadenza.
Fiducia tra realm		Specifica una relazione di fiducia tra realm tra un KDC dedicato al cluster in cluster che utilizzano questa configurazione di sicurezza e un KDC in un altro realm Kerberos. Le entità principali (in genere gli utenti) di un altro realm vengono autenticate nei cluster che utilizzano questa configurazione. È necessaria una configura zione aggiuntiva nell'altro realm Kerberos. Per ulteriori informazioni, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.
Proprietà di fiducia tra realm	Realm (Dominio)	Specifica il nome di realm Kerberos dell'altro realm della relazione di fiducia. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.

Parametro		Descrizione
	Dominio	Specifica il nome di dominio dell'altro realm della relazione di fiducia.
	Server amministratore	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server di amministrazione nell'altro realm della relazione di fiducia. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :749`).
	Server KDC	Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server KDC nell'altro realm della relazione di fiducia. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :88`).
KDC esterno		Specifica che i cluster KDC esterni vengono utilizzati dal cluster.

Parametro			Descrizione
Proprietà del KDC esterno	Server amministratore		Specifica il nome di dominio completo (FQDN) o l'indirizzo IP del server amministratore esterno. Generalmente, il server amministratore e il server KDC vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma comunicano su porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 749, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :749`).
	Server KDC		Specifica il nome di dominio completo (FQDN) del server KDC esterno. Generalmente, il server KDC e il server amministratore vengono eseguiti sullo stesso computer con lo stesso nome di dominio completo (FQDN), ma utilizzano porte diverse. Se non viene specificata alcuna porta, si utilizza la porta 88, ossia l'impostazione predefinita di Kerberos. Facoltativamente, puoi indicare la porta (ad esempio, `domain.example.com :88`).
	Integrazione con Active Directory		Specifica che l'autenticazione dell'entità Kerberos principale è integrata con un dominio Microsoft Active Directory.
	Proprietà di integrazione con Active Directory	Realm di Active Directory	Specifica il nome del realm Kerberos del dominio Active Directory. Per convenzione, i nomi del realm Kerberos sono uguali al nome di dominio, ma utilizzano solo lettere maiuscole.
		Dominio Active Directory	Specifica il nome del dominio di Active Directory.

Parametro			Descrizione
		Server Active Directory	Specifica il nome di dominio completo (FQDN) del controller di dominio Microsoft Active Directory.

Impostazioni Kerberos per i cluster

Puoi specificare le impostazioni Kerberos quando crei un cluster utilizzando la console Amazon EMR, AWS CLI o l'API EMR.

Utilizza i seguenti riferimenti per comprendere le impostazioni di configurazione di cluster disponibili per l'architettura Kerberos scelta. Vengono visualizzate le impostazioni della console di Amazon EMR. Per le corrispondenti opzioni con la CLI, consulta Esempi di configurazione.

Parametro	Descrizione
Realm (Dominio)	Il nome di realm Kerberos per il cluster. La convenzione Kerberos consiste a impostare un nome identico al nome di dominio, ma in maiuscolo. Ad esempio, per il dominio `ec2.internal` , utilizza `EC2.INTERNAL` come nome di realm.
Password amministratore KDC	La password utilizzata nel cluster per `kadmin` o `kadmin.local` . Queste sono interfacce a riga di comando per il sistema di amministr azione Kerberos V5, che gestisce principali Kerberos, policy sulle password e keytab per il cluster.
Password del principale del trust tra realm (facoltativa)	Richiesta quando si stabilisce un trust tra realm. La password del principale inter-rea lm, che deve essere identica in tutti i realm. Utilizzare una password complessa.

Parametro	Descrizione
Utente di aggiunta al dominio Active Directory (facoltativo)	Richiesto durante l'utilizzo di Active Directory in un trust tra realm. Questo è il nome di accesso utente per un account Active Directory con l'autorizzazione per aggiungere computer al dominio. Amazon EMR utilizza questa identità per aggiungere il cluster al dominio. Per ulteriori informazioni, consulta the section called “Fase 3: aggiunta di account al dominio per il cluster EMR”.
Password di aggiunta al dominio Active Directory (facoltativa)	La password per l'utente di aggiunta al dominio Active Directory. Per ulteriori informazioni, consulta the section called “Fase 3: aggiunta di account al dominio per il cluster EMR”.

Esempi di configurazione

Gli esempi seguenti illustrano le configurazioni di sicurezza e le configurazioni dei cluster per scenari comuni. AWS CLI i comandi sono mostrati per brevità.

KDC locale

I seguenti comandi creano un cluster con un KDC dedicato in esecuzione sul nodo primario. È necessaria una configurazione aggiuntiva sul cluster. Per ulteriori informazioni, consulta Configurazione di un cluster Amazon EMR per utenti HDFS autenticati da Kerberos e connessioni SSH.

Crea una configurazione di sicurezza

```
aws emr create-security-configuration --name LocalKDCSecurityConfig \
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ClusterDedicatedKdc",\
"ClusterDedicatedKdcConfiguration": {"TicketLifetimeInHours": 24 }}}}'
```

Crea cluster

```
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 dedicato del cluster con trust tra realm Active Directory

I seguenti comandi creano un cluster con un KDC dedicato in esecuzione sul nodo primario e un trust tra realm con un dominio di Active Directory. È necessaria una configurazione aggiuntiva sul cluster e in Active Directory. Per ulteriori informazioni, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.

Crea una configurazione di sicurezza

```
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"}}}}}'
```

Crea cluster

```
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 esterno su un altro cluster

I seguenti comandi creano un cluster che fa riferimento a un KDC dedicato sul nodo primario di un altro cluster per autenticare i principali. È necessaria una configurazione aggiuntiva sul cluster. Per ulteriori informazioni, consulta Configurazione di un cluster Amazon EMR per utenti HDFS autenticati da Kerberos e connessioni SSH.

Crea una configurazione di sicurezza

```
aws emr create-security-configuration --name ExtKDCOnDifferentCluster \
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ExternalKdc", \
"ExternalKdcConfiguration": {"KdcServerType": "Single", \
"AdminServer": "MasterDNSOfKDCMaster:749", \
"KdcServer": "MasterDNSOfKDCMaster:88"}}}}'
```

Crea cluster

```
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 esterno del cluster con trust tra realm Active Directory

I seguenti comandi creano un cluster senza KDC. Il cluster fa riferimento a un KDC dedicato in esecuzione sul nodo primario di un altro cluster per autenticare i principali. Il server KDC ha un trust tra realm con controller di dominio Active Directory. È necessaria una configurazione aggiuntiva sul nodo primario con il KDC. Per ulteriori informazioni, consulta Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory.

Crea una configurazione di sicurezza

```
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"}}}}}'
```

Crea cluster

```
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
```

Configurazione di un cluster Amazon EMR per utenti HDFS autenticati da Kerberos e connessioni SSH

Amazon EMR crea client utente autenticati in Kerberos per le applicazioni eseguite sul cluster, ad esempio, l'utente `hadoop`, l'utente `spark` e altri utenti. Puoi anche aggiungere utenti autenticati durante i processi del cluster mediante Kerberos. Gli utenti autenticati possono quindi connettersi al cluster con le relative credenziali Kerberos e lavorare con le applicazioni. Per consentire l'autenticazione dell'utente nel cluster, sono necessarie le seguenti configurazioni:

- Nel cluster deve esistere un account Linux corrispondente all'entità principale di Kerberos nel KDC. Amazon EMR esegue questa operazione automaticamente nelle architetture che si integrano con Active Directory.
- È necessario creare una directory utente HDFS nel nodo primario per ogni utente e assegnare all'utente le autorizzazioni per la directory.
- È necessario configurare il servizio SSH in modo che GSSAPI sia abilitato per il nodo primario. Inoltre, gli utenti devono disporre di un client SSH con GSSAPI abilitato.

Aggiunta di utenti Linux ed entità principali Kerberos al nodo primario

Se non utilizzi Active Directory, dovrai creare gli account Linux sul nodo primario del cluster e aggiungere i principali per tali utenti Linux al KDC. Questo include un principale nel KDC per il nodo primario. Oltre ai principali dell'utente, il KDC in esecuzione sul nodo primario necessita di un principale per l'host locale.

Se l'architettura include l'integrazione con Active Directory, i principali e gli utenti Linux sul server locale KDC, se applicabili, vengono creati automaticamente. Puoi ignorare questa fase. Per ulteriori informazioni, consultare Fiducia tra realm e KDC esterno: KDC del cluster su un cluster diverso con la relazione di fiducia tra realm Active Directory.

> **Important**
>
> Il KDC, insieme al database dei principali, viene perso quando il nodo primario termina perché quest'ultimo utilizza l'archiviazione temporanea. Se si creano utenti per connessioni

SSH, è consigliabile stabilire un trust tra realm con un KDC esterno configurato per la disponibilità elevata. In alternativa, se si creano utenti per connessioni SSH utilizzando account Linux, è possibile automatizzare il processo di creazione dell'account utilizzando operazioni di bootstrap e script in modo che possa essere ripetuto quando si crea un nuovo cluster.

Inviare una fase al cluster dopo averla creata oppure quando viene creato il cluster è il modo più semplice per aggiungere utenti e principali di KDC. In alternativa, è possibile connettersi al nodo primario utilizzando una EC2 key pair come `hadoop` utente predefinito per eseguire i comandi. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

L'esempio seguente invia uno script Bash `configureCluster.sh` a un cluster già esistente, facendo riferimento al relativo ID di cluster. Lo script viene salvato in 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"]
```

L'esempio seguente mostra il contenuto dello script `configureCluster.sh`. Lo script gestisce anche la creazione di directory utente HDFS e l'abilitazione di GSSAPI per SSH, argomenti che saranno trattati nelle sezioni di seguito.

```
#!/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
```

Aggiunta di directory HDFS utente

Per consentire agli utenti di accedere al cluster ed eseguire processi Hadoop, è necessario aggiungere directory utente HDFS per i relativi account Linux e concedere a ciascun utente la proprietà della sua directory.

Inviare una fase al cluster dopo averla creata oppure quando viene creato il cluster è il modo più semplice per creare directory HDFS. In alternativa, è possibile connettersi al nodo primario utilizzando una EC2 key pair come `hadoop` utente predefinito per eseguire i comandi. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

L'esempio seguente invia uno script Bash `AddHDFSUsers.sh` a un cluster già esistente, facendo riferimento al relativo ID di cluster. Lo script viene salvato in 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"]
```

L'esempio seguente mostra il contenuto dello 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
```

Abilitazione di GSSAPI per SSH

Perché gli utenti autenticati in Kerberos possano connettersi al nodo primario tramite SSH, per il servizio SSH deve essere abilitato il metodo di autenticazione GSSAPI. A questo scopo, esegui i seguenti comandi dalla riga di comando del nodo primario oppure utilizza una fase per eseguirla come script. Dopo aver riconfigurato SSH, devi riavviare il servizio.

```
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
```

Utilizzo di SSH per connettersi a cluster kerberizzati con Amazon EMR

Questa sezione mostra la procedura per la connessione di un utente autenticato su Kerberos al nodo primario di un cluster EMR.

Ogni computer usato per una connessione SSH deve avere installati un client SSH e applicazioni client Kerberos. Con ogni probabilità nei computer Linux sono inclusi per impostazione predefinita. Ad esempio, OpenSSH è installato nella maggior parte dei sistemi operativi Linux, Unix e macOS. È possibile verificare la presenza di un client SSH digitando ssh nella riga di comando. Se il computer non riconosce il comando, installa un client SSH per la connessione al nodo primario. Il progetto OpenSSH fornisce un'implementazione gratuita della suite completa di strumenti SSH. Per ulteriori informazioni, consulta il sito Web OpenSSH. Gli utenti di Windows possono utilizzare applicazioni come PuTTY come client SSH.

Per ulteriori informazioni sulle connessioni SSH, vedi Connettiti a un cluster Amazon EMR.

SSH usa GSSAPI per autenticare i client Kerberos ed è necessario abilitare l'autenticazione GSSAPI per il servizio SSH nel nodo primario del cluster. Per ulteriori informazioni, consulta Abilitazione di GSSAPI per SSH. Anche i client SSH devono utilizzare GSSAPI.

Negli esempi seguenti, per *MasterPublicDNS* utilizzare il valore visualizzato per Master public DNS nella scheda Riepilogo del riquadro dei dettagli del cluster, ad esempio. *ec2-11-222-33-44.compute-1.amazonaws.com*

Prerequisito per krb5.conf (non Active Directory)

Quando si utilizza una configurazione senza l'integrazione di Active Directory, oltre al client SSH e alle applicazioni client Kerberos, ogni computer client deve disporre di una copia del file `/etc/krb5.conf` corrispondente al file `/etc/krb5.conf` sul nodo primario del cluster.

Per copiare il file krb5.conf

1. Usa SSH per connetterti al nodo primario utilizzando una EC2 key pair e l'`hadoop`utente predefinito, ad esempio. `hadoop@`*MasterPublicDNS* Per istruzioni dettagliate, vedi Connettiti a un cluster Amazon EMR.
2. Dal nodo primario, copia i contenuti del file `/etc/krb5.conf`. Per ulteriori informazioni, consulta Connettiti a un cluster Amazon EMR.
3. Su ogni computer client utilizzato per la connessione al cluster, creare un file `/etc/krb5.conf` identico basato sulla copia creata nella fase precedente.

Utilizzo di Kinit e SSH

Ogni volta che si connette da un computer client con credenziali di Kerberos, l'utente deve prima rinnovare i ticket Kerberos per il proprio utente sul computer client. Inoltre, il client SSH deve essere configurato per utilizzare l'autenticazione GSSAPI.

Per utilizzare SSH per la connessione a un cluster EMR con Kerberos

1. Utilizzare `kinit` rinnovare i ticket Kerberos come nell'esempio seguente

   ```
   kinit user1
   ```

2. Utilizzare un client `ssh` insieme al principale creato nel KDC dedicato al cluster o al nome utente di Active Directory. Assicurarsi che l'autenticazione GSSAPI sia abilitata come illustrato negli esempi seguenti.

   Esempio: utenti Linux

   L'opzione `-K` specifica l'autenticazione GSSAPI.

```
ssh -K user1@MasterPublicDNS
```

Esempio: utenti Windows (PuTTY)

Assicurarsi che l'opzione di autenticazione GSSAPI per la sessione sia abilitata come illustrato:

Tutorial: configura un KDC dedicato al cluster con Amazon EMR

Questo argomento illustra come creare un cluster con un key distribution center (KDC) dedicato, aggiungendo manualmente account Linux a tutti i nodi cluster, aggiungendo entità principali Kerberos al KDC sul nodo primario e verificando che un client Kerberos sia installato nei computer client.

Per maggiori informazioni sul supporto di Amazon EMR per Kerberos e KDC, nonché sui collegamenti alla documentazione di MIT Kerberos, consulta Utilizzo di Kerberos per l'autenticazione con Amazon EMR.

Fase 1: Creazione del cluster che utilizza Kerberos

1. Creare una configurazione di sicurezza che attiva Kerberos. L'esempio seguente mostra un `create-security-configuration` comando che utilizza il AWS CLI che specifica la configurazione di sicurezza come struttura JSON in linea. È anche possibile fare riferimento a un file salvato in locale.

   ```
   aws emr create-security-configuration --name MyKerberosConfig \
   --security-configuration '{"AuthenticationConfiguration": {"KerberosConfiguration":
   {"Provider": "ClusterDedicatedKdc", "ClusterDedicatedKdcConfiguration":
    {"TicketLifetimeInHours": 24}}}}'
   ```

2. Creare un cluster che fa riferimento alla configurazione di sicurezza, stabilisce attributi Kerberos per il cluster e aggiunge account Linux utilizzando un'operazione di bootstrap. L'esempio seguente illustra l'utilizzo di un comando `create-cluster` con l' AWS CLI. Il comando fa riferimento alla configurazione di sicurezza creata precedentemente, `MyKerberosConfig`, nonché a uno script semplice, `createlinuxusers.sh`, come un'operazione di bootstrap, che deve essere creato e caricato in Amazon S3 prima di creare il cluster.

   ```
   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
```

Il codice seguente mostra il contenuto dello script `createlinuxusers.sh`, che aggiunge user1, user2 e user3 a ogni nodo nel cluster. Nella fase successiva, si aggiungeranno questi utenti come principali KDC.

```
#!/bin/bash
sudo adduser user1
sudo adduser user2
sudo adduser user3
```

Fase 2: Aggiunta di entità principali al KDC, creazione di directory utente HDFS e configurazione di SSH

Il KDC in esecuzione sul nodo primario necessita di un principale aggiunto per l'host locale e per ciascun utente creato nel cluster. È anche possibile creare directory HDFS per ogni utente che deve connettersi al cluster ed eseguire processi Hadoop. Analogamente, configurare il servizio SSH per attivare l'autenticazione GSSAPI, necessaria per Kerberos. Dopo aver attivato GSSAPI, riavviare il servizio SSH.

Il modo più semplice di eseguire queste operazioni è di inviare una fase al cluster. L'esempio seguente invia uno script Bash `configurekdc.sh` al cluster creato nella fase precedente, facendo riferimento al relativo ID di cluster. Lo script viene salvato in Amazon S3. In alternativa, è possibile connettersi al nodo primario utilizzando una EC2 key pair per eseguire i comandi o inviare il passaggio durante la creazione del cluster.

```
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"]
```

Il codice seguente mostra il contenuto dello 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
```

Gli utenti aggiunti ora dovrebbero essere in grado di connettersi al cluster tramite SSH. Per ulteriori informazioni, consulta Utilizzo di SSH per connettersi a cluster kerberizzati con Amazon EMR.

Tutorial: Configurazione di un trust tra realm con un controller di dominio Active Directory

Quando configuri un trust tra realm, autorizzi i principali (di solito utenti) di un altro realm Kerberos a eseguire l'autenticazione a componenti dell'applicazione sul cluster EMR. Il KDC (Key Distribution Center) dedicato al cluster stabilisce una relazione di fiducia con un altro KDC utilizzando un principio cross-realm esistente in entrambi. KDCs Il nome e la password del principale corrispondono in modo preciso.

Un trust cross-realm richiede che KDCs possano raggiungersi tramite la rete e risolvere i rispettivi nomi di dominio. Di seguito vengono descritti i passaggi per stabilire una relazione di trust tra domini con un controller di dominio Microsoft AD in esecuzione come EC2 istanza, insieme a un esempio di configurazione di rete che fornisce la connettività e la risoluzione dei nomi di dominio richieste. Qualsiasi configurazione di rete che consenta il traffico di rete richiesto tra di loro è accettabile. KDCs

Facoltativamente, dopo aver stabilito un trust tra realm con Active Directory utilizzando un KDC su un cluster, è possibile creare un altro cluster utilizzando un'altra configurazione di sicurezza per fare riferimento al KDC del primo cluster come KDC esterno. Per un esempio di configurazione di sicurezza e di impostazione del cluster, consultare KDC esterno del cluster con trust tra realm Active Directory.

Per maggiori informazioni sul supporto di Amazon EMR per Kerberos e KDC, nonché sui collegamenti alla documentazione di MIT Kerberos, consulta Utilizzo di Kerberos per l'autenticazione con Amazon EMR.

> **Important**
>
> Amazon EMR non supporta trust cross-realm con. AWS Directory Service for Microsoft Active Directory

Fase 1: Configurazione di VPC e sottorete

Fase 2: Avvio e installazione del controller di dominio Active Directory

Fase 3: aggiunta di account al dominio per il cluster EMR

Fase 4: Configurazione di un trust in entrata sul controller di dominio Active Directory

Fase 5: Utilizzo di un set di opzioni DHCP per specificare il controller di dominio Active Directory come server DNS di VPC

Fase 6: avvio di un cluster EMR che utilizza Kerberos

Fase 7: creazione di utenti HDFS e impostazione di autorizzazioni sul cluster per account Active Directory

Fase 1: Configurazione di VPC e sottorete

Le fasi seguenti descrivono la creazione di un VPC e di una sottorete di modo che il KDC dedicato al cluster possa accedere al controller di dominio Active Directory e risolvere il relativo nome di dominio.

In queste fasi, la risoluzione dei nomi di dominio viene fornita facendo riferimento al controller di dominio Active Directory come server dei nomi di dominio nel set di opzioni DHCP. Per ulteriori informazioni, consulta Fase 5: Utilizzo di un set di opzioni DHCP per specificare il controller di dominio Active Directory come server DNS di VPC.

Il KDC e il controller di dominio Active Directory devono essere in grado di risolvere uno il nome di dominio dell'altro. Ciò consente ad Amazon EMR di aggiungere computer al dominio e di configurare automaticamente gli account Linux e i parametri SSH corrispondenti sulle istanze del cluster.

Se Amazon EMR non è in grado di risolvere il nome di dominio, è possibile fare riferimento al trust utilizzando l'indirizzo IP del controller di dominio Active Directory. Tuttavia, è necessario aggiungere manualmente gli account Linux, aggiungere i principali corrispondenti al KDC dedicato al cluster e configurare SSH.

Per configurare VPC e sottorete

1. Creazione di un Amazon VPC con una singola sottorete pubblica Per ulteriori informazioni, consulta Fase 1: Creazione del VPC nella Guida alle nozioni di base su Amazon VPC.

   > **Important**
   >
   > Quando utilizzi un controller di dominio Microsoft Active Directory, scegli un blocco CIDR per il cluster EMR in modo che IPv4 tutti gli indirizzi abbiano una lunghezza inferiore a nove caratteri (ad esempio, 10.0.0.0/16). Questo perché i nomi DNS dei computer del cluster vengono utilizzati quando i computer si uniscono alla directory di Active Directory. AWS assegna nomi host DNS in base IPv4 all'indirizzo in modo che indirizzi IP più lunghi possano dare come risultato nomi DNS più lunghi di 15 caratteri. Active Directory ha un limite di 15 caratteri per la registrazione di nomi di computer aggiunti e tronca i nomi più lunghi, condizione che può causare errori imprevedibili.

2. Rimuovere il set di opzioni DHCP di default assegnato al VPC. Per ulteriori informazioni, consulta Modifica di un VPC per non utilizzare opzioni DHCP. Successivamente, aggiungere un nuovo set che specifica il controller di dominio Active Directory come server DNS.
3. Confermare che il supporto DNS è attivato per il VPC, ovvero, che gli hostname DNS e la risoluzione DNS sono attivati (impostazione predefinita). Per ulteriori informazioni, consulta Aggiornamento del supporto DNS per il VPC.

4. Confermare che il VPC dispone di un gateway Internet associato (impostazione predefinita). Per ulteriori informazioni, consulta l'argomento relativo alla creazione e all'associazione di un gateway Internet.

   > **Note**
   >
   > Un gateway Internet è utilizzato in questo esempio in quanto si sta creando un nuovo controller di dominio per il VPC. È tuttavia possibile che un gateway Internet non sia necessario per l'applicazione. Il solo requisito è che il KDC dedicato al cluster possa accedere al controller di dominio Active Directory.

5. Creare una tabella di routing personalizzata, aggiungere un instradamento che porta al gateway Internet, quindi associarlo alla sottorete. Per ulteriori informazioni, consulta Creazione di una tabella di routing personalizzata.
6. Quando si avvia l' EC2 istanza per il controller di dominio, deve disporre di un IPv4 indirizzo pubblico statico per potersi connettere tramite RDP. Il modo più semplice per farlo è configurare la sottorete per l'assegnazione automatica di indirizzi pubblici. IPv4 che non è l'impostazione predefinita quando si crea una sottorete. Per ulteriori informazioni, consulta Modifica dell'attributo di IPv4 indirizzamento pubblico della sottorete. Eventualmente, è possibile assegnare l'indirizzo all'avvio dell'istanza. Per ulteriori informazioni, consulta Assegnazione di un IPv4 indirizzo pubblico durante l'avvio dell'istanza.
7. Al termine, prendi nota del tuo VPC e della sottorete. IDs Questi ID saranno utilizzati in seguito quando si avvia il controller di dominio Active Directory e il cluster.

Fase 2: Avvio e installazione del controller di dominio Active Directory

1. Avvia un' EC2 istanza basata sull'AMI di base di Microsoft Windows Server 2016. Consigliamo un tipo di istanza m4.xlarge o superiore. Per ulteriori informazioni, consulta Launching an Marketplace AWS Instance nella Amazon EC2 User Guide.
2. Prendi nota dell'ID di gruppo del gruppo di sicurezza associato all'EC2 istanza. Servirà per Fase 6: avvio di un cluster EMR che utilizza Kerberos. Usiamo*`sg-012xrlmdomain345`*. In alternativa, è possibile specificare gruppi di sicurezza diversi per il cluster EMR e per questa istanza che consente il traffico tra di essi. Per ulteriori informazioni, consulta i gruppi EC2 di sicurezza Amazon per le istanze Linux nella Amazon EC2 User Guide.
3. Connect all' EC2 istanza tramite RDP. Per ulteriori informazioni, consulta Connessione all'istanza Windows nella Amazon EC2 User Guide.

4. Avvia Server Manager per installare e configurare il ruolo dei servizi di dominio Active Directory sul server. Alzare il livello del server a controller di dominio e assegnare un nome di dominio (l'esempio utilizzato qui è *ad.domain.com*). Annotare il nome di dominio in quanto sarà necessario in un secondo momento quando si crea la configurazione di sicurezza EMR e il cluster. Per informazioni sulla configurazione di Active Directory, segui le istruzioni contenute in Come configurare Active Directory (AD) in Windows Server 2016.

   L'istanza viene riavviata al termine della procedura.

Fase 3: aggiunta di account al dominio per il cluster EMR

Esegui una connessione RDP al controller di dominio Active Directory per creare account in utenti e computer Active Directory per ogni utente del cluster. Per ulteriori informazioni, consulta Create a User Account in Active Directory Users and Computers sul sito Microsoft Learn. Annotare User logon name (Nome di accesso utente) di ogni utente. Questi nomi saranno necessari in seguito per la configurazione del cluster.

Inoltre, crea un account con privilegi sufficienti per aggiungere computer al dominio. Questo account viene specificato quando si crea un cluster. Amazon EMR lo utilizza per aggiungere istanze di cluster al dominio. Questo account e la relativa password sono specificati in Fase 6: avvio di un cluster EMR che utilizza Kerberos. Per delegare privilegi di aggiunta di computer all'account, è consigliabile creare un gruppo con privilegi di aggiunta e quindi assegnare l'utente al gruppo. Per istruzioni, consulta Delega dei privilegi di aggiunta di directory nella Guida per l'amministrazione di AWS Directory Service .

Fase 4: Configurazione di un trust in entrata sul controller di dominio Active Directory

I comandi di esempio seguenti creano un trust in Active Directory, ovvero un trust tra realm unidirezionale, in entrata e non transitivo con il KDC dedicato al cluster. L'esempio che utilizziamo per il realm del cluster è *EC2.INTERNAL*. Sostituisci il nome *KDC-FQDN* con il nome DNS pubblico indicato per il nodo primario di Amazon EMR che ospita il KDC. Il parametro `passwordt` specifica la cross-realm principal password (password del principale inter-realm), definita insieme al realm del cluster quando si crea un cluster. Il nome di realm è derivato dal nome di dominio di default in `us-east-1` per il cluster. `Domain` è il dominio di Active Directory in cui si crea il trust, che è minuscolo per convenzione. L'esempio utilizza *ad.domain.com*.

Aprire il prompt dei comandi di Windows con privilegi di amministratore e digitare i comandi seguenti per creare la relazione di trust sul controller di dominio 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
```

Fase 5: Utilizzo di un set di opzioni DHCP per specificare il controller di dominio Active Directory come server DNS di VPC

Ora che il controller di dominio Active Directory è configurato, è necessario configurare il VPC per utilizzarlo come server dei nomi di dominio per la risoluzione dei nomi in VPC. A questo proposito, associare un set di opzioni DHCP. Specifica il Domain name (Nome di dominio) come nome di dominio del cluster, ad esempio `ec2.internal` se il cluster si trova in us-east-1 o *`region`*`.compute.internal` per le altre Regioni. Per i server con nomi di dominio, è necessario specificare l'indirizzo IP del controller di dominio Active Directory (che deve essere raggiungibile dal cluster) come prima voce, seguita da AmazonProvidedDNS (ad esempio, DNS). ***`xx.xx.xx.xx`*** AmazonProvided Per ulteriori informazioni, consulta Modifica dei set di opzioni DHCP.

Fase 6: avvio di un cluster EMR che utilizza Kerberos

1. In Amazon EMR, crea una configurazione di sicurezza che specifica il controller di dominio Active Directory creato nelle fasi precedenti. Un esempio di comando è riportato di seguito. Sostituire il dominio, *`ad.domain.com`*, con il nome del dominio specificato in Fase 2: Avvio e installazione del controller di dominio 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. Creare il cluster con gli attributi seguenti:

   - Utilizzare l'opzione `--security-configuration` per specificare la configurazione di sicurezza creata. *`MyKerberosConfig`*Nell'esempio utilizziamo.
   - Utilizzare la proprietà `SubnetId` di `--ec2-attributes option` per specificare la sottorete creata in Fase 1: Configurazione di VPC e sottorete. Usiamo *`step1-subnet`* nell'esempio.
   - Utilizza `AdditionalMasterSecurityGroups` e `AdditionalSlaveSecurityGroups` dell'opzione `--ec2-attributes` per specificare che il gruppo di sicurezza associato al controller di dominio AD da Fase 2: Avvio e installazione del controller di dominio Active Directory è associato al nodo primario del cluster, nonché ai nodi principali e attività. Usiamo *`sg-012xrlmdomain345`* nell'esempio.

   Utilizzare `--kerberos-attributes` per specificare i seguenti attributi Kerberos specifici del cluster:

   - Il realm per il cluster specificato nella configurazione del controller di dominio Active Directory.
   - La password del principale del trust tra realm specificata come `passwordt` in Fase 4: Configurazione di un trust in entrata sul controller di dominio Active Directory.
   - Una `KdcAdminPassword`, che può essere utilizzata per amministrare il KDC dedicato al cluster.
   - La password e il nome di accesso utente dell'account Active Directory con privilegi di aggiunta di computer creati in Fase 3: aggiunta di account al dominio per il cluster EMR.

   L'esempio seguente avvia un cluster che utilizza 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
```

Fase 7: creazione di utenti HDFS e impostazione di autorizzazioni sul cluster per account Active Directory

Quando si configura una relazione di trust con Active Directory, Amazon EMR crea utenti Linux sul cluster per ogni account Active Directory. Ad esempio, il nome di accesso utente `LiJuan` in Active Directory dispone di un account Linux `lijuan`. I nomi utente Active Directory possono contenere lettere maiuscole, ma Linux non supporta la combinazione di maiuscole e minuscole di Active Directory.

Per consentire agli utenti di accedere al cluster ed eseguire processi Hadoop, è necessario aggiungere directory utente HDFS per i relativi account Linux e concedere a ciascun utente la proprietà della sua directory. A questo proposito, consigliamo di eseguire uno script salvato in Amazon S3 come fase di cluster. In alternativa, è possibile eseguire i comandi nello script illustrato di seguito dalla riga di comando sul nodo primario. Usa la EC2 key pair che hai specificato quando hai creato il cluster per connetterti al nodo primario tramite SSH come utente Hadoop. Per ulteriori informazioni, consulta Usa una coppia di EC2 chiavi per le credenziali SSH per Amazon EMR.

Esegui il comando seguente per aggiungere un passaggio al cluster che esegue uno 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"]
```

Il contenuto del file *AddHDFSUsers.sh* è il seguente.

```
#!/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
```

Gruppi Active Directory mappati a gruppi Hadoop

Amazon EMR utilizza System Security Services Daemon (SSD) per mappare gruppi Active Directory a gruppi Hadoop. Per confermare le mappature dei gruppi, dopo aver eseguito l'accesso al nodo primario come descritto in Utilizzo di SSH per connettersi a cluster kerberizzati con Amazon EMR, puoi utilizzare il comando `hdfs groups` per confermare che i gruppi Active Directory a cui l'account Active Directory appartiene siano stati mappati a gruppi Hadoop per l'utente Hadoop corrispondente sul cluster. È inoltre possibile verificare le mappature di gruppi di altri utenti specificando uno o più nomi utente con il comando, ad esempio `hdfs groups lijuan`. Per ulteriori informazioni, consulta groups nella Apache HDFS Commands Guide.

Utilizzo di server Active Directory o LDAP per l'autenticazione con Amazon EMR

Con i rilasci 6.12.0 e successivi di Amazon EMR, puoi utilizzare il protocollo LDAP over SSL (LDAPS) per avviare un cluster che si integra in modo nativo con il tuo server di identità aziendale. LDAP (Lightweight Directory Access Protocol) è un protocollo applicativo aperto e indipendente dal fornitore che accede e mantiene i dati. LDAP è comunemente usato per l'autenticazione degli utenti su server di identità aziendali ospitati su applicazioni come Active Directory (AD) e OpenLDAP. Con questa integrazione nativa, puoi utilizzare il tuo server LDAP per autenticare gli utenti su Amazon EMR.

I punti salienti dell'integrazione LDAP di Amazon EMR includono:

- Amazon EMR configura le applicazioni supportate per l'autenticazione con LDAP per tuo conto.
- Amazon EMR configura e mantiene la sicurezza per le applicazioni supportate con il protocollo Kerberos. Non è necessario immettere comandi o script.
- Puoi ottenere il controllo granulare degli accessi (FGAC) attraverso l'autorizzazione di Apache Ranger per il database e le tabelle di Hive Metastore. Per ulteriori informazioni, consulta Integrazione di Amazon EMR con Apache Ranger.
- Quando richiedi le credenziali LDAP per accedere a un cluster, ottieni un controllo granulare degli accessi (FGAC) su chi può accedere ai cluster EMR tramite SSH.

Le pagine seguenti forniscono una panoramica concettuale, i prerequisiti e le fasi per avviare un cluster EMR con l'integrazione LDAP di Amazon EMR.

Argomenti

- Panoramica di LDAP con Amazon EMR

- Componenti LDAP per Amazon EMR
- Supporto e considerazioni sulle applicazioni con LDAP per Amazon EMR
- Configurazione e avvio di un cluster EMR con LDAP
- Esempi di utilizzo di LDAP con Amazon EMR

Panoramica di LDAP con Amazon EMR

Lightweight Directory Access Protocol (LDAP) è un protocollo software utilizzato dagli amministratori di rete per gestire e controllare l'accesso ai dati autenticando gli utenti all'interno di una rete aziendale. Il protocollo LDAP memorizza le informazioni in una struttura di directory gerarchica ad albero. Per ulteriori informazioni, consulta Basic LDAP Concepts su LDAP.com.

All'interno della rete aziendale, molte applicazioni potrebbero utilizzare il protocollo LDAP per autenticare gli utenti. Con l'integrazione LDAP di Amazon EMR, i cluster EMR possono utilizzare in modo nativo lo stesso protocollo LDAP con una configurazione di sicurezza aggiuntiva.

Esistono due implementazioni principali del protocollo LDAP supportate da Amazon EMR: Active Directory e OpenLDAP. Sebbene siano possibili altre implementazioni, la maggior parte si adatta agli stessi protocolli di autenticazione di Active Directory o OpenLDAP.

Active Directory (AD)

Active Directory (AD) è un servizio di directory di Microsoft per reti di dominio Windows. AD è incluso nella maggior parte dei sistemi operativi Windows Server e può comunicare con i client tramite i protocolli LDAP e LDAPS. Per l'autenticazione, Amazon EMR tenta di associare un utente alla tua istanza AD utilizzando lo User Principal Name (UPN) come nome distinto e password. L'UPN utilizza il formato standard `username@domain_name`.

OpenLDAP

OpenLDAP è un'implementazione gratuita e open-source del protocollo LDAP. Per l'autenticazione, Amazon EMR tenta di associare un utente alla tua istanza OpenLDAP con il nome di dominio completo (FQDN) come nome distinto e password. L'FQDN utilizza il formato standard `username_attribute=username,LDAP_user_search_base`. In genere, il valore di `username_attribute` è `uid` e il valore `LDAP_user_search_base` contiene gli attributi della struttura ad albero che conduce all'utente. Ad esempio, `ou=People,dc=example,dc=com`.

Altre implementazioni gratuite e open-source del protocollo LDAP in genere seguono un FQDN simile a OpenLDAP per i nomi distinti dei loro utenti.

Componenti LDAP per Amazon EMR

Puoi utilizzare il tuo server LDAP per autenticarti con Amazon EMR e qualsiasi applicazione che l'utente utilizza direttamente sul cluster EMR tramite i seguenti componenti.

Agente segreto

L'agente segreto è un processo in cluster che autentica tutte le richieste degli utenti. L'agente segreto crea l'associazione utente al server LDAP per conto delle applicazioni supportate sul cluster EMR. L'agente segreto viene eseguito come utente `emrsecretagent` e scrive log nella directory `/emr/secretagent/log`. Questi log forniscono dettagli sullo stato della richiesta di autenticazione di ciascun utente e sugli eventuali errori che potrebbero emergere durante l'autenticazione dell'utente.

Daemon dei servizi di sicurezza del sistema (SSSD)

SSSD è un daemon che viene eseguito su ogni nodo di un cluster EMR abilitato per LDAP. SSSD crea e gestisce un utente UNIX per sincronizzare l'identità aziendale remota su ciascun nodo. Le applicazioni basate su YARN come Hive e Spark richiedono l'esistenza di un utente UNIX locale su ogni nodo che esegue una query per un utente.

Supporto e considerazioni sulle applicazioni con LDAP per Amazon EMR

Questo argomento elenca le applicazioni supportate, le funzionalità supportate e le funzionalità non supportate.

Applicazioni supportate con LDAP per Amazon EMR

> **Important**
>
> Le applicazioni elencate in questa pagina sono le uniche applicazioni supportate da Amazon EMR per LDAP. Per garantire la sicurezza dei cluster, è possibile includere solo applicazioni compatibili con LDAP quando si crea un cluster EMR con il protocollo LDAP abilitato. Se tenti di installare altre applicazioni non supportate, Amazon EMR rifiuta la richiesta di un nuovo cluster.

I rilasci 6.12 e successivi di Amazon EMR supportano l'integrazione LDAP con le seguenti applicazioni:

- Apache Livy
- Da Apache Hive a 2 () HiveServer HS2
- Trino
- Presto
- Hue

Le seguenti applicazioni possono essere installate in un cluster EMR e possono essere configurate per soddisfare le esigenze di sicurezza:

- Apache Spark
- Apache Hadoop

Funzionalità supportate con LDAP per Amazon EMR

Puoi utilizzare le seguenti funzionalità di Amazon EMR con l'integrazione LDAP:

> **Note**
>
> Per proteggere le credenziali LDAP, devi utilizzare la crittografia in transito per proteggere il flusso di dati all'interno e all'esterno del cluster. Per maggiori informazioni sulla crittografia in transito, consulta Crittografa i dati inattivi e in transito con Amazon EMR.

- Crittografia dei dati in transito (obbligatoria) e a riposo
- Gruppi di istanze, parchi istanze e istanze Spot
- Riconfigurazione delle applicazioni in un cluster in esecuzione
- Server-Side Encryption (SSE) EMRFS

Caratteristiche non supportate

Considera le seguenti limitazioni quando utilizzi l'integrazione LDAP di Amazon EMR:

- Amazon EMR disabilita i passaggi per i cluster con LDAP abilitato.
- Amazon EMR non supporta ruoli di runtime e AWS Lake Formation integrazioni per cluster con LDAP abilitato.

- Amazon EMR non supporta LDAP con StartTLS.
- Amazon EMR non supporta la modalità ad alta disponibilità (cluster con più nodi primari) per i cluster con LDAP abilitato.
- Non puoi ruotare le credenziali o i certificati di associazione per i cluster con LDAP abilitato. Se uno di questi campi è stato ruotato, ti consigliamo di avviare un nuovo cluster con le credenziali o i certificati di associazione aggiornati.
- È necessario utilizzare basi di ricerca esatte con LDAP. La base di ricerca LDAP per utenti e gruppi non supporta i filtri di ricerca LDAP.

Configurazione e avvio di un cluster EMR con LDAP

Questa sezione spiega come configurare Amazon EMR per l'uso con l'autenticazione LDAP.

Argomenti

- Aggiungi AWS Secrets Manager autorizzazioni al ruolo dell'istanza Amazon EMR
- Creazione della configurazione di sicurezza di Amazon EMR per l'integrazione LDAP
- Avvio di un cluster EMR che si autentica con LDAP

Aggiungi AWS Secrets Manager autorizzazioni al ruolo dell'istanza Amazon EMR

Amazon EMR utilizza un ruolo di servizio IAM per eseguire operazioni per tuo conto per il provisioning e la gestione dei cluster. Il ruolo di servizio per EC2 le istanze cluster, chiamato anche profilo di EC2 istanza per Amazon EMR, è un tipo speciale di ruolo di servizio che Amazon EMR assegna a EC2 ogni istanza di un cluster al momento del lancio.

Per definire le autorizzazioni per un cluster EMR per interagire con i dati di Amazon S3 e AWS altri servizi, definisci un profilo di istanza EC2 Amazon personalizzato anziché quando avvii `EMR_EC2_DefaultRole` il cluster. Per ulteriori informazioni, consultare Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2 e Personalizza i ruoli IAM con Amazon EMR.

Aggiungi le seguenti istruzioni al profilo di EC2 istanza predefinito per consentire ad Amazon EMR di etichettare le sessioni e accedere ai certificati LDAP AWS Secrets Manager che archivia.

```
    {
      "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**
>
> Le richieste dei cluster daranno esito negativo se dimentichi il carattere jolly * alla fine del nome segreto quando imposti le autorizzazioni di Secrets Manager. Il carattere jolly rappresenta le versioni segrete.
> È inoltre necessario limitare l'ambito della AWS Secrets Manager policy ai soli certificati necessari al cluster per il provisioning delle istanze.

Creazione della configurazione di sicurezza di Amazon EMR per l'integrazione LDAP

Prima di avviare un cluster EMR con integrazione LDAP, segui i passaggi riportati Crea una configurazione di sicurezza con la console Amazon EMR o con AWS CLI per creare una configurazione di sicurezza Amazon EMR per il cluster. Completa le seguenti configurazioni nel blocco `LDAPConfiguration` in `AuthenticationConfiguration` o nei campi corrispondenti nella sezione Configurazioni di sicurezza della console Amazon EMR:

`EnableLDAPAuthentication`

Opzione console: Protocollo di autenticazione: LDAP

Per utilizzare l'integrazione LDAP, imposta questa opzione su `true` o selezionala come protocollo di autenticazione quando crei un cluster nella console. Per impostazione predefinita, `EnableLDAPAuthentication` è `true` quando crei una configurazione di sicurezza nella console Amazon EMR.

`LDAPServerURL`

Opzione console: posizione del server LDAP

La posizione del server LDAP, incluso il prefisso: `ldaps://`*`location_of_server`*.

`BindCertificateARN`

Opzione console: certificato SSL LDAP

L' AWS Secrets Manager ARN che contiene il certificato per firmare il certificato SSL utilizzato dal server LDAP. Se il server LDAP è firmato da un'autorità di certificazione (CA) pubblica, puoi fornire un AWS Secrets Manager ARN con un file vuoto. Per ulteriori informazioni su come archiviare il certificato in Secrets Manager, consulta Archiviazione dei certificati TLS in AWS Secrets Manager.

`BindCredentialsARN`

Opzione console: credenziali di associazione del server LDAP

Un AWS Secrets Manager ARN che contiene le credenziali di associazione dell'utente amministratore LDAP. Le credenziali sono archiviate come oggetto JSON. Esiste una sola coppia chiave-valore in questo segreto; la chiave nella coppia è il nome utente e il valore è la password. Ad esempio, `{"uid=admin,cn=People,dc=example,dc=com": "AdminPassword1"}`. Questo campo è facoltativo, a meno che non abiliti l'accesso SSH per il cluster EMR. In molte configurazioni, le istanze di Active Directory richiedono credenziali di associazione per consentire a SSSD di sincronizzare gli utenti.

`LDAPAccessFilter`

Opzione console: filtro di accesso LDAP

Specifica il sottoinsieme di oggetti all'interno del server LDAP che possono effettuare l'autenticazione. Ad esempio, se si desidera concedere l'accesso a tutti gli utenti con la classe di oggetti `posixAccount` nel server LDAP, il filtro di accesso viene definito come `(objectClass=posixAccount)`.

`LDAPUserSearchBase`

Opzione console: base di ricerca utenti LDAP

La base di ricerca a cui appartengono gli utenti all'interno del server LDAP. Ad esempio, `cn=People,dc=example,dc=com`.

`LDAPGroupSearchBase`

Opzione console: base di ricerca per gruppi LDAP

La base di ricerca a cui appartengono i gruppi all'interno del server LDAP. Ad esempio, `cn=Groups,dc=example,dc=com`.

`EnableSSHLogin`

Opzione console: accesso SSH

Specifica se consentire o meno l'autenticazione tramite password con credenziali LDAP. Ti consigliamo di abilitare questa opzione. Le coppie di chiavi rappresentano un percorso più sicuro per consentire l'accesso ai cluster EMR. Questo campo è facoltativo e, per impostazione predefinita, corrisponde a `false`.

`LDAPServerType`

Opzione console: tipo di server LDAP

Specifica il tipo di server LDAP a cui si connette Amazon EMR. Le opzioni supportate sono Active Directory e OpenLDAP. Altri tipi di server LDAP potrebbero funzionare, ma Amazon EMR non supporta ufficialmente altri tipi di server. Per ulteriori informazioni, consulta Componenti LDAP per Amazon EMR.

`ActiveDirectoryConfigurations`

Un blocco secondario necessario per le configurazioni di sicurezza che utilizzano il tipo di server Active Directory.

`ADDomain`

Opzione console: dominio Active Directory

Il nome di dominio utilizzato per creare lo User Principal Name (UPN) per l'autenticazione degli utenti con configurazioni di sicurezza che utilizzano il tipo di server Active Directory.

Considerazioni sulle configurazioni di sicurezza con LDAP e Amazon EMR

- Per creare una configurazione di sicurezza con l'integrazione LDAP di Amazon EMR, devi utilizzare la crittografia in transito. Per informazioni sulla crittografia in transito, consulta Crittografa i dati inattivi e in transito con Amazon EMR.

- Non è possibile definire la configurazione Kerberos nella stessa configurazione di sicurezza. Amazon EMR fornisce un KDC dedicato all'automazione e gestisce la password di amministratore per questo KDC. Gli utenti non possono accedere alla password di amministratore.
- Non è possibile definire ruoli di runtime IAM AWS Lake Formation nella stessa configurazione di sicurezza.
- Nel valore dell'`LDAPServerURL` deve essere compreso il protocollo `ldaps://`.
- Il `LDAPAccessFilter` non può essere vuoto.

Utilizzo di LDAP con l'integrazione di Apache Ranger per Amazon EMR

Con l'integrazione LDAP per Amazon EMR, puoi integrare ulteriormente con Apache Ranger. Quando trasferisci gli utenti LDAP in Ranger, puoi associare questi utenti a un server di policy Apache Ranger per eseguire l'integrazione con Amazon EMR e altre applicazioni. Per fare ciò, definisci il campo `RangerConfiguration` all'interno di `AuthorizationConfiguration` nella configurazione di sicurezza che usi con il cluster LDAP. Per ulteriori informazioni su come impostare la configurazione di sicurezza, consulta Creazione di una configurazione di sicurezza EMR.

Quando usi LDAP con Amazon EMR, non devi necessariamente fornire una `KerberosConfiguration` con l'integrazione Amazon EMR per Apache Ranger.

Avvio di un cluster EMR che si autentica con LDAP

Utilizza i seguenti passaggi per avviare un cluster EMR con LDAP o Active Directory.

1. Configurazione dell'ambiente:

   - Assicurati che i nodi del tuo cluster EMR possano comunicare con Amazon S3 e. AWS Secrets Manager Per ulteriori informazioni su come modificare il ruolo del profilo dell'EC2 istanza per comunicare con questi servizi, consulta. Aggiungi AWS Secrets Manager autorizzazioni al ruolo dell'istanza Amazon EMR
   - Se prevedi di eseguire il tuo cluster EMR in una sottorete privata, dovresti utilizzare uno degli endpoint AWS PrivateLink Amazon VPC o utilizzare la traduzione degli indirizzi di rete (NAT) per configurare il VPC per comunicare con S3 e Secrets Manager. Per ulteriori informazioni, consulta Endpoint AWS PrivateLink e VPC e Istanze NAT nella Guida alle operazioni di base di Amazon VPC.
   - Assicurati che ci sia connettività di rete tra il tuo cluster Amazon EMR e il server LDAP. I cluster EMR devono accedere al server LDAP tramite la rete. I nodi primario, principale e

attività per il cluster comunicano con il server LDAP per la sincronizzazione dei dati degli utenti. Se il tuo server LDAP funziona su Amazon EC2, aggiorna il gruppo EC2 di sicurezza per accettare il traffico dal cluster EMR. Per ulteriori informazioni, consulta Aggiungi AWS Secrets Manager autorizzazioni al ruolo dell'istanza Amazon EMR.

2. Crea una configurazione di sicurezza di Amazon EMR per l'integrazione LDAP. Per ulteriori informazioni, consulta Creazione della configurazione di sicurezza di Amazon EMR per l'integrazione LDAP.
3. Ora che hai eseguito la configurazione, segui i passaggi indicati in Avvio di un cluster Amazon EMR per avviare il cluster con le seguenti impostazioni:

   - Seleziona il rilascio 6.12 o successivo di Amazon EMR. Consigliamo di utilizzare l'ultimo rilascio di Amazon EMR.
   - Specifica o seleziona solo le applicazioni per il cluster che supportano LDAP. Per un elenco di applicazioni supportate da LDAP con Amazon EMR, consulta Supporto e considerazioni sulle applicazioni con LDAP per Amazon EMR.
   - Applica la configurazione di sicurezza che hai creato nel passaggio precedente.

Esempi di utilizzo di LDAP con Amazon EMR

Dopo aver effettuato il provisioning di un cluster EMR che utilizza l'integrazione LDAP, puoi fornire le credenziali LDAP a qualsiasi applicazione supportata tramite il meccanismo di autenticazione integrato di nome utente e password. In questa pagina sono riportati alcuni esempi.

Utilizzo dell'autenticazione LDAP con Apache Hive

Example - Apache Hive

Il seguente comando di esempio avvia una sessione di Apache Hive tramite HiveServer 2 e Beeline:

```
beeline -u "jdbc:hive2://$HOSTNAME:10000/default;ssl=true;sslTrustStore=
$TRUSTSTORE_PATH;trustStorePassword=$TRUSTSTORE_PASS"  -n LDAP_USERNAME -
p LDAP_PASSWORD
```

Utilizzo dell'autenticazione LDAP con Apache Livy

Example - Apache Livy

Il comando di esempio seguente avvia una sessione Livy tramite cURL. Sostituisci *ENCODED-KEYPAIR* con una stringa codificata in Base64 per `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
```

Utilizzo dell'autenticazione LDAP con Presto

Example - Presto

Il seguente comando di esempio avvia una sessione Presto tramite la CLI Presto:

```
presto-cli --user "LDAP_USERNAME" --password --catalog hive
```

Dopo aver eseguito questo comando, quando richiesto, inserisci la password LDAP.

Utilizzo dell'autenticazione LDAP con Trino

Example - Trino

Il seguente comando di esempio avvia una sessione Trino tramite la CLI di Trino:

```
trino-cli --user "LDAP_USERNAME" --password --catalog hive
```

Dopo aver eseguito questo comando, quando richiesto, inserisci la password LDAP.

Utilizzo dell'autenticazione LDAP con Hue

Puoi accedere all'interfaccia utente di Hue tramite un tunnel SSH creato sul cluster oppure puoi impostare un server proxy per trasmettere pubblicamente la connessione a Hue. Poiché Hue non funziona in modalità HTTPS per impostazione predefinita, ti consigliamo di utilizzare un livello di crittografia aggiuntivo per garantire che la comunicazione tra i client e l'interfaccia utente di Hue sia crittografata con HTTPS. In questo modo si riduce la possibilità di esporre accidentalmente le credenziali utente in formato di testo normale.

Per utilizzare l'interfaccia utente di Hue, apri l'interfaccia utente di Hue nel browser e inserisci nome utente e password LDAP per accedere. Se le credenziali sono corrette, Hue effettua l'accesso e utilizza la tua identità per eseguire l'autenticazione con tutte le applicazioni supportate.

Utilizzo di SSH per l'autenticazione tramite password e ticket Kerberos per altre applicazioni

> **Important**
>
> Non è consigliabile utilizzare l'autenticazione tramite password su SSH in un cluster EMR.

È possibile utilizzare le credenziali LDAP per accedere a un cluster EMR tramite SSH. A tale scopo, imposta la configurazione di `EnableSSHLogin` su `true` nella configurazione di sicurezza di Amazon EMR utilizzata per avviare il cluster. Quindi, usa il seguente comando per accedere al cluster tramite SSH una volta avviato:

```
ssh username@EMR_PRIMARY_DNS_NAME
```

Dopo aver eseguito questo comando, quando richiesto, inserisci la password LDAP.

Amazon EMR include uno script su cluster che consente agli utenti di generare un file keytab Kerberos e un ticket da utilizzare con le applicazioni supportate che non accettano direttamente le credenziali LDAP. Alcune di queste applicazioni includono Spark `spark-submit` SQL e. PySpark

Esegui `ldap-kinit` e segui le istruzioni. Se l'autenticazione ha esito positivo, il file keytab Kerberos viene visualizzato nella directory home con un ticket Kerberos valido. Utilizza il ticket Kerberos per eseguire le applicazioni come faresti in qualsiasi ambiente con Kerberos.

Integra Amazon EMR con AWS IAM Identity Center

Con le versioni 6.15.0 e successive di Amazon EMR, puoi utilizzare identità da per AWS IAM Identity Center autenticarti con un cluster Amazon EMR. Le sezioni seguenti forniscono una panoramica concettuale, i prerequisiti e le fasi necessari per avviare un cluster EMR con l'integrazione del Centro identità.

Argomenti

- Panoramica
- Funzionalità e vantaggi
- Guida introduttiva all' AWS IAM Identity Center integrazione per Amazon EMR
- Considerazioni e limitazioni per Amazon EMR con l'integrazione del Centro identità

Panoramica

Identity Center è l'approccio consigliato per l'autenticazione e l'autorizzazione della forza lavoro per organizzazioni di qualsiasi dimensione e tipo. AWS Con Identity Center, puoi creare e gestire le identità degli utenti o connettere la tua fonte di identità esistente, tra cui Microsoft Active Directory, Okta, Ping Identity JumpCloud, Google Workspace e Microsoft Entra ID (precedentemente Azure AD). AWS

La propagazione affidabile delle identità è una AWS IAM Identity Center funzionalità che gli amministratori di Connected Servizi AWS possono utilizzare per concedere e controllare l'accesso ai dati del servizio. L'accesso a questi dati si basa su attributi utente come le associazioni di gruppo. La configurazione di una propagazione affidabile delle identità richiede la collaborazione tra gli amministratori di connected Servizi AWS e gli amministratori di IAM Identity Center. Per ulteriori informazioni, consulta Prerequisiti e considerazioni.

Funzionalità e vantaggi

L'integrazione di Amazon EMR con il Centro identità IAM offre i seguenti vantaggi:

- Amazon EMR fornisce le credenziali per inoltrare l'identità del tuo Centro identità a un cluster EMR.
- Amazon EMR configura tutte le applicazioni supportate per l'autenticazione con le credenziali del cluster.
- Amazon EMR configura e mantiene la sicurezza per le applicazioni supportate con il protocollo Kerberos senza la necessità di tuoi comandi o script.
- La capacità di applicare l'autorizzazione a livello di prefisso di Amazon S3 con le identità del Centro identità sui prefissi S3 gestiti da S3 Access Grants.
- La capacità di applicare l'autorizzazione a livello di tabella con le identità di Identity Center sulle tabelle Glue gestite AWS Lake Formation . AWS

Guida introduttiva all' AWS IAM Identity Center integrazione per Amazon EMR

Questa sezione ti aiuta a configurare Amazon EMR per l'integrazione con. AWS IAM Identity Center

Argomenti

- Crea un'istanza del Centro identità

- Crea un ruolo IAM per il Centro identità
- Aggiungi le autorizzazioni per i servizi non integrati con IAM Identity Center
- Crea una configurazione di sicurezza abilitata per il Centro identità
- Crea e avvia un cluster abilitato per il Centro identità
- Configura Lake Formation per un cluster EMR abilitato per il Centro identità IAM
- Utilizzo di S3 Access Grants su un cluster EMR abilitato per il Centro identità IAM

> **Note**
>
> Per utilizzare Identity Center, è necessario abilitare l'integrazione con EMR, Lake Formation o S3 Access Grants. Puoi anche usarli entrambi. Se nessuno dei due è abilitato, l'integrazione con Identity Center non è supportata.

Crea un'istanza del Centro identità

Se non ne hai ancora una, crea un'istanza del Centro identità nella Regione AWS in cui desideri avviare il cluster EMR. Un'istanza del Centro identità può esistere solo in una singola regione per un Account AWS.

Utilizzate il AWS CLI comando seguente per creare una nuova istanza denominata*MyInstance*:

```
aws sso-admin create-instance --name MyInstance
```

Crea un ruolo IAM per il Centro identità

Con cui integrare Amazon EMR AWS IAM Identity Center, crea un ruolo IAM che si autentichi con Identity Center dal cluster EMR. Dietro le quinte, Amazon EMR utilizza le credenziali SigV4 per inoltrare l'identità del Centro identità a servizi downstream come AWS Lake Formation. Il tuo ruolo dovrebbe avere anche le autorizzazioni necessarie per richiamare i servizi downstream.

Quando crei il ruolo, utilizza la seguente policy di autorizzazione:

```
{
  "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 policy di attendibilità per questo ruolo consente al ruolo InstanceProfile di assumere quel ruolo.

```
{
    "Sid": "AssumeRole",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::12345678912:role/EMR_EC2_DefaultRole"
    },
    "Action": [
        "sts:AssumeRole",
        "sts:SetContext"
    ]
}
```

Se il ruolo non dispone di credenziali affidabili e accede a una tabella protetta da Lake Formation, Amazon EMR imposta `principalId` automaticamente il ruolo assunto su. *`userID`*`-untrusted` Di seguito è riportato un frammento di un evento che mostra il. 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"
        ...
```

Aggiungi le autorizzazioni per i servizi non integrati con IAM Identity Center

AWS credenziali che utilizzano Trusted Identity Propagation, le politiche IAM definite nel ruolo IAM per tutte le chiamate effettuate a servizi non integrati con IAM Identity Center. Ciò include, ad esempio, il. AWS Key Management Service Il tuo ruolo dovrebbe anche definire eventuali autorizzazioni IAM per tutti i servizi a cui tenteresti di accedere. I servizi integrati IAM Identity Center attualmente supportati includono AWS Lake Formation Amazon S3 Access Grants.

Per ulteriori informazioni su Trusted Identity Propagation, consulta Trusted Identity Propagation tra le applicazioni.

Crea una configurazione di sicurezza abilitata per il Centro identità

Per avviare un cluster EMR con l'integrazione del Centro identità IAM, utilizza il seguente comando di esempio per creare una configurazione di sicurezza Amazon EMR con il Centro identità abilitato. Ogni configurazione è spiegata di seguito.

```
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`**: (obbligatorio) abilita l'integrazione del Centro identità.
- **`IdentityCenterInstanceARN`**— (opzionale) L'ARN dell'istanza di Identity Center. Se questo non è incluso, l'ARN dell'istanza IAM Identity Center esistente viene cercato come parte della fase di configurazione.
- **`IAMRoleForEMRIdentityCenterApplicationARN`**: (obbligatorio) il ruolo IAM che procura i token del Centro identità dal cluster.
- **`IdentityCenterApplicationAssignmentRequired`** : (booleano) determina se sarà richiesta un'assegnazione per utilizzare l'applicazione del Centro identità. Questo campo è facoltativo. Se non viene fornito un valore, il valore predefinito è`false`.
- **`AuthorizationConfiguration`**/**`LakeFormationConfiguration`**— Facoltativamente, configura l'autorizzazione:
  - **`IAMConfiguration`**— Consente l'utilizzo della funzionalità EMR Runtimes Roles in aggiunta all'identità TIP. Se abiliti questa configurazione, a te (o al AWS servizio chiamante) verrà richiesto di specificare un IAM Runtime Role in ogni chiamata a EMR Steps o EMR. `GetClusterSessionCredentials` APIs Se il cluster EMR viene utilizzato con SageMaker Unified Studio, questa opzione è necessaria se è abilitata anche la Trusted Identity Propagation.
  - **`EnableLakeFormation`**: abilita l'autorizzazione di Lake Formation sul cluster.

Per abilitare l'integrazione del Centro identità con Amazon EMR, devi specificare `EncryptionConfiguration` e. `IntransitEncryptionConfiguration`

Crea e avvia un cluster abilitato per il Centro identità

Ora che hai configurato il ruolo IAM che esegue l'autenticazione con il Centro identità e hai creato una configurazione di sicurezza Amazon EMR con il Centro identità abilitato, puoi creare e avviare il tuo cluster con riconoscimento dell'identità. Per i passaggi per avviare il cluster con la configurazione di sicurezza richiesta, consulta Specificare una configurazione di sicurezza per un cluster Amazon EMR.

Le seguenti sezioni descrivono come configurare un cluster abilitato per Identity Center con le opzioni di sicurezza supportate da Amazon EMR:

- Utilizzo di S3 Access Grants su un cluster EMR abilitato per il Centro identità IAM
- Configura Lake Formation per un cluster EMR abilitato per il Centro identità IAM

Configura Lake Formation per un cluster EMR abilitato per il Centro identità IAM

Puoi integrarti AWS Lake Formationcon il tuo cluster EMR AWS IAM Identity Center abilitato.

Innanzitutto, assicurati di avere un'istanza del Centro identità configurata nella stessa regione del cluster. Per ulteriori informazioni, consulta Crea un'istanza del Centro identità. Puoi trovare l'ARN dell'istanza nella console del Centro identità IAM quando visualizzi i dettagli dell'istanza o utilizzare il seguente comando per visualizzare i dettagli di tutte le istanze dalla CLI:

```
aws sso-admin list-instances
```

Quindi usa l'ARN e l'ID del tuo AWS account con il seguente comando per configurare Lake Formation in modo che sia compatibile 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"
}
```

Ora, chiama `put-data-lake-settings` e abilita `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
    }
}
```

Infine, concedi le autorizzazioni complete per la tabella all'ARN dell'identità per l'utente che accede al cluster EMR. L'ARN contiene l'ID utente del Centro identità. Vai al Centro identità nella console, seleziona Utenti e poi l'utente per visualizzarne le impostazioni delle Informazioni generali.

Copia l'ID utente e incollalo nel seguente ARN per *user-id*:

```
arn:aws:identitystore:::user/user-id
```

> **Note**
>
> Le query sul cluster EMR funzionano solo se l'identità del Centro identità IAM ha accesso completo alla tabella protetta di Lake Formation. Se l'identità non ha accesso completo alla tabella, la query avrà esito negativo.

Per concedere all'utente l'accesso completo alla tabella, utilizza il seguente comando:

```
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"
    ]
}
```

Aggiungere l'applicazione ARN all'integrazione di IDC for Lake Formation

Per interrogare le risorse abilitate a Lake Formation, è necessario aggiungere l'ARN dell'applicazione IDC. A tale scopo, seguire queste fasi:

1. Sulla console, scegli AWS Lake Formation.
2. Seleziona l'integrazione con IAM Identity Center e l'integrazione dell'applicazione Lake Formation abbinando l'ARN dell'applicazione. L'ARN verrà visualizzato nell'elenco degli ID dell'applicazione.

Utilizzo di S3 Access Grants su un cluster EMR abilitato per il Centro identità IAM

Puoi integrare S3 Access Grants con il tuo cluster AWS IAM Identity Center EMR abilitato.

Utilizza S3 Access Grants per autorizzare l'accesso ai tuoi set di dati dai cluster che utilizzano il Centro identità. Crea autorizzazioni per aumentare quelle impostate per utenti, gruppi, ruoli IAM o per una directory aziendale. Per ulteriori informazioni, consulta Utilizzo di S3 Access Grants con Amazon EMR.

Argomenti

- Crea un'istanza e una posizione S3 Access Grants
- Crea autorizzazioni per le identità del Centro identità

Crea un'istanza e una posizione S3 Access Grants

Se non ne hai ancora una, crea un'istanza S3 Access Grants nella Regione AWS in cui desideri avviare il cluster EMR.

Usa il seguente AWS CLI comando per creare una nuova istanza denominata: *MyInstance*

```
aws s3control-access-grants create-access-grants-instance \
--account-id 12345678912 \
--identity-center-arn "identity-center-instance-arn" \
```

Quindi, crea una posizione S3 Access Grants, sostituendo i valori rossi con:

```
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**
>
> Definisci il parametro `iam-role-arn` come ARN `accessGrantRole`.

Crea autorizzazioni per le identità del Centro identità

Infine, crea le autorizzazioni per le identità che hanno accesso al tuo cluster:

```

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"
```

Output di esempio:

```
{
"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/*"
}
```

Considerazioni e limitazioni per Amazon EMR con l'integrazione del Centro identità

Considera i seguenti punti quando utilizzi il Centro identità IAM con Amazon EMR:

- La propagazione affidabile delle identità tramite Identity Center è supportata su Amazon EMR 6.15.0 e versioni successive e solo con Apache Spark. Inoltre, la Trusted Identity Propagation tramite Identity Center utilizzando la funzionalità EMR Runtime Roles è supportata su Amazon EMR 7.8.0 e versioni successive e solo con Apache Spark.
- Per abilitare i cluster EMR con propagazione di identità affidabile, è necessario utilizzare AWS CLI per creare una configurazione di sicurezza con propagazione delle identità affidabile abilitata e utilizzare tale configurazione di sicurezza all'avvio del cluster. Per ulteriori informazioni, consulta Crea una configurazione di sicurezza abilitata per il Centro identità.
- I controlli di accesso granulari AWS Lake Formation che utilizzano Trusted Identity Propagation sono disponibili per i cluster Amazon EMR nella versione 7.2.0 e successive di EMR. Tra le versioni EMR 6.15.0 e 7.1.0, è disponibile solo il controllo degli accessi a livello di tabella, basato su Lake Formation. AWS
- Con i cluster Amazon EMR che utilizzano Trusted Identity Propagation, le operazioni che supportano il controllo degli accessi basato su Lake Formation con Apache Spark includono SELECT, ALTER TABLE, INSERT INTO e DROP TABLE.
- La propagazione affidabile delle identità con Amazon EMR è supportata nei seguenti paesi: Regioni AWS
  - `af-south-1`: Africa (Città del Capo)
  - `ap-east-1`: Asia Pacifico (Hong Kong)
  - `ap-northeast-1`: Asia Pacifico (Tokyo)
  - `ap-northeast-2`: Asia Pacifico (Seoul)
  - `ap-northeast-3`: Asia Pacifico (Osaka-Locale)

- `ap-south-1`: Asia Pacifico (Mumbai)
- `ap-south-2`— Asia Pacifico (Hyderabad)
- `ap-southeast-1`: Asia Pacifico (Singapore)
- `ap-southeast-2`: Asia Pacifico (Sydney)
- `ap-southeast-3`: Asia Pacifico (Giacarta)
- `ap-southeast-4`— Asia Pacifico (Melbourne)
- `ca-central-1`: Canada (Centrale)
- `eu-central-1`: Europa (Francoforte)
- `eu-central-2`— Europa (Zurigo)
- `eu-north-1`: Europa (Stoccolma)
- `eu-south-1`: Europa (Milano)
- `eu-south-2`— Europa (Spagna)
- `eu-west-1`: Europa (Irlanda)
- `eu-west-2`: Europa (Londra)
- `eu-west-3`: Europa (Parigi)
- `il-central-1`— Israele (Tel Aviv)
- `me-central-1`— Medio Oriente (Emirati Arabi Uniti)
- `me-south-1`: Medio Oriente (Bahrein)
- `sa-east-1`: Sud America (San Paolo)
- `us-east-1`: Stati Uniti orientali (Virginia settentrionale)
- `us-east-2`: Stati Uniti orientali (Ohio)
- `us-west-1`: Stati Uniti occidentali (California settentrionale)
- `us-west-2`: Stati Uniti occidentali (Oregon)

Integra Amazon EMR con AWS Lake Formation

AWS Lake Formation è un servizio gestito che ti aiuta a scoprire, catalogare, pulire e proteggere i dati in un data lake Amazon Simple Storage Service (S3). Lake Formation fornisce un accesso granulare a livello di colonna a database e tabelle nel Glue Data Catalog. AWS Per ulteriori informazioni, consulta What is AWS Lake Formation?

Con Amazon EMR versione 6.7.0 e successive, puoi applicare il controllo degli accessi basato su Lake Formation ai processi Spark, Hive e Presto inviati ai cluster Amazon EMR. Per l'integrazione con Lake Formation, devi creare un cluster EMR con un ruolo di runtime. Un ruolo di runtime è un ruolo AWS Identity and Access Management (IAM) che puoi associare ai processi o alle query di Amazon EMR. Amazon EMR utilizza quindi questo ruolo per accedere AWS alle risorse. Per ulteriori informazioni, consulta Ruoli di runtime per le fasi di Amazon EMR.

Funzionamento di Amazon EMR con Lake Formation

Dopo aver integrato Amazon EMR con Lake Formation, puoi eseguire query sui cluster Amazon EMR con l'`Step`API o con AI Studio. SageMaker Quindi, Lake Formation fornisce l'accesso ai dati tramite credenziali temporanee per Amazon EMR. Questo processo è denominato distribuzione di credenziali. Per ulteriori informazioni, consulta What is AWS Lake Formation?

Di seguito è riportata una panoramica generale sul modo in cui Amazon EMR ottiene l'accesso ai dati protetti dalle policy di sicurezza Lake Formation.

1. Un utente invia una query Amazon EMR per i dati in Lake Formation.
2. Amazon EMR richiede le credenziali temporanee da Lake Formation per consentire all'utente di accedere ai dati.
3. Lake Formation restituisce le credenziali temporanee.
4. Amazon EMR invia la richiesta di query per recuperare dati da Amazon S3.
5. Amazon EMR riceve i dati da Amazon S3, li filtra e restituisce i risultati in base alle autorizzazioni utente definite in Lake Formation.

Per ulteriori informazioni sull'aggiunta di utenti e gruppi ai policy di Lake Formation, consulta Concessione delle autorizzazioni Data Catalog.

Prerequisiti

Prima di integrare Amazon EMR e Lake Formation, è necessario soddisfare i seguenti requisiti:

- Attiva l'autorizzazione dei ruoli di runtime sul cluster Amazon EMR.
- Usa il AWS Glue Data Catalog come archivio di metadati.
- Definisci e gestisci le autorizzazioni in Lake Formation per accedere a database, tabelle e colonne in AWS Glue Data Catalog. Per ulteriori informazioni, consulta What is AWS Lake Formation?

Argomenti

- Abilitazione di Amazon EMR con Lake Formation
- Apache Hudi e Lake Formation con Amazon EMR
- Apache Iceberg e Lake Formation con Amazon EMR
- Delta Lake e Lake Formation su Amazon EMR
- Considerazioni su Amazon EMR con Lake Formation
- Utilizzo delle viste del catalogo dati di Glue in Amazon EMR
- Utilizzo di un ruolo di runtime del job Lake Formation con accesso completo alla tabella

Abilitazione di Amazon EMR con Lake Formation

Con Amazon EMR 6.15.0 e versioni successive, quando esegui job Spark su Amazon EMR EC2 su cluster che accedono ai dati nel AWS Glue Data Catalog, AWS Lake Formation puoi utilizzare per applicare autorizzazioni a livello di tabella, riga, colonna e cella su tabelle basate su Hudi, Iceberg o Delta Lake.

Questa sezione illustra come creare una configurazione di sicurezza e impostare Lake Formation per l'utilizzo con Amazon EMR. Descrive inoltre come avviare un cluster con la configurazione di sicurezza creata per Lake Formation.

Passaggio 1: Impostazione di un ruolo di runtime per il cluster EMR

Per utilizzare un ruolo di runtime per il cluster EMR, devi prima creare una configurazione di sicurezza. Con una configurazione di sicurezza puoi applicare opzioni di sicurezza, autorizzazione e autenticazione coerenti in tutti i tuoi cluster.

1. Crea un file denominato `lf-runtime-roles-sec-cfg.json` con la seguente configurazione di sicurezza.

```
{
    "AuthorizationConfiguration": {
        "IAMConfiguration": {
            "EnableApplicationScopedIAMRole": true,
            "ApplicationScopedIAMRoleConfiguration": {
                "PropagateSourceIdentity": true
            }
        },
        "LakeFormationConfiguration": {
```

```
            "AuthorizedSessionTagValue": "Amazon EMR"
        }
    },
    "EncryptionConfiguration": {
     "EnableAtRestEncryption": false,
            "EnableInTransitEncryption": true,
            "InTransitEncryptionConfiguration": {
            "TLSCertificateConfiguration": {<certificate-configuration>}
        }
    }
}
```

L'esempio seguente illustra come utilizzare un file zip con certificati in Amazon S3 per la configurazione dei certificati:

- Un file zip con certificati in Amazon S3 viene utilizzato come fornitore di chiavi. (Vedi Fornitura di certificati per la crittografia di dati in transito con Amazon EMR per i requisiti dei certificati).

```
"TLSCertificateConfiguration": {
 "CertificateProviderType": "PEM",
 "S3Object": "s3://MyConfigStore/artifacts/MyCerts.zip"
 }
```

L'esempio seguente illustra come utilizzare un provider di chiavi personalizzate per la configurazione dei certificati:

- Viene utilizzato un provider di chiavi personalizzato. (Vedi Fornitura di certificati per la crittografia di dati in transito con Amazon EMR per i requisiti del certificato).

```
"TLSCertificateConfiguration": {
 "CertificateProviderType": "Custom",
 "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
 "CertificateProviderClass": "com.mycompany.MyCertProvider"
                         }
```

2. Quindi, per assicurarti che il tag di sessione possa autorizzare Lake Formation, imposta la proprietà `LakeFormationConfiguration/AuthorizedSessionTagValue` su `Amazon EMR`.
3. Utilizza il seguente comando per creare una configurazione di sicurezza Amazon EMR.

```
aws emr create-security-configuration \
--name 'iamconfig-with-iam-lf' \
--security-configuration file://lf-runtime-roles-sec-cfg.json
```

In alternativa, puoi utilizzare la console Amazon EMR per creare una configurazione di sicurezza con impostazioni personalizzate.

Fase 2: avvio di un cluster Amazon EMR

Ora puoi avviare un cluster EMR con la configurazione di sicurezza creata nel passaggio precedente. Per ulteriori informazioni sulle configurazioni di sicurezza, consulta le sezioni Usa le configurazioni di sicurezza per configurare la sicurezza dei cluster Amazon EMR e Ruoli di runtime per le fasi di Amazon EMR.

Passaggio 3a: configura le autorizzazioni a livello di tabella basate su Lake Formation con i ruoli di runtime di Amazon EMR

Se non hai bisogno di un controllo granulare degli accessi a livello di colonna, riga o cella, puoi impostare le autorizzazioni a livello di tabella con Catalogo dati Glue. Per abilitare l'accesso a livello di tabella, vai alla AWS Lake Formation console e seleziona l'opzione Impostazioni di integrazione delle applicazioni dalla sezione Amministrazione nella barra laterale. Quindi, abilita la seguente opzione e scegli Salva:

Consenti ai motori esterni di accedere ai dati nelle posizioni Amazon S3 con accesso completo alla tabella

Passaggio 3a: imposta le autorizzazioni a livello di colonna, riga o cella basate su Lake Formation con i ruoli di runtime di Amazon EMR

Per applicare le autorizzazioni a livello di tabella e colonna con Lake Formation, l'amministratore del data lake per Lake Formation deve impostare `Amazon EMR` come valore per la configurazione del tag di sessione, `AuthorizedSessionTagValue`. Lake Formation utilizza questo tag di sessione per autorizzare i chiamanti e fornire l'accesso al data lake. Puoi impostare questo tag di sessione nella sezione External data filtering (Filtraggio dati esterni) della console Lake Formation. Sostituiscila *123456789012* con il tuo ID. Account AWS

Fase 4: Configurazione delle sovvenzioni AWS Glue and Lake Formation per i ruoli di runtime di Amazon EMR

Per continuare con la configurazione del controllo degli accessi basato su Lake Formation con i ruoli di runtime di Amazon EMR, devi configurare le sovvenzioni AWS Glue and Lake Formation per i ruoli di runtime di Amazon EMR. Per permettere ai ruoli di runtime IAM di interagire con Lake Formation, concedi loro l'accesso con `lakeformation:GetDataAccess` e `glue:Get*`.

Le autorizzazioni di Lake Formation controllano l'accesso alle risorse di AWS Glue Data Catalog, alle sedi Amazon S3 e ai dati sottostanti in tali sedi. Le autorizzazioni IAM controllano l'accesso a Lake Formation and AWS Glue APIs e alle risorse. Anche se disponi dell'autorizzazione di Lake Formation per accedere a una tabella nel catalogo dati (SELECT), l'operazione fallisce se non disponi dell'autorizzazione IAM per l'API `glue:Get*`. Per maggiori dettagli sul controllo degli accessi in Lake Formation, consulta la sezione Lake Formation access control overview (Panoramica del controllo degli accessi a Lake Formation).

1. Crea il file `emr-runtime-roles-lake-formation-policy.json` con i seguenti contenuti.

   JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationManagedAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:Get*",
        "glue:Create*",
        "glue:Update*"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

2. Crea la policy IAM correlata.

```
aws iam create-policy \
--policy-name emr-runtime-roles-lake-formation-policy \
--policy-document file://emr-runtime-roles-lake-formation-policy.json
```

3. Per assegnare questa policy ai ruoli di runtime IAM, segui i passaggi in Managing AWS Lake Formation permissions.

Ora puoi utilizzare i ruoli di runtime e Lake Formation per applicare le autorizzazioni a livello di tabella e colonna. Puoi anche utilizzare un'identità di origine per controllare le azioni e monitorare le operazioni con AWS CloudTrail.

Per ogni ruolo IAM che intendi utilizzare come ruolo di runtime, imposta la seguente policy di attendibilità, sostituendo `EMR_EC2_DefaultRole` con il ruolo del profilo dell'istanza. Per modificare la policy di attendibilità di un ruolo IAM, consulta Modifica di una policy di attendibilità del ruolo.

```
{
   "Sid":"AllowAssumeRole",
   "Effect":"Allow",
   "Principal":{
     "AWS":"arn:aws:iam::<AWS_ACCOUNT_ID>:role/EMR_EC2_DefaultRole"
   },
   "Action":[
        "sts:AssumeRole",
        "sts:TagSession"
       ]
 }
```

Per un end-to-end esempio dettagliato, consulta Introducing runtime roles for Amazon EMR steps.

Per informazioni su come integrare Iceberg e AWS Glue Data Catalog per una gerarchia multicatalogo, vedi Configurare Spark per accedere a una gerarchia multicatalogo in Glue Data Catalog. AWS

Apache Hudi e Lake Formation con Amazon EMR

Le versioni 6.15.0 e successive di Amazon EMR includono il supporto per il controllo granulare degli accessi basato su Apache Hudi durante la lettura e la scrittura di dati AWS Lake Formation con Spark SQL. Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con

Apache Hudi. Con questa funzionalità, puoi eseguire query istantanee sulle copy-on-write tabelle per interrogare l'istantanea più recente della tabella in un determinato commit o istante di compattazione.

Attualmente, un cluster Amazon EMR abilitato a Lake Formation deve recuperare la colonna del tempo di commit di Hudi per eseguire query incrementali e query sui viaggi nel tempo. Non supporta la sintassi e la funzione di Spark. `timestamp as of` `Spark.read()` La sintassi corretta è. `select * from table where _hoodie_commit_time <= point_in_time` Per ulteriori informazioni, consulta Point in time Time-Travel queries sulla tabella Hudi.

La seguente matrice di supporto elenca alcune funzioni principali di Apache Hudi con Lake Formation:

	Copia su scrittura	unisci in lettura
Query snapshot: Spark SQL	✓	✓
Query ottimizzate per la lettura: Spark SQL	✓	✓
Query incrementali	✓	✓
Query temporali	✓	✓
Tabelle dei metadati	✓	✓
Comandi **`INSERT`** DML	✓	✓
Comandi DDL		
Query su origini dati Spark		
Scritture di origini dati Spark		

Esecuzione di query per tabelle Hudi

Questa sezione mostra come eseguire le query supportate sopra descritte su un cluster abilitato per Lake Formation. La tabella deve essere una tabella di catalogo registrata.

1. Per avviare la shell (interprete di comandi) Spark, utilizza i comandi seguenti.

```
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
```

Se desideri che Lake Formation utilizzi il server di registrazione per gestire il tuo catalogo Spark, imposta su `spark.sql.catalog.<managed_catalog_name>.lf.managed` true.

2. Per interrogare l'ultima istantanea delle copy-on-write tabelle, usa i seguenti comandi.

```
SELECT * FROM my_hudi_cow_table
```

```
spark.read.table("my_hudi_cow_table")
```

3. Per eseguire query sui dati compatti più recenti delle tabelle `MOR`, puoi eseguire query sulla tabella ottimizzata per la lettura che presenta il suffisso con `_ro`:

```
SELECT * FROM my_hudi_mor_table_ro
```

```
spark.read.table("my_hudi_mor_table_ro")
```

> **Note**
>
> Le prestazioni delle letture possono essere più lente nei cluster Lake Formation a causa di ottimizzazioni non supportate. Queste funzioni includono l'elenco dei file basato sui metadati Hudi e i dati ignorati. È preferibile testare le prestazioni dell'applicazione per accertarsi che soddisfi i requisiti.

Apache Iceberg e Lake Formation con Amazon EMR

Le versioni 6.15.0 e successive di Amazon EMR includono il supporto per il controllo granulare degli accessi basato su Apache Iceberg durante la lettura e la scrittura di dati AWS Lake Formation con Spark SQL. Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con Apache Iceberg. Con questa funzionalità, puoi eseguire query istantanee sulle copy-on-write tabelle per interrogare l'istantanea più recente della tabella in un determinato commit o istante di compattazione.

Se desideri utilizzare il formato Iceberg, imposta le seguenti configurazioni. Sostituisci *DB_LOCATION* con il percorso Amazon S3 in cui si trovano le tabelle Iceberg e sostituisci i segnaposto dell'ID regione e account con i tuoi valori.

```
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
```

Se desideri che Lake Formation utilizzi il server di registrazione per gestire il tuo catalogo Spark, imposta su `spark.sql.catalog.<managed_catalog_name>.lf.managed` true.

Inoltre, fai attenzione a NON passare le seguenti impostazioni per il ruolo di assunzione:

```
--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 seguente matrice di supporto elenca alcune funzionalità principali di Apache Iceberg con Lake Formation:

	Copia su scrittura	unisci in lettura
Query snapshot: Spark SQL	✓	✓
Query ottimizzate per la lettura: Spark SQL	✓	✓
Query incrementali	✓	✓
Query temporali	✓	✓
Tabelle dei metadati	✓	✓
Comandi **`INSERT`** DML	✓	✓
Comandi DDL		
Query su origini dati Spark		
Scritture di origini dati Spark		

Delta Lake e Lake Formation su Amazon EMR

Le versioni 6.15.0 e successive di Amazon EMR includono il supporto per il controllo granulare degli accessi basato su Delta Lake durante la lettura e la scrittura di dati AWS Lake Formation con Spark SQL. Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con Delta Lake. Con questa funzionalità, puoi eseguire query istantanee sulle copy-on-write tabelle per interrogare l'istantanea più recente della tabella in un determinato commit o compattazione.

Per usare Delta Lake con Lake Formation, esegui il seguente 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
```

Se desideri che Lake Formation utilizzi il server di registrazione per gestire il tuo catalogo Spark, imposta su `spark.sql.catalog.<managed_catalog_name>.lf.managed` true.

La seguente matrice di supporto elenca alcune funzioni principali di Delta Lake con Lake Formation:

	Copia su scrittura	unisci in lettura
Query snapshot: Spark SQL	✓	✓
Query ottimizzate per la lettura: Spark SQL	✓	✓
Query incrementali	Non supportato	Non supportato
Query temporali	Non supportato	Non supportato
Tabelle dei metadati	✓	✓
Comandi **`INSERT`** DML	✓	✓
Comandi DDL		
Query su origini dati Spark		
Scritture di origini dati Spark		

Creazione di una tabella Delta Lake in AWS Glue Data Catalog

Amazon EMR con Lake Formation non supporta i comandi DDL e la creazione di tabelle Delta. Segui questi passaggi per creare tabelle nel AWS Glue Data Catalog.

1. Usa l'esempio seguente per creare una tabella Delta. Assicurati che la tua posizione S3 esista.

```
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. Per vedere i dettagli della tua tabella, vai a https://console.aws.amazon.com/glue/.
3. Nella barra di navigazione a sinistra, espandi Data Catalog, scegli Tabelle, quindi scegli la tabella che hai creato. In Schema, dovresti vedere che la tabella Delta che hai creato con Spark memorizza tutte le colonne in un tipo di dati di `array<string>` AWS Glue.
4. Per definire filtri a livello di colonna e cella in Lake Formation, rimuovi la `col` colonna dallo schema, quindi aggiungi le colonne presenti nello schema della tabella. In questo esempio, aggiungi le colonne `xy`, e. `z`

Considerazioni su Amazon EMR con Lake Formation

Considera quanto segue quando usi Amazon EMR con. AWS Lake Formation

Amazon EMR with Lake Formation è disponibile in tutte le regioni disponibili.

- Amazon EMR supporta il controllo granulare degli accessi tramite Lake Formation solo per le tabelle Apache Hive e Apache Iceberg. I formati Apache Hive includono Parquet, ORC e XSv.
- Non puoi smettere di lavorare `DynamicResourceAllocation` per Lake Formation.
- Puoi usare Lake Formation solo con i job Spark.
- Amazon EMR with Lake Formation supporta solo una singola sessione Spark per tutta la durata di un job.
- Amazon EMR with Lake Formation supporta solo query tabellari tra account condivise tramite link a risorse.
- Quanto segue non è supportato:
  - Set di dati distribuiti resilienti (RDD)
  - Streaming Spark
  - Scrivi con le autorizzazioni concesse da Lake Formation
  - Controllo degli accessi per le colonne annidate
- Amazon EMR blocca funzionalità che potrebbero compromettere il completo isolamento dei driver di sistema, tra cui:
  - UDTs, Hive UDFs e qualsiasi funzione definita dall'utente che includa classi personalizzate
  - Origini dati personalizzate
  - Fornitura di vasetti aggiuntivi per l'estensione, il connettore o il metastore Spark

  - Comando `ANALYZE TABLE`
- Per applicare i controlli di accesso `EXPLAIN PLAN` e le operazioni DDL, ad esempio non esporre informazioni riservate`DESCRIBE TABLE`.
- Amazon EMR limita l'accesso ai log Spark dei driver di sistema sulle applicazioni abilitate per Lake Formation. Poiché il driver di sistema viene eseguito con autorizzazioni elevate, gli eventi e i log generati dal driver di sistema possono includere informazioni riservate. Per impedire a utenti o codici non autorizzati di accedere a questi dati sensibili, Amazon EMR disabilita l'accesso ai registri dei driver di sistema.

  I log dei profili di sistema vengono sempre conservati nello storage gestito: si tratta di un'impostazione obbligatoria che non può essere disabilitata. Questi registri vengono archiviati in modo sicuro e crittografati utilizzando una chiave KMS gestita dal cliente o una chiave KMS gestita. AWS

  Se la tua applicazione Amazon EMR si trova in una sottorete privata con endpoint VPC per Amazon S3 e alleghi una policy di endpoint per controllare l'accesso, prima che i tuoi lavori possano inviare dati di log a Managed AWS Amazon S3, devi includere le autorizzazioni dettagliate in Managed Storage nella tua policy VPC all'endpoint gateway S3. Per la risoluzione dei problemi relativi alle richieste, contatta l'assistenza. AWS
- Se hai registrato una posizione in una tabella con Lake Formation, il percorso di accesso ai dati passa attraverso le credenziali archiviate di Lake Formation indipendentemente dall'autorizzazione IAM per il ruolo di job runtime di Amazon EMR. Se configuri erroneamente il ruolo registrato con la posizione della tabella, i lavori inviati che utilizzano il ruolo con l'autorizzazione S3 IAM per la posizione della tabella avranno esito negativo.
- La scrittura su una tabella Lake Formation utilizza l'autorizzazione IAM anziché le autorizzazioni concesse da Lake Formation. Se il tuo ruolo di job runtime dispone delle autorizzazioni S3 necessarie, puoi utilizzarlo per eseguire operazioni di scrittura.

Di seguito sono riportate considerazioni e limitazioni relative all'utilizzo di Apache Iceberg:

- È possibile utilizzare Apache Iceberg solo con il catalogo delle sessioni e non con i cataloghi con nomi arbitrari.
- Le tabelle Iceberg registrate in Lake Formation supportano solo le tabelle di metadati`history`,`metadata_log_entries`,, `snapshots` `filesmanifests`, e. `refs` Amazon EMR nasconde le colonne che potrebbero contenere dati sensibili, ad esempio`partitions`, `path`

e. `summaries` Questa limitazione non si applica alle tabelle Iceberg che non sono registrate in Lake Formation.

- Le tabelle che non vengono registrate in Lake Formation supportano tutte le stored procedure Iceberg. Le `migrate` procedure `register_table` and non sono supportate per nessuna tabella.
- Ti consigliamo di utilizzare Iceberg DataFrameWriter V2 anziché V1.
- EMR 7.10 offre un modo per tornare a RecordServer utilizzare funzionalità supportate da RecordServer, ma non ancora supportate, da FGAC nativo, come il writeback alle tabelle registrate di Lake Formation. Per tornare indietro, specifica le seguenti configurazioni all'avvio del cluster.

```
{
    "Classification": "spark-defaults",
    "Properties": {
      "spark.emr.lakeformation.legacy.enabled": "true"
    }
  },
  {
    "Classification": "yarn-site",
    "Properties": {
      "spark.emr.lakeformation.legacy.enabled": "true"
    }
  }
```

Utilizzo delle viste del catalogo dati di Glue in Amazon EMR

> **Note**
>
> La creazione e AWS la gestione delle viste di Glue Data Catalog da utilizzare con EMR on EC2 è disponibile con Amazon EMR versione 7.10.0 e successive.

È possibile creare e gestire viste nel AWS Glue Data Catalog da utilizzare con EMR attivo. EC2 Queste sono note comunemente come viste del AWS Glue Data Catalog. Queste viste sono utili perché supportano più motori di query SQL, quindi puoi accedere alla stessa vista su AWS servizi diversi, come EMR on e EC2 Amazon Amazon Athena Redshift.

Creando una vista nel Data Catalog, puoi utilizzare la concessione di risorse e i controlli di accesso basati su tag AWS Lake Formation per concedere l'accesso ad essa. Utilizzando questo metodo di controllo degli accessi, non è necessario configurare un accesso aggiuntivo alle tabelle a cui hai fatto

riferimento durante la creazione della vista. Questo metodo di concessione delle autorizzazioni si chiama definer semantics e queste viste sono chiamate definer views. Per ulteriori informazioni sul controllo degli accessi in Lake Formation, consulta Concessione e revoca delle autorizzazioni sulle risorse del Data Catalog nella AWS Lake Formation Developer Guide.

Le visualizzazioni del catalogo dati sono utili per i seguenti casi d'uso:

- Controllo granulare degli accessi: è possibile creare una visualizzazione che limiti l'accesso ai dati in base alle autorizzazioni necessarie all'utente. Ad esempio, puoi utilizzare le viste nel Catalogo dati per impedire ai dipendenti che non lavorano nel reparto delle risorse umane di visualizzare le informazioni di identificazione personale (PII).
- Definizione completa della vista: applicando filtri alla visualizzazione nel Data Catalog, ti assicuri che i record di dati disponibili in una vista del Data Catalog siano sempre completi.
- Sicurezza avanzata: la definizione della query utilizzata per creare la vista deve essere completa. Questo vantaggio significa che le visualizzazioni del Data Catalog sono meno suscettibili ai comandi SQL di attori malintenzionati.
- Condivisione semplice dei dati: condividi i dati con altri AWS account senza spostare i dati. Per ulteriori informazioni, consulta Condivisione dei dati tra account in Lake Formation.

Creazione di una vista di Catalogo Dati

Esistono diversi modi per creare una visualizzazione del catalogo dati. Questi includono l'utilizzo di AWS CLI o Spark SQL. Seguono alcuni esempi.

Using SQL

Di seguito viene illustrata la sintassi per la creazione di una vista del catalogo dati. Nota il tipo di `MULTI DIALECT` visualizzazione. Ciò distingue la vista del catalogo dati dalle altre viste. Il `SECURITY` predicato è specificato come. `DEFINER` Ciò indica una vista del catalogo dati con `DEFINER` semantica.

```
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;
```

Di seguito è riportato un esempio di `CREATE` dichiarazione, che segue la sintassi:

```
CREATE PROTECTED MULTI DIALECT VIEW catalog_view
SECURITY DEFINER
AS
SELECT order_date, sum(totalprice) AS price
FROM source_table
GROUP BY order_date
```

È inoltre possibile creare una vista in modalità dry-run, utilizzando SQL, per testare la creazione della vista, senza creare effettivamente la risorsa. L'utilizzo di questa opzione comporta una «esecuzione a secco» che convalida l'input e, se la convalida ha esito positivo, restituisce il JSON dell'oggetto della tabella AWS Glue che rappresenterà la vista. In questo caso, la visualizzazione effettiva non viene creata.

```
CREATE [ OR REPLACE ] PROTECTED MULTI DIALECT VIEW view_name
SECURITY DEFINER
[ SHOW VIEW JSON ]
AS view-sql
```

Using the AWS CLI

> **Note**
>
> Quando si utilizza il comando CLI, l'SQL utilizzato per creare la vista non viene analizzato. Ciò può causare un caso in cui la vista viene creata, ma le query non hanno esito positivo. Assicurati di testare la sintassi SQL prima di creare la vista.

Utilizzate il seguente comando CLI per creare una vista:

```
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"
        }
      ]
    }
  }
}'
```

Operazioni di visualizzazione supportate

I seguenti frammenti di comandi mostrano vari modi di lavorare con le viste del catalogo dati:

- CREA VISTA

  Crea una vista del catalogo di dati. Di seguito è riportato un esempio che mostra la creazione di una vista da una tabella esistente:

  ```
  CREATE PROTECTED MULTI DIALECT VIEW catalog_view
  SECURITY DEFINER AS SELECT * FROM my_catalog.my_database.source_table
  ```

- MODIFICA VISTA

  Sintassi disponibile:

  - `ALTER VIEW view_name [FORCE] ADD DIALECT AS query`

- `ALTER VIEW view_name [FORCE] UPDATE DIALECT AS query`
- `ALTER VIEW view_name DROP DIALECT`

È possibile utilizzare l'`FORCE ADD DIALECT`opzione per forzare l'aggiornamento dello schema e degli oggetti secondari secondo il nuovo dialetto del motore. Tieni presente che questa operazione può causare errori di query se non lo utilizzi anche `FORCE` per aggiornare altri dialetti del motore. Di seguito viene mostrato un esempio:

```
ALTER VIEW catalog_view FORCE ADD DIALECT
AS
SELECT order_date, sum(totalprice) AS price
FROM source_table
GROUP BY orderdate;
```

Quanto segue mostra come modificare una vista per aggiornare il dialetto:

```
ALTER VIEW catalog_view UPDATE DIALECT AS
SELECT count(*) FROM my_catalog.my_database.source_table;
```

- DESCRIVI LA VISTA

  Sintassi disponibile per descrivere una vista:

  - `SHOW COLUMNS {FROM|IN} view_name [{FROM|IN} database_name]`— Se l'utente dispone delle autorizzazioni AWS Glue and Lake Formation necessarie per descrivere la vista, può elencare le colonne. Di seguito sono riportati un paio di comandi di esempio per la visualizzazione delle colonne:

    ```
    SHOW COLUMNS FROM my_database.source_table;
    SHOW COLUMNS IN my_database.source_table;
    ```

  - `DESCRIBE view_name`— Se l'utente dispone delle autorizzazioni AWS Glue and Lake Formation necessarie per descrivere la vista, può elencare le colonne della vista insieme ai relativi metadati.

- RILASCIA LA VISTA

  Sintassi disponibile:

  - `DROP VIEW [ IF EXISTS ] view_name`

L'esempio seguente mostra un'DROPistruzione che verifica l'esistenza di una vista prima di eliminarla:

```
DROP VIEW IF EXISTS catalog_view;
```

- MOSTRA CREA VISTA
  - `SHOW CREATE VIEW view_name`— Mostra l'istruzione SQL che crea la vista specificata. Di seguito è riportato un esempio che mostra la creazione di una vista del catalogo di dati:

```
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)
```

- MOSTRA VISUALIZZAZIONI

  Elenca tutte le viste del catalogo, ad esempio viste regolari, visualizzazioni multidialettali (MDV) e MDV senza dialetto Spark. La sintassi disponibile è la seguente:

  - `SHOW VIEWS [{ FROM | IN } database_name] [LIKE regex_pattern]`:

    Di seguito viene illustrato un comando di esempio per mostrare le viste:

```
SHOW VIEWS IN marketing_analytics LIKE 'catalog_view*';
```

Per ulteriori informazioni sulla creazione e la configurazione delle viste del catalogo dati, consulta Building AWS Glue Data Catalog views nella Developer Guide. AWS Lake Formation

Interrogazione di una vista di Catalogo Dati

Dopo aver creato una vista del catalogo dati, puoi interrogarla utilizzando un job Amazon EMR Spark con controllo granulare degli accessi AWS Lake Formation abilitato. Il ruolo di job runtime deve

disporre dell'SELECTautorizzazione Lake Formation per la visualizzazione Data Catalog. Non è necessario concedere l'accesso alle tabelle sottostanti a cui si fa riferimento nella vista.

Dopo aver impostato tutto, puoi interrogare la tua vista. Ad esempio, dopo aver creato un'applicazione Amazon EMR in EMR Studio, puoi eseguire la seguente query per accedere a una vista.

```
SELECT * from my_database.catalog_view LIMIT 10;
```

Una funzione utile è la. `invoker_principal` Restituisce l'identificatore univoco del ruolo EMRS Job Runtime. Questo può essere usato per controllare l'output della vista, in base al principio di invocazione. Puoi usarlo per aggiungere una condizione nella tua vista che perfeziona i risultati della query, in base al ruolo chiamante. Il ruolo di job runtime deve disporre dell'autorizzazione all'azione `LakeFormation:GetDataLakePrincipal` IAM per utilizzare questa funzione.

```
select invoker_principal();
```

È possibile aggiungere questa funzione a una `WHERE` clausola, ad esempio, per perfezionare i risultati delle query.

Considerazioni e limitazioni

Quando si creano viste del catalogo dati, si applica quanto segue:

- Puoi creare viste del catalogo dati solo con Amazon EMR 7.10 e versioni successive.
- Il definitore della vista del catalogo dati deve avere `SELECT` accesso alle tabelle di base sottostanti a cui si accede dalla vista. La creazione della vista Data Catalog non riesce se una tabella base specifica ha dei filtri Lake Formation imposti sul ruolo di definizione.
- Le tabelle di base non devono avere l'autorizzazione del `IAMAllowedPrincipals` data lake in Lake Formation. Se presente, si verifica l'errore Multi Dialect views può fare riferimento solo a tabelle senza le autorizzazioni IAMAllowed Principal.
- La posizione Amazon S3 della tabella deve essere registrata come posizione del data lake Lake Formation. Se la tabella non è registrata, si verifica l'errore Le viste multidialettali possono fare riferimento solo alle tabelle gestite da Lake Formation. Per informazioni su come registrare le sedi Amazon S3 in Lake Formation, consulta Registrazione di una sede Amazon S3 nella Developer Guide. AWS Lake Formation
- Puoi creare solo viste di `PROTECTED` Data Catalog. `UNPROTECTED`le visualizzazioni non sono supportate.

- Non è possibile fare riferimento alle tabelle di un altro AWS account in una definizione di visualizzazione del catalogo dati. Inoltre, non puoi fare riferimento a una tabella nello stesso account che si trova in una regione separata.
- Per condividere i dati tra un account o un'area geografica, l'intera visualizzazione deve essere condivisa tra account e regioni, utilizzando i link alle risorse di Lake Formation.
- Le funzioni definite dall'utente (UDFs) non sono supportate.
- È possibile utilizzare viste basate sulle tabelle Iceberg. Sono supportati anche i formati a tabella aperta Apache Hudi e Delta Lake.
- Non puoi fare riferimento ad altre viste nelle viste del Catalogo dati.
- Lo schema di visualizzazione di AWS Glue Data Catalog viene sempre memorizzato in lettere minuscole. Ad esempio, se si utilizza un'istruzione DDL per creare una vista del Glue Data Catalog con una colonna denominata`Castle`, la colonna creata nel Glue Data Catalog verrà composta in minuscolo, to. `castle` Se poi si specifica il nome della colonna in una query DML come `Castle` o`CASTLE`, EMR Spark renderà il nome in minuscolo per eseguire la query. Tuttavia, l'intestazione della colonna viene visualizzata utilizzando il maiuscolo/minuscolo specificato nella query.

  Se si desidera che una query abbia esito negativo nel caso in cui il nome di colonna specificato nella query DML non corrisponda al nome della colonna nel Glue Data Catalog, è possibile impostare`spark.sql.caseSensitive=true`.

Utilizzo di un ruolo di runtime del job Lake Formation con accesso completo alla tabella

Gli argomenti di questa sezione spiegano come accedere alle tabelle del catalogo Glue Data protette da Lake Formation dai job EMR Spark o dalle sessioni interattive con accesso completo alla tabella.

Accesso non filtrato a Lake Formation per EMR su EC2

Con le versioni 7.8.0 e successive di Amazon EMR, puoi sfruttare AWS Lake Formation with Glue Data Catalog, dove il ruolo di job runtime dispone di autorizzazioni complete per la tabella. Questa funzionalità consente di leggere e scrivere su tabelle protette da Lake Formation dal batch EMR Spark e dai lavori interattivi.

Utilizzo di Lake Formation con accesso completo ai tavoli

È possibile accedere alle tabelle del catalogo Glue Data protette da AWS Lake Formation dai job EMR Spark o dalle sessioni interattive in cui il ruolo di runtime del job ha accesso completo alla

tabella. Non è necessario abilitare AWS Lake Formation sul cluster EMR. Quando un job Spark è configurato per Full Table Access (FTA), le credenziali di AWS Lake Formation vengono utilizzate per leggere/scrivere i dati S3 per le tabelle registrate di Lake AWS Formation, mentre le credenziali del ruolo di runtime del lavoro verranno utilizzate per le tabelle read/write non registrate con Lake Formation. AWS

> **Note**
>
> Il controllo granulare degli accessi di Lake Formation non è richiesto o supportato per un job per eseguire Full Table Access (FTA).

> **Note**
>
> L'accesso completo alla tabella non richiede l'uso di un ruolo di runtime EMR e funziona sia con il ruolo di runtime che con il ruolo di servizio per le istanze del cluster EC2 . Per ulteriori informazioni, consulta Ruolo di servizio per EC2 le istanze del cluster (profilo di EC2 istanza) o Ruoli di runtime per le fasi di Amazon EMR.

Passaggio 1: abilitare l'accesso completo alla tabella in Lake Formation

Per utilizzare la modalità Full Table Access (FTA), devi consentire ai motori di query di terze parti di accedere ai dati senza la convalida dei tag di sessione IAM in AWS Lake Formation. Per abilitarla, segui i passaggi indicati in Integrazione delle applicazioni per l'accesso completo alla tabella.

> **Note**
>
> Quando si accede a tabelle con più account, è necessario abilitare l'accesso completo alla tabella sia nell'account produttore che in quello consumatore. Allo stesso modo, quando si accede alle tabelle interregionali, questa impostazione deve essere abilitata sia nelle regioni produttrici che in quelle di consumo.

Fase 2: Configurazione delle autorizzazioni IAM per il ruolo Job Runtime

Per l'accesso in lettura o scrittura ai dati sottostanti, oltre alle autorizzazioni di Lake Formation, un ruolo di job runtime richiede l'autorizzazione `lakeformation:GetDataAccess` IAM. Con questa autorizzazione, Lake Formation concede la richiesta di credenziali temporanee per accedere ai dati.

Di seguito è riportato un esempio di politica su come fornire le autorizzazioni IAM per accedere a uno script in Amazon S3, caricare i log su S3, le autorizzazioni dell'API AWS Glue e l'autorizzazione per accedere 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": [
        "*"
      ]
    }
  ]
}
```

Passaggio 2.1 Configurare i permessi di Lake Formation

- I job Spark che leggono dati da S3 richiedono l'autorizzazione Lake Formation SELECT.
- Spark fa in modo che write/delete i dati in S3 richiedano l'autorizzazione Lake Formation ALL (SUPER).
- I lavori Spark che interagiscono con il catalogo Glue Data richiedono l'autorizzazione DESCRIBE, ALTER, DROP, a seconda dei casi.

Per ulteriori informazioni, consulta Concessione delle autorizzazioni sulle risorse di Data Catalog.

Passaggio 3: inizializza una sessione Spark per l'accesso completo alla tabella utilizzando Lake Formation

Prerequisiti

AWS Glue Data Catalog deve essere configurato come metastore per accedere alle tabelle di Lake Formation.

Imposta le seguenti impostazioni per configurare il catalogo Glue come metastore:

```
--conf spark.sql.catalogImplementation=hive
--conf
 spark.hive.metastore.client.factory.class=com.amazonaws.glue.catalog.metastore.AWSGlueDataCata
```

Per abilitare Data Catalog for EMR su EC2, consulta Use AWS Glue Data Catalog with Spark su Amazon EMR.

Accesso alle tabelle

Per accedere alle tabelle registrate con AWS Lake Formation, è necessario impostare le seguenti configurazioni durante l'inizializzazione di Spark per configurare Spark per utilizzare le credenziali di Lake Formation AWS .

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.A` Configurare EMR Filesystem (EMRFS) per utilizzare le credenziali di Lake AWS Formation S3 per le tabelle registrate di Lake Formation. Se la tabella non è registrata, utilizza le credenziali del ruolo di runtime del job.
- `spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject=truespark.hadoop.fs.s3.fo` configura EMRFS per utilizzare l'intestazione del tipo di contenuto application/x-directory anziché il suffisso durante la creazione di cartelle S3. `$folder$` Questo è necessario per leggere le tabelle di Lake Formation, poiché le credenziali di Lake Formation non consentono la lettura delle cartelle delle tabelle con il suffisso $folder$.
- `spark.sql.catalog.skipLocationValidationOnCreateTable.enabled=true`: Configura Spark per saltare la convalida del vuoto della posizione della tabella prima della creazione. Ciò è necessario per le tabelle registrate di Lake Formation, poiché le credenziali di Lake Formation per verificare la posizione vuota sono disponibili solo dopo la creazione della

tabella Glue Data Catalog. Senza questa configurazione, le credenziali del ruolo di runtime del job convalideranno la posizione della tabella vuota.

- `spark.sql.catalog.createDirectoryAfterTable.enabled=true`: configura Spark per creare la cartella Amazon S3 dopo la creazione della tabella nel metastore Hive. Questo è necessario per le tabelle registrate di Lake Formation, poiché le credenziali di Lake Formation per creare la cartella S3 sono disponibili solo dopo la creazione della tabella Glue Data Catalog.
- `spark.sql.catalog.dropDirectoryBeforeTable.enabled=true`: Configura Spark per eliminare la cartella S3 prima dell'eliminazione della tabella nel metastore Hive. Ciò è necessario per le tabelle registrate di Lake Formation, poiché le credenziali di Lake Formation per eliminare la cartella S3 non sono disponibili dopo l'eliminazione della tabella dal Glue Data Catalog.
- `spark.sql.catalog.<catalog>.glue.lakeformation-enabled=true`: Configura il catalogo Iceberg per utilizzare le credenziali di AWS Lake Formation S3 per le tabelle registrate di Lake Formation. Se la tabella non è registrata, usa le credenziali di ambiente predefinite.

Configura la modalità di accesso completo alla tabella in SageMaker Unified Studio

Per accedere alle tabelle registrate di Lake Formation dalle sessioni interattive di Spark nei JupyterLab notebook, è necessario utilizzare la modalità di autorizzazione alla compatibilità. Usa il comando `%%configure` magico per configurare la configurazione di Spark. Scegli la configurazione in base al tipo di tabella:

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
    }
}
```

Sostituisci i segnaposto:

- `S3_DATA_LOCATION`: Il percorso del tuo bucket S3
- `REGION`: AWS regione (ad es. us-east-1)
- `ACCOUNT_ID`: L'ID del tuo account AWS

> **Note**
>
> È necessario impostare queste configurazioni prima di eseguire qualsiasi operazione Spark sul notebook.

Operazioni supportate

Queste operazioni utilizzeranno le credenziali di AWS Lake Formation per accedere ai dati della tabella.

- CREATE TABLE
- ALTER TABLE
- INSERT INTO
- INSERT OVERWRITE
- UPDATE
- MERGE INTO
- DELETE FROM

- ANALIZZA LA TABELLA
- TABELLA DI RIPARAZIONE
- DROP TABLE
- Query su origini dati Spark
- Scritture di origini dati Spark

> **Note**
>
> Le operazioni non elencate sopra continueranno a utilizzare le autorizzazioni IAM per accedere ai dati delle tabelle.

Considerazioni

- Se una tabella Hive viene creata utilizzando un lavoro per cui non è abilitato l'accesso completo alla tabella e non viene inserito alcun record, le letture o scritture successive da un lavoro con accesso completo alla tabella avranno esito negativo. Questo perché EMR Spark senza accesso completo alla tabella aggiunge il `$folder$` suffisso al nome della cartella della tabella. Per risolvere questo problema, puoi:
  - Inserire almeno una riga nella tabella da un lavoro in cui l'FTA non è abilitato.
  - Configura il lavoro che non ha FTA abilitato in modo che non utilizzi il `$folder$` suffisso nel nome della cartella in S3. Ciò può essere ottenuto impostando la configurazione di Spark. `spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject=true`
  - Crea una cartella S3 nella posizione della tabella `s3://path/to/table/table_name` utilizzando la console AWS S3 o la CLI AWS S3.
- Full Table Access funziona esclusivamente con EMR Filesystem (EMRFS). Il file system S3A non è compatibile.
- L'accesso completo alla tabella è supportato per le tabelle Hive e Iceberg. Il supporto per le tabelle Hudi e Delta non è stato ancora aggiunto.
- I lavori che fanno riferimento alle tabelle con le regole Lake Formation Fine-Grained Access Control (FGAC) o Glue Data Catalog Views avranno esito negativo.
- L'accesso completo alla tabella non supporta Spark Streaming.
- Quando si scrive Spark DataFrame su una tabella Lake Formation, è supportata solo la modalità APPEND: `df.write.mode("append").saveAsTable(table_name)`

Integrazione di Amazon EMR con Apache Ranger

A partire da Amazon EMR 5.32.0, è possibile avviare un cluster che si integra nativamente con Apache Ranger. Apache Ranger è un framework open source che consente di abilitare, monitorare e gestire la sicurezza completa dei dati attraverso la piattaforma Hadoop. Per ulteriori informazioni, consulta Apache Ranger. L'integrazione nativa consente di utilizzare Apache Ranger per imporre un controllo granulare di accesso ai dati su Amazon EMR.

Questa sezione fornisce una panoramica delle nozioni di base dell'integrazione di Amazon EMR con Apache Ranger. Include inoltre i prerequisiti e le fasi necessari per avviare un cluster Amazon EMR integrato con Apache Ranger.

L'integrazione nativa di Amazon EMR con Apache Ranger offre i seguenti vantaggi principali:

- Controllo dell'accesso granulare ai database e alle tabelle del metastore Hive, che consente di definire policy di filtraggio dei dati a livello di database, tabella e colonna per le applicazioni Apache Spark e Apache Hive. Il filtraggio a livello di riga e il mascheramento dei dati sono supportati con le applicazioni Hive.
- La possibilità di utilizzare le policy Hive esistenti direttamente con Amazon EMR per le applicazioni Hive.
- Controllo dell'accesso ai dati di Amazon S3 a livello di prefisso e oggetto, che consente di definire policy di filtraggio dei dati per l'accesso ai dati S3 utilizzando il file system EMR.
- La possibilità di utilizzare CloudWatch Logs per il controllo centralizzato.
- Amazon EMR installa e gestisce i plug-in Apache Ranger per conto tuo.

Apache Ranger con Amazon EMR

Apache Ranger è un framework che consente di abilitare, monitorare e gestire la sicurezza completa dei dati attraverso la piattaforma Hadoop.

Apache Ranger è dotata delle seguenti caratteristiche:

- Amministrazione della sicurezza centralizzata per gestire tutte le attività relative alla sicurezza in un'interfaccia utente centrale o utilizzando REST. APIs
- Autorizzazione granulare per eseguire un'operazione specifica o un'operazione con un componente o uno strumento Hadoop, gestito tramite uno strumento di amministrazione centrale.
- Un metodo di autorizzazione standardizzato per tutti i componenti Hadoop.

- Supporto avanzato per vari metodi di autorizzazione.
- Verifica centralizzata dell'accesso degli utenti e delle operazioni amministrative (relative alla sicurezza) all'interno di tutti i componenti di Hadoop.

Apache Ranger utilizza due componenti chiave per l'autorizzazione:

- Server di amministrazione delle policy di Apache Ranger: questo server consente di definire le policy di autorizzazione per le applicazioni Hadoop. Durante l'integrazione con Amazon EMR, puoi definire e applicare policy per Apache Spark e Hive per accedere al Hive Metastore e ai dati di Amazon S3 File System EMR (EMRFS). Per l'integrazione con Amazon EMR, puoi configurare un server di amministrazione delle policy di Apache Ranger esistente.
- Plug-in Apache Ranger: questo plug-in convalida l'accesso di un utente rispetto alle policy di autorizzazione definite nel server di amministrazione delle policy di Apache Ranger. Amazon EMR installa e configura automaticamente il plug-in Apache Ranger per ogni applicazione Hadoop selezionata nella configurazione di Apache Ranger.

Argomenti

- Architettura dell'integrazione di Amazon EMR con Apache Ranger
- Componenti Amazon EMR da utilizzare con Apache Ranger

Architettura dell'integrazione di Amazon EMR con Apache Ranger

Componenti Amazon EMR da utilizzare con Apache Ranger

Amazon EMR consente il controllo granulare degli accessi con Apache Ranger attraverso i seguenti componenti. Fai riferimento al diagramma dell'architettura per una rappresentazione grafica dei componenti Amazon EMR con i plug-in Apache Ranger.

Secret agent (Agente segreto): l'agente segreto archivia in modo sicuro i segreti e distribuisce i segreti ad altri componenti o applicazioni Amazon EMR. I segreti possono includere credenziali utente temporanee, chiavi di crittografia o ticket Kerberos. L'agente segreto viene eseguito su ogni nodo del cluster e intercetta le chiamate al servizio metadati dell'istanza. Per le richieste alle credenziali del ruolo del profilo dell'istanza, l'agente segreto vende le credenziali in base all'utente richiedente e alle risorse richieste dopo aver autorizzato la richiesta con il plug-in EMRFS S3 Ranger. L'agente segreto viene eseguito come *`emrsecretagent`*utente e scrive i log nella directory /.

emr/secretagent/log Il processo si basa su un set specifico di regole `iptables` per funzionare. È importante assicurarsi che `iptables` non sia disabilitato. Se si personalizza la configurazione `iptables`, le regole della tabella NAT devono essere mantenute e lasciate invariate.

EMR record server (Server di registro EMR): il server di registro riceve le richieste di accesso ai dati da Spark. In seguito autorizza le richieste inoltrando le risorse richieste al plug-in Spark Ranger per Amazon EMR. Il server di registro legge i dati da Amazon S3 e restituisce i dati filtrati a cui l'utente è autorizzato ad accedere in funzione della policy di Ranger. Il server di registrazione viene eseguito su ogni nodo del cluster come utente emr_record_server e scrive i log nella directory/-record-server. var/log/emr

Considerazioni sull'utilizzo di Amazon EMR con Apache Ranger

Applicazioni supportate per Amazon EMR con Apache Ranger

Attualmente, l'integrazione tra Amazon EMR e Apache Ranger in cui EMR installa i plug-in Ranger supporta le seguenti applicazioni:

- Apache Spark (disponibile con EMR 5.32+ e EMR 6.3+)
- Apache Hive (disponibile con EMR 5.32+ e EMR 6.3+)
- S3 Access tramite EMRFS (disponibile con EMR 5.32+ e EMR 6.3+)

Le seguenti applicazioni possono essere installate in un cluster EMR e possono essere configurate per soddisfare le esigenze di sicurezza:

- Apache Hadoop (disponibile con EMR 5.32+ e EMR 6.3+ compresi YARN e HDFS)
- Apache Livy (disponibile con EMR 5.32+ e EMR 6.3+)
- Apache Zeppelin (disponibile con EMR 5.32+ e EMR 6.3+)
- Apache Hue (disponibile con EMR 5.32+ e EMR 6.3+)
- Ganglia (disponibile con EMR 5.32+ e EMR 6.3+)
- HCatalog (Disponibile con EMR 5.32+ ed EMR 6.3+)
- Mahout (disponibile con EMR 5.32+ e EMR 6.3+)
- MXNet (Disponibile con EMR 5.32+ ed EMR 6.3+)
- TensorFlow (Disponibile con EMR 5.32+ ed EMR 6.3+)
- Tez (disponibile con EMR 5.32+ e EMR 6.3+)
- Trino (disponibile con EMR 6.7+)

- ZooKeeper (Disponibile con EMR 5.32+ ed EMR 6.3+)

> **Important**
>
> Le applicazioni elencate sopra sono le uniche applicazioni attualmente supportate. Per garantire la sicurezza del cluster, è consentito creare un cluster EMR con solo le applicazioni nell'elenco precedente quando Apache Ranger è abilitato.
> Altre applicazioni non sono attualmente supportate. Per garantire la sicurezza del cluster, il tentativo di installare altre applicazioni causerà il rifiuto del cluster.
> AWS I formati Glue Data Catalog e Open table come Apache Hudi, Delta Lake e Apache Iceberg non sono supportati.

Funzionalità di Amazon EMR supportate con Apache Ranger

Le seguenti funzionalità di Amazon EMR sono supportate quando utilizzi Amazon EMR con Apache Ranger:

- Crittografia dei dati su disco e in transito.
- Autenticazione Kerberos (obbligatoria)
- Gruppi di istanze, parchi istanze e istanze Spot
- Riconfigurazione delle applicazioni in un cluster in esecuzione
- Server-Side Encryption (SSE) EMRFS

> **Note**
>
> Le impostazioni di crittografia di Amazon EMR regolano SSE. Per ulteriori informazioni, consulta Opzioni di crittografia.

Limiti dell'applicazione

Ci sono diverse limitazioni da tenere a mente quando si integra Amazon EMR e Apache Ranger:

- Al momento non è possibile utilizzare la console per creare una configurazione di sicurezza che specifichi l'opzione di integrazione AWS Ranger in. AWS GovCloud (US) Region La configurazione della sicurezza può essere eseguita utilizzando la CLI.

- Kerberos deve essere installato nel cluster.
- Le applicazioni UIs (interfacce utente) come l'interfaccia utente di YARN Resource Manager, l'interfaccia utente HDFS e l' NameNode interfaccia utente Livy non sono impostate con l'autenticazione per impostazione predefinita.
- Le autorizzazioni predefinite HDFS `umask` sono configurate in modo che gli oggetti creati siano impostati su `world wide readable` per impostazione predefinita.
- Amazon EMR non supporta la modalità ad alta disponibilità (principale multipla) con Apache Ranger.
- Per ulteriori limitazioni, consulta le limitazioni per ogni applicazione.

> **Note**
>
> Le impostazioni di crittografia di Amazon EMR regolano SSE. Per ulteriori informazioni, consulta Opzioni di crittografia.

Limitazioni del plug-in

Ogni plug-in ha limitazioni specifiche. Per le limitazioni del plug-in Apache Hive, consulta Limitazioni del plug-in Apache Hive. Per le limitazioni del plug-in Apache Spark, consulta Limitazioni del plug-in Apache Spark. Per le limitazioni del plug-in EMRFS S3, consulta Limitazioni del plug-in EMRFS S3.

Configurazione di Amazon EMR per Apache Ranger

Prima di installare Apache Ranger, consulta le informazioni contenute in questa sezione per verificare che Amazon EMR sia configurato correttamente.

Argomenti

- Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR
- Ruoli IAM per l'integrazione nativa con Apache Ranger
- Creazione di una configurazione di sicurezza EMR
- Archiviazione dei certificati TLS in AWS Secrets Manager
- Avvia un cluster EMR con Apache Ranger
- Configurazione dei cluster Amazon EMR abilitati da Zeppelin per Apache Ranger

- Problemi noti per l'integrazione con Amazon EMR

Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR

Per l'integrazione di Amazon EMR, i plug-in dell'applicazione Apache Ranger devono comunicare con il server Admin utilizzando TLS/SSL.

Prerequisito: SSL abilitato con il server Admin Ranger

Apache Ranger su Amazon EMR richiede una comunicazione SSL bidirezionale tra i plug-in e il server Admin Ranger. Per garantire che i plugin comunichino con il server Apache Ranger tramite SSL, abilita il seguente attributo all'interno del ranger-admin-site file.xml sul server Ranger Admin.

```
<property>
    <name>ranger.service.https.attrib.ssl.enabled</name>
    <value>true</value>
</property>
```

Inoltre, sono necessarie le seguenti configurazioni.

```
<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>
```

Certificati TLS per l'integrazione di Apache Ranger con Amazon EMR

L'integrazione di Apache Ranger con Amazon EMR richiede che il traffico dai nodi di Amazon EMR al server Admin Ranger sia crittografato utilizzando TLS e che i plug-in Ranger si autentichino al server Apache Ranger utilizzando l'autenticazione TLS reciproca bidirezionale. Il servizio Amazon EMR richiede il certificato pubblico del server Admin Ranger (specificato nell'esempio precedente) e il certificato privato.

Certificati del plug-in Apache Ranger

I certificati TLS pubblici del plug-in Apache Ranger devono essere accessibili al server Admin Apache Ranger per convalidare quando i plug-in si connettono. Esistono tre metodi diversi per eseguire questa operazione.

Metodo 1: Configurazione di un truststore nel server Admin Apache Ranger

Compila le seguenti configurazioni in ranger-admin-site .xml per configurare un truststore.

```
<property>
    <name>ranger.truststore.file</name>
    <value><LOCATION TO TRUSTSTORE></value>
</property>

<property>
    <name>ranger.truststore.password</name>
    <value><PASSWORD FOR TRUSTSTORE></value>
</property>
```

Metodo 2: Caricamento del certificato in Java cacerts truststore

Se il tuo server Admin Ranger non specifica un truststore nelle sue opzioni JVM, puoi inserire i certificati pubblici del plug-in nell'archivio cacerts predefinito.

Metodo 3: Creazione di un truststore specificato come parte delle opzioni JVM

In {RANGER_HOME_DIRECTORY}/ews/ranger-admin-services.sh, modifica JAVA_OPTS per includere "-Djavax.net.ssl.trustStore=*<TRUSTSTORE_LOCATION>*" e "-Djavax.net.ssl.trustStorePassword=*<TRUSTSTORE_PASSWORD>*". Ad esempio, aggiungi la seguente riga dopo il JAVA_OPTS esistente.

```
JAVA_OPTS=" ${JAVA_OPTS} -Djavax.net.ssl.trustStore=${RANGER_HOME}/truststore/
truststore.jck -Djavax.net.ssl.trustStorePassword=changeit"
```

> **Note**
>
> Questa specifica può esporre la password truststore se un utente è in grado di accedere al server Admin Apache Ranger e vedere i processi in esecuzione, ad esempio quando si utilizza il comando ps.

Utilizzo di certificati autofirmati

I certificati autofirmati non sono consigliati come certificati. I certificati autofirmati potrebbero non essere revocati o non essere conformi ai requisiti di sicurezza interni.

Installazione della definizione del servizio per l'integrazione di Ranger con Amazon EMR

Una definizione di servizio viene utilizzata dal server Admin Ranger per descrivere gli attributi delle policy per un'applicazione. Le policy vengono quindi archiviate in un repository di policy per il download dei client.

Per poter configurare le definizioni dei servizi, le chiamate REST devono essere effettuate al server Admin Ranger. Consulta Apache Ranger Public APIsv2 per informazioni APIs richieste nella sezione seguente.

Installazione della definizione del servizio di Apache Spark

Per installare la definizione del servizio di Apache Spark, consulta Plugin Apache Spark per l'integrazione di Ranger con Amazon EMR.

Installazione della definizione del servizio EMRFS

Per installare la definizione del servizio S3 per Amazon EMR, consulta Plugin EMRFS S3 per l'integrazione di Ranger con Amazon EMR.

Utilizzo della definizione del servizio Hive

Apache Hive può utilizzare la definizione del servizio Ranger esistente fornita con Apache Ranger 2.0 e versioni successive. Per ulteriori informazioni, consulta Plugin Apache Hive per l'integrazione di Ranger con Amazon EMR.

Regole del traffico di rete per l'integrazione con Amazon EMR

Quando Apache Ranger è integrato con il cluster EMR, il cluster deve comunicare con server aggiuntivi e AWS.

Tutti i nodi Amazon EMR, inclusi i nodi principali e attività, devono essere in grado di comunicare con i server Admin Apache Ranger per scaricare le policy. Se il tuo amministratore Apache Ranger è in esecuzione su Amazon EC2, devi aggiornare il gruppo di sicurezza per poter prelevare il traffico dal cluster EMR.

Oltre a comunicare con il server Ranger Admin, tutti i nodi devono essere in grado di comunicare con i seguenti servizi: AWS

- Amazon S3
- AWS KMS (se si utilizza EMRFS SSE-KMS)
- Amazon CloudWatch
- AWS STS

Se prevedi di eseguire il cluster EMR all'interno di una sottorete privata, configura il VPC per comunicare con questi servizi utilizzando AWS PrivateLink ed endpoint VPC nella Guida per l'utente di Amazon VPC oppure utilizzando un'istanza NAT (Network Address Translation) nella Guida per l'utente di Amazon VPC.

Ruoli IAM per l'integrazione nativa con Apache Ranger

L'integrazione tra Amazon EMR e Apache Ranger si basa su tre ruoli chiave che è necessario creare prima di avviare il cluster:

- Un profilo di EC2 istanza Amazon personalizzato per Amazon EMR
- Un ruolo IAM per Apache Ranger Engines
- Un ruolo IAM per altri servizi AWS

Questa sezione fornisce una panoramica di questi ruoli e delle policy che è necessario includere per ogni ruolo IAM. Per informazioni sulla creazione di questi ruoli, consulta Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR.

EC2 profilo di istanza per Amazon EMR

Amazon EMR utilizza un ruolo di servizio IAM per eseguire operazioni per tuo conto per il provisioning e la gestione dei cluster. Il ruolo di servizio per EC2 le istanze di cluster, chiamato anche profilo di EC2 istanza per Amazon EMR, è un tipo speciale di ruolo di servizio assegnato a EC2 ogni istanza di un cluster al momento del lancio.

Per definire le autorizzazioni per l'interazione del cluster EMR con i dati di Amazon S3 e con il metastore Hive protetto da Apache Ranger e AWS altri servizi, definisci un profilo di istanza EC2 personalizzato da utilizzare al posto di quando avvii il cluster. `EMR_EC2_DefaultRole`

Per ulteriori informazioni, consultare Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2 e Personalizza i ruoli IAM con Amazon EMR.

È necessario aggiungere le seguenti istruzioni al profilo di EC2 istanza predefinito per Amazon EMR per poter etichettare le sessioni e accedere ai certificati TLS AWS Secrets Manager che archiviano.

```
    {
      "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**
>
> Per le autorizzazioni di Secrets Manager, non dimenticare il carattere jolly ("*") alla fine del nome segreto, altrimenti le richieste non avranno esito positivo. Il carattere jolly vale per le versioni segrete.

> **Note**
>
> Limita l'ambito della AWS Secrets Manager policy ai soli certificati necessari per il provisioning.

Ruolo IAM per Apache Ranger

Questo ruolo fornisce ai motori di esecuzione attendibili, ad esempio Apache Hive e Amazon EMR Record Serve, le credenziali per accedere ai dati Amazon S3. Utilizza solo questo ruolo per accedere ai dati di Amazon S3, incluse le chiavi KMS, se utilizzi S3 SSE-KMS.

Questo ruolo deve essere creato con la policy minima riportata nell'esempio seguente.

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

L'asterisco «*» alla fine della risorsa di CloudWatch registro deve essere incluso per consentire la scrittura nei flussi di registro.

Note

Se utilizzi la visualizzazione di coerenza EMRFS o la crittografia S3-SSE, aggiungi le autorizzazioni alle tabelle DynamoDB e alle chiavi del servizio di gestione delle chiavi in modo che i motori di esecuzione possano interagire con tali motori.

Il ruolo IAM per Apache Ranger viene assunto dall'Instance Profile Role. EC2 Utilizza l'esempio seguente per creare una policy di fiducia che consenta di assumere il ruolo IAM di Apache Ranger dal ruolo del profilo dell' EC2 istanza.

```
   {
     "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"]
   }
```

Ruolo IAM per altri AWS servizi per l'integrazione con Amazon EMR

Questo ruolo fornisce agli utenti che non sono motori di esecuzione affidabili le credenziali per interagire con AWS i servizi, se necessario. Non utilizzare questo ruolo IAM per consentire l'accesso ai dati di Amazon S3, a meno che non si tratti di dati che devono essere accessibili a tutti gli utenti.

Questo ruolo verrà assunto dall' EC2 Instance Profile Role. Utilizza l'esempio seguente per creare una policy di fiducia che consenta di assumere il ruolo IAM per Apache Ranger dal ruolo del profilo dell' EC2 istanza.

```
   {
     "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"]
   }
```

Convalida le tue autorizzazioni per l'integrazione di Amazon EMR con Apache Ranger

Consulta Risoluzione dei problemi di Apache Ranger per istruzioni sulla convalida delle autorizzazioni.

Creazione di una configurazione di sicurezza EMR

Creazione di una configurazione di sicurezza di Amazon EMR per Apache Ranger

Prima di lanciare un cluster Amazon EMR integrato con Apache Ranger, crea una configurazione di sicurezza.

Console

Creazione di una configurazione di sicurezza che specifichi l'opzione di integrazione AWS Ranger

1. Nella console di Amazon EMR, seleziona Security configurations (Configurazioni di sicurezza) e in seguito Create (Crea).
2. Digitare un nome in Name (Nome) per la configurazione di sicurezza. Questo nome è utilizzato per specificare la configurazione di sicurezza al momento della creazione di un cluster.
3. In AWS Ranger Integration (Integrazione Ranger), seleziona Enable fine-grained access control managed by Apache Ranger (Abilita controllo granulare degli accessi gestito da Apache Ranger).
4. Seleziona lo IAM role for Apache Ranger (Ruolo IAM per Apache Ranger) da applicare. Per ulteriori informazioni, consulta Ruoli IAM per l'integrazione nativa con Apache Ranger.
5. Seleziona il ruolo IAM per altri servizi AWS da applicare.
6. Configura i plugin per connetterti al server di amministrazione Ranger inserendo l'ARN di Secrets Manager per il server di amministrazione e l'indirizzo.
7. Seleziona le applicazioni per configurare i plug-in Ranger. Inserisci l'ARN di Secrets Manager che contiene il certificato TLS privato per il plug-in.

   Se Apache Spark o Apache Hive non vengono configurati e vengono selezionati come applicazione per il cluster, la richiesta ha esito negativo.
8. Configurare altre opzioni per la configurazione di sicurezza come appropriato e scegliere Create (Crea). È necessario abilitare l'autenticazione Kerberos utilizzando il KDC dedicato al cluster o esterno.

> **Note**
>
> Al momento non è possibile utilizzare la console per creare una configurazione di sicurezza che specifichi l'opzione di integrazione AWS Ranger in. AWS GovCloud (US) Region La configurazione della sicurezza può essere eseguita utilizzando la CLI.

CLI

Creazione di una configurazione di sicurezza per l'integrazione di Apache Ranger

1. Sostituiscila *<ACCOUNT ID>* con l'ID del tuo AWS account.
2. Sostituisci *<REGION>* con la Regione in cui si trova la risorsa.
3. Specifica un valore per `TicketLifetimeInHours` per determinare il periodo di validità di un ticket Kerberos emesso dal KDC.
4. Specifica l'indirizzo del server Admin Ranger per `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>_"
               }
            }
         }
      }
   }
}
```

PolicyRespositoryNames Sono i nomi dei servizi specificati nell'amministratore di Apache Ranger.

Crea una configurazione di sicurezza Amazon EMR con il seguente comando. Sostituisci security-configuration con un nome a tua scelta. Seleziona questa configurazione per nome quando crei il cluster.

```
aws emr create-security-configuration \
--security-configuration file://./security-configuration.json \
--name security-configuration
```

Configurazione di caratteristiche di sicurezza aggiuntive

Per integrare Amazon EMR con Apache Ranger in modo sicuro, configura le seguenti caratteristiche di sicurezza di EMR:

- Abilita l'autenticazione Kerberos utilizzando il KDC dedicato al cluster o esterno. Per istruzioni, consultare Utilizzo di Kerberos per l'autenticazione con Amazon EMR.
- (Facoltativo) Abilita la crittografia in transito o inattiva. Per ulteriori informazioni, consulta Opzioni di crittografia per Amazon EMR.

Per ulteriori informazioni, consulta Sicurezza in Amazon EMR.

Archiviazione dei certificati TLS in AWS Secrets Manager

I plug-in Ranger installati su un cluster Amazon EMR e il server Admin Ranger devono comunicare tramite TLS per garantire che i dati delle policy e le altre informazioni inviate non possano essere letti se vengono intercettati. EMR impone inoltre che i plug-in eseguano l'autenticazione al server Admin Ranger fornendo il proprio certificato TLS ed eseguendo l'autenticazione TLS bidirezionale. Questa configurazione richiedeva la creazione di quattro certificati: due coppie di certificati TLS privati e pubblici. Per istruzioni sull'installazione del certificato nel server Admin Ranger, consulta Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR. Per completare la configurazione, i plug-in Ranger installati nel cluster EMR necessitano di due certificati: il certificato TLS pubblico del server di amministrazione e il certificato privato che il plug-in utilizzerà per autenticarsi sul server Admin Ranger. Per fornire questi certificati TLS, devono essere presenti in AWS Secrets Manager e forniti in una configurazione di sicurezza EMR.

> **Note**
>
> Si consiglia, per quanto non sia necessario, di creare una coppia di certificati per ciascuna delle applicazioni per limitare l'impatto in caso di compromissione di uno dei certificati del plug-in.

> **Note**
>
> È necessario monitorare e ruotare i certificati prima della data di scadenza.

Formato del certificato

L'importazione dei certificati su AWS Secrets Manager è la stessa indipendentemente dal fatto che si tratti del certificato del plug-in privato o del certificato di amministrazione Ranger pubblico. Prima di importare i certificati TLS, questi devono essere in formato PEM 509x.

Un esempio di certificato pubblico è nel formato:

```
-----BEGIN CERTIFICATE-----
...Certificate Body...
-----END CERTIFICATE-----
```

Un esempio di certificato privato è nel formato:

```
-----BEGIN PRIVATE KEY-----
...Private Certificate Body...
-----END PRIVATE KEY-----
-----BEGIN CERTIFICATE-----
...Trust Certificate Body...
-----END CERTIFICATE-----
```

Anche il certificato privato deve contenere un certificato di affidabilità.

Puoi verificare che i certificati siano nel formato corretto eseguendo il seguente comando:

```
openssl x509 -in <PEM FILE> -text
```

Importazione di un certificato in AWS Secrets Manager

Quando crei il tuo segreto nel Secrets Manager, seleziona Other type of secrets (Altro tipo di segreti) in secret type (Tipo di segreto) e incolla il certificato codificato PEM nel campo Plaintext (Testo normale).

Avvia un cluster EMR con Apache Ranger

Prima di avviare un cluster Amazon EMR con Apache Ranger, assicurati che ogni componente soddisfi i seguenti requisiti minimi di versione:

- Amazon EMR versione 5.32.0 o successive, o 6.3.0 o successive. Consigliamo di utilizzare la versione del rilascio di Amazon EMR più recente.

- Server Admin Apache Ranger 2.x.

Completa questa procedura:

- Installa Apache Ranger se non lo hai già fatto. Per ulteriori informazioni, consulta Installazione di Apache Ranger 0.5.0.
- Assicurati che ci sia connettività di rete tra il tuo cluster Amazon EMR e il server Admin Apache Ranger. Per informazioni, consultare Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR.
- Crea i ruoli IAM necessari. Consultare Ruoli IAM per l'integrazione nativa con Apache Ranger.
- Crea una configurazione di sicurezza EMR per l'installazione di Apache Ranger. Per ulteriori informazioni, consulta Creazione di una configurazione di sicurezza EMR.

Configurazione dei cluster Amazon EMR abilitati da Zeppelin per Apache Ranger

L'argomento illustra come configurare Apache Zeppelin per un cluster Amazon EMR abilitato per Apache Ranger in modo da poter utilizzare Zeppelin come notebook per l'esplorazione interattiva dei dati. Zeppelin è incluso nel rilascio di Amazon EMR versione 5.0.0 e successive. Le versioni precedenti includono Zeppelin come applicazione sandbox. Per ulteriori informazioni, consulta Versioni del rilascio 4.x di Amazon EMR nella Guida ai rilasci di Amazon EMR.

Per impostazione predefinita, Zeppelin è configurato con un log-in e una password di default che non sono sicuri in un ambiente multi-tenant.

Per configurare Zeppelin, completa la procedura riportata di seguito.

1. Modifica del meccanismo di autenticazione.

   Modifica del file `shiro.ini` per implementare il meccanismo di autenticazione preferito. Zeppelin supporta Active Directory, LDAP, PAM e Knox SSO. Consulta Autenticazione Apache Shiro per Apache Zeppelin per ulteriori informazioni.
2. Configurazione di Zeppelin per rappresentare l'utente finale

   Consentire a Zeppelin di rappresentare l'utente finale fa sì che i processi inviati da Zeppelin vengano eseguiti come utente finale. Aggiungi la seguente configurazione a `core-site.xml`:

   ```
   [
   ```

```
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.zeppelin.hosts": "*",
      "hadoop.proxyuser.zeppelin.groups": "*"
    },
    "Configurations": [
    ]
  }
]
```

Quindi, aggiungi la seguente configurazione a `hadoop-kms-site.xml` in `/etc/hadoop/conf`:

```
[
  {
    "Classification": "hadoop-kms-site",
    "Properties": {
      "hadoop.kms.proxyuser.zeppelin.hosts": "*",
      "hadoop.kms.proxyuser.zeppelin.groups": "*"
    },
    "Configurations": [
    ]
  }
]
```

Puoi anche aggiungere queste configurazioni al tuo cluster Amazon EMR tramite la console seguendo la procedura descritta in Riconfigurazione di un gruppo di istanze nella console.

3. Consenti a Zeppelin di eseguire il comando sudo come utente finale

   Crea un file `/etc/sudoers.d/90-zeppelin-user` che contenga quanto segue:

   ```
   zeppelin ALL=(ALL) NOPASSWD:ALL
   ```

4. Modifica le impostazioni degli interpreti per eseguire i processi degli utenti nei propri processi.

   Per tutti gli interpreti, configurali per creare un'istanza degli interpreti "per user" (per utente) nei processi "isolated" (isolati).

5. Modifica **`zeppelin-env.sh`**

Aggiungi quanto segue a `zeppelin-env.sh` in modo che Zeppelin inizi ad avviare interpreti come utente finale:

```
ZEPPELIN_IMPERSONATE_USER=`echo ${ZEPPELIN_IMPERSONATE_USER} | cut -d @ -f1`
export ZEPPELIN_IMPERSONATE_CMD='sudo -H -u ${ZEPPELIN_IMPERSONATE_USER} bash -c'
```

Aggiungi quanto segue a `zeppelin-env.sh` per modificare le autorizzazioni predefinite del notebook in sola lettura solo per l'autore:

```
export ZEPPELIN_NOTEBOOK_PUBLIC="false"
```

Infine, aggiungete quanto segue `zeppelin-env.sh` per includere il percorso della RecordServer classe EMR dopo la prima `CLASSPATH` istruzione:

```
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. Riavvia Zeppelin.

   Per riavviare Zeppelin, esegui il comando seguente:

   ```
   sudo systemctl restart zeppelin
   ```

Problemi noti per l'integrazione con Amazon EMR

Problemi noti

C'è un problema noto all'interno del rilascio di Amazon EMR 5.32 in cui le autorizzazioni per `hive-site.xml` sono state modificate in modo che solo gli utenti privilegiati possano leggerlo, in quanto potrebbero esserci credenziali memorizzate al suo interno. Ciò potrebbe impedire a Hue di leggere `hive-site.xml` e fare in modo che le pagine Web vengano ricaricate continuamente. Se si verificano problemi, aggiungi la seguente configurazione per risolvere il problema:

```
[
  {
```

```
    "Classification": "hue-ini",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "desktop",
        "Properties": {
          "server_group":"hive_site_reader"
         },
        "Configurations":[
        ]
      }
    ]
  }
]
```

Sussiste un problema noto per cui il plug-in EMRFS S3 per Apache Ranger attualmente non supporta la caratteristica Security Zone di Apache Ranger. Le restrizioni per il controllo degli accessi definite utilizzando la funzione Security Zone non vengono applicate ai cluster Amazon EMR.

Applicazione UIs

Per impostazione predefinita, l'interfaccia utente di un'applicazione non esegue l'autenticazione. Ciò include l'ResourceManager interfaccia utente, l'interfaccia NodeManager utente, l'interfaccia utente Livy, tra gli altri. Inoltre, qualsiasi utente che abbia la possibilità di accedere a UIs è in grado di visualizzare le informazioni sui lavori di tutti gli altri utenti.

Se questo comportamento non è desiderato, è necessario assicurarsi che venga utilizzato un gruppo di sicurezza per limitare l'accesso all'applicazione UIs da parte degli utenti.

Autorizzazioni HDFS predefinite

Per impostazione predefinita, agli oggetti creati dagli utenti in HDFS vengono concesse autorizzazioni leggibili a livello mondiale. Ciò può tradursi nella leggibilità dei dati da parte degli utenti che non dovrebbero avervi accesso. Per modificare questo comportamento in modo che le autorizzazioni predefinite dei file siano impostate per la lettura e la scrittura solo dall'autore del processo, esegui questa procedura.

Quando si crea il cluster EMR, fornisci la seguente configurazione:

```
[
  {
```

```
    "Classification": "hdfs-site",
    "Properties": {
      "dfs.namenode.acls.enabled": "true",
      "fs.permissions.umask-mode": "077",
      "dfs.permissions.superusergroup": "hdfsadmingroup"
    }
  }
]
```

Inoltre, esegui la seguente operazione di bootstrap:

```
--bootstrap-actions Name='HDFS UMask Setup',Path=s3://elasticmapreduce/hdfs/umask/
umask-main.sh
```

Plugin Apache Ranger per scenari di integrazione Amazon EMR

I plug-in Apache Ranger convalidano l'accesso di un utente rispetto alle policy di autorizzazione definite nel server amministratore delle policy di Apache Ranger.

Argomenti

- Plugin Apache Hive per l'integrazione di Ranger con Amazon EMR
- Plugin Apache Spark per l'integrazione di Ranger con Amazon EMR
- Plugin EMRFS S3 per l'integrazione di Ranger con Amazon EMR
- Plugin Trino per l'integrazione di Ranger con Amazon EMR

Plugin Apache Hive per l'integrazione di Ranger con Amazon EMR

Apache Hive è un motore di esecuzione popolare all'interno dell'ecosistema Hadoop. Amazon EMR fornisce un plug-in Apache Ranger per essere in grado di fornire controlli di accesso granulare per Hive. Il plug-in è compatibile con il server open source Apache Ranger Admin versione 2.0 e successive.

Argomenti

- Funzionalità supportate
- Installazione della configurazione del servizio
- Considerazioni

- Limitazioni

Funzionalità supportate

Il plug-in Apache Ranger per Hive su EMR supporta tutte le funzionalità del plug-in open source, che include database, tabella, controlli di accesso a livello di colonna, filtraggio riga e mascheramento dei dati. Per una tabella dei comandi Hive e delle autorizzazioni Ranger associate, consulta Comandi Hive per la mappatura delle autorizzazioni Ranger.

Installazione della configurazione del servizio

Il plugin Apache Hive è compatibile con la definizione del servizio Hive esistente all'interno di Apache Hive Hadoop SQL.

Se non disponi di un'istanza del servizio in Hadoop SQL, come mostrato sopra, puoi crearne una. Fai clic sul pulsante + accanto ad Hadoop SQL.

1. Nome del servizio (se visualizzato): immetti il nome del servizio. Il valore suggerito è **`amazonemrhive`**. Prendi nota di questo nome del servizio: è necessario quando si crea una configurazione di sicurezza EMR.
2. Nome visualizzato: immetti il nome da visualizzare per il servizio. Il valore suggerito è **`amazonemrhive`**.

Le proprietà di configurazione di Apache Hive vengono utilizzate per stabilire una connessione al server di amministrazione Apache Ranger con un 2 per HiveServer implementare il completamento automatico durante la creazione delle politiche. Non è necessario che le proprietà seguenti siano accurate se non si dispone di un processo HiveServer 2 persistente e possono essere compilate con qualsiasi informazione.

- Nome utente: inserisci un nome utente per la connessione JDBC a un'istanza di un'istanza HiveServer 2.
- Password: inserisci la password per il nome utente sopra.
- jdbc.driver. ClassName: Immettere il nome della classe JDBC per la connettività Apache Hive. Puoi utilizzare il valore predefinito.

- jdbc.url: Immettere la stringa di connessione JDBC da utilizzare per la connessione a 2. HiveServer
- Nome comune per certificato: il campo CN all'interno del certificato utilizzato per connettersi al server Admin da un plug-in client. Questo valore deve corrispondere al campo CN nel certificato TLS creato per il plug-in.

Il pulsante Test Connection verifica se i valori sopra riportati possono essere utilizzati per connettersi correttamente all'istanza 2. HiveServer Una volta che il servizio è stato creato correttamente, il Service Manager dovrebbe avere il seguente aspetto:

Considerazioni

Server dei metadati Hive

Il server dei metadati Hive è accessibile solo dai motori attendibili, in particolare Hive e `emr_record_server`, come misura di protezione da accessi non autorizzati. Il server dei metadati Hive è accessibile anche da tutti i nodi del cluster. La porta 9083 richiesta consente a tutti i nodi di accedere al nodo principale.

Autenticazione

Per impostazione predefinita, Apache Hive è configurato per l'autenticazione tramite Kerberos come configurato nella configurazione di sicurezza EMR. HiveServer2 può essere configurato per autenticare gli utenti anche tramite LDAP. Consulta Implementazione dell'autenticazione LDAP per Hive su un cluster Amazon EMR multi-tenant per maggiori informazioni.

Limitazioni

Di seguito sono riportate le attuali limitazioni per il plug-in Apache Hive su Amazon EMR 5.x:

- I ruoli Hive non sono attualmente supportati. Le istruzioni Grant (Concedi) e Revoke (Revoca) non sono supportate.
- La CLI di Hive non è supportata. JDBC/Beeline è l'unico modo autorizzato per connettere Hive.
- `hive.server2.builtin.udf.blacklist`la configurazione deve essere compilata con UDFs ciò che ritieni non sicuro.

Plugin Apache Spark per l'integrazione di Ranger con Amazon EMR

Amazon EMR ha integrato EMR per fornire un controllo granulare degli RecordServer accessi per SparkSQL. EMR RecordServer è un processo privilegiato in esecuzione su tutti i nodi di un cluster abilitato per Apache Ranger. Quando un driver o un executor Spark esegue un'istruzione SparkSQL, tutte le richieste di metadati e dati passano attraverso. RecordServer Per ulteriori informazioni su EMR RecordServer, consulta la Componenti Amazon EMR da utilizzare con Apache Ranger pagina.

Argomenti

- Funzionalità supportate
- Ridistribuire la definizione del servizio per utilizzare le istruzioni INSERT, ALTER o DDL
- Installazione della definizione del servizio
- Creazione di policy SparkSQL
- Considerazioni
- Limitazioni

Funzionalità supportate

Azione SQL statement/Ranger	STATUS	Rilascio di EMR supportato
SELECT	Supportato	A partire da 5.32
SHOW DATABASES	Supportato	A partire da 5.32
SHOW COLUMNS	Supportato	A partire da 5.32
SHOW TABLES	Supportato	A partire da 5.32
SHOW TABLE PROPERTIES (MOSTRA PROPRIETÀ TABELLA)	Supportato	A partire da 5.32
DESCRIBE TABLE	Supportato	A partire da 5.32

Azione SQL statement/Ranger	STATUS	Rilascio di EMR supportato
INSERT OVERWRITE	Supportato	A partire da 5.34 e 6.4
INSERT INTO	Supportato	A partire da 5.34 e 6.4
ALTER TABLE	Supportato	A partire da 6.4
CREATE TABLE	Supportato	A partire da 5.35 e 6.7
CREATE DATABASE	Supportato	A partire da 5.35 e 6.7
DROP TABLE	Supportato	A partire da 5.35 e 6.7
DROP DATABASE	Supportato	A partire da 5.35 e 6.7
DROP VIEW	Supportato	A partire da 5.35 e 6.7
CREATE VIEW	Non supportato	

Quando si utilizza SparkSQL sono supportate le seguenti caratteristiche:

- Controllo granulare degli accessi sulle tabelle all'interno del metastore Hive; le policy possono essere create a livello di database, tabella e colonna.
- Le policy di Apache Ranger possono includere policy di concessione e di negazione a utenti e gruppi.
- Gli eventi di controllo vengono inviati ai CloudWatch registri.

Ridistribuire la definizione del servizio per utilizzare le istruzioni INSERT, ALTER o DDL

> **Note**
>
> A partire da Amazon EMR 6.4, puoi utilizzare Spark SQL con le istruzioni: INSERT INTO, INSERT OVERWRITE o ALTER TABLE. A partire da Amazon EMR 6.7, puoi utilizzare Spark

SQL per creare o rimuovere database e tabelle. Se si dispone di un'installazione esistente sul server Apache Ranger con le definizioni del servizio Apache Spark implementate, utilizzare il codice seguente per ridistribuire le definizioni del servizio.

```
# 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>'
```

Installazione della definizione del servizio

L'installazione della definizione del servizio Apache Spark di EMR richiede l'installazione del server Admin Ranger. Consultare Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR.

Segui queste fasi per installare la definizione del servizio Apache Spark:

Fase 1: SSH nel server Admin Apache Ranger.

Ad esempio:

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

Fase 2: Download della definizione del servizio e del plug-in del server Admin Apache Ranger

In una directory temporanea, scarica la definizione del servizio. Questa definizione del servizio è supportata dalle versioni Ranger 2.x.

```
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
```

Fase 3: Installazione del plug-in Apache Spark per 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
```

Fase 4: Registrazione della definizione del servizio Apache Spark per 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'
```

Se questo comando viene eseguito correttamente, viene visualizzato un nuovo servizio nell'interfaccia utente del server Admin Ranger denominato "AMAZON-EMR-SPARK", come mostrato nell'immagine seguente (Ranger versione 2.0).

Fase 5: Creare un'istanza dell'applicazione AMAZON-EMR-SPARK

Nome del servizio (se visualizzato): il nome del servizio che verrà utilizzato. Il valore suggerito è **`amazonemrspark`**. Prendi nota di questo nome del servizio dal momento che sarà necessario quando crei una configurazione di sicurezza EMR.

Nome visualizzato: il nome da visualizzare per questa istanza. Il valore suggerito è **`amazonemrspark`**.

Nome comune per certificato: il campo CN all'interno del certificato utilizzato per connettersi al server Admin da un plug-in client. Questo valore deve corrispondere al campo CN nel certificato TLS creato per il plug-in.

> **Note**
>
> Il certificato TLS per questo plug-in dovrebbe essere stato registrato nel truststore sul server Admin Ranger. Per ulteriori dettagli, consulta Certificati TLS per l'integrazione di Apache Ranger con Amazon EMR.

Creazione di policy SparkSQL

Quando si crea una nuova policy, i campi da compilare sono i seguenti:

Nome policy: il nome della policy.

Etichetta policy: un'etichetta che è possibile inserire in questa policy.

Database: il database a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutti i database.

Tabella: le tabelle a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutte le tabelle.

Colonna Spark EMR: le colonne a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutte le colonne.

Description (Descrizione): la descrizione di questa policy.

Per specificare gli utenti e i gruppi, immetti gli utenti e i gruppi riportati di seguito per concedere le autorizzazioni. È inoltre possibile specificare delle esclusioni per consentire e negare le condizioni.

Dopo aver specificato le condizioni di autorizzazione e negazione, fai clic su Save (Salva).

Considerazioni

Ogni nodo all'interno del cluster EMR deve essere in grado di connettersi al nodo principale sulla porta 9083.

Limitazioni

Di seguito sono riportate le attuali limitazioni per il plug-in Apache Spark:

- Record Server si connette sempre ad HMS in esecuzione su un cluster Amazon EMR. Configura HMS per la connessione alla modalità remota, se necessario. Non inserire i valori di configurazione all'interno del file di configurazione Hive-site.xml di Apache Spark.
- Le tabelle create utilizzando le origini dati Spark su CSV o Avro non sono leggibili utilizzando EMR. RecordServer Utilizza Hive per creare e scrivere dati, per poi leggerli utilizzando Record.
- Le tabelle Delta Lake, Hudi e Iceberg non sono supportate.
- Gli utenti devono avere accesso al database predefinito. Questo è un requisito per Apache Spark.
- Il server Admin Ranger non supporta il completamento automatico.
- Il plug-in SparkSQL per Amazon EMR non supporta i filtri di riga o il mascheramento dei dati.
- Quando si utilizza ALTER TABLE con Spark SQL, una posizione di partizione deve essere la directory figlio di una posizione di tabella. L'inserimento di dati in una partizione in cui la posizione della partizione è diversa da quella della tabella non è supportata.

Plugin EMRFS S3 per l'integrazione di Ranger con Amazon EMR

Per semplificare la fornitura di controlli degli accessi agli oggetti in S3 in un cluster multi-tenant, il plug-in EMRFS S3 fornisce controlli degli accessi ai dati all'interno di S3 quando vi si accede tramite EMRFS. È possibile consentire l'accesso alle risorse S3 a livello di utente e gruppo.

Per raggiungere questo obiettivo, quando l'applicazione tenta di accedere ai dati all'interno di S3, EMRFS invia una richiesta di credenziali al processo Secret Agent, in cui la richiesta viene autenticata e autorizzata rispetto a un plug-in Apache Ranger. Se la richiesta è autorizzata, l'agente segreto assume il ruolo IAM per Apache Ranger Engines con una policy limitata per generare credenziali che hanno accesso solo alla policy Ranger che ha consentito l'accesso. Le credenziali vengono quindi passate a EMRFS per accedere a S3.

Argomenti

- Funzionalità supportate
- Installazione della configurazione del servizio
- Creazione di policy EMRFS S3
- Note sull'utilizzo delle policy EMRFS S3
- Limitazioni

Funzionalità supportate

Il plug-in EMRFS S3 fornisce l'autorizzazione a livello di archiviazione. È possibile creare policy che consentano l'accesso di utenti e gruppi a bucket e prefissi S3. L'autorizzazione è gestita solo rispetto a EMRFS.

Installazione della configurazione del servizio

Per installare la definizione del servizio EMRFS, devi configurare il server di amministrazione Ranger. Per configurare il server, vedere. Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR

Attenersi alla seguente procedura per installare la definizione del servizio EMRFS.

Fase 1: SSH nel server Admin Apache Ranger.

Per esempio:

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

Fase 2: Scarica la definizione del servizio EMRFS.

In una directory temporanea, scarica la definizione del servizio Amazon EMR. Questa definizione del servizio è supportata dalle versioni Ranger 2.x.

```
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/version-2.0/
ranger-servicedef-amazon-emr-emrfs.json
```

Fase 3: Registrare la definizione del servizio 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'
```

Se questo comando viene eseguito correttamente, viene visualizzato un nuovo servizio nell'interfaccia utente del server Admin Ranger denominato "AMAZON-EMR-S3", come mostrato nell'immagine seguente (Ranger versione 2.0).

Fase 4: Creare un'istanza dell'applicazione. AMAZON-EMR-EMRFS

Crea un'istanza della definizione del servizio.

- Fai clic sul segno + accanto a AMAZON-EMR-EMRFS.

Riempi i seguenti campi:

Nome del servizio (se visualizzato): il valore suggerito è**`amazonemrs3`**. Prendi nota di questo nome del servizio dal momento che sarà necessario quando crei una configurazione di sicurezza EMR.

Nome visualizzato: immetti il nome da visualizzare per il servizio. Il valore suggerito è **`amazonemrs3`**.

Nome comune per certificato: il campo CN all'interno del certificato utilizzato per connettersi al server Admin da un plug-in client. Questo valore deve corrispondere al campo CN nel certificato TLS creato per il plug-in.

> **Note**
>
> Il certificato TLS per questo plug-in dovrebbe essere stato registrato nel truststore sul server Admin Ranger. Per ulteriori dettagli, consulta Certificati TLS per l'integrazione di Apache Ranger con Amazon EMR.

Quando viene creato il servizio, il Service Manager include "AMAZON-EMR-EMRFS", come mostrato nell'immagine seguente.

Creazione di policy EMRFS S3

Per creare una nuova policy, nella pagina Create Policy (Crea policy) del Service Manager compila i campi riportati di seguito.

Nome policy: il nome della policy.

Etichetta policy: un'etichetta che è possibile inserire in questa policy.

Risorsa S3: una risorsa che inizia con il bucket e il prefisso facoltativo. Per ulteriori informazioni sulle best practice, consulta Note sull'utilizzo delle policy EMRFS S3. Le risorse nel server Admin Ranger non devono contenere **`s3://`**, **`s3a://`** o **`s3n://`**.

È possibile specificare utenti e gruppi per concedere le autorizzazioni. È inoltre possibile specificare delle esclusioni per consentire e negare le condizioni.

> **Note**
>
> Sono consentite al massimo tre risorse per ogni policy. L'aggiunta di più di tre risorse può causare un errore quando questa policy viene utilizzata in un cluster EMR. L'aggiunta di più di tre policy consente di visualizzare un promemoria relativo al limite di policy.

Note sull'utilizzo delle policy EMRFS S3

Quando si creano policy S3 all'interno di Apache Ranger, occorre tenere a mente alcune considerazioni sull'utilizzo.

Autorizzazioni a più oggetti S3

È possibile utilizzare policy ricorrenti ed espressioni con caratteri jolly per concedere autorizzazioni a più oggetti S3 con prefissi comuni. Le policy ricorrenti forniscono autorizzazioni a tutti gli oggetti con un prefisso comune. Le espressioni con caratteri jolly selezionano più prefissi. Insieme, forniscono le autorizzazioni a tutti gli oggetti con più prefissi comuni, come mostrato nei seguenti esempi.

Example Utilizzo di una policy ricorrente

Supponiamo che necessiti di autorizzazioni per elencare tutti i file parquet in un bucket S3 organizzato come segue.

```
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
```

Innanzitutto, considera i file parquet con il prefisso `s3://sales-reports/americas/year=2000`. Puoi concedere GetObject le autorizzazioni a tutti in due modi:

Utilizzo di policy non ricorrenti: un'opzione consiste nell'utilizzare due policy non ricorrenti separate, una per la directory e l'altra per i file.

La prima policy concede l'autorizzazione al prefisso `s3://sales-reports/americas/year=2020` (non è presente il `/` finale).

```
- S3 resource = "sales-reports/americas/year=2000"
- permission = "GetObject"
- user = "analyst"
```

La seconda policy utilizza l'espressione jolly per concedere autorizzazioni a tutti i file con prefisso `sales-reports/americas/year=2020/` (nota il `/` finale).

```
- S3 resource = "sales-reports/americas/year=2020/*"
- permission = "GetObject"
- user = "analyst"
```

Utilizzo di una policy ricorrente: un'alternativa più conveniente consiste nell'utilizzare una singola policy ricorrente e concedere l'autorizzazione ricorrente al prefisso.

```
 - S3 resource = "sales-reports/americas/year=2020"
- permission = "GetObject"
- user = "analyst"
- is recursive = "True"
```

Finora, solo i file parquet con il prefisso `s3://sales-reports/americas/year=2000` sono stati inclusi. È ora possibile includere anche i file parquet con un prefisso diverso, `s3://sales-`

`reports/americas/year=2020`, nella stessa policy ricorrente introducendo un'espressione con caratteri jolly come segue.

```
- S3 resource = "sales-reports/americas/year=20?0"
- permission = "GetObject"
- user = "analyst"
- is recursive = "True"
```

Politiche PutObject e DeleteObject autorizzazioni

La scrittura di politiche `PutObject` e `DeleteObject` autorizzazioni per i file su EMRFS richiede un'attenzione speciale perché, a differenza delle autorizzazioni, richiedono GetObject autorizzazioni ricorsive aggiuntive concesse al prefisso.

Example Politiche PutObject DeleteObject e autorizzazioni

Ad esempio, l'eliminazione del file non `annual-summary.parquet` richiede solo l' DeleteObject autorizzazione per il file effettivo.

```
- S3 resource = "sales-reports/americas/year=2020/annual-summary.parquet"
- permission = "DeleteObject"
- user = "analyst"
```

ma anche una policy che conceda autorizzazioni ricorrenti `GetObject` e `PutObject` per il suo prefisso.

Allo stesso modo, la modifica del file `annual-summary.parquet` richiede non solo un'autorizzazione `PutObject` per il file effettivo,

```
- S3 resource = "sales-reports/americas/year=2020/annual-summary.parquet"
- permission = "PutObject"
- user = "analyst"
```

ma anche una policy che conceda l'autorizzazione ricorrente `GetObject` per il suo prefisso.

```
- S3 resource = "sales-reports/americas/year=2020"
- permission = "GetObject"
- user = "analyst"
- is recursive = "True"
```

Caratteri jolly nelle policy

Sono disponibili due aree in cui è possibile specificare i caratteri jolly. Quando si specifica una risorsa S3, è possibile utilizzare i valori "*" e "?". "*" fornisce la corrispondenza con un percorso S3 e corrisponde a tutto ciò che viene dopo il prefisso. Per esempio, la seguente policy.

```
S3 resource = "sales-reports/americas/*"
```

Questo corrisponde ai seguenti percorsi 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
```

Il carattere jolly "?" corrisponde a un singolo carattere. Ad esempio, per la policy:

```
S3 resource = "sales-reports/americas/year=201?/"
```

Questo corrisponde ai seguenti percorsi S3.

```
sales-reports/americas/year=2019/
sales-reports/americas/year=2018/
sales-reports/americas/year=2017/
```

Caratteri jolly negli utenti

Esistono due caratteri jolly incorporati quando si assegnano utenti per consentire l'accesso agli utenti. Il primo è il carattere jolly "{USER}" che fornisce l'accesso a tutti gli utenti. Il secondo carattere jolly è "{OWNER}" che fornisce l'accesso diretto o l'accesso al proprietario di un particolare oggetto. Tuttavia, il carattere jolly "{USER}" non è attualmente supportato.

Limitazioni

Di seguito sono riportate le attuali limitazioni per il plug-in EMRFS S3:

- Apache Ranger può avere al massimo tre policy.
- L'accesso a S3 deve essere realizzato tramite EMRFS e può essere utilizzato con le applicazioni correlate ad Hadoop. Il seguente elemento non è supportato:

- Librerie Boto3

- AWS SDK e AWK CLI

- Connettore open source S3A

- Le policy di negazione di Apache Ranger non sono supportate.
- Le operazioni su S3 con chiavi con crittografia CSE-KMS non sono attualmente supportate.
- Il supporto tra Regioni non è supportato.
- La caratteristica Security Zone di Apache Ranger non è supportata. Le restrizioni per il controllo degli accessi definite utilizzando la funzione Security Zone non vengono applicate ai cluster Amazon EMR.
- L'utente Hadoop non genera alcun evento di controllo poiché Hadoop accede sempre al profilo dell'istanza. EC2
- Si consiglia di disabilitare la visualizzazione di coerenza Amazon EMR. S3 ha un'elevata coerenza, pertanto tale visualizzazione non è più necessaria. Per maggiori informazioni, consulta Forte coerenza di Amazon S3.
- Il plug-in EMRFS S3 effettua numerose chiamate STS. Si consiglia di eseguire test di carico su un account di sviluppo e monitorare il volume delle chiamate STS. Si consiglia inoltre di effettuare una richiesta STS per aumentare i limiti del servizio. AssumeRole
- Il server Ranger Admin non supporta il completamento automatico.

Plugin Trino per l'integrazione di Ranger con Amazon EMR

Trino (precedentemente PrestoSQL) è un motore di query SQL che consente di eseguire query su più origini dati, come HDFS, archiviazione di oggetti, database relazionali e database NoSQL. Elimina la necessità di migrare i dati in una posizione centrale e consente di interrogare i dati ovunque si trovino. Amazon EMR mette a disposizione un plugin Apache Ranger per fornire un controllo granulare degli accessi a Trino. Il plug-in è compatibile con il server open source Apache Ranger Admin versione 2.0 e successive.

Argomenti

- Funzionalità supportate
- Installazione della configurazione del servizio
- Creazione di policy Trino

- Considerazioni
- Limitazioni

Funzionalità supportate

Il plugin Apache Ranger per Trino su Amazon EMR supporta tutte le funzionalità del motore di query Trino, che è protetto da un controllo granulare degli accessi comprendente controlli degli accessi a livello di database, tabella e colonna, filtraggio di riga e mascheramento dei dati. Le policy di Apache Ranger possono includere policy di concessione e di negazione a utenti e gruppi. Gli eventi di controllo vengono inoltre inviati ai log. CloudWatch

Installazione della configurazione del servizio

L'installazione della definizione del servizio Trino richiede la configurazione del server Admin Ranger. Per configurare il server Admin Ranger, consulta Configura un server di amministrazione Ranger per l'integrazione con Amazon EMR.

Attenersi alla seguente procedura per installare la definizione del servizio Trino.

1. SSH nel server Admin Apache Ranger.

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

2. Disinstalla il plugin del server Presto, se presente. Esegui il comando seguente. Se viene visualizzato l'errore "Service not found" (Servizio non trovato), il plugin del server Presto non è stato installato sul server. Passare alla fase successiva.

```
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. Scarica la definizione del servizio e il plugin del server Admin Apache Ranger. In una directory temporanea, scarica la definizione del servizio. Questa definizione del servizio è supportata dalle versioni Ranger 2.x.

```
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/
version-2.0/ranger-servicedef-amazon-emr-trino.json
```

4. Registra la definizione del servizio Apache Trino per 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'
```

Se questo comando viene eseguito correttamente, viene visualizzato un nuovo servizio nell'interfaccia utente di Admin Apache Ranger denominato `TRINO`, come mostrato nell'immagine seguente.

5. Crea un'istanza dell'applicazione `TRINO` inserendo le informazioni seguenti.

   Service Name (Nome del servizio): il nome del servizio che verrà utilizzato. Il valore suggerito è `amazonemrtrino`. Prendi nota di questo nome del servizio dal momento che sarà necessario durante la creazione di una configurazione di sicurezza per Amazon EMR.

   Nome visualizzato: il nome da visualizzare per questa istanza. Il valore suggerito è `amazonemrtrino`.

   jdbc.driver. ClassName: Il nome della classe JDBC per la connettività Trino. Puoi usare il valore predefinito.

   jdbc.url: la stringa di connessione JDBC da utilizzare per la connessione a un coordinatore Trino.

   Nome comune per certificato: il campo CN all'interno del certificato utilizzato per connettersi al server Admin da un plug-in client. Questo valore deve corrispondere al campo CN nel certificato TLS creato per il plug-in.

   Il certificato TLS per questo plugin dovrebbe essere stato registrato nel trust store sul server Admin Ranger. Per ulteriori informazioni, consulta la sezione Certificati TLS.

Creazione di policy Trino

Quando crei una nuova policy, compila i campi seguenti.

Nome policy: il nome della policy.

Etichetta policy: un'etichetta che è possibile inserire in questa policy.

Catalog (Catalogo): il catalogo a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutti i cataloghi.

Schema (Schema): gli schemi a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutti gli schemi.

Tabella: le tabelle a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutte le tabelle.

Column (Colonna): le colonne a cui viene applicata questa policy. Il carattere jolly "*" rappresenta tutte le colonne.

Description (Descrizione): la descrizione di questa policy.

Esistono altri tipi di policy, ad esempio Trino User (Utente Trino) (per l'accesso alla rappresentazione di utenti), Trino System/Session Property (Proprietà sistema/sessione Trino) (per modificare le proprietà del sistema o della sessione del motore), Functions/Procedures (Funzioni/procedure) (per consentire chiamate a funzioni o procedure) e URL (per consentire l'accesso in lettura/scrittura al motore nei percorsi dati).

Per concedere autorizzazioni a utenti e gruppi specifici, inserirli. È inoltre possibile specificare delle esclusioni per consentire e negare le condizioni.

Dopo aver specificato le condizioni di autorizzazione e negazione, scegli Save (Salva).

Considerazioni

Quando si creano policy Trino all'interno di Apache Ranger, occorre tenere a mente alcune considerazioni sull'utilizzo.

Server dei metadati Hive

Il server dei metadati Hive è accessibile solo dai motori attendibili, in particolare il motore Trino, come misura di protezione da accessi non autorizzati. Il server dei metadati Hive è accessibile anche da tutti i nodi del cluster. La porta 9083 richiesta consente a tutti i nodi di accedere al nodo principale.

Autenticazione

Per impostazione predefinita, Trino è configurato per l'autenticazione utilizzando Kerberos, come specificato nella configurazione di sicurezza di Amazon EMR.

Crittografia in transito obbligatoria

Il plugin Trino richiede che la crittografia in transito sia abilitata nella configurazione di sicurezza di Amazon EMR. Per abilitare la crittografia, consulta Crittografia in transito.

Limitazioni

Di seguito sono riportate le attuali limitazioni per il plugin Trino:

- Il server Admin Ranger non supporta il completamento automatico.

Risoluzione dei problemi di Apache Ranger

Seguono alcuni dei problemi comunemente diagnosticati relativi all'utilizzo di Apache Ranger.

Raccomandazioni

- Test utilizzando un cluster a nodo principale singolo: il provisioning dei cluster a nodo master singolo è più rapido rispetto a un cluster a più nodi e può ridurre il tempo per ogni iterazione di test.
- Imposta la modalità di sviluppo sul cluster. Quando avvii il cluster EMR, imposta il parametro `--additional-info"` su:

  ```
  '{"clusterType":"development"}'
  ```

  Questo parametro può essere impostato solo tramite AWS CLI o AWS SDK e non è disponibile tramite la console Amazon EMR. Quando questo contrassegno è impostato e il master non riesce a eseguire il provisioning, il servizio Amazon EMR mantiene il cluster attivo per un po' di tempo prima di disattivarlo. Questo arco temporale è molto utile per sondare i vari file di log prima che il cluster venga terminato.

Insuccesso del provisioning del cluster EMR

Esistono diversi motivi alla base del mancato avvio di un cluster Amazon EMR. Di seguito sono riportati alcuni modi per diagnosticare il problema.

Verifica dei log di provisioning EMR

Amazon EMR utilizza Puppet per installare e configurare le applicazioni in un cluster. Esaminando i registri verranno forniti dettagli sulla presenza o meno di errori durante la fase di provisioning di un cluster. I log sono accessibili sul cluster o su S3 se sono configurati per essere inviati a S3.

I log vengono archiviati in `/var/log/provision-node/apps-phase/0/{UUID}/puppet.log` sul disco e `s3://<LOG LOCATION>/<CLUSTER ID>/node/<EC2 INSTANCE ID>/provision-node/apps-phase/0/{UUID}/puppet.log.gz`..

Messaggi di errore comuni

Messaggio di errore	Causa
Puppet (err): avvio di Systemd non riuscito! emr-record-server log journalctl per: emr-recor d-server	Impossibile avviare EMR Record Server. Consulta "Log di EMR Record Server" di seguito.
Puppet (err): avvio di Systemd non riuscito! emr-record-server log journalctl per emrsecret agent:	Impossibile avviare EMR Secret Agent. Consulta "Controllo dei log Secret Agent" di seguito.
/Stage [main] /ranger_plugins: :ranger_hive_ (avviso): 140408606197664:ERROR:0906D 06C:Routine PEM: PEM_READ_BIO:NESSU NA RIGA DI AVVIO:PEM_LIB.C:707:IN ATTESA: QUALSIASI CHIAVE plugin/Ra nger_plugins::Prepare_two_way_tls[configure 2-way TLS in Hive plugin]/Exec[create keystore and truststore for Ranger Hive plugin]/returns PRIVATA	Il certificato TLS privato in Secrets Manager per il certificato del plug-in Apache Ranger non è nel formato corretto o non è un certificato privato. Consulta Certificati TLS per l'integra zione di Apache Ranger con Amazon EMR per i formati dei certificati.
/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	Il EC2 ruolo del profilo di istanza non dispone delle autorizzazioni corrette per recuperare i certificati TLS da Secrets Agent.

Messaggio di errore	Causa
XX non è autorizzato a eseguire: secretsma nager: GetSecretValue on resource: arn:aws:S ecretsManager:US-EAST-1:xxxxxxxx:secret: - XXXXXXX AdminServer	

Controlla i registri SecretAgent

I log di Secret Agent si trovano in `/emr/secretagent/log/` su un nodo EMR o nella directory `s3://<LOG LOCATION>/<CLUSTER ID>/node/<EC2 INSTANCE ID>/daemons/secretagent/` in S3.

Messaggi di errore comuni

Messaggio di errore	Causa
Eccezione nel thread «main» com.amazo naws.services.securitytoken.model. AWSSecuri tyTokenServiceException: User: arn:aws:s ts: :xxxxxxxxxx:Assumed- role/EMR_ EC2_DefaultRole/i -XXXXXXXXXXX non è autorizzato a eseguire: sts: AssumeRol e on resource: arn:aws:iam: :XXXXXXXX XX:role/* (Service:; Codice di stato: 403; Codice di errore:; Codice di errore:; ID richiesta : XXXXXXXX-XXXXXX-XXXXXXXXXX; Proxy: nullo RangerPluginDataAccessRole) AWSSecurity TokenService AccessDenied	L'eccezione precedente indica che il ruolo del profilo dell' EC2 istanza EMR non dispone delle autorizzazioni per assumere il ruolo. RangerPlu ginDataAccessRole Consultare Ruoli IAM per l'integrazione nativa con Apache Ranger.
ERROR qtp54617902-149: Web App Exception Occurred javax.ws.rs. NotAllowedException: Metodo HTTP 405 non consentito	Puoi ignorare questi errori senza correre rischi.

Controllo dei log di EMR Record Server (per SparkSQL)

<CLUSTER ID>I log dell'EMR Record Server sono disponibili in/var/log/emr-record-server/ su un nodo EMR oppure possono essere trovati nella directory s3:<LOG LOCATION>////node/ < INSTANCE ID>/daemons//in S3. EC2 emr-record-server

Messaggi di errore comuni

Messaggio di errore	Causa
InstanceMetadataServiceResourceFetcherI log di EMR Record Server sono disponibili in/-record-server/ su un nodo EMR oppure possono essere trovati nella directory s3:////node/ < INSTANCE ID>/daemons//in S3. ----sep----:105 - [] Impossibile recuperare il token com.amazonaws. SdkClientException : connessione all'endpoint del servizio non riuscita	L'EMR SecretAgent non è riuscito a comparire o presenta un problema. Ispeziona SecretAgent i log per individuare eventuali errori e lo script del pupazzo per determinare se ci sono errori di provisioning.

Le query falliscono inaspettatamente per l'integrazione di Ranger con Amazon EMR

Controlla i log dei plugin Apache Ranger (registri di Apache Hive, EMR, EMR, ecc.RecordServer) SecretAgent

Questa sezione è comune a tutte le applicazioni che si integrano con il plug-in Ranger, come Apache Hive, EMR Record Server ed EMR. SecretAgent

Messaggi di errore comuni

Messaggio di errore	Causa
ERROR:272 PolicyRefresher - [] (PolicyRefresherserviceName=policy-repository): impossibile trovare il servizio. Will clean up local cache of policies (-1)	Questo messaggio di errore indica che il nome del servizio fornito nella configurazione di sicurezza EMR non corrisponde a un repositor

Messaggio di errore	Causa
	y delle policy di servizio nel server Admin Ranger.

Se all'interno del server Ranger Admin il tuo AMAZON-EMR-SPARK servizio è simile al seguente, dovresti inserirlo come nome del servizio. **`amazonemrspark`**

Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR

I gruppi di sicurezza fungono da firewall virtuali per le EC2 istanze del cluster per controllare il traffico in entrata e in uscita. Ogni gruppo di sicurezza ha un set di regole che controlla il traffico in entrata verso le istanze e un set di regole per controllare il traffico in uscita. Per ulteriori informazioni, consulta i gruppi EC2 di sicurezza Amazon per le istanze Linux nella Amazon EC2 User Guide.

È possibile utilizzare due classi di gruppi di sicurezza con Amazon EMR: gruppi di sicurezza gestiti da Amazon EMR e gruppi di sicurezza aggiuntivi.

A ogni cluster sono associati gruppi di sicurezza gestiti. È possibile utilizzare i gruppi di sicurezza gestiti predefiniti creati da Amazon EMR oppure specificare gruppi di sicurezza gestiti personalizzati. In entrambi i casi, Amazon EMR aggiunge automaticamente le regole ai gruppi di sicurezza gestiti di cui un cluster ha bisogno per comunicare tra istanze e servizi del cluster. AWS

I gruppi di sicurezza aggiuntivi sono facoltativi. È possibile specificarli in aggiunta ai gruppi di sicurezza gestiti per definire l'accesso alle istanze del cluster. Altri gruppi di sicurezza contengono solo regole definite. Amazon EMR non li modifica.

Le regole che Amazon EMR crea nei gruppi di sicurezza gestiti consentono al cluster di comunicare tra i componenti interni. Per consentire agli utenti e alle applicazioni di accedere a un cluster dall'esterno del cluster, è possibile modificare le regole per i gruppi di sicurezza gestiti oppure è possibile creare ulteriori gruppi di sicurezza con regole aggiuntive o entrambe le cose.

> **Important**
>
> Modificare le regole per i gruppi di sicurezza gestiti può avere conseguenze impreviste. Potresti inavvertitamente bloccare il traffico necessario per il corretto funzionamento dei

cluster e causare errori in quanto i nodi non sono raggiungibili. Pianifica attentamente e verifica le configurazioni per i gruppi di sicurezza prima dell'implementazione.

Puoi, quindi, specificare i gruppi di sicurezza solo al momento della creazione di un cluster. Essi non possono essere aggiunti a un cluster o alle istanze del cluster mentre il cluster è in esecuzione, ma è possibile modificare, aggiungere o rimuovere regole dai gruppi di sicurezza esistenti. Le regole diventano effettive non appena vengono salvate.

Per impostazione predefinita, i gruppi di sicurezza sono restrittivi Se non si aggiunge una regola che consente il traffico, il traffico viene rifiutato. Se è presente più di una regola che si applica allo stesso traffico e alla stessa origine, viene applicata la regola più permissiva. Ad esempio, se si dispone di una regola che consente SSH dall'indirizzo IP 192.0.2.12/32 e un'altra regola che consente l'accesso a tutto il traffico TCP dall'intervallo 192.0.2.0/24, ha precedenza la regola che consente tutto il traffico TCP dall'intervallo che include 192.0.2.12. In questo caso, il client all'intervallo 192.0.2.12 può avere più accesso di quanto si desideri.

Important

Presta attenzione quando modifichi le regole per i gruppi di sicurezza per aprire le porte. Assicurati di aggiungere regole che permettano solo il traffico da client affidabili e autenticati per i protocolli e le porte necessari per eseguire i tuoi carichi di lavoro.

È possibile configurare il blocco degli accessi pubblici per Amazon EMR in ogni Regione utilizzata per impedire la creazione del cluster quando una regola consente l'accesso pubblico su una porta non aggiunta a un elenco di eccezioni. Per AWS gli account creati dopo luglio 2019, l'accesso pubblico a blocchi di Amazon EMR è attivo per impostazione predefinita. Per AWS gli account che hanno creato un cluster prima di luglio 2019, l'accesso pubblico a blocchi di Amazon EMR è disattivato per impostazione predefinita. Per ulteriori informazioni, consulta Utilizzo del blocco dell'accesso pubblico di Amazon EMR.

Argomenti

- Utilizzo dei gruppi di sicurezza gestiti da Amazon EMR
- Utilizzo di gruppi di sicurezza aggiuntivi per un cluster Amazon EMR
- Specifica dei gruppi di sicurezza gestiti da Amazon EMR e aggiuntivi
- Specificazione dei gruppi EC2 di sicurezza per i notebook EMR

- Utilizzo del blocco dell'accesso pubblico di Amazon EMR

> **Note**
>
> Amazon EMR mira a utilizzare alternative inclusive per termini di settore potenzialmente offensivi o non inclusivi come "master" e "slave". Siamo passati a una nuova terminologia per promuovere un'esperienza più inclusiva e facilitare la comprensione dei componenti del servizio.
> Ora descriviamo i "nodi" come istanze e descriviamo i tipi di istanze Amazon EMR come primarioi, principale e attività. Durante la transizione, potresti ancora trovare riferimenti precedenti a termini obsoleti, come quelli relativi ai gruppi di sicurezza per Amazon EMR.

Utilizzo dei gruppi di sicurezza gestiti da Amazon EMR

> **Note**
>
> Amazon EMR mira a utilizzare alternative inclusive per termini di settore potenzialmente offensivi o non inclusivi come "master" e "slave". Siamo passati a una nuova terminologia per promuovere un'esperienza più inclusiva e facilitare la comprensione dei componenti del servizio.
> Ora descriviamo i "nodi" come istanze e descriviamo i tipi di istanze Amazon EMR come primarioi, principale e attività. Durante la transizione, potresti ancora trovare riferimenti precedenti a termini obsoleti, come quelli relativi ai gruppi di sicurezza per Amazon EMR.

I diversi gruppi di sicurezza gestiti sono associati all'istanza primaria e alle istanze principali e attività in un cluster. Quando si crea un cluster in una sottorete privata, è necessario un ulteriore gruppo di sicurezza gestito per l'accesso ai servizi. Per ulteriori informazioni sul ruolo dei gruppi di sicurezza gestiti per quanto riguarda la configurazione di rete, consulta Opzioni Amazon VPC all'avvio di un cluster.

Quando vengono specificati i gruppi di sicurezza gestiti per un cluster, è necessario utilizzare lo stesso tipo di gruppo di sicurezza, predefinito o personalizzato, per tutti i gruppi di sicurezza gestiti. Ad esempio, non è possibile specificare un gruppo di sicurezza personalizzato per l'istanza primaria, quindi non specificare un gruppo di sicurezza personalizzato per le istanze principali e le istanze attività.

Se utilizzi i gruppi di sicurezza gestiti predefiniti, non è necessario specificarli quando si crea un cluster. Amazon EMR utilizza automaticamente le impostazioni predefinite. Inoltre, se le impostazioni predefinite non esistono ancora nel VPC del cluster, Amazon EMR le crea. Amazon EMR le crea anche se vengono specificate in modo esplicito e non esistono ancora.

Puoi modificare le regole nei gruppi di sicurezza gestiti dopo la creazione dei cluster. Quando crei un nuovo cluster, Amazon EMR controlla le regole nei gruppi di sicurezza gestiti specificati e quindi crea tutte le regole in entrata mancanti necessarie per il nuovo cluster, oltre alle regole che possono essere state aggiunte in precedenza. Salvo diversamente specificato, ciascuna regola per gruppi di sicurezza gestiti da Amazon EMR predefiniti viene anche aggiunta a gruppi di sicurezza gestiti da Amazon EMR personalizzati da te specificati.

I gruppi di sicurezza gestiti predefiniti sono i seguenti:

- ElasticMapReduce-primario

  Per le regole in questo gruppo di sicurezza, consulta Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria (sottoreti pubbliche).
- ElasticMapReduce-nucleo

  Per le regole in questo gruppo di sicurezza, consulta Gruppo di sicurezza gestito da Amazon EMR per le istanze principali e attività (sottoreti pubbliche).
- ElasticMapReduce-Primario-Privato

  Per le regole in questo gruppo di sicurezza, consulta Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria (sottoreti private).
- ElasticMapReduce-Privato principale

  Per le regole in questo gruppo di sicurezza, consulta Gruppo di sicurezza gestito da Amazon EMR per le istanze principali e attività (sottoreti private).
- ElasticMapReduce-ServiceAccess

  Per le regole in questo gruppo di sicurezza, consulta Gruppo di sicurezza gestito da Amazon EMR per l'accesso ai servizi (sottoreti private).

Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria (sottoreti pubbliche)

Il gruppo di sicurezza gestito predefinito per l'istanza principale nelle sottoreti pubbliche ha il nome di gruppo -primary. ElasticMapReduce Include le seguenti regole: Se specifichi un gruppo di

sicurezza gestito personalizzato, Amazon EMR aggiungerà le stesse regole al gruppo di sicurezza personalizzato.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Regole in entrata				
Tutto ICMP-IPv4	Tutti	N/D	L'ID del gruppo di sicurezza gestito dell'istanza primaria. In altre parole, lo stesso gruppo di sicurezza in cui appare la regola.	Tali regole riflessive consentono il traffico in entrata da qualsiasi istanza associata al gruppo di sicurezza specificato. Utilizzando `ElasticMapReduce-primary` predefinito per più cluster consente ai nodi principali e di task di tali cluster di comunicare tra loro su ICMP o su qualsiasi porta TCP o UDP. Specifica i gruppi di sicurezza gestiti personalizzati per limitare l'accesso tra cluster.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID del gruppo di sicurezza gestito specificato per nodi principali e di task.	Tali regole consentono il traffico ICMP in entrata e il traffico su qualsiasi porta TCP o UDP da qualsiasi istanza principale e di attività associata al gruppo di sicurezza specificato, anche se le istanze si trovano in cluster diversi.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		
Personalizza	TCP	8443	Vari intervalli di indirizzi IP Amazon	Tali regole consentono al cluster manager di comunicare con il nodo primario.

Per concedere a fonti attendibili l'accesso SSH al gruppo di sicurezza primario con la console

Per modificare i gruppi di sicurezza, devi disporre dell'autorizzazione per gestire i gruppi di sicurezza per il VPC in cui si trova il cluster. Per ulteriori informazioni, consulta Modifica delle autorizzazioni per un utente e la politica di esempio che consente la gestione dei gruppi di EC2 sicurezza nella Guida per l'utente IAM.

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. Scegli Cluster. Scegli l'ID del cluster che desideri modificare.
3. Nel riquadro Rete e sicurezza, espandi il menu a discesa dei gruppi di EC2 sicurezza (firewall).
4. In Nodo primario, scegli il tuo gruppo di sicurezza.
5. Sceglere Edit inbound rules (Modifica regole in entrata).
6. Verifica la presenza di una regola in entrata che consenta l'accesso pubblico con le seguenti impostazioni. Se esiste, scegli Elimina per rimuoverla.

   - Tipo

     SSH
   - Porta

     22
   - Origine

     Personalizzata 0.0.0.0/0

   > **Warning**
   >
   > Prima di dicembre 2020, esisteva una regola preconfigurata per consentire il traffico in entrata sulla porta 22 da tutte le fonti. Questa regola è stata creata per semplificare le connessioni SSH iniziali al nodo primario. Consigliamo vivamente di rimuovere questa regola in entrata e limitare il traffico alle origini affidabili.

7. Scorri fino alla fine dell'elenco di regole e seleziona Aggiungi regola.
8. Per Tipo, seleziona SSH.

   Selezionando SSH si inserisce in automatico TCP per Protocollo e 22 per Intervallo porta.

9. Per origine, seleziona Il mio IP per aggiungere in automatico il tuo indirizzo IP come indirizzo di origine. Puoi anche aggiungere un intervallo di indirizzi IP affidabili del client Custom (Personalizzato), o creare regole aggiuntive per altri client. In molti ambienti di rete, gli indirizzi IP vengono allocati in modo dinamico, perciò, nel futuro, potrebbe essere necessario aggiornare gli indirizzi IP per client affidabili.
10. Scegli Save (Salva).
11. Facoltativamente, scegli l'altro gruppo di sicurezza in Core e task nodes nel pannello Rete e sicurezza e ripeti i passaggi precedenti per consentire al client SSH l'accesso ai nodi principali e task.

Gruppo di sicurezza gestito da Amazon EMR per le istanze principali e attività (sottoreti pubbliche)

Il gruppo di sicurezza gestito predefinito per le istanze core e task nelle sottoreti pubbliche ha il nome di gruppo -core. ElasticMapReduce Il gruppo di sicurezza gestito predefinito ha le regole seguenti e Amazon EMR aggiunge le stesse regole se si specifica un gruppo di sicurezza gestito personalizzato.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Regole in entrata				
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID del gruppo di sicurezza gestito specificato per le istanze principali e di attività. In altre parole, lo stesso gruppo di sicurezza in cui appare la regola.	Tali regole riflessive consentono il traffico in entrata da qualsiasi istanza associata al gruppo di sicurezza specificato. Utilizzando `ElasticMapReduce-core` predefinito per più cluster consente alle istanze principali e di attività di tali cluster di comunicare tra loro su ICMP o su qualsiasi porta TCP o UDP. Specifica i gruppi di sicurezza gestiti personalizzati per limitare l'accesso tra cluster.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID del gruppo di sicurezza gestito dell'istanza primaria.	Tali regole consentono tutto il traffico ICMP in entrata e il traffico su qualsiasi porta TCP o UDP da tutte le istanze primarie associate al gruppo di sicurezza specificato, anche se le istanze si trovano in cluster diversi.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		

Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria (sottoreti private)

Il gruppo di sicurezza gestito predefinito per l'istanza principale nelle sottoreti private ha il nome di gruppo -Primary-Private. ElasticMapReduce Il gruppo di sicurezza gestito predefinito ha le regole seguenti e Amazon EMR aggiunge le stesse regole se si specifica un gruppo di sicurezza gestito personalizzato.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Regole in entrata				
Tutto ICMP-IPv4	Tutti	N/D	L'ID del gruppo di sicurezza gestito dell'istanza primaria. In altre parole, lo stesso gruppo di sicurezza in cui appare la regola.	Tali regole riflessive consentono il traffico in entrata da qualsiasi istanza associata al gruppo di sicurezza specificato e raggiungibile dall'interno di una sottorete privata. Utilizzando `ElasticMapReduce-Primary-Private` predefinito per più cluster consente ai nodi principali e di task di tali cluster di comunicare tra loro su ICMP o su qualsiasi
Tutte le regole TCP	TCP	Tutti		

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Tutte le regole UDP	UDP	Tutti		porta TCP o UDP. Specifica i gruppi di sicurezza gestiti personalizzati per limitare l'accesso tra cluster.
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID gruppo del gruppo di sicurezza gestito specificato per i nodi principali e di task.	Tali regole consentono il traffico ICMP in entrata e il traffico su qualsiasi porta TCP o UDP da qualsiasi istanza principale e di attività associata al gruppo di sicurezza specificato e raggiungibile dalla sottorete privata, anche se le istanze si trovano in cluster diversi.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		
HTTPS (8443)	TCP	8443	L'ID del gruppo di sicurezza gestito per l'accesso ai servizi in una sottorete privata.	Questa regola consente al cluster manager di comunicare con il nodo primario.
Regole in uscita				
Tutto il traffico	Tutti	Tutti	0.0.0.0/0	Fornisce l'accesso in uscita a Internet.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
TCP personalizzato	TCP	9443	L'ID del gruppo di sicurezza gestito per l'accesso ai servizi in una sottorete privata.	Tieni presente che la regola in uscita predefinita "Tutto il traffico" sopra descritta è un requisito minimo per Amazon EMR 5.30.0 e versioni successive. **Note** Amazon EMR non aggiunge questa regola quando si utilizza un gruppo di sicurezza gestito personalizzato.
TCP personalizzato	TCP	80 (http) o 443 (https)	L'ID del gruppo di sicurezza gestito per l'accesso ai servizi in una sottorete privata.	Se la regola in uscita predefinita "Tutto il traffico" sopra descritta viene rimossa, tieni presente che si tratta di un requisito minimo per Amazon EMR 5.30.0 e versioni successive e per la connessione ad Amazon S3 tramite https. **Note** Amazon EMR non aggiunge questa regola quando si utilizza un gruppo di sicurezza gestito personalizzato.

Gruppo di sicurezza gestito da Amazon EMR per le istanze principali e attività (sottoreti private)

Il gruppo di sicurezza gestito predefinito per le istanze core e task nelle sottoreti private ha il nome di gruppo -Core-Private. ElasticMapReduce Il gruppo di sicurezza gestito predefinito ha le regole seguenti e Amazon EMR aggiunge le stesse regole se si specifica un gruppo di sicurezza gestito personalizzato.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Regole in entrata				
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID del gruppo di sicurezza gestito specificato per le istanze principali e di attività. In altre parole, lo stesso gruppo di sicurezza in cui appare la regola.	Tali regole riflessive consentono il traffico in entrata da qualsiasi istanza associata al gruppo di sicurezza specificato. Utilizzando `ElasticMapReduce-core` predefinito per più cluster consente alle istanze principali e di attività di tali cluster di comunicare tra loro su ICMP o su qualsiasi porta TCP o UDP. Specifica i gruppi di sicurezza gestiti personalizzati per limitare l'accesso tra cluster.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		
Tutti gli ICMP-IPV4	Tutti	N/D	L'ID del gruppo di sicurezza gestito dell'istanza primaria.	Tali regole consentono tutto il traffico ICMP in entrata e il traffico su qualsiasi porta TCP o UDP da tutte le istanze primarie associate al gruppo di sicurezza specificato, anche se le istanze si trovano in cluster diversi.
Tutte le regole TCP	TCP	Tutti		
Tutte le regole UDP	UDP	Tutti		
HTTPS (8443)	TCP	8443	L'ID del gruppo di sicurezza gestito per l'accesso ai servizi in una sottorete privata.	Questa regola consente al cluster manager di comunicare con i nodi principali e di task.
Regole in uscita				

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Tutto il traffico	Tutti	Tutti	0.0.0.0/0	Consulta Modifica di regole in uscita qui di seguito.
TCP personalizzato	TCP	80 (http) o 443 (https)	L'ID del gruppo di sicurezza gestito per l'accesso ai servizi in una sottorete privata.	Se la regola in uscita predefinita "Tutto il traffico" sopra descritta viene rimossa, tieni presente che si tratta di un requisito minimo per Amazon EMR 5.30.0 e versioni successive per la connessione ad Amazon S3 tramite https. **Note** Amazon EMR non aggiunge questa regola quando si utilizza un gruppo di sicurezza gestito personalizzato.

Modifica di regole in uscita

Per impostazione predefinita, Amazon EMR crea questo gruppo di sicurezza con regole in uscita che consentono tutto il traffico in uscita su protocolli e porte. L'autorizzazione di tutto il traffico in uscita è selezionata perché diverse applicazioni Amazon EMR e cliente eseguibili su cluster Amazon EMR potrebbero richiedere regole di uscita diverse. Amazon EMR non è in grado di prevedere queste impostazioni specifiche durante la creazione di gruppi di sicurezza predefiniti. È possibile definire l'ambito in uscita nei gruppi di sicurezza in modo da includere solo le regole che si adattano ai casi d'uso e alle policy di sicurezza. Questo gruppo di sicurezza richiede almeno le seguenti regole in uscita, ma alcune applicazioni potrebbero richiedere un'uscita aggiuntiva.

Tipo	Protocollo	Intervallo porte	Destinazione	Informazioni
Tutte le regole TCP	TCP	Tutti	pl- *xxxxxxxx*	Elenco dei prefissi Amazon S3 gestiti `com.amazonaws.` *MyRegion*`.s3`.
All Traffic	Tutti	Tutti	sg- *xxxxxxxxx xxxxxxxx*	L'ID del gruppo di sicurezza `ElasticMapReduce-Core-Private` .
All Traffic	Tutti	Tutti	sg- *xxxxxxxxx xxxxxxxx*	L'ID del gruppo di sicurezza `ElasticMapReduce-Primary-Private` .
TCP personalizzato	TCP	9443	sg- *xxxxxxxxx xxxxxxxx*	L'ID del gruppo di sicurezza `ElasticMapReduce-ServiceAccess` .

Gruppo di sicurezza gestito da Amazon EMR per l'accesso ai servizi (sottoreti private)

Il gruppo di sicurezza gestito predefinito per l'accesso ai servizi nelle sottoreti private ha il nome del gruppo di -. ElasticMapReduce ServiceAccess Dispone di regole in entrata e di regole in uscita che consentono il traffico su HTTPS (porta 8443, porta 9443) per altri i gruppi di sicurezza gestiti nelle sottoreti private. Tali regole consentono al cluster manager di comunicare con il nodo primario e con i nodi principale e attività. Le stesse regole sono necessarie se si utilizzano gruppi di sicurezza personalizzati.

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
Regole in ingresso richieste per i cluster Amazon EMR con Amazon EMR versione 5.30.0 e successive.				
TCP personalizzato	TCP	9443	L'ID del gruppo di sicurezza	

Tipo	Protocollo	Intervallo porte	Origine	Informazioni
			gestito dell'istanza primaria.	Questa regola consente la comunicazione tra il gruppo di sicurezza dell'istanza primaria e il gruppo di sicurezza di accesso al servizio.
Regole in uscita richieste per tutti i cluster Amazon EMR				
TCP personalizzato	TCP	8443	L'ID del gruppo di sicurezza gestito dell'istanza primaria.	Tali regole consentono al cluster manager di comunicare con il nodo primario e con i nodi principale e attività.
TCP personalizzato	TCP	8443	L'ID del gruppo di sicurezza gestito specificato per le istanze principali e di attività.	Tali regole consentono al cluster manager di comunicare con il nodo primario e con i nodi principale e attività.

Utilizzo di gruppi di sicurezza aggiuntivi per un cluster Amazon EMR

Sia se si utilizzano i gruppi di sicurezza gestiti predefiniti o se si specificano gruppi di sicurezza gestiti personalizzati, è possibile utilizzare i gruppi di sicurezza aggiuntivi. I gruppi di sicurezza aggiuntivi offrono la flessibilità necessaria per definire l'accesso tra diversi cluster e dai client, risorse e applicazioni esterne.

Si consideri lo scenario riportato di seguito come un esempio: Disponi di più cluster che devono comunicare reciprocamente ma desideri consentire l'accesso SSH in entrata all'istanza primaria solo a un determinato sottoinsieme di cluster. Per eseguire questa operazione, puoi utilizzare lo stesso set di gruppi di sicurezza gestiti per i cluster. Quindi crei gruppi di sicurezza aggiuntivi che consentono l'accesso SSH in entrata da client affidabili e specifichi i gruppi di sicurezza aggiuntivi per l'istanza primaria per ogni cluster nel sottoinsieme.

Puoi applicare fino a 15 gruppi di sicurezza aggiuntivi per l'istanza principale, 15 per le istanze core e task e 15 per l'accesso al servizio (nelle sottoreti private). Se necessario, puoi specificare lo stesso gruppo di sicurezza aggiuntivo per le istanze primarie, principali, attività e per l'accesso ai servizi. Il

numero massimo di gruppi di sicurezza e di regole nell'account è soggetto ai limiti dell'account. Per ulteriori informazioni, consulta Limiti del gruppo di sicurezza nella Guida per l'utente di Amazon VPC.

Specifica dei gruppi di sicurezza gestiti da Amazon EMR e aggiuntivi

È possibile specificare i gruppi di sicurezza utilizzando l' AWS Management Console AWS CLI API Amazon EMR o l'API Amazon EMR. Se non specifichi i gruppi di sicurezza, Amazon EMR crea i gruppi di sicurezza predefiniti. Specificare i gruppi di sicurezza aggiuntivi è facoltativo. Puoi assegnare i gruppi di sicurezza aggiuntivi per le istanze primarie, principali, attività e per l'accesso ai servizi (solo sottoreti private).

Console

Per specificare i gruppi di sicurezza con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Rete, seleziona la freccia accanto ai gruppi EC2 di sicurezza (firewall) per espandere questa sezione. In Primary node (Nodo primario) e Core and task nodes (Nodi core e attività), i gruppi di sicurezza gestiti da Amazon EMR predefiniti sono selezionati per impostazione predefinita. Se utilizzi una sottorete privata, hai la possibilità di selezionare un gruppo di sicurezza anche per Service access (Accesso al servizio).
4. Per modificare il gruppo di sicurezza gestito da Amazon EMR, utilizza il menu a discesa Choose security groups (Scegli gruppi di sicurezza) per selezionare un'opzione diversa dall'elenco di opzioni Amazon EMR-managed security group (Gruppo di sicurezza gestito da Amazon EMR). Hai un gruppo di sicurezza gestito da Amazon EMR sia per Primary node (Nodo primario) sia per Core and task nodes (Nodi core e attività).
5. Per aggiungere gruppi di sicurezza personalizzati, utilizza lo stesso menu a discesa Choose security groups (Scegli gruppi di sicurezza) per selezionare fino a quattro gruppi di sicurezza personalizzati dall'elenco di opzioni Custom security group (Gruppi di sicurezza personalizzati). È possibile avere fino a quattro gruppi di sicurezza personalizzati sia per Primary node (Nodo primario) sia per Core and task nodes (Nodi core e attività).
6. Scegli qualsiasi altra opzione applicabile al cluster.
7. Per avviare il cluster, scegli Create cluster (Crea cluster).

Specifica dei gruppi di sicurezza con la AWS CLI

Per specificare i gruppi di sicurezza utilizzando il AWS CLI `create-cluster` comando con i seguenti parametri dell'`--ec2-attributes`opzione:

Parametro	Descrizione
`EmrManagedPrimarySecurityGroup`	Utilizza questo parametro per specificare un gruppo di sicurezza gestito personalizzato per l'istanza primaria. Se è specificato questo parametro, è necessario specificare anche `EmrManagedCoreSecurityGroup` . Per i cluster in sottoreti private deve essere specifica to anche `ServiceAccessSecurityGroup` .
`EmrManagedCoreSecurityGroup`	Utilizza questo parametro per specificare un gruppo di sicurezza gestito personalizzato per le istanze principali e di attività. Se è specifica to questo parametro, è necessario specificare anche `EmrManagedPrimarySecurityGr oup` . Per i cluster in sottoreti private deve essere specificato anche `ServiceAc cessSecurityGroup` .
`ServiceAccessSecurityGroup`	Utilizza questo parametro per specificare un gruppo di sicurezza gestito personalizzato per l'accesso ai servizi, che si applica solo ai cluster nelle sottoreti private. Il gruppo di sicurezza specificato come `ServiceAc cessSecurityGroup` non deve essere utilizzato per alcun altro scopo e deve essere riservato anche ad Amazon EMR. Se è specificato questo parametro, è necessario specificare anche `EmrManagedPrimaryS ecurityGroup` .

Parametro	Descrizione
`AdditionalPrimarySecurityGroups`	Utilizza questo parametro per specificare fino a quattro gruppi di sicurezza aggiuntivi per l'istanza primaria.
`AdditionalCoreSecurityGroups`	Utilizza questo parametro per specificare fino a quattro gruppi di sicurezza aggiuntivi per le istanze principali e di attività.

Example Specifica di gruppi di sicurezza gestiti da Amazon EMR e gruppi di sicurezza aggiuntivi personalizzati

L'esempio seguente specifica gruppi di sicurezza gestiti da Amazon EMR personalizzati per un cluster in una sottorete privata, più gruppi di sicurezza aggiuntivi per l'istanza primaria e un singolo gruppo di sicurezza aggiuntivo per le istanze principale e attività.

> **Note**
>
> I caratteri di continuazione della riga Linux (\) sono inclusi per questioni di leggibilità. Possono essere rimossi o utilizzati nei comandi Linux. Per Windows, rimuoverli o sostituirli con un accento circonflesso (^).

```
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
```

Per ulteriori informazioni, consulta create-cluster nella Guida di riferimento ai comandi della AWS CLI .

Specificazione dei gruppi EC2 di sicurezza per i notebook EMR

Al momento della creazione di un notebook EMR, per controllare il traffico di rete tra il notebook EMR e il cluster Amazon EMR vengono utilizzati due gruppi di sicurezza quando si usa l'editor del notebook. I gruppi di sicurezza predefiniti dispongono di regole minime che consentono solo il traffico di rete tra il servizio EMR Notebooks e i cluster a cui sono collegati i notebook.

Un notebook EMR utilizza Apache Livy per comunicare con il cluster tramite un proxy utilizzando la porta TCP 18888. Quando si creano gruppi di sicurezza personalizzati con regole su misura per l'ambiente, è possibile limitare il traffico di rete in modo che solo un sottoinsieme di notebook sia in grado di eseguire il codice all'interno dell'editor del notebook su determinati cluster. Il cluster utilizza la sicurezza personalizzata in aggiunta ai gruppi di sicurezza predefiniti per il cluster. Per ulteriori informazioni, consulta Controllo del traffico di rete con i gruppi di sicurezza nella Guida alla gestione di Amazon EMR e Specificazione dei gruppi EC2 di sicurezza per i notebook EMR.

Gruppo di EC2 sicurezza predefinito per l'istanza principale

Il gruppo EC2 di sicurezza predefinito per l'istanza primaria è associato all'istanza principale in aggiunta ai gruppi di sicurezza del cluster per l'istanza primaria.

Nome del gruppo: ElasticMapReduceEditors-Livy

Regole

- In entrata

  Consenti la porta TCP 18888 da qualsiasi risorsa nel gruppo di EC2 sicurezza predefinito per EMR Notebooks
- In uscita

  Nessuno

Gruppo EC2 di sicurezza predefinito per EMR Notebooks

Il gruppo EC2 di sicurezza predefinito per il notebook EMR è associato all'editor di notebook per ogni notebook EMR a cui è assegnato.

Nome del gruppo: -Editor ElasticMapReduceEditors

Regole

- In entrata

  Nessuno

- In uscita

  Consenti la porta TCP 18888 a tutte le risorse del gruppo di EC2 sicurezza predefinito per EMR Notebooks.

Gruppo EC2 di sicurezza personalizzato per Notebooks EMR quando si associano notebook a repository Git

Per collegare un repository Git al notebook, il gruppo di sicurezza per il notebook EMR deve includere una regola in uscita per consentire al notebook di instradare il traffico a Internet. Si consiglia di creare un nuovo gruppo di sicurezza per questo scopo. L'aggiornamento del gruppo di sicurezza predefinito ElasticMapReduceEditors-Editor può fornire le stesse regole in uscita ad altri notebook collegati a questo gruppo di sicurezza.

Regole

- In entrata

  Nessuno

- In uscita

  Consenti al notebook di instradare il traffico a Internet tramite il cluster, come illustrato nell'esempio seguente: Ad esempio, viene utilizzato il valore 0.0.0.0/0. È possibile modificare questa regola per specificare gli indirizzi IP dei repository basati su Git.

Tipo	Protocollo	Intervallo porte	Destinazione
Regola TCP personalizzata	TCP	18888	SG-
HTTPS	TCP	443	0.0.0.0/0

Utilizzo del blocco dell'accesso pubblico di Amazon EMR

Il blocco dell'accesso pubblico (BPA) di Amazon EMR impedisce l'avvio di un cluster in una sottorete pubblica se il cluster dispone di una configurazione di sicurezza che consente il traffico in entrata da indirizzi IP pubblici su una porta.

> **Important**
>
> Il blocco dell'accesso pubblico è abilitato per impostazione predefinita. Per aumentare la protezione degli account, ti consigliamo di mantenerlo abilitato.

Informazioni sul blocco dell'accesso pubblico

Puoi utilizzare la configurazione a livello di account del blocco dell'accesso pubblico per gestire a livello centrale l'accesso della rete pubblica ai cluster Amazon EMR.

Quando un tuo utente Account AWS avvia un cluster, Amazon EMR verifica le regole delle porte nel gruppo di sicurezza per il cluster e le confronta con le regole del traffico in entrata. Se il gruppo di sicurezza ha una regola in entrata che apre le porte agli indirizzi IP pubblici IPv4 0.0.0.0/0 o IPv6 : :/0 e tali porte non sono specificate come eccezioni per il tuo account, Amazon EMR non consente all'utente di creare il cluster.

Se un utente modifica le regole del gruppo di sicurezza per un cluster in esecuzione in una sottorete pubblica in modo da avere una regola di accesso pubblico che viola la configurazione BPA del tuo account, Amazon EMR revoca la nuova regola se dispone dell'autorizzazione per farlo. Se Amazon EMR non dispone dell'autorizzazione per revocare la regola, crea un evento nel pannello di controllo di AWS Health che descrive la violazione. Per concedere l'autorizzazione alla revoca della regola ad Amazon EMR, consulta Configura Amazon EMR per revocare le regole dei gruppi di sicurezza.

Il blocco dell'accesso pubblico è abilitato per impostazione predefinita per tutti i cluster di ogni Regione AWS per il tuo Account AWS. Il BPA si applica all'intero ciclo di vita di un cluster, ma non si applica ai cluster creati in sottoreti private. È possibile configurare eccezioni alla regola BPA; la porta 22 è un'eccezione per impostazione predefinita. Per ulteriori informazioni sull'impostazione delle eccezioni, consulta Configurazione del blocco degli accessi pubblici.

Configurazione del blocco degli accessi pubblici

È possibile aggiornare i gruppi di sicurezza e la configurazione del blocco dell'accesso pubblico negli account in qualsiasi momento.

Puoi attivare e disattivare le impostazioni di accesso pubblico a blocchi (BPA) con AWS Management Console, the AWS Command Line Interface (AWS CLI) e l'API Amazon EMR. Le impostazioni si applicano a tutto l'account su base individuale. Region-by-Region Per preservare la sicurezza del cluster, si consiglia di mantenere abilitato il blocco dell'accesso pubblico.

Console

Per configurare l'accesso pubblico a blocchi con la console

1. Accedi a AWS Management Console, quindi apri la console Amazon EMR su https://console.aws.amazon.com /emr.
2. Nella barra di navigazione in alto, seleziona la Regione che desideri configurare, se non è già selezionata.
3. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Blocca accesso pubblico.
4. In Block public access settings (Impostazioni blocco accesso pubblico) completare la procedura seguente.

A…	Esegui questa operazione…
Attivare o disattivare il blocco degli accessi pubblici	Scegli Edit (Modifica), scegli Turn on (Attiva) oppure Turn off (Disattiva) seconda dei casi, quindi scegli Save (Salva).
Modificare le porte nell'elenco delle eccezioni	1. Scegli Edit (Modifica) e cerca la sezione Port range exceptions (Eccezioni dell'intervallo di porte). 2. Per aggiungere porte all'elenco delle eccezioni, scegliere Add a port range (Aggiungi un intervallo di porte) e immettere una nuova porta o un nuovo intervallo di porte. Ripetere l'operazione per ogni porta o intervallo di porte da aggiungere. 3.

A…	Esegui questa operazione…
	Per rimuovere una porta o un intervall o di porte, scegli Remove (Rimuovi) accanto alla voce nell'elenco degli intervalli di porte. 4. Scegli Save (Salva).

AWS CLI

Per configurare l'accesso pubblico a blocchi utilizzando il AWS CLI

- Utilizzare il comando `aws emr put-block-public-access-configuration` per configurare il blocco degli accessi pubblici come mostrato negli esempi seguenti.

A…	Esegui questa operazione…
Attivare il blocco degli accessi pubblici	Impostare `BlockPublicSecurityGroupRules` su `true` come mostrato nell'esempio seguente. Per consentire l'avvio del cluster, nessun gruppo di sicurezza associato a un cluster può avere una regola in entrata che consente l'accesso pubblico. `aws emr put-block-public-access-configuration --block-public-access-configuration BlockPublicSecurityGroupRules=true`

A…	Esegui questa operazione…
Disattivare il blocco degli accessi pubblici	Impostare `BlockPublicSecurityGroupRules` su `false` come mostrato nell'esempio seguente. I gruppi di sicurezza associati a un cluster possono avere regole in entrata che consentono l'accesso pubblico su qualsiasi porta. Questa configurazione non è consigliata. `aws emr put-block-public-access-configuration --block-public-access-configuration BlockPublicSecurityGroupRules=false`
Attivare il blocco degli accessi pubblici e specificare le porte come eccezioni	L'esempio seguente attiva il blocco degli accessi pubblici e specifica la porta 22 e le porte 100-101 come eccezioni. Ciò consente la creazione di cluster se per un gruppo di sicurezza associato è specificata una regola in entrata che consente l'accesso pubblico sulla porta 22, sulla porta 100 o sulla porta 101. `aws emr put-block-public-access-configuration --block-public-access-configuration '{ "BlockPublicSecurityGroupRules": true, "PermittedPublicSecurityGroupRuleRanges": [ { "MinRange": 22, "MaxRange": 22 }, { "MinRange": 100, "MaxRange": 101 } ] }'`

Configura Amazon EMR per revocare le regole dei gruppi di sicurezza

Amazon EMR necessita dell'autorizzazione per revocare le regole dei gruppi di sicurezza e conformarsi alla configurazione del blocco dell'accesso pubblico. Puoi utilizzare uno dei seguenti approcci per concedere ad Amazon EMR l'autorizzazione necessaria:

- Consigliato Collega la policy gestita da `AmazonEMRServicePolicy_v2` al ruolo di servizio. Per ulteriori informazioni, consulta Ruolo di servizio per Amazon EMR (ruolo EMR).
- Crea una nuova policy inline che consenta l'operazione `ec2:RevokeSecurityGroupIngress` sui gruppi di sicurezza. Per ulteriori informazioni su come modificare una policy di autorizzazione dei ruoli, consulta Modifica di una policy di autorizzazioni dei ruoli con la console IAM, l'API AWS e la AWS CLI nella Guida per l'utente IAM.

Risoluzione delle violazioni dell'accesso pubblico

Se si verifica una violazione del blocco dell'accesso pubblico, puoi mitigarla con una delle seguenti azioni:

- Se desideri accedere a un'interfaccia Web sul tuo cluster, usa una delle opzioni descritte in Visualizzazione di interfacce Web ospitate su cluster Amazon EMR per accedere all'interfaccia tramite SSH (porta 22).
- Per consentire il traffico verso il cluster da indirizzi IP specifici anziché dall'indirizzo IP pubblico, aggiungi una regola del gruppo di sicurezza. Per ulteriori informazioni, consulta Aggiungere regole a un gruppo di sicurezza nella Amazon EC2 Getting Started Guide.
- (Non consigliato) Puoi configurare le eccezioni BPA di Amazon EMR per includere la porta o l'intervallo di porte desiderati. Quando specifichi un'eccezione BPA, introduci un rischio con una porta non protetta. Se intendi specificare un'eccezione, devi rimuoverla appena non è più necessaria. Per ulteriori informazioni, consulta Configurazione del blocco degli accessi pubblici.

Identificazione dei cluster associati alle regole dei gruppi di sicurezza

Potrebbe essere necessario identificare tutti i cluster associati a una determinata regola dei gruppi di sicurezza o trovare la regola dei gruppi di sicurezza per un determinato cluster.

- Se conosci il gruppo di sicurezza, puoi identificare i cluster associati, se trovi le interfacce di rete per tale gruppo di sicurezza. Per ulteriori informazioni, consulta Come posso trovare le risorse associate a un gruppo EC2 di sicurezza Amazon? su AWS re:Post. Le EC2 istanze

Amazon collegate a queste interfacce di rete verranno contrassegnate con l'ID del cluster a cui appartengono.

- Se desideri trovare i gruppi di sicurezza per un cluster noto, segui i passaggi riportati in Visualizza lo stato e i dettagli del cluster Amazon EMR. Puoi trovare i gruppi di sicurezza per il cluster nel riquadro Rete e sicurezza della console o nel campo `Ec2InstanceAttributes` della AWS CLI.

Convalida della conformità per Amazon EMR

I revisori di terze parti valutano la sicurezza e la conformità di Amazon EMR nell'ambito di AWS diversi programmi di conformità. Questi includono SOC, PCI, FedRAMP, HIPAA e altri.

Per un elenco dei AWS servizi che rientrano nell'ambito di specifici programmi di conformità, consulta la sezione AWS Servizi rientranti nell'ambito del programma di conformità. Per informazioni generali, consultare Programmi per la conformità di AWS.

È possibile scaricare report di audit di terze parti utilizzando AWS Artifact. Per ulteriori informazioni, consulta Scaricamento dei report in AWS Artifact.

La tua responsabilità di conformità durante l'utilizzo di Amazon EMR è determinata dalla riservatezza dei dati, dagli obiettivi di conformità dell'azienda e dalle leggi e normative applicabili. Se l'uso di Amazon EMR è soggetto alla conformità a standard come HIPAA, PCI o FedRAMP, fornisce risorse per aiutarti a: AWS

- Guide introduttive su sicurezza e conformità: queste guide all'implementazione illustrano considerazioni sull'architettura e forniscono passaggi per implementare ambienti di base incentrati sulla sicurezza e la conformità. AWS
- Whitepaper Architecting for HIPAA Security and Compliance: questo white paper descrive in che modo le aziende possono utilizzare per creare applicazioni conformi allo standard HIPAA. AWS
- AWS risorse per la conformità: questa raccolta di cartelle di lavoro e guide potrebbe riguardare il settore e la località in cui operi.
- AWS Config— Questo AWS servizio valuta la conformità delle configurazioni delle risorse alle pratiche interne, alle linee guida del settore e alle normative.
- AWS Security Hub— Questo AWS servizio offre una visione completa dello stato di sicurezza dell'utente e consente AWS di verificare la conformità agli standard e alle best practice del settore della sicurezza.

Resilienza in Amazon EMR

L'infrastruttura AWS globale è costruita attorno a AWS regioni e zone di disponibilità. AWS Le regioni forniscono più zone di disponibilità fisicamente separate e isolate, collegate con reti a bassa latenza, ad alto throughput e altamente ridondanti. Con le zone di disponibilità, è possibile progettare e gestire applicazioni e database che eseguono il failover automatico tra zone di disponibilità senza interruzioni. Le zone di disponibilità sono più disponibili, tolleranti ai guasti e scalabili rispetto alle infrastrutture tradizionali a data center singolo o multiplo.

Per ulteriori informazioni su AWS regioni e zone di disponibilità, consulta infrastruttura globale.AWS

Oltre all'infrastruttura AWS globale, Amazon EMR offre diverse funzionalità per supportare le tue esigenze di resilienza e backup dei dati.

- Integrazione con Amazon S3 tramite EMRFS
- Supporto per più nodi master

Sicurezza dell'infrastruttura in Amazon EMR

In quanto servizio gestito, Amazon EMR è protetto dalla sicurezza di rete AWS globale. Per informazioni sui servizi di AWS sicurezza e su come AWS protegge l'infrastruttura, consulta AWS Cloud Security. Per progettare il tuo AWS ambiente utilizzando le migliori pratiche per la sicurezza dell'infrastruttura, vedi Infrastructure Protection in Security Pillar AWS Well-Architected Framework.

Utilizzi chiamate API AWS pubblicate per accedere ad Amazon EMR attraverso la rete. I client devono supportare quanto segue:

- Transport Layer Security (TLS). È richiesto TLS 1.2 ed è consigliato TLS 1.3.
- Suite di cifratura con Perfect Forward Secrecy (PFS), ad esempio Ephemeral Diffie-Hellman (DHE) o Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). La maggior parte dei sistemi moderni, come Java 7 e versioni successive, supporta tali modalità.

Inoltre, le richieste devono essere firmate utilizzando un ID chiave di accesso e una chiave di accesso segreta associata a un principale IAM. O puoi utilizzare AWS Security Token Service (AWS STS) per generare credenziali di sicurezza temporanee per sottoscrivere le richieste.

Argomenti

- Connessione ad Amazon EMR utilizzando un endpoint VPC di interfaccia

Connessione ad Amazon EMR utilizzando un endpoint VPC di interfaccia

Puoi connetterti direttamente ad Amazon EMR utilizzando un'interfaccia VPC endpoint (AWS PrivateLink) nel tuo Virtual Private Cloud (VPC) anziché collegarti tramite Internet. Quando utilizzi un endpoint VPC di interfaccia, la comunicazione tra il tuo VPC e Amazon EMR avviene interamente all'interno della rete. AWS Ogni endpoint VPC è rappresentato da una o più interfacce di rete elastiche (ENIs) con indirizzi IP privati nelle sottoreti VPC.

L'endpoint VPC di interfaccia collega il tuo VPC direttamente ad Amazon EMR senza un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. AWS Direct Connect Le istanze presenti nel tuo VPC non richiedono indirizzi IP pubblici per comunicare con l'API di Amazon EMR.

Per utilizzare Amazon EMR tramite il VPC, devi connetterti da un'istanza che si trova all'interno del VPC o connettere la rete privata al VPC utilizzando Amazon Virtual Private Network (VPN) o AWS Direct Connect. Per informazioni su Amazon VPN, consulta Connessioni VPN nella Guida per l'utente di Amazon Virtual Private Cloud. Per informazioni su AWS Direct Connect, consulta Creare una connessione nella Guida per l'utente.AWS Direct Connect

Puoi creare un endpoint VPC di interfaccia per connetterti ad Amazon EMR utilizzando la AWS console o i comandi (). AWS Command Line Interface AWS CLI Per ulteriori informazioni, consulta Creazione di un endpoint di interfaccia.

Se dopo aver creato un endpoint VPC di interfaccia abiliti nomi host DNS privati per l'endpoint, l'endpoint Amazon EMR predefinito restituisce il tuo endpoint VPC. L'endpoint del nome del servizio predefinito per Amazon EMR è nel formato seguente.

```
elasticmapreduce.Region.amazonaws.com
```

Se non abiliti nomi host DNS privati, Amazon VPC fornisce un nome di endpoint DNS che puoi utilizzare nel formato seguente:

```
VPC_Endpoint_ID.elasticmapreduce.Region.vpce.amazonaws.com
```

Per ulteriori informazioni, consultare Endpoint VPC di interfaccia (AWS PrivateLink) nella Guida per l'utente di Amazon VPC.

Amazon EMR supporta l'esecuzione di chiamate a tutte le sue operazioni API all'interno del VPC.

Puoi collegare le policy di endpoint VPC a un endpoint VPC per controllare l'accesso per le entità principali IAM. Puoi inoltre associare i gruppi di sicurezza a un endpoint VPC per controllare l'accesso in ingresso e in uscita in base all'origine e alla destinazione del traffico di rete, ad esempio un intervallo di indirizzi IP. Per ulteriori informazioni, consulta Controllo dell'accesso ai servizi con endpoint VPC.

Creazione di una policy di endpoint VPC per Amazon EMR

Puoi creare una policy per gli endpoint di Amazon VPC per Amazon EMR per specificare quanto segue:

- Il principale che può o non può eseguire azioni
- Le azioni che possono essere eseguite
- Le risorse sui cui si possono eseguire le azioni

Per ulteriori informazioni, consultare Controllo degli accessi ai servizi con endpoint VPC in Guida per l'utente di Amazon VPC.

Example — Policy degli endpoint VPC per negare tutti gli accessi da un account specifico AWS

La seguente politica degli endpoint VPC nega all' AWS account *123456789012* tutti gli accessi alle risorse che utilizzano l'endpoint.

```
{
    "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": "*"
        },
        {
            "Action": "*",
            "Effect": "Deny",
            "Resource": "*",
            "Principal": {
                "AWS": [
                    "123456789012"
                ]
```

```
            }
        }
    ]
}
```

Example - Policy di endpoint VPC per consentire l'accesso VPC solo a un'entità principale IAM (utente) specificata

La seguente politica degli endpoint VPC consente l'accesso completo solo a un utente *lijuan* nell'account. AWS *123456789012* A tutte le altre entità principali IAM viene negato l'accesso utilizzando l'endpoint.

```
{
    "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": {
                "AWS": [
                    "arn:aws:iam::123456789012:user/lijuan"
                ]
            }
        }]
}
```

Example - Policy di endpoint VPC per consentire operazioni EMR di sola lettura

La seguente policy sugli endpoint VPC consente solo *123456789012* all' AWS account di eseguire le azioni Amazon EMR specificate.

Le operazioni specificate forniscono l'accesso di sola lettura equivalente per Amazon EMR. Tutte le altre azioni sul VPC vengono negate per l'account specificato. A tutti gli altri account viene negato l'accesso. Per un elenco delle operazioni Amazon EMR, consulta Operazioni, risorse e chiavi di condizione per 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 - Policy di endpoint VPC per negare l'accesso a un cluster specificato

La seguente policy sugli endpoint VPC consente l'accesso completo a tutti gli account e i principali, ma nega qualsiasi accesso AWS dell'account *123456789012* alle azioni eseguite sul cluster Amazon EMR con ID cluster. *j-A1B2CD34EF5G* Altre operazioni Amazon EMR che non supportano le autorizzazioni a livello di risorsa per i cluster sono comunque consentite. Per l'elenco delle operazioni Amazon EMR e dei tipi di risorse corrispondenti, consulta Operazioni, risorse e chiavi di condizione per 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"
                ]
            }
        }
    ]
}
```

Gestione dei cluster Amazon EMR

Dopo aver avviato il cluster, puoi connetterti ad esso e gestirlo. Amazon EMR offre una raccolta di strumenti che puoi utilizzare a tale scopo. Questa sezione fornisce indicazioni per la connessione al cluster, per dargli da fare, per monitorare il lavoro in corso e per la risoluzione dei problemi.

Argomenti

- Connettiti a un cluster Amazon EMR
- Invia il lavoro a un cluster Amazon EMR
- Visualizza e monitora un cluster Amazon EMR mentre esegue il lavoro
- Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione
- Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa
- Clonare un cluster Amazon EMR utilizzando la console
- Automatizza i cluster Amazon EMR ricorrenti con AWS Data Pipeline

Connettiti a un cluster Amazon EMR

Quando si esegue un cluster Amazon EMR, spesso è sufficiente eseguire un'applicazione per analizzare i dati e raccogliere l'output da un bucket Amazon S3. In altri casi, potresti voler interagire con il nodo primario mentre il cluster è in esecuzione. Ad esempio, potresti volerti connettere al nodo primario per eseguire query interattive, controllare i file di log, eseguire il debug di un problema con il cluster, monitorare le prestazioni utilizzando un'applicazione come Ganglia in esecuzione sul nodo primario e così via. Le sezioni seguenti descrivono le tecniche che è possibile utilizzare per connettersi al nodo primario.

In un cluster EMR, il nodo primario è un' EC2 istanza Amazon che coordina le EC2 istanze in esecuzione come nodi task e core. Il nodo primario espone un nome DNS pubblico che è possibile utilizzare per connettersi ad esso. Per impostazione predefinita, Amazon EMR crea regole del gruppo di sicurezza per il nodo primario e i nodi core e attività, che determinano la modalità di accesso ai nodi.

> **Note**
>
> È possibile connettersi al nodo primario solo mentre il cluster è in esecuzione. Quando il cluster termina, l' EC2 istanza che funge da nodo primario viene terminata e non è più

disponibile. Per connettersi al nodo primario, è necessario anche autenticarsi al cluster. Puoi utilizzare Kerberos per l'autenticazione o specificare una chiave privata Amazon EC2 key pair all'avvio del cluster. Per ulteriori informazioni sulla configurazione di Kerberos e sulla connessione, vedere Utilizzo di Kerberos per l'autenticazione con Amazon EMR. Quando avvii un cluster dalla console, la EC2 chiave privata Amazon key pair viene specificata nella sezione Sicurezza e accesso della pagina Crea cluster.

Per impostazione predefinita, il gruppo di sicurezza ElasticMapReduce -master non consente l'accesso SSH in entrata. Potrebbe essere necessario aggiungere una regola in entrata che consenta l'accesso SSH (porta TCP 22) dalle sorgenti a cui si desidera accedere. Per ulteriori informazioni sulla modifica delle regole dei gruppi di sicurezza, consulta Adding rules to a security group nella Amazon EC2 User Guide.

Important

Non modificare le regole rimanenti nel gruppo di sicurezza ElasticMapReduce -master. La modifica di queste regole può interferire con il funzionamento del cluster.

Argomenti

- Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata
- Connect al nodo primario del cluster Amazon EMR tramite SSH

Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata

Prima di effettuare la connessione a un cluster Amazon EMR, devi autorizzare il traffico SSH in entrata (porta 22) da parte di client affidabili come l'indirizzo IP del computer. A tale scopo, modifica le regole del gruppo di sicurezza gestito per i nodi a cui desideri connetterti. Ad esempio, le seguenti istruzioni mostrano come aggiungere una regola in entrata per l'accesso SSH al gruppo di sicurezza predefinito -master. ElasticMapReduce

Per ulteriori informazioni sull'utilizzo dei gruppi di sicurezza tramite Amazon EMR, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Console

Per concedere a fonti attendibili l'accesso SSH al gruppo di sicurezza primario con la console

Per modificare i gruppi di sicurezza, devi disporre dell'autorizzazione per gestire i gruppi di sicurezza per il VPC in cui si trova il cluster. Per ulteriori informazioni, consulta Modifica delle autorizzazioni per un utente e la politica di esempio che consente la gestione dei gruppi di EC2 sicurezza nella Guida per l'utente IAM.

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli il cluster che desideri aggiornare. Si apre la pagina dei dettagli del cluster. La scheda Properties (Proprietà) in questa pagina sarà preselezionata.
3. In Rete nella scheda Proprietà, seleziona la freccia accanto ai gruppi EC2 di sicurezza (firewall) per espandere questa sezione. In Primary node (Nodo primario), seleziona il collegamento al gruppo di sicurezza. Verrà aperta la EC2 console.
4. Seleziona la scheda Inbound rules (Regole in entrata), quindi scegli Edit inbound rules (Modifica regola in entrata).
5. Verifica la presenza di una regola in entrata che consenta l'accesso pubblico con le seguenti impostazioni. Se esiste, scegli Elimina per rimuoverla.

   - Tipo

     SSH
   - Porta

     22
   - Origine

     Personalizzata 0.0.0.0/0

> **Warning**
>
> Prima di dicembre 2020, il gruppo di sicurezza ElasticMapReduce -master disponeva di una regola preconfigurata per consentire il traffico in entrata sulla porta 22 da tutte le fonti. Questa regola è stata creata per semplificare le connessioni SSH iniziali al

nodo primario. Consigliamo vivamente di rimuovere questa regola in entrata e limitare il traffico alle origini affidabili.

6. Scorri fino alla fine dell'elenco di regole e seleziona Aggiungi regola.
7. Per Tipo, seleziona SSH. Questa selezione inserisce automaticamente TCP per Protocol (Protocollo) e 22 per Port Range (Intervallo porte).
8. Per origine, seleziona Il mio IP per aggiungere in automatico il tuo indirizzo IP come indirizzo di origine. Puoi anche aggiungere un intervallo di indirizzi IP affidabili del client Custom (Personalizzato), o creare regole aggiuntive per altri client. In molti ambienti di rete, gli indirizzi IP vengono allocati in modo dinamico, perciò, nel futuro, potrebbe essere necessario aggiornare gli indirizzi IP per client affidabili.
9. Scegli Save (Salva).
10. Facoltativamente, torna al passaggio 3, scegli Core and task nodes (Nodi core e attività) e ripeti i passaggi da 4 a 8. Ciò garantisce l'accesso al client SSH ai nodi core e attività.

Connect al nodo primario del cluster Amazon EMR tramite SSH

Secure Shell (SSH) è un protocollo di rete che può essere utilizzato per creare una connessione sicura a un computer remoto. Dopo aver effettuato il collegamento, il terminale del computer locale si comporta come se fosse in esecuzione sul computer remoto. I comandi emessi localmente vengono eseguiti sul computer remoto e l'output del comando dal computer remoto viene visualizzato nella finestra del terminale.

Quando usi SSH con AWS, ti connetti a un' EC2 istanza, che è un server virtuale in esecuzione nel cloud. Quando si lavora con Amazon EMR, l'uso più comune di SSH è quello di connettersi all' EC2 istanza che funge da nodo primario del cluster.

L'utilizzo di SSH per connettersi al nodo primario consente di monitorare e interagire con il cluster. È possibile eseguire comandi Linux sul nodo primario, eseguire applicazioni come Hive e Pig in modo interattivo, sfogliare le directory, leggere i file di log e così via. Si può anche creare un tunnel nella connessione SSH per visualizzare le interfacce Web ospitate sul nodo primario. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Per connettersi al nodo primario utilizzando SSH, è necessario il nome DNS pubblico del nodo primario. Inoltre, il gruppo di sicurezza associato al nodo primario deve avere una regola in entrata che consenta il traffico SSH (porta TCP 22) da un'origine che include il client da cui viene creata

la connessione SSH. Potrebbe essere necessario aggiungere una regola per consentire una connessione SSH dal client. Per ulteriori informazioni sulla modifica delle regole dei gruppi di sicurezza, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR Adding rules to a security group nella Amazon EC2 User Guide.

Recupero del nome DNS pubblico del nodo primario

È possibile recuperare il nome DNS pubblico primario utilizzando la console Amazon EMR e la AWS CLI.

Console

Recupero del nome DNS pubblico del nodo primario con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi seleziona il cluster in cui desideri recuperare il nome DNS pubblico.
3. Prendi nota del valore Primary node public DNS (DNS pubblico del nodo primario) che compare nella sezione Summary (Riepilogo) della pagina dei dettagli del cluster.

CLI

Per recuperare il nome DNS pubblico del nodo primario con AWS CLI

1. Per recuperare l'identificatore del cluster, digitare il seguente comando:

```
aws emr list-clusters
```

L'output elenca i cluster, incluso il cluster. IDs Prendere nota dell'ID del cluster per il cluster a cui ci si connette.

```
"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. Per elencare le istanze del cluster, incluso il nome DNS pubblico del cluster, inserisci uno dei seguenti comandi. Sostituisci *j-2AL4XXXXXX5T9* con l'ID del cluster restituito dal comando precedente.

```
aws emr list-instances --cluster-id j-2AL4XXXXXX5T9
```

O:

```
aws emr describe-cluster --cluster-id j-2AL4XXXXXX5T9
```

L'output elenca le istanze del cluster, compresi i nomi DNS e gli indirizzi IP. Prendere nota del valore per `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": "###.##.#.###"
```

Per ulteriori informazioni, consulta Comandi di Amazon EMR nella AWS CLI.

Connect al nodo primario tramite SSH e una chiave EC2 privata Amazon su Linux, Unix e Mac OS X

Per creare una connessione SSH autenticata con un file di chiave privata, devi specificare la chiave privata Amazon EC2 key pair quando avvii un cluster. Per ulteriori informazioni sull'accesso alla tua coppia di chiavi, consulta Amazon EC2 key pair nella Amazon EC2 User Guide.

Molto probabilmente il tuo computer Linux include un client SSH per impostazione predefinita. Ad esempio, OpenSSH è installato nella maggior parte dei sistemi operativi Linux, Unix e macOS. È possibile verificare la presenza di un client SSH digitando ssh nella riga di comando. Se il computer non riconosce il comando, installa un client SSH per la connessione al nodo primario. Il progetto OpenSSH fornisce un'implementazione gratuita della suite completa di strumenti SSH. Per ulteriori informazioni, consulta il sito Web OpenSSH.

Le seguenti istruzioni illustrano l'apertura di una connessione SSH al nodo primario Amazon EMR su Linux, Unix e Mac OS X.

Per configurare le autorizzazioni per i file chiave privati della coppia di chiavi

Prima di poter utilizzare la chiave privata Amazon EC2 key pair per creare una connessione SSH, devi impostare le autorizzazioni sul `.pem` file in modo che solo il proprietario della chiave abbia il permesso di accedere al file. Ciò è necessario per creare una connessione SSH utilizzando il terminale o il. AWS CLI

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Individuare il file `.pem`. Queste istruzioni presuppongono che il file sia denominato `mykeypair.pem` e che venga memorizzato nella directory principale dell'utente corrente.
3. Digitare il seguente comando per impostare le autorizzazioni. Sostituisci *~/mykeypair.pem* con il percorso completo e il nome del file della chiave privata della tua coppia di chiavi. Ad esempio `C:/Users/<username>/.ssh/mykeypair.pem`.

   ```
   chmod 400 ~/mykeypair.pem
   ```

   Se non si impostano le autorizzazioni sul file `.pem`, si riceverà un errore che indica che il file chiave non è protetto e la chiave verrà rifiutata. Per connettersi, è sufficiente impostare le autorizzazioni sul file chiave privata della coppia di chiavi la prima volta che lo si utilizza.

Connessione al nodo primario utilizzando il terminale

1. Apri una finestra del terminale. In Mac OS X, selezionare Applications > Utilities > Terminal (Applicazioni > Utility > Terminale). In altre distribuzioni Linux, la finestra terminal si trova generalmente in Applications > Accessories > Terminal (Applicazioni > Accessori >; Terminale).
2. Per stabilire una connessione al nodo primario, inserisci il seguente comando. Sostituiscilo *ec2-###-##-##-###.compute-1.amazonaws.com* con il nome DNS pubblico principale del cluster e *~/mykeypair.pem* sostituiscilo con il percorso completo e il nome del `.pem` file. Ad esempio `C:/Users/<username>/.ssh/mykeypair.pem`.

   ```
   ssh hadoop@ec2-###-##-##-###.compute-1.amazonaws.com -i ~/mykeypair.pem
   ```

   > **Important**
   >
   > Quando si esegue la connessione al nodo primario di Amazon EMR occorre utilizzare il nome di login `hadoop`, altrimenti potrebbe verificarsi un errore simile a `Server refused our key`.

3. Un avviso indica che non è possibile verificare l'autenticità dell'host a cui ci si connette. Per continuare, digitare `yes`.
4. Una volta terminato il lavoro sul nodo primario, inserisci il seguente comando per chiudere la connessione SSH.

   ```
   exit
   ```

Se riscontri problemi nell'utilizzo di SSH per la connessione al nodo primario, consulta la sezione Troubleshoot connecting to your instance (Risoluzione dei problemi di connessione all'istanza).

Connessione al nodo primario tramite SSH su Windows

Gli utenti Windows possono utilizzare un client SSH come PuTTY per connettersi al nodo primario. Prima di connetterti al nodo primario di Amazon EMR, devi scaricare e installare PuTTY e Pu. TTYgen Questi strumenti possono essere scaricati dalla pagina di download di PuTTY.

PuTTY non supporta nativamente il formato di file della chiave privata della coppia di chiavi `.pem` () generato da Amazon. EC2 Utilizzate Pu TTYgen per convertire il file chiave nel formato PuTTY

richiesto (). `.ppk` È necessario convertire la chiave privata nel formato `.ppk` prima di tentare una connessione al nodo primario tramite PuTTY.

Per ulteriori informazioni sulla conversione della chiave, consulta Convertire la chiave privata utilizzando Pu TTYgen nella Amazon EC2 User Guide.

Connessione al nodo primario usando PuTTY

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Aprire `putty.exe`. È anche possibile avviare PuTTY dall'elenco dei programmi di Windows.
3. Se necessario, nell'elenco Category (Categoria) selezionare Session (Sessione).
4. Per Nome host (o indirizzo IP), digita. `hadoop@` *`MasterPublicDNS`* Ad esempio: `hadoop@`*`ec2-###-##-##-###.compute-1.amazonaws.com`*.
5. Nell'elenco Category (Categoria) scegliere Connection > SSH (Connessione > SSH), Auth (Autenticazione).
6. Per Private key file for authentication (File chiave privata per autenticazione), scegliere Browse (Sfoglia) e selezionare il file `.ppk` generato.
7. Seleziona Open (Apri), quindi Yes (Sì) per ignorare l'avviso di sicurezza di PuTTY.

> **Important**
>
> Quando accedi al nodo primario, inserisci `hadoop` nel caso in cui venga richiesto un nome utente.

8. Una volta terminato il lavoro sul nodo primario puoi chiudere la connessione SSH chiudendo PuTTY.

> **Note**
>
> Per evitare il timeout della connessione SSH, scegliere Connection (Connessione) nell'elenco Category (Categoria) e selezionare l'opzione Enable TCP_keepalives (Abilita TCP_keepalives). Se si dispone di una sessione SSH attiva in PuTTY, è possibile modificare le impostazioni aprendo il contesto (clic con il tasto destro del mouse) per la barra del titolo di PuTTY e scegliendo Change Settings (Modifica impostazioni).

Se riscontri problemi nell'utilizzo di SSH per la connessione al nodo primario, consulta la sezione Troubleshoot connecting to your instance (Risoluzione dei problemi di connessione all'istanza).

Connessione al nodo primario tramite la AWS CLI

Puoi creare una connessione SSH con il nodo primario utilizzando Windows e Linux, Unix e Mac OS X. Indipendentemente dalla piattaforma, sono necessari il nome DNS pubblico del nodo primario e la chiave privata Amazon EC2 key pair. AWS CLI Se utilizzi Linux, Unix o Mac OS X, devi anche impostare le autorizzazioni sul file della chiave privata (`.pemo.ppk`) come mostrato in. AWS CLI Per configurare le autorizzazioni per i file chiave privati della coppia di chiavi

Per connettersi al nodo primario utilizzando il AWS CLI

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Per recuperare l'identificatore del cluster, digitare:

   ```
   aws emr list-clusters
   ```

   L'output elenca i cluster, incluso il cluster IDs. Prendere nota dell'ID del cluster per il cluster a cui ci si connette.

   ```
   "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. Inserisci il seguente comando per aprire una connessione SSH al nodo primario. Nell'esempio seguente, sostituitelo *`j-2AL4XXXXXX5T9`* con l'ID del cluster e *`~/mykeypair.key`* sostituitelo con il percorso completo e il nome del `.pem` file (per Linux, Unix e Mac OS X) o `.ppk` del file (per Windows). Ad esempio `C:\Users\<username>\.ssh\mykeypair.pem`.

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

4. Quando hai finito di lavorare sul nodo principale, chiudi la AWS CLI finestra.

   Per ulteriori informazioni, consulta Comandi Amazon EMR nella AWS CLI. Se riscontri problemi nell'utilizzo di SSH per la connessione al nodo primario, consulta la sezione Troubleshoot connecting to your instance (Risoluzione dei problemi di connessione all'istanza).

Porte dei servizi Amazon EMR

> **Note**
>
> Di seguito sono riportate le interfacce e le porte dei servizi per i componenti su Amazon EMR. Questo non è l'elenco completo delle porte di servizio. I servizi non predefiniti, come le porte SSL e diversi tipi di protocolli, non sono elencati.

> **Important**
>
> Presta attenzione quando modifichi le regole per i gruppi di sicurezza per aprire le porte. Assicurati di aggiungere regole che permettano solo il traffico da client affidabili e autenticati per i protocolli e le porte necessari per eseguire i tuoi carichi di lavoro.

Componente	Descrizione del servizio	Servizio in esecuzione per impostazione predefinita	Porta	Chiave di configurazione
Hadoop	HTTP KMS REST API	Sì	9600	hadoop.kms.http.port
HDFS	Interfaccia utente Web di Namenode	Sì	9870	dfs.namenode.http-address

Componente	Descrizione del servizio	Servizio in esecuzione per impostazione predefinita	Porta	Chiave di configurazione
	NameNode RPC	Sì	8020	dfs.namenode.rpc-indirizzo-pc
	DataNode Interfaccia utente Web	Sì	9864	dfs.datanode.http.address
	Datanode HTTP per il trasferimento di dati	Sì	9866	dfs.datanode.address
	Datanode RPC per il trasferimento dei dati	Sì	9867	dfs.datanode.ipc.address
Hive	HiveServer2 Parsimonia	Sì	10000	hive.server2.thrift.port
	HiveServer2 HTTP	No	10001	hive.server2.thrift.http.port
	HiveServer2 Interfaccia utente Web	Sì	10002	hive.server2.webui.port
	Hive Metastore	Sì	9083	hive.metastore.port/metastore.thrift.port
	Web HCat	No	5011	templeton.port
	Servizio di gestione daemon (RPC) LLAP	No	15004	hive.llap.management.rpc.port

Componente	Descrizione del servizio	Servizio in esecuzione per impostazione predefinita	Porta	Chiave di configurazione
	Porta shuffle YARN per shuffle LLAP-daemon-hosted	No	15551	hive.llap.daemon.yarn.shuffle.port
	RPC daemon LLAP	No	Dinamico	hive.llap.daemon.rpc.port
	Interfaccia utente Web daemon LLAP	No	15002	hive.llap.daemon.web.port
	Servizio di output daemon LLAP	No	15003	hive.llap.daemon.output.service.port
Oozie		Sì	11000	
Tez	Interfaccia utente Tez	Sì	8080	
YARN	Shuffle	Sì	13562	mapreduce.shuffle.port
	Localizzatore RPC	Sì	8040	yarn.nodemanager.localizer.address
		Sì	8041	
	Indirizzo di NM Webapp	Sì	8042	yarn.nodemanager.webapp.address

Componente	Descrizione del servizio	Servizio in esecuzione per impostazione predefinita	Porta	Chiave di configurazione
	Applicazione web RM	Sì	8088	yarn.resourcemanager.webapp.address
		Sì	8025	
	Pianificatore	Sì	8030	yarn.resourcemanager.scheduler.address
	Interfaccia del gestore delle applicazioni	Sì	8032	yarn.resourcemanager.address
	Interfaccia amministratore di RM	Sì	8033	yarn.resourcemanager.admin.address
	JobHistory Interfaccia utente Web del server	Sì	1988	mapreduce.jobhistory.webapp.address
	JobHistory Interfaccia utente Web di amministrazione del server	Sì	10033	mapreduce.jobhistory.admin.address
	JobHistory Server (RPC)	Sì	10020	mapreduce.jobhistory.address

Componente	Descrizione del servizio	Servizio in esecuzione per impostazione predefinita	Porta	Chiave di configurazione
	Application Timeline Server (RPC)	Sì	10200	yarn.timeline-service.indirizzo
	Interfaccia utente Web HTTP di Application Timeline Server	Sì	8188	yarn.timeline-service.webapp.address
	Interfaccia utente Web HTTPS di Application Timeline Server	No	8190	yarn.timeline-service.webapp.https.address
		Sì	20888	
Zookeeper	Porta client	Sì	2181	
		Sì	37301	
		Sì	8341	

Visualizzazione di interfacce Web ospitate su cluster Amazon EMR

Important

È possibile configurare un gruppo di sicurezza personalizzato per consentire l'accesso in entrata a queste interfacce Web. Tenere presente che qualsiasi porta su cui si consente il traffico in entrata rappresenta una potenziale vulnerabilità per la sicurezza. Esaminare attentamente i gruppi di sicurezza personalizzati per assicurarsi di ridurre al minimo le vulnerabilità. Per ulteriori informazioni, consulta Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Hadoop e altre applicazioni installate sul cluster EMR pubblicano interfacce utente come siti Web ospitati sul nodo primario. Per motivi di sicurezza, quando utilizzi i gruppi di sicurezza gestiti da Amazon EMR, questi siti Web sono disponibili solo sul server Web locale del nodo primario. Per questo motivo, devi connetterti al nodo primario per visualizzare le interfacce Web. Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH. Hadoop pubblica anche le interfacce utente come siti Web ospitati sui nodi di task e principali. Questi siti web sono disponibili anche solo sui server web locali dei nodi.

La tabella seguente elenca le interfacce Web che puoi visualizzare sulle istanze di cluster. Queste interfacce di Hadoop sono disponibili su tutti i cluster. Per le interfacce dell'istanza master, *master-public-dns-name* sostituiscile con il DNS pubblico principale elencato nella scheda Riepilogo del cluster nella console Amazon EMR. Per le interfacce di base e di istanza task, sostituiscile *coretask-public-dns-name* con il nome DNS pubblico elencato per l'istanza. Per trovare un'istanza Nome DNS pubblico, nella console Amazon EMR, scegli il cluster dall'elenco, quindi la scheda Hardware, l'ID del gruppo di istanze che contiene l'istanza a cui connetterti, infine annota il Nome DNS pubblico indicato per l'istanza.

Nome dell'interfaccia	URI
Flink History Server (EMR versione 5.33 e successive)	http: //:8082/ *master-public-dns-name*
Ganglia	http://ganglia/ *master-public-dns-name*
Hadoop HDFS (versione NameNode EMR precedente alla 6.x)	https: *master-public-dns-name* //50470/
Hadoop HDFS NameNode	http: //:50070/ *master-public-dns-name*
Hadoop HDFS DataNode	http: //:50075/ *coretask-public-dns-name*
Hadoop HDFS (NameNode EMR versione 6.x)	*master-public-dns-name* https://:9870/
Hadoop HDFS (versione DataNode EMR precedente alla 6.x)	https: *coretask-public-dns-name* //:50475/

Nome dell'interfaccia	URI
Hadoop HDFS (DataNode EMR versione 6.x)	https: *coretask-public-dns-name* //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*
Scintilla HistoryServer	http: //:18080/ *master-public-dns-name*
Tez	http://8080/tez-ui *master-public-dns-name*
FILATO NodeManager	http: //:8042/ *coretask-public-dns-name*
FILATO ResourceManager	http: //:8088/ *master-public-dns-name*
Zeppelin	http: //:8890/ *master-public-dns-name*

Poiché sul nodo primario sono disponibili diverse interfacce specifiche per l'applicazione che non sono disponibili sui nodi core e attività, le istruzioni di questo documento sono specifiche per il nodo primario Amazon EMR. È possibile accedere alle interfacce Web sui nodi core e attività nello stesso modo in cui si accede alle interfacce Web sul nodo primario.

Esistono diversi modi per accedere alle interfacce Web sul nodo primario. Il metodo più semplice e veloce è utilizzare SSH per connettersi al nodo primario e usare il browser basato su testo, Lynx, per visualizzare i siti Web nel proprio client SSH. Tuttavia, Lynx è un browser basato su testo con un'interfaccia utente limitata che non è in grado di visualizzare la grafica. L'esempio seguente mostra come aprire l' ResourceManager interfaccia Hadoop utilizzando Lynx (Lynx viene fornito anche quando si accede al URLs nodo primario tramite SSH).

```
lynx http://ip-###-##-##-###.us-west-2.compute.internal:8088/
```

Esistono altre due opzioni per accedere alle interfacce Web sul nodo primario che forniscono caratteristiche complete per il browser. Seleziona una delle seguenti opzioni:

- Opzione 1 (raccomandata per gli utenti più esperti): utilizzare un client SSH per connettersi al nodo primario, configurare il tunneling SSH con l'inoltro porta locale e utilizzare un browser Internet per aprire le interfacce Web ospitate sul nodo primario. Questo metodo consente di configurare l'accesso all'interfaccia web senza utilizzare un proxy SOCKS.
- Opzione 2 (consigliata per i nuovi utenti): utilizza un client SSH per connetterti al nodo primario, configura il tunneling SSH con il port forwarding dinamico e configura il tuo browser Internet per utilizzare un componente aggiuntivo come FoxyProxy Firefox o Chrome per gestire le impostazioni del proxy SOCKS. SwitchyOmega Questo metodo consente di filtrare automaticamente in URLs base a modelli di testo e di limitare le impostazioni del proxy ai domini che corrispondono al formato del nome DNS del nodo primario. Per ulteriori informazioni su come configurare FoxyProxy Firefox e Google Chrome, consulta. Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR

> **Note**
>
> Se modifichi la porta in cui un'applicazione viene eseguita tramite la configurazione del cluster, il collegamento ipertestuale alla porta non verrà aggiornato nella console di Amazon EMR. Questo avviene perché la console non dispone di una caratteristica per leggere la configurazione `server.port`.

Con Amazon EMR versione 5.25.0 o successiva, puoi accedere all'interfaccia utente del server della cronologia Spark dalla console senza configurare un proxy Web tramite una connessione SSH. Per ulteriori informazioni, consulta Accesso con un clic a Spark History Server persistente.

Argomenti

- Opzione 1: configura un tunnel SSH verso il nodo primario di Amazon EMR utilizzando il port forwarding locale
- Opzione 2, parte 1: impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta dinamico
- Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR

Opzione 1: configura un tunnel SSH verso il nodo primario di Amazon EMR utilizzando il port forwarding locale

Per connetterti al server Web locale sul nodo primario, crea un tunnel SSH tra il computer e il nodo primario. È noto anche come inoltro porta. Se non desideri utilizzare un proxy SOCKS, puoi impostare un tunnel SSH per il nodo primario utilizzando l'inoltro porta locale. Con l'inoltro porta locale, si specificano le porte locali inutilizzate che vengono impiegate per inoltrare il traffico a specifiche porte remote sul server Web locale del nodo primario.

La configurazione di un tunnel SSH utilizzando l'inoltro porta locale richiede il nome DNS pubblico del nodo primario e il file di chiave privata della coppia di chiavi. Per ulteriori informazioni sull'individuazione del nome DNS pubblico master, vedere Recupero del nome DNS pubblico del nodo primario. Per ulteriori informazioni sull'accesso alla tua coppia di chiavi, consulta Amazon EC2 key pair nella Amazon EC2 User Guide. Per ulteriori informazioni sui siti da visualizzare nel nodo primario, consulta la sezione Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta locale con OpenSSH

Per impostare un tunnel SSH utilizzando l'inoltro porta locale nel terminale

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Apri una finestra del terminale. In Mac OS X, selezionare Applications > Utilities > Terminal (Applicazioni > Utility > Terminale). In altre distribuzioni Linux, la finestra terminal si trova generalmente in Applications > Accessories > Terminal (Applicazioni > Accessori >; Terminale).
3. Digita il seguente comando per aprire un tunnel SSH sul tuo computer locale. Questo comando di esempio accede all'interfaccia ResourceManager Web inoltrando il traffico sulla porta locale 8157 (una porta locale non utilizzata scelta a caso) alla porta 8088 sul server Web locale del nodo master.

   Nel comando, sostituisci *~/mykeypair.pem* con la posizione e il nome del file e sostituiscilo *ec2-###-##-##-###.compute-1.amazonaws.com* con il nome `.pem` DNS pubblico principale del cluster. Per accedere a un'interfaccia web diversa, sostituiscila `8088` con il numero di porta appropriato. Ad esempio, sostituisci `8088` con `8890` per l'interfaccia Zeppelin.

   ```
   ssh -i ~/mykeypair.pem -N -L 8157:ec2-###-##-##-
   ###.compute-1.amazonaws.com:8088 hadoop@ec2-###-##-##-###.compute-1.amazonaws.com
   ```

`-L` indica l'uso dell'inoltro porta locale che consente di specificare una porta locale utilizzata per inoltrare i dati alla porta remota identificata sul server Web locale del nodo master.

Dopo l'emissione di questo comando, il terminale rimane aperto e non risponde.

4. Per aprire l'interfaccia ResourceManager web nel browser, digita `http://localhost:8157/` nella barra degli indirizzi.
5. Una volta terminato il lavoro con le interfacce Web sul nodo primario, chiudi le finestre del terminale.

Opzione 2, parte 1: impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta dinamico

Per connetterti al server Web locale sul nodo primario, crea un tunnel SSH tra il computer e il nodo primario. È noto anche come inoltro porta. Se crei il tunnel SSH utilizzando l'inoltro porte dinamico, tutto il traffico instradato verso una specifica porta locale inutilizzata viene inoltrato al server Web locale sul nodo primario. In questo modo viene creato un proxy SOCKS. È quindi possibile configurare il browser Internet per utilizzare un componente aggiuntivo come FoxyProxy o SwitchyOmega per gestire le impostazioni del proxy SOCKS.

L'utilizzo di un componente aggiuntivo per la gestione dei proxy consente di filtrare automaticamente in URLs base a modelli di testo e di limitare le impostazioni del proxy ai domini che corrispondono al formato del nome DNS pubblico del nodo primario. Il componente aggiuntivo del browser gestisce automaticamente l'attivazione e la disattivazione del proxy quando si passa dalla visualizzazione di siti Web ospitati sul nodo primario a quella di siti Web su Internet.

Prima di iniziare, devi disporre del nome DNS pubblico del nodo primario e del file di chiave privata della coppia di chiavi. Per informazioni sull'individuazione del nome DNS pubblico primario, consulta Recupero del nome DNS pubblico del nodo primario. Per ulteriori informazioni sull'accesso alla tua coppia di chiavi, consulta Amazon EC2 key pair nella Amazon EC2 User Guide. Per ulteriori informazioni sui siti da visualizzare nel nodo primario, consulta la sezione Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta dinamico su OpenSSH

Impostazione di un tunnel SSH utilizzando l'inoltro porta dinamico con OpenSSH

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.

2. Apri una finestra del terminale. In Mac OS X, selezionare Applications > Utilities > Terminal (Applicazioni > Utility > Terminale). In altre distribuzioni Linux, la finestra terminal si trova generalmente in Applications > Accessories > Terminal (Applicazioni > Accessori >; Terminale).
3. Digitare il seguente comando per aprire un tunnel SSH sulla macchina locale. Sostituiscilo *~/mykeypair.pem* con la posizione e il nome del `.pem` file, *8157* sostituiscilo con un numero di porta locale non utilizzato e *ec2-###-##-##-###.compute-1.amazonaws.com* sostituiscilo con il nome DNS pubblico principale del cluster.

   ```
   ssh -i ~/mykeypair.pem -N -D 8157 hadoop@ec2-###-##-##-###.compute-1.amazonaws.com
   ```

   Dopo l'emissione di questo comando, il terminale rimane aperto e non risponde.

   > **Note**
   >
   > `-D` indica l'uso dell'inoltro porta dinamico, che consente di specificare una porta locale utilizzata per inoltrare i dati verso tutte le porte remote sul server Web locale del nodo primario. L'inoltro dinamico della porta crea un proxy SOCKS locale ascolto sulla porta specificato nel comando.

4. Dopo che il tunnel è attivo, configurare un proxy SOCKS per il browser. Per ulteriori informazioni, consulta Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR.
5. Una volta terminato il lavoro con le interfacce Web sul nodo primario, chiudi la finestra del terminale.

Configura un tunnel SSH utilizzando il port forwarding dinamico con AWS CLI

È possibile creare una connessione SSH con il nodo primario utilizzando Windows e Linux, Unix e Mac OS X. Se si utilizza il nodo AWS CLI su Linux, Unix o Mac OS X, è necessario impostare le autorizzazioni sul file come mostrato in. AWS CLI `.pem` Per configurare le autorizzazioni per i file chiave privati della coppia di chiavi Se si utilizza AWS CLI su Windows, PuTTY deve apparire nella variabile di ambiente path o è possibile che venga visualizzato un errore come OpenSSH o PuTTY non disponibile.

Per configurare un tunnel SSH utilizzando il port forwarding dinamico con AWS CLI

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Crea una connessione SSH con il nodo primario come illustrato nella Connessione al nodo primario tramite la AWS CLI.
3. Per recuperare l'identificatore del cluster, digitare:

   ```
   aws emr list-clusters
   ```

   L'output elenca i cluster, incluso il cluster. IDs Prendere nota dell'ID del cluster per il cluster a cui ci si connette.

   ```
   "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. Per aprire un tunnel SSH con il nodo primario utilizzando l'inoltro alla porta dinamico, digita il seguente comando. Nell'esempio seguente, sostituitelo *j-2AL4XXXXXX5T9* con l'ID del cluster e *~/mykeypair.key* sostituitelo con la posizione e il nome del `.pem` file (per Linux, Unix e Mac OS X) o `.ppk` del file (per Windows).

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

   > **Note**
   >
   > Il comando socks configura automaticamente l'inoltro dinamico della porta sulla porta locale 8157. Al momento, questa impostazione non può essere modificata.

5. Dopo che il tunnel è attivo, configurare un proxy SOCKS per il browser. Per ulteriori informazioni, consulta Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR.
6. Quando hai finito di lavorare con le interfacce web sul nodo principale, chiudi la AWS CLI finestra.

   Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Configurazione di un tunnel SSH per il nodo primario utilizzando PuTTY

Gli utenti Windows possono usare un client SSH come PuTTY per creare un tunnel SSH al nodo primario. Prima di connetterti al nodo primario di Amazon EMR, devi scaricare e installare PuTTY e Pu. TTYgen Questi strumenti possono essere scaricati dalla pagina di download di PuTTY.

PuTTY non supporta nativamente il formato di file della chiave privata della coppia di chiavi `.pem` () generato da Amazon. EC2 Utilizzate Pu TTYgen per convertire il file chiave nel formato PuTTY richiesto (). `.ppk` È necessario convertire la chiave privata nel formato `.ppk` prima di tentare una connessione al nodo primario tramite PuTTY.

Per ulteriori informazioni sulla conversione della chiave, consulta Convertire la chiave privata utilizzando Pu TTYgen nella Amazon EC2 User Guide.

Impostazione di un tunnel SSH utilizzando l'inoltro porta dinamico su PuTTY

1. Assicurati di aver consentito il traffico SSH in entrata. Per istruzioni, consultare Prima di connetterti ad Amazon EMR: autorizza il traffico in entrata.
2. Fai doppio clic su `putty.exe` per avviare PuTTY. È anche possibile avviare PuTTY dall'elenco dei programmi di Windows.

   > **Note**
   >
   > Se disponi già di una sessione SSH attiva con il nodo primario, puoi aggiungere un tunnel facendo clic con il pulsante destro del mouse sulla barra del titolo di PuTTY e scegliendo Change Settings (Modifica impostazioni).

3. Se necessario, nell'elenco Category (Categoria) selezionare Session (Sessione).

4. Nel campo Nome host, digita. **hadoop@** *MasterPublicDNS* Ad esempio: **hadoop@***ec2-###-##-##-###.compute-1.amazonaws.com*.
5. Nell'elenco Category (Categoria), espandere Connection > SSH (Connessione > SSH), quindi scegliere Auth (Autenticazione).
6. Per Private key file for authentication (File chiave privata per autenticazione), scegliere Browse (Sfoglia) e selezionare il file `.ppk` generato.

   > **Note**
   >
   > PuTTY non supporta nativamente il formato di file della chiave privata della coppia di chiavi `.pem` () generato da Amazon. EC2 Utilizzate Pu TTYgen per convertire il file chiave nel formato PuTTY richiesto (). `.ppk` È necessario convertire la chiave privata nel formato `.ppk` prima di tentare una connessione al nodo primario tramite PuTTY.

7. Nell'elenco Category (Categoria), espandere Connection > SSH (Connessione > SSH), quindi scegliere Tunnels (Tunnel).
8. Nel campo Source port (Porta sorgente), digita `8157` (un numero di porta locale inutilizzato), quindi seleziona Add (Aggiungi).
9. Lasciare vuoto il campo Destination (Destinazione).
10. Selezionare le opzioni Dynamic (Dinamico) e Auto (Automatico).
11. Scegliere Open (Apri).
12. Scegliere Yes (Sì) per ignorare l'avviso di sicurezza PuTTY.

    > **Important**
    >
    > Quando accedi al nodo primario, inserisci `hadoop` nel caso in cui venga richiesto un nome utente.

13. Dopo che il tunnel è attivo, configurare un proxy SOCKS per il browser. Per ulteriori informazioni, consulta Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR.
14. Una volta terminato il lavoro con le interfacce Web sul nodo primario, chiudi la finestra di PuTTY.

Opzione 2, parte 2: configura le impostazioni proxy per visualizzare i siti Web ospitati sul nodo primario del cluster Amazon EMR

Se si utilizza un tunnel SSH con inoltro dinamico delle porte, è necessario utilizzare un add-on per la gestione dei proxy SOCKS per controllare le impostazioni proxy nel browser. L'utilizzo di uno strumento di gestione proxy SOCKS consente di filtrare automaticamente in URLs base a modelli di testo e di limitare le impostazioni del proxy ai domini che corrispondono alla forma del nome DNS pubblico del nodo primario. Il componente aggiuntivo del browser gestisce automaticamente l'attivazione e la disattivazione del proxy quando si passa dalla visualizzazione di siti Web ospitati sul nodo primario a quella di siti Web su Internet. Per gestire le impostazioni del proxy, configura il browser per utilizzare un componente aggiuntivo come o. FoxyProxy SwitchyOmega

Per ulteriori informazioni sulla creazione di un tunnel SSH, consulta Opzione 2, parte 1: impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta dinamico. Per ulteriori informazioni sulle interfacce Web, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Includi le seguenti impostazioni quando configuri il componente aggiuntivo proxy:

- Utilizza localhost come indirizzo host.
- Utilizza lo stesso numero di porta locale che hai selezionato per stabilire il tunnel SSH con il nodo primario in Opzione 2, parte 1: impostazione di un tunnel SSH sul nodo primario utilizzando l'inoltro porta dinamico. Ad esempio, port*8157*. Questa porta deve anche corrispondere al numero di porta utilizzato in PuTTY o in un qualsiasi altro emulatore di terminale utilizzato per il collegamento.
- Specifica il valore del protocollo SOCKS v5. SOCKS v5 consente di impostare facoltativamente l'autorizzazione utente.
- URL Patterns (Modelli URL)

  I seguenti modelli URL dovrebbero essere nella whitelist e specificati con un tipo di modello jolly:

  - I modelli *ec2*.*compute*.amazonaws.com* e *10*.amazonaws.com* corrispondono al nome DNS pubblico dei cluster nelle regioni US.
  - I modelli *ec2*.compute* e *10*.compute* corrispondono al nome DNS pubblico dei cluster in tutte le altre regioni.
  - Un 10. * modello per fornire l'accesso ai file di JobTracker registro in Hadoop. Modificare questo filtro se entra in conflitto con il piano di accesso della rete.
  - I modelli *.ec2.internal* e *.compute.internal* devono corrispondere ai nomi DNS privati (interni) dei cluster rispettivamente nella regione `us-east-1` e in tutte le altre regioni.

Esempio: configurazione per Firefox FoxyProxy

L'esempio seguente mostra una configurazione FoxyProxy Standard (versione 7.5.1) per Mozilla Firefox.

FoxyProxy fornisce una serie di strumenti di gestione dei proxy. Ti consente di utilizzare un server proxy per URLs i modelli di corrispondenza corrispondenti ai domini utilizzati dalle EC2 istanze Amazon nel tuo cluster Amazon EMR.

Per installare e configurare FoxyProxy utilizzando Mozilla Firefox

1. In Firefox, vai su https://addons.mozilla.org/, cerca FoxyProxy Standard e segui le istruzioni per aggiungerlo FoxyProxy a Firefox.
2. Mediante un editor di testo, crea un file JSON denominato `foxyproxy-settings.json` dalla configurazione di esempio riportata di seguito:

```
{
  "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. Apri la pagina Manage Your Extensions (Gestisci estensioni) di Firefox (vai a about:addons, quindi seleziona Extensions (Estensioni).
4. Scegli FoxyProxy Standard, quindi scegli il pulsante Altre opzioni (il pulsante che assomiglia a un puntino di sospensione).
5. Seleziona Options (Opzioni) dal menu a discesa.
6. Seleziona Import Settings (Importa impostazioni) dal menu a sinistra.
7. Nella pagina Impostazioni di importazione, scegli Importa impostazioni in Importa impostazioni dalla versione FoxyProxy 6.0+, individua la posizione del **`foxyproxy-settings.json`** file che hai creato, seleziona il file e scegli Apri.
8. Scegli OK quando il sistema richiede di sovrascrivere le impostazioni esistenti e salvare la nuova configurazione.

Esempio: configura per chrome SwitchyOmega

L'esempio seguente mostra come configurare l' SwitchyOmegaestensione per Google Chrome. SwitchyOmega consente di configurare, gestire e passare da un proxy all'altro.

Per installare e configurare SwitchyOmega utilizzando Google Chrome

1. Vai alla https://chrome.google.com/webstore/categoria/estensioni, cerca Proxy SwitchyOmega e aggiungilo a Chrome.
2. Seleziona New profile (Nuovo profilo) e immetti `emr-socks-proxy` come nome del profilo.
3. Seleziona PAC profile (Profilo PAC) e in seguito Create (Crea). I file Proxy Auto-Configuration (PAC) (Configurazione automatica proxy (PAC)) consentono di definire un elenco di autorizzazioni per le richieste del browser che devono essere inoltrate a un server proxy Web.

4. Nel campo Script PAC, sostituisci il contenuto con lo script seguente che definisce quale URLs deve essere inoltrato tramite il tuo server proxy web. Se hai specificato un numero di porta diverso durante la configurazione del tunnel SSH, sostituiscilo *8157* con il tuo numero di porta.

```
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. In Actions (Operazioni), scegli Apply changes (Applica modifiche) per salvare le impostazioni del proxy.
6. Nella barra degli strumenti di Chrome, scegli SwitchyOmega e seleziona il `emr-socks-proxy` profilo.

Accesso a un'interfaccia Web nel browser

Per aprire un'interfaccia Web, inserisci il nome DNS pubblico del nodo primario o core seguito dal numero di porta dell'interfaccia scelta nella barra degli indirizzi del browser. L'esempio seguente mostra l'URL da inserire per connetterti a HistoryServer Spark.

```
http://master-public-dns-name:18080/
```

Per istruzioni su come recuperare il nome DNS pubblico di un nodo, consulta Recupero del nome DNS pubblico del nodo primario. Per un elenco completo delle interfacce web URLs, vediVisualizzazione di interfacce Web ospitate su cluster Amazon EMR.

Invia il lavoro a un cluster Amazon EMR

Questa sezione descrive i metodi che puoi utilizzare per l'invio di lavoro a un cluster Amazon EMR. Per inviare un lavoro, puoi aggiungere fasi oppure puoi inviare processi Hadoop al nodo primario in modo interattivo.

Quando invii fasi a un cluster, considera le seguenti regole di comportamento:

- Un ID fase può contenere fino a 256 caratteri.
- Puoi avere fino a 256 fasi PENDING e RUNNING in un cluster.
- Puoi inviare processi al nodo primario in modo interattivo anche se hai 256 fasi attive in esecuzione sul cluster. Puoi inviare un numero illimitato di fasi durante l'intero ciclo di vita di un cluster di lunga durata, ma soltanto 256 fasi possono essere RUNNING (IN ESECUZIONE) o PENDING (IN SOSPESO) in un determinato momento.
- Con Amazon EMR versioni 4.8.0 e successive, a eccezione della versione 5.0.0, puoi annullare le fasi in sospeso. Per ulteriori informazioni, consulta Annulla i passaggi quando invii il lavoro a un cluster Amazon EMR.
- Con Amazon EMR versioni 5.28.0 e successive, puoi annullare sia le fasi in sospeso che quelle in esecuzione. Puoi inoltre scegliere di eseguire più fasi in parallelo per migliorare l'utilizzo del cluster e risparmiare sui costi. Per ulteriori informazioni, consulta Considerazioni sull'esecuzione di più passaggi in parallelo quando invii un lavoro ad Amazon EMR.

> **Note**
>
> Per prestazioni ottimali, ti consigliamo di archiviare azioni di bootstrap personalizzate, script e altri file che desideri utilizzare con Amazon EMR in un bucket Amazon S3 che si trova nello stesso cluster. Regione AWS

Argomenti

- Aggiunta di fasi a un cluster con la console di gestione Amazon EMR
- Aggiungere passaggi a un cluster Amazon EMR con AWS CLI
- Considerazioni sull'esecuzione di più passaggi in parallelo quando invii un lavoro ad Amazon EMR
- Visualizzazione dei passaggi dopo l'invio del lavoro a un cluster Amazon EMR
- Annulla i passaggi quando invii il lavoro a un cluster Amazon EMR

Aggiunta di fasi a un cluster con la console di gestione Amazon EMR

Utilizza le procedure seguenti per aggiungere fasi a un cluster utilizzando la AWS Management Console. Per informazioni dettagliate su come inviare fasi per applicazioni Big Data specifiche, consulta le seguenti sezioni della Guida ai rilasci di Amazon EMR:

- Invio di una fase JAR personalizzata
- Invio di una fase di streaming Hadoop
- Invio di una fase Spark
- Invio di una fase Pig
- Esecuzione di un comando o di uno script come fase
- Passaggio di valori in fasi per l'esecuzione di script Hive

Aggiungere fasi durante la creazione del cluster

Da AWS Management Console, puoi aggiungere passaggi quando crei un cluster.

Console

Per aggiungere passaggi quando si crea un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. In Steps (Fasi), scegli Add step (Aggiungi fase). Inserisci i valori appropriati nei campi della finestra di dialogo Add step (Aggiungi fase). Per informazioni sulla formattazione degli argomenti delle fasi, consulta Aggiungere argomenti di fase. Le opzioni variano a seconda del tipo di fase. Per aggiungere la fase e chiudere la finestra di dialogo, scegli Aggiungi fase.
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

Aggiungere fasi a un cluster in esecuzione

Con AWS Management Console, puoi aggiungere passaggi a un cluster con l'opzione di terminazione automatica disabilitata.

Console

Per aggiungere passaggi a un cluster in esecuzione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.

2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Steps (Fasi) della pagina dei dettagli del cluster, scegli Add step (Aggiungi fase). Per clonare una fase esistente, scegli il menu a discesa Actions (Operazioni) e seleziona Clone step (Clona fase).
4. Inserisci i valori appropriati nei campi della finestra di dialogo Add step (Aggiungi fase). Le opzioni variano a seconda del tipo di fase. Per aggiungere la fase e uscire dalla finestra di dialogo, scegli Add step (Aggiungi fase).

Modificare il livello di concorrenza delle fasi in un cluster in esecuzione

Con AWS Management Console, puoi modificare il livello di concorrenza dei passaggi in un cluster in esecuzione.

> **Note**
>
> Puoi eseguire più fasi in parallelo solo con Amazon EMR versione 5.28.0 e successive.

Console

Per modificare la concorrenza dei passaggi in un cluster in esecuzione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare. Il cluster deve essere in esecuzione per modificare l'attributo di simultaneità.
3. Nella scheda Steps (Fasi) della pagina dei dettagli del cluster, cerca la sezione Attributes (Attributi). Seleziona Edit (Modifica) per modificare la simultaneità. Inserisci un valore compreso tra 1 e 256.

Aggiungere argomenti di fase

Quando usi AWS Management Console per aggiungere un passo al tuo cluster, puoi specificare gli argomenti per quel passaggio nel campo Argomenti. È necessario separare gli argomenti con spazi bianchi e circondare gli argomenti di stringa costituiti da caratteri e spazio bianco con virgolette.

Example : Argomenti corretti

Gli argomenti di esempio seguenti sono formattati correttamente per l' AWS Management Console argomento stringa finale racchiuso tra virgolette.

```
bash -c "aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh ."
```

Puoi inoltre inserire ogni argomento su una riga separata per la leggibilità come mostrato nell'esempio seguente.

```
bash
-c
"aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh ."
```

Example : Argomenti errati

I seguenti argomenti di esempio sono formattati in modo improprio per la AWS Management Console. Notate che l'argomento finale della stringa, `aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh .`, contiene spazi bianchi e non è circondato da virgolette.

```
bash -c aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh .
```

Aggiungere passaggi a un cluster Amazon EMR con AWS CLI

Le seguenti procedure illustrano come aggiungere fasi a un cluster appena creato e a un cluster in esecuzione con la AWS CLI. In entrambi gli esempi, viene utilizzato il sottocomando `--steps` per aggiungere fasi al cluster.

Per aggiungere fasi durante la creazione del cluster

- Digitare il comando seguente per creare un cluster e aggiungere una fase Apache Pig. Assicurati di sostituirlo *myKey* con il nome della tua coppia di EC2 chiavi 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**
>
> L'elenco di argomenti cambia in funzione del tipo di fase.

Per impostazione predefinita, il livello di concorrenza delle fasi è 1. Puoi impostare il livello di simultaneità delle fasi utilizzando il parametro `StepConcurrencyLevel` quando crei un cluster.

L'output è un identificatore di cluster simile a quanto segue.

```
{
    "ClusterId": "j-2AXXXXXXGAPLF"
}
```

Per aggiungere una fase a un cluster in esecuzione

- Digitare il comando seguente per aggiungere una fase a un cluster in esecuzione. Sostituisci *j-2AXXXXXXGAPLF* con l'ID del tuo cluster.

```
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"}]'
```

L'output è un identificatore di fase simile a quanto segue.

```
{
    "StepIds": [
 "s-Y9XXXXXXAPMD"
    ]
}
```

Per modificarli StepConcurrencyLevel in un cluster in esecuzione

1. In un cluster in esecuzione, puoi modificare il `StepConcurrencyLevel` con l'API `ModifyCluster`. Ad esempio, digita il comando seguente per aumentare `StepConcurrencyLevel` fino a `10`. Sostituisci *`j-2AXXXXXXGAPLF`* con l'ID del tuo cluster.

   ```
   aws emr modify-cluster --cluster-id j-2AXXXXXXGAPLF --step-concurrency-level 10
   ```

2. L'output è simile a quello riportato di seguito.

   ```
   {
   "StepConcurrencyLevel": 10
   }
   ```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta il AWS CLI Command Reference.

Considerazioni sull'esecuzione di più passaggi in parallelo quando invii un lavoro ad Amazon EMR

L'esecuzione di più passaggi in parallelo quando invii un lavoro ad Amazon EMR richiede decisioni preliminari sulla pianificazione delle risorse e sulle aspettative relative al comportamento del cluster. Questi sono descritti in dettaglio qui.

- Le fasi in esecuzione in parallelo possono essere completate in qualsiasi ordine, ma le fasi in sospeso nella coda passano allo stato in esecuzione nell'ordine in cui sono state inviate.
- Quando si seleziona un livello di simultaneità delle fasi per il cluster, è necessario considerare se il tipo di istanza del nodo primario soddisfa o meno i requisiti di memoria dei carichi di lavoro degli utenti. Il processo di esecuzione della fase principale viene eseguito sul nodo primario per ogni fase. L'esecuzione di più fasi in parallelo richiede più memoria e utilizzo della CPU del nodo primario rispetto all'esecuzione di una fase alla volta.
- Per ottenere la pianificazione complessa e la gestione delle risorse delle fasi simultanee, è possibile utilizzare caratteristiche di programmazione YARN come `FairScheduler` o `CapacityScheduler`. Ad esempio, è possibile utilizzare `FairScheduler` con un set `queueMaxAppsDefault` per impedire l'esecuzione di più di un certo numero di processi contemporaneamente.

- Il livello di concorrenza delle fasi è soggetto alle configurazioni dei gestori delle risorse. Ad esempio, se YARN è configurato con un solo parallelismo di `5`, allora è possibile avere solo cinque applicazioni YARN in esecuzione in parallelo anche se `StepConcurrencyLevel` è impostato su `10`. Per ulteriori informazioni sulla configurazione dei gestori delle risorse, consulta Configurazione delle applicazioni nella Guida al rilascio di Amazon EMR.
- Non è possibile aggiungere una fase con un `ActionOnFailure` diverso da CONTINUE (CONTINUA) mentre il livello di concorrenza della fase del cluster è maggiore di 1.
- Se il livello di concorrenza della fase di un cluster è maggiore di uno, la caratteristica `ActionOnFailure` della fase non si attiverà.
- Se un cluster ha un livello di concorrenza della fase 1 ma ha più fasi in esecuzione, `TERMINATE_CLUSTER ActionOnFailure` potrebbe attivarsi, ma `CANCEL_AND_WAIT ActionOnFailure` non lo farà. Questo caso limite si verifica quando il livello di concorrenza della fase del cluster era maggiore di uno, ma si è abbassato durante l'esecuzione di più fasi.
- È possibile utilizzare la scalabilità automatica di EMR per aumentare o ridurre in base alle risorse YARN evitando per evitare conflitti tra risorse. Per ulteriori informazioni, consulta Utilizzo del dimensionamento automatico con una policy personalizzata per i gruppi di istanze nella Guida alla gestione di Amazon EMR.
- Quando si riduce il livello di concorrenza delle fasi, EMR consente di completare tutte le fasi in esecuzione prima di ridurne il numero. Se le risorse sono esaurite perché il cluster esegue troppi passaggi simultanei, è consigliabile annullare manualmente tutte le fasi in esecuzione per liberare risorse.

Visualizzazione dei passaggi dopo l'invio del lavoro a un cluster Amazon EMR

Puoi vedere fino a 10.000 passaggi completati da Amazon EMR negli ultimi sette giorni. Puoi anche visualizzare 1.000 passaggi completati da Amazon EMR in qualsiasi momento. Questo totale include le fasi inviate dall'utente e le fasi di sistema.

Se invii nuovi passaggi una volta che il cluster raggiunge il limite di record di 1.000 passaggi, Amazon EMR elimina i passaggi inattivi inviati dall'utente il cui stato è COMPLETATO, ANNULLATO o FALLITO per più di sette giorni. Se invii passaggi oltre il limite di 10.000 passaggi, Amazon EMR elimina i record dei passaggi inattivi inviati dall'utente indipendentemente dalla loro durata di inattività. Amazon EMR non rimuove questi record dai file di registro. Amazon EMR li rimuove dalla AWS

console e non vengono restituiti quando usi l'API AWS CLI o per recuperare le informazioni sul cluster. I record relativi alle fasi di sistema non vengono mai rimossi.

Le informazioni sulle fasi che puoi visualizzare dipendono dal meccanismo utilizzato per recuperare le informazioni del cluster. La tabella seguente indica le informazioni sulle fasi restituite da ciascuna delle opzioni disponibili.

Opzione	DescribeJobFlow o --describe --jobflow	ListSteps o list-steps
SDK	256 fasi	Fino a 10.000 passaggi
CLI di Amazon EMR	256 fasi	N/A
AWS CLI	N/A	Fino a 10.000 passaggi
API	256 fasi	Fino a 10.000 passaggi

Annulla i passaggi quando invii il lavoro a un cluster Amazon EMR

Puoi annullare i passaggi in sospeso e in esecuzione di Amazon EMR o Amazon EMR quando invii il lavoro al tuo cluster. AWS Management Console AWS CLI API.

Console

Per annullare i passaggi con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi seleziona il cluster che desideri aggiornare.
3. Nella scheda Steps (Fasi) della pagina dei dettagli del cluster, seleziona la casella di controllo accanto alla fase da annullare. Scegli il menu a discesa Actions (Operazioni) e seleziona Cancel steps (Annulla fasi).
4. Nella finestra di dialogo Cancel the step (Annulla fase), scegli se annullare la fase e attenderne l'uscita oppure annullare la fase e forzarne l'uscita. Quindi scegli Conferma.

5. Lo stato delle fasi nella tabella Steps (Fasi) diventa `CANCELLED`.

CLI

Per annullare utilizzando il AWS CLI

- Utilizzare il comando `aws emr cancel-steps`, specificando il cluster e le fasi da annullare. L'esempio seguente illustra un comando dell' AWS CLI per annullare due fasi.

```
aws emr cancel-steps --cluster-id j-2QUAXXXXXXXXX \
--step-ids s-3M8DXXXXXXXXX s-3M8DXXXXXXXXX \
--step-cancellation-option SEND_INTERRUPT
```

Con Amazon EMR versione 5.28.0, è possibile scegliere una delle due seguenti opzioni di annullamento per il parametro `StepCancellationOption` durante l'annullamento delle fasi.

- `SEND_INTERRUPT`: questa è l'opzione predefinita. Quando riceve una richiesta di annullamento fase, EMR invia un segnale `SIGTERM` alla fase. Aggiungi un gestore di segnale `SIGTERM` alla logica di fase per rilevare questo segnale e terminare i processi discendenti della fase o attendi che questi vengano completati.
- `TERMINATE_PROCESS`: quando questa opzione è selezionata, EMR invia un segnale `SIGKILL` alla fase e a tutti i relativi processi discendenti, i quali vengono terminati immediatamente.

Considerazioni per l'annullamento delle fasi

- L'annullamento di una fase in esecuzione o in sospeso rimuove tale fase dal numero di fasi attive.
- L'annullamento di una fase in esecuzione non consente l'avvio di una fase in sospeso, presupponendo l'assenza di modifiche a `stepConcurrencyLevel`.
- L'annullamento di una fase in esecuzione non attiva il `ActionOnFailure` della fase.
- Per EMR 5.32.0 e versioni successive, `SEND_INTERRUPT StepCancellationOption` invia un segnale `SIGTERM` al processo figlio della fase. Sarebbe opportuno rilevare questo segnale ed eseguire una pulizia e uno spegnimento corretti. `TERMINATE_PROCESS StepCancellationOption` invia un segnale `SIGKILL` al processo figlio della fase e a tutti i relativi processi discendenti; tuttavia, i processi asincroni non sono interessati.

Visualizza e monitora un cluster Amazon EMR mentre esegue il lavoro

Amazon EMR fornisce vari strumenti utili per raccogliere informazioni sul cluster. Puoi accedere a le informazioni relative al cluster dalla console, dall'interfaccia a riga di comando (CLI) o a livello di codice. Le interfacce Web Hadoop standard e i file di log sono disponibili sul nodo primario. Puoi anche utilizzare servizi di monitoraggio come CloudWatch Ganglia per monitorare le prestazioni del tuo cluster.

La cronologia delle applicazioni è disponibile anche dalla console utilizzando l'applicazione «persistente» UIs per Spark History Server a partire da Amazon EMR 5.25.0. Con Amazon EMR 6.x, sono disponibili anche il server della timeline YARN persistente e le interfacce utente Tez. Questi servizi sono ospitati fuori dal cluster, quindi è possibile accedere alla cronologia delle applicazioni per 30 giorni dopo la chiusura del cluster, senza la necessità di una connessione SSH o di un proxy Web. Consulta Visualizzazione della cronologia dell'applicazione

Argomenti

- Visualizza lo stato e i dettagli del cluster Amazon EMR
- Debug avanzato in fasi con Amazon EMR
- Visualizza la cronologia delle applicazioni Amazon EMR
- Visualizza i file di log di Amazon EMR
- Visualizza le istanze di cluster in Amazon EC2
- CloudWatch eventi e metriche di Amazon EMR
- Visualizza i parametri delle applicazioni cluster utilizzando Ganglia con Amazon EMR
- Registrazione delle chiamate AWS API EMR utilizzando AWS CloudTrail
- Best practice per l'osservabilità dell'EMR

Visualizza lo stato e i dettagli del cluster Amazon EMR

Dopo la creazione di un cluster, puoi monitorarne lo stato e ottenere informazioni dettagliate sull'esecuzione dello stesso e sugli errori che possono verificarsi, anche dopo la terminazione del cluster. Per scopi di riferimento, Amazon EMR conserva i metadati relativi ai cluster terminati per due mesi; dopo questo periodo, i metadati vengono eliminati. Non puoi eliminare cluster dalla cronologia dei cluster, ma mediante AWS Management Console puoi utilizzare il comando Filter (Filtro) e con

AWS CLI puoi utilizzare opzioni con il comando `list-clusters` per concentrare l'attenzione sui cluster che ti interessano.

Puoi accedere alla cronologia dell'applicazione archiviata nel cluster per una settimana dal momento in cui viene registrata, indipendentemente dal fatto che il cluster sia in esecuzione o terminato. Inoltre, le interfacce utente delle applicazioni persistenti archiviano la cronologia delle applicazioni fuori cluster per 30 giorni dopo la chiusura di un cluster. Consulta Visualizzazione della cronologia dell'applicazione

Per ulteriori informazioni sugli stati del cluster, ad esempio Waiting (In attesa) e Running (In esecuzione), consulta Comprensione del ciclo di vita del cluster.

Visualizzazione dei dettagli del cluster con la AWS Management Console

L'elenco dei cluster in https://console.aws.amazon.com/emr elenca tutti i cluster del tuo account e della tua AWS regione, inclusi i cluster terminati. L'elenco mostra quanto segue per ogni cluster: il nome e l'ID, i dettagli sullo stato e lo stato, l'ora di creazione, il tempo trascorso di esecuzione del cluster e le ore di istanza normalizzate accumulate per tutte le istanze del cluster. EC2 Questo elenco è il punto di partenza per monitorare lo stato dei cluster. È concepito per eseguire il drill-down dei dettagli di ogni cluster per scopi di analisi e di risoluzione dei problemi.

Console

Per visualizzare le informazioni sul cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri visualizzare.
3. Utilizza il pannello Summary (Riepilogo) per visualizzare informazioni di base sulla configurazione del cluster, ad esempio lo stato del cluster, le applicazioni open source che Amazon EMR ha installato nel cluster e la versione di Amazon EMR utilizzata per creare il cluster. Consulta ciascuna scheda in Summary (Riepilogo) per visualizzare le informazioni come descritto nella sezione seguente.

Visualizza i dettagli del cluster utilizzando il AWS CLI

I seguenti esempi illustrano come recuperare informazioni dettagliate sul cluster mediante l' AWS CLI. Per ulteriori informazioni sui comandi disponibili, consulta la Guida di riferimento ai comandi

della AWS CLI per Amazon EMR. Puoi utilizzare il comando describe-cluster per visualizzare dettagli a livello del cluster, tra cui stato, configurazione hardware e software, impostazioni VPC, operazioni di bootstrap, gruppi di istanze e così via. Per ulteriori informazioni sugli stati del cluster, consulta Comprensione del ciclo di vita del cluster. L'esempio seguente illustra l'utilizzo del comando `describe-cluster`, seguito da esempi del comando list-clusters.

Example Visualizzazione dello stato del cluster

Per utilizzare il comando `describe-cluster`, è necessario l'ID del cluster. Questo esempio dimostra come utilizzare per ottenere un elenco di cluster creati entro un determinato intervallo di date e quindi utilizzare uno dei cluster IDs restituiti per elencare ulteriori informazioni sullo stato di un singolo cluster.

Il comando seguente descrive il cluster*j-1K48XXXXXXHCB*, che sostituisci con il tuo ID del cluster.

```
aws emr describe-cluster --cluster-id j-1K48XXXXXXHCB
```

L'output del comando è simile a quanto segue:

```
{
    "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 Visualizzazione dei cluster per data di creazione

Per recuperare i cluster creati entro un determinato intervallo di tempo, utilizza il comando `list-clusters` con i parametri `--created-after` e `--created-before`.

Il comando seguente elenca tutti i cluster creati tra il 9 ottobre 2019 e il 12 ottobre 2019.

```
aws emr list-clusters --created-after 2019-10-09T00:12:00 --created-
before 2019-10-12T00:12:00
```

Example Visualizzazione dei cluster per stato

Per elencare i cluster per stato, utilizza il comando `list-clusters` con il parametro `--cluster-states`. Lo stato dei cluster validi può essere: STARTING, BOOTSTRAPPING, RUNNING, WAITING, TERMINATING, TERMINATED e TERMINATED_WITH_ERRORS.

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

Puoi anche utilizzare i parametri seguenti per elencare tutti i cluster negli stati specificati.

- `--active` filtra i cluster il cui stato è STARTING, BOOTSTRAPPING, RUNNING, WAITING o TERMINATING.
- `--terminated` filtra i cluster il cui stato è TERMINATED.
- Il parametro `--failed` filtra i cluster il cui stato è TERMINATED_WITH_ERRORS.

I seguenti comandi restituiscono lo stesso risultato.

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

```
aws emr list-clusters --terminated
```

Per ulteriori informazioni sugli stati del cluster, consulta Comprensione del ciclo di vita del cluster.

Debug avanzato in fasi con Amazon EMR

Se una fase di Amazon EMR ha esito negativo e hai inviato il lavoro utilizzando l'operazione API Step con un'AMI versione 5.x o successiva, in alcuni casi Amazon EMR può identificare e restituire la causa principale dell'errore nella fase insieme al nome del file di log pertinente e una parte della traccia di stack dell'applicazione mediante l'API. Ad esempio, è possibile identificare i seguenti errori:

- Un errore Hadoop comune come una directory di output già esistente, una directory di input inesistente o un'applicazione con memoria insufficiente.
- Errori Java, ad esempio un'applicazione compilata con una versione incompatibile di Java o eseguita con una classe principale introvabile.
- Un problema di accesso agli oggetti archiviati in Amazon S3.

Queste informazioni sono disponibili utilizzando le operazioni e API. DescribeStepListSteps Il FailureDetailscampo StepSummaryrestituito da tali operazioni. Per accedere alle FailureDetails informazioni, usa la AWS CLI, la console o AWS l'SDK.

Console

La nuova console Amazon EMR non offre il debug delle fasi. Tuttavia, è possibile visualizzare i dettagli della terminazione del cluster con la procedura seguente.

Per visualizzare i dettagli degli errori con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi seleziona il cluster che desideri visualizzare.
3. Prendi nota del valore Status (Stato) nella sezione Summary (Riepilogo) della pagina dei dettagli del cluster. Se lo stato è Terminated with errors (Terminato con errori), passa con il mouse sul testo per visualizzare i dettagli degli errori del cluster.

CLI

Per visualizzare i dettagli degli errori con AWS CLI

- Per ottenere i dettagli sugli errori relativi a un passaggio con AWS CLI, utilizzare il `describe-step` comando.

```
aws emr describe-step --cluster-id j-1K48XXXXXHCB --step-id s-3QM0XXXXXM1W
```

L'output risulterà simile al seguente:

```
{
  "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"
  }
}
```

Visualizza la cronologia delle applicazioni Amazon EMR

È possibile visualizzare i dettagli delle applicazioni Spark History Server e Timeline Server di YARN dallla pagina dei dettagli del cluster nella console. La cronologia dell'applicazione di Amazon EMR facilita la risoluzione dei problemi e l'analisi dei processi attivi e della cronologia dei processi.

> **Note**
>
> Per aumentare la sicurezza delle applicazioni off-console che potresti utilizzare con Amazon EMR, i domini di hosting delle applicazioni sono registrati nella Public Suffix List (PSL). Alcuni esempi di questi domini di hosting includono: `emrstudio-prod.us-east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Per maggiore sicurezza, se hai bisogno di impostare cookie sensibili nel nome di dominio predefinito, consigliamo di utilizzare i cookie con un prefisso `__Host-`. Questa pratica ti aiuterà a difendere il tuo dominio dai tentativi CSRF (cross-site request forgery). Per ulteriori informazioni, consulta la pagina Set-Cookie in Mozilla Developer Network.

La sezione Application user interfaces (Interfacce utente delle applicazioni) della scheda Applications (Applicazioni) offre diverse opzioni di visualizzazione, a seconda dello stato del cluster e delle applicazioni installate nel cluster.

- Accesso dall'esterno del cluster alle interfacce utente delle applicazioni persistenti: a partire dalla versione 5.25.0 di Amazon EMR, i collegamenti alle interfacce utente delle applicazioni persistenti sono disponibili per l'interfaccia utente di Spark e per Spark History Service. Con Amazon EMR versione 5.30.1 e successive, anche Tez UI e il timeline server di YARN hanno interfacce utente di applicazioni persistenti. Il server timeline YARN e l'interfaccia utente Tez sono applicazioni open source che forniscono parametri e strumenti visivi per i cluster attivi e terminati. L'interfaccia utente Spark fornisce dettagli sulle fasi e le attività del pianificatore, le dimensioni di RDD e l'utilizzo della memoria, le informazioni ambientali e le informazioni sugli esecutori attivi. UIs Le applicazioni persistenti vengono eseguite fuori dal cluster, quindi le informazioni e i log del cluster sono disponibili per 30 giorni dopo la chiusura di un'applicazione. A differenza delle interfacce utente delle applicazioni on-cluster, le applicazioni persistenti UIs non richiedono la configurazione di un proxy Web tramite una connessione SSH.
- Interfacce utente delle applicazioni nel cluster: esistono diverse interfacce utente della cronologia di applicazioni che possono essere eseguite in un cluster. Le interfacce utente nel cluster sono

ospitate sul nodo master e richiedono l'impostazione di una connessione SSH al server Web. Le interfacce utente delle applicazioni nel cluster mantengono la cronologia delle applicazioni per una settimana dopo la fine di un'applicazione. Per ulteriori informazioni e istruzioni su come configurare un tunnel SSH, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

A eccezione di Spark History Server, del Timeline Server di YARN e delle applicazioni Hive, la cronologia delle applicazioni nel cluster può essere visualizzata solo mentre il cluster è in esecuzione.

Visualizza le interfacce utente persistenti delle applicazioni in Amazon EMR

A partire da Amazon EMR versione 5.25.0, è possibile connettersi ai dettagli dell'applicazione Spark History Server persistente ospitata fuori cluster utilizzando la pagina Summary (Riepilogo) o la scheda Application user interfaces (Interfacce utente applicazioni) nella console. Le interfacce utente delle applicazioni persistenti Timeline Server di YARN e Tez sono disponibili a partire dalla versione Amazon EMR 5.30.1. L'accesso al collegamento con un clic alla cronologia delle applicazioni persistenti offre i seguenti vantaggi:

- È possibile analizzare e risolvere rapidamente i processi attivi e la cronologia dei processi senza configurare un proxy Web tramite una connessione SSH.
- È possibile accedere alla cronologia delle applicazioni e ai file di log pertinenti per cluster attivi e chiusi. I log sono disponibili per 30 giorni dopo la fine dell'applicazione.

Accedi ai dettagli del cluster nella console e seleziona la scheda Applicazioni. Seleziona l'interfaccia utente dell'applicazione che desideri dopo l'avvio del cluster. L'interfaccia utente dell'applicazione si apre in una nuova scheda del browser. Per ulteriori informazioni, consulta Monitoraggio e strumentazione.

È possibile visualizzare i log dei container YARN tramite i collegamenti sul server cronologia Spark, sul server della timeline YARN e sull'interfaccia utente Tez.

> **Note**
>
> Per accedere ai log del container YARN da Spark History Server e dall'interfaccia utente Tez è necessario abilitare la registrazione ad Amazon S3 per il cluster. Se la registrazione non è abilitata, i collegamenti ai log del container YARN non funzioneranno.

Raccolta di log

Per abilitare l'accesso con un solo clic alle interfacce utente delle applicazioni persistenti, Amazon EMR raccoglie due tipi di log:

- Ilog di eventi dell'applicazione vengono raccolti in un bucket di sistema EMR. I log di eventi vengono crittografati mentre sono inattivi utilizzando la crittografia lato server con chiavi gestite da Amazon S3 (SSE-S3). Se utilizzi una sottorete privata per il tuo cluster, assicurati di includere il bucket di sistema corretto ARNs nell'elenco delle risorse della politica di Amazon S3 per la sottorete privata. Per ulteriori informazioni, consulta Policy Amazon S3 minima per sottorete privata.
- I log del container YARN vengono raccolti in un bucket Amazon S3 di proprietà. È necessario abilitare la registrazione per il cluster per accedere ai log del container YARN. Per ulteriori informazioni, consulta Configurazione della registrazione e del debug di cluster.

Se è necessario disabilitare questa caratteristica per motivi di privacy, è possibile arrestare il daemon utilizzando uno script bootstrap quando si crea un cluster, come illustrato nell'esempio seguente.

```
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"]
```

Dopo aver eseguito questo script bootstrap, Amazon EMR non raccoglierà i log di eventi di Spark History Server o del Timeline Server di YARN nel bucket di sistema EMR. Nessuna informazione sulla cronologia delle applicazioni sarà disponibile nella scheda Interfacce utente applicazione e si perderà l'accesso a tutte le interfacce utente dell'applicazione dalla console.

File di log degli eventi Spark di grandi dimensioni

In alcuni casi, i job Spark a esecuzione prolungata, come lo streaming Spark, e i job di grandi dimensioni, come le query SQL Spark, possono generare registri di eventi di grandi dimensioni. Con i registri degli eventi di grandi dimensioni, puoi utilizzare rapidamente lo spazio su disco sulle istanze di calcolo e riscontrare errori durante il caricamento di Persistent. `OutOfMemory` UIs Per evitare

questi problemi, si consiglia di abilitare la caratteristica di rolling e compattazione del log eventi Spark. Questa caratteristica è disponibile solo nella versione Amazon EMR mr-6.1.0 e successive. Per ulteriori dettagli su rolling e compattazione, consulta Applicazione della compattazione ai file di log degli eventi nella documentazione di Spark.

Per attivare la funzione di rolling e compattazione del log degli eventi di Spark, attiva le seguenti impostazioni di configurazione di Spark.

- `spark.eventLog.rolling.enabled`— Attiva la rolling del log degli eventi in base alle dimensioni. Questa impostazione è disattivata per impostazione predefinita.
- `spark.eventLog.rolling.maxFileSize`— Quando la rolling è attivata, specifica la dimensione massima del file di log degli eventi prima che venga eseguito il rollover. Il valore predefinito è 128 MB.
- `spark.history.fs.eventLog.rolling.maxFilesToRetain`— Specifica il numero massimo di file di log eventi non compatti da mantenere. Per impostazione predefinita, tutti i file di log degli eventi vengono mantenuti. Imposta un numero inferiore per compattare i registri degli eventi più vecchi. Il valore più basso è 1.

Nota che la compattazione tenta di escludere eventi con file di log eventi obsoleti, come i seguenti. Se elimina gli eventi, non li vedrai più nell'interfaccia utente di Spark History Server.

- Eventi per i lavori terminati e relativi eventi relativi alla fase o all'attività.
- Eventi per esecutori licenziati.
- Eventi per le interrogazioni SQL completate e i relativi eventi relativi a job, stage e attività.

Per avviare un cluster con funzionalità di rolling e compattazione abilitate

1. Creare un file `spark-configuration.json` con la seguente configurazione.

```
[
   {
     "Classification": "spark-defaults",
        "Properties": {
           "spark.eventLog.rolling.enabled": true,
           "spark.history.fs.eventLog.rolling.maxFilesToRetain": 1
        }
   }
```

```
]
```

2. Crea un cluster con la configurazione Spark Rolling Compaction come segue.

```
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
```

Autorizzazioni per la visualizzazione delle interfacce utente persistenti delle applicazioni

L'esempio seguente mostra le autorizzazioni di ruolo necessarie per l'accesso alle interfacce utente persistenti delle applicazioni. Per i cluster con ruolo di runtime abilitato, ciò consentirà agli utenti di accedere solo alle applicazioni inviate con la stessa identità utente e lo stesso ruolo di runtime.

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"
    }
  ]
}
```

L'esempio seguente mostra le autorizzazioni di ruolo necessarie per rimuovere le restrizioni sulla visualizzazione delle applicazioni nelle interfacce utente persistenti delle applicazioni per i cluster abilitati al ruolo di runtime.

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-
XXXXXXXXXXXX",
        "arn:aws:elasticmapreduce:us-east-1:123456789012:persistent-app-ui/*"
      ],
      "Condition": {
        "StringEqualsIfExists": {
          "elasticmapreduce:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/YourExecutionRoleName"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEGetpersistentappuipresigneduRl"
    }
  ]
}
```

Considerazioni e limitazioni

L'accesso con un solo clic alle interfacce utente dell'applicazione persistente presenta attualmente le seguenti limitazioni.

- Ci sarà un ritardo di almeno due minuti quando i dettagli dell'applicazione vengono visualizzati nell'interfaccia utente di Spark History Server.
- Questa caratteristica è attiva solo quando la directory dei log di eventi per l'applicazione si trova in HDFS. Per impostazione predefinita, Amazon EMR archivia i log di eventi in una directory di HDFS. Se si modifica la directory predefinita in un file system diverso, ad esempio Amazon S3, questa funzionalità non sarà attiva.
- Al momento, questa caratteristica non è disponibile per i cluster EMR con più nodi principali o per i cluster EMR integrati con AWS Lake Formation.
- Per abilitare l'accesso con un clic alle interfacce utente persistenti delle applicazioni, devi disporre dell'autorizzazione per Amazon `CreatePersistentAppUI` EMR `DescribePersistentAppUI` e per `GetPersistentAppUIPresignedURL` le azioni. Se neghi l'autorizzazione a un responsabile IAM per queste azioni, occorrono circa cinque minuti prima che la modifica dell'autorizzazione si propaghi.
- Se un cluster è un cluster abilitato al ruolo di runtime, quando accede allo Spark History Server dall'interfaccia utente della Persistent App, l'utente potrà accedere a un job Spark solo se il job Spark viene inviato da un ruolo di runtime.

- Se un cluster è un cluster abilitato al ruolo di runtime, ogni utente può accedere solo a un'applicazione inviata con la stessa identità utente e lo stesso ruolo di runtime.
- L'`AccessAllEventLogs`azione per Amazon EMR è necessaria per visualizzare tutte le applicazioni nelle interfacce utente persistenti delle applicazioni per i cluster abilitati al ruolo di runtime.
- Se si riconfigurano le applicazioni in un cluster in esecuzione, la cronologia delle applicazioni non sarà disponibile tramite l'interfaccia utente dell'applicazione.
- Per ciascuna di esse Account AWS, il limite predefinito per l'applicazione UIs attiva è 200.
- Di seguito Regioni AWS, puoi accedere all'applicazione UIs dalla console con Amazon EMR 6.14.0 e versioni successive:
  - Asia Pacific (Giacarta) (ap-southeast-3)
  - Europa (Spagna) (eu-south-2)
  - Asia Pacifico (Melbourne) (ap-southeast-4)
  - Israele (Tel Aviv) (il-central-1)
  - Medio Oriente (EAU) (me-central-1)
- Di seguito Regioni AWS, puoi accedere all'applicazione UIs dalla console con Amazon EMR 5.25.0 e versioni successive:
  - Stati Uniti orientali (Virginia settentrionale) (us-east-1)
  - Stati Uniti occidentali (Oregon) (us-west-2)
  - Asia Pacifico (Mumbai) (ap-south-1)
  - Asia Pacifico (Seoul) (ap-northeast-2)
  - Asia Pacifico (Singapore) (ap-southeast-1)
  - Asia Pacifico (Sydney) (ap-southeast-2)
  - Asia Pacifico (Tokyo) (ap-northeast-1)
  - Canada (Centrale) (ca-central-1)
  - Sud America (San Paolo) (sa-east-1)
  - Europa (Francoforte) (eu-central-1)
  - Europa (Irlanda) (eu-west-1)
  - Europa (Londra) (eu-west-2)
  - Europe (Parigi) (eu-west-3)
  - Europa (Stoccolma) (eu-north-1)

- Cina (Pechino) cn-north-1
- Cina (Ningxia) cn-nordovest-1

Visualizza una cronologia delle applicazioni di alto livello in Amazon EMR

> **Note**
>
> Ti consigliamo di utilizzare l'interfaccia persistente dell'applicazione per un'esperienza utente migliore, in grado di conservare la cronologia delle applicazioni fino a un massimo di 30 giorni. La cronologia delle applicazioni di alto livello descritta in questa pagina non è disponibile nella nuova console Amazon EMR https://console.aws.amazon.com(/emr). Per ulteriori informazioni, consulta Visualizza le interfacce utente persistenti delle applicazioni in Amazon EMR.

Con Amazon EMR rilasci da 5.8.0 a 5.36.0 e rilasci della serie 6.x fino a 6.8.0, puoi visualizzare la cronologia di applicazioni di alto livello dalla scheda Application user interfaces (Interfacce utente delle applicazioni) nella vecchia console Amazon EMR. Una Application user interface (Interfaccia utente dell'applicazione) di Amazon EMR conserva il riepilogo della cronologia delle applicazioni per 7 giorni dopo il completamento di un'applicazione.

Considerazioni e limitazioni

Considera i seguenti limiti quando utilizzi la scheda Interfacce utente applicazioni nella precedente console di Amazon EMR.

- Puoi accedere alla funzionalità di cronologia delle applicazioni di alto livello solo se utilizzi Amazon EMR versioni da 5.8.0 a 5.36.0 e versioni della serie 6.x fino a 6.8.0. A partire dal 23 gennaio 2023, Amazon EMR interromperà la cronologia delle applicazioni di alto livello per tutte le versioni. Se utilizzi Amazon EMR versione 5.25.0 o successive, ti consigliamo di utilizzare invece l'interfaccia utente persistente dell'applicazione.
- La caratteristica della cronologia delle applicazioni di alto livello non supporta le applicazioni Spark Streaming.
- L'accesso con un clic alle interfacce utente delle applicazioni persistenti non è attualmente disponibile per i cluster Amazon EMR con più nodi principali o per i cluster Amazon EMR integrati con AWS Lake Formation.

Esempio: Visualizzazione di una cronologia delle applicazioni di alto livello

Nella sequenza seguente viene illustrato un drill-down nei dettagli del processo attraverso un'applicazione Spark o YARN utilizzando la scheda Application user interfaces (Interfacce utente delle applicazioni) nella pagina dei dettagli del cluster nella vecchia console.

Per visualizzare i dettagli del cluster, seleziona un Name (Nome) del cluster dall'elenco Clusters (Cluster). Per visualizzare informazioni sui log del container YARN, è necessario abilitare la registrazione per il cluster. Per ulteriori informazioni, consulta Configurazione della registrazione e del debug di cluster. Per la cronologia delle applicazioni Spark, le informazioni fornite nella tabella di riepilogo sono solo un sottoinsieme delle informazioni disponibili tramite l'interfaccia utente di Spark History Server.

Nella scheda Application user interfaces (Interfacce utente applicazioni), sotto High-level application history (Cronologia delle applicazioni di alto livello), è possibile espandere una riga per visualizzare il riepilogo diagnostico per un'applicazione Spark o selezionare un Application ID (ID applicazione) per visualizzare i dettagli di un'applicazione diversa.

Quando si seleziona un Application ID (ID applicazione), l'interfaccia utente cambia per mostrare i dettagli della YARN application (Applicazione YARN) relativi a tale applicazione. Nella scheda Jobs (Processi) dei dettagli della YARN application (Applicazione YARN), è possibile scegliere la Description (Descrizione) di un processo per visualizzarne i dettagli.

Nella pagina dei dettagli del processo, puoi espandere le informazioni sulle singole fasi del processo e selezionare il collegamento Description (Descrizione) per vedere i dettagli della fase.

Nella pagina dei dettagli della fase è possibile visualizzare i parametri principali delle attività e degli executor della fase. È inoltre possibile visualizzare i log delle attività e degli executor utilizzando i collegamenti View logs (Visualizza log).

Visualizza i file di log di Amazon EMR

Amazon EMR e Hadoop producono entrambi file di log che comunicano lo stato sul cluster. Per impostazione predefinita, questi file vengono scritti nel nodo primario nella directory `/mnt/var/log/`. A seconda di come il cluster è stato configurato quando è stato avviato, questi log possono anche essere archiviati in Amazon S3 ed essere visualizzati tramite lo strumento di debug grafico.

Esistono molti tipi di log scritti nel nodo primario. Amazon EMR scrive log di stato della fase, dell'operazione di bootstrap e dell'istanza. Apache Hadoop scrive i log per comunicare l'elaborazione di processi, attività e tentativi di attività. Hadoop registra inoltre i log dei suoi daemon. Per ulteriori informazioni sui log scritti da Hadoop, vai a http://hadoop.apache. org/docs/stable/hadoop-project-dist/hadoop-common/ClusterSetup.html.

Visualizzazione di file di log sul nodo primario

Nella tabella seguente vengono elencati alcuni dei file di log che si trovano sul nodo primario.

Ubicazione	Descrizione
/emr/instance-controller/log/bootstrap-azioni	Log scritti durante l'elaborazione delle operazioni di bootstrap.
/-state-pusher mnt/var/log/hadoop	Log scritti dal processo pusher dello stato di Hadoop.
/emr/instance-controller/log	Log del controller istanze.
/emr/instance-state	Log degli stati istanza. Contengono informazioni su CPU, stato della memoria e thread del garbage collector del nodo.
/emr/service-nanny	Log scritti dal processo nanny del servizio.
/mnt/var/log/*application*	Log specifici di un'applicazione, ad esempio Hadoop, Spark, o Hive.
/mnt/var/log/hadoop/steps/*N*	Log di fase contenenti informazioni sull'elaborazione della fase. Il valore di *N* indica lo StepiD assegnato da Amazon EMR. Ad esempio, un cluster ha due fasi: `s-1234ABCDEFGH` e `s-5678IJKLMNOP`. La prima fase si trova in `/mnt/var/log/hadoop/steps/s-1234ABCDEFGH/` e la seconda in `/mnt/var/log/hadoop/steps/s-5678IJKLMNOP/`.

Ubicazione	Descrizione
	I log di fase scritti da Amazon EMR sono i seguenti. • controller: informazioni sull'elaborazione della fase. Se la fase ha esito negativo durante il caricamento, puoi trovare la traccia di stack in questo log. • syslog: descrive l'esecuzione dei processi Hadoop nella fase. • stderr: il canale di errore standard di Hadoop durante l'elaborazione della fase. • stdout: il canale di uscita standard di Hadoop durante l'elaborazione della fase.

Visualizzazione dei file di log sul nodo primario con la AWS CLI.

1. Utilizza SSH per connetterti al nodo primario come descritto in Connect al nodo primario del cluster Amazon EMR tramite SSH.
2. Passare alla directory contenente le informazioni sul file di log che si desidera visualizzare. La tabella precedente fornisce un elenco dei tipi di file di log che sono disponibili e dove è possibile trovarli. L'esempio seguente mostra il comando per passare al log della fase con un ID, `s-1234ABCDEFGH`.

   ```
   cd /mnt/var/log/hadoop/steps/s-1234ABCDEFGH/
   ```

3. Utilizzare un visualizzatore file preferito per visualizzare il file di log. L'esempio seguente utilizza il comando `less` Linux per visualizzare i file di log `controller`.

   ```
   less controller
   ```

Visualizzazione dei file di log archiviati in Amazon S3

Per impostazione predefinita, i cluster Amazon EMR avviati utilizzando la console archiviano automaticamente i file di log in Amazon S3. Puoi specificare il tuo percorso dei log, oppure consentire

alla console di generare automaticamente un percorso dei log per te. Per cluster avviati utilizzando la CLI o l'API, occorre configurare l'archiviazione dei log Amazon S3 manualmente.

Quando Amazon EMR è configurato per archiviare i file di log su Amazon S3, archivia i file nella posizione S3 specificata, nella cartella*cluster-id*//, *cluster-id* dove si trova l'ID del cluster.

Nella tabella seguente vengono elencati alcuni dei file di log che si trovano su Amazon S3.

Ubicazione	Descrizione
//nodo/ *cluster-id*	Log dei nodi, inclusi log di operazioni di bootstrap, stato istanza e applicazioni per il nodo. I log di ogni nodo sono archiviati in una cartella etichettata con l'identificatore dell'istanza di quel nodo. EC2
//nodo//*cluster-id instance-id application*	I log creati da ogni applicazione o daemon associato a un'applicazione. Ad esempio, il log del server Hive si trova in *cluster-id* /node/*instance-id* /hive/hive-server.log .
//passi//*cluster-id step-id*	Log di fase contenenti informazioni sull'elaborazione della fase. Il valore di *step-id* indica l'ID della fase assegnato da Amazon EMR. Ad esempio, un cluster ha due fasi: s-1234ABCDEFGH e s-5678IJKLMNOP . La prima fase si trova in /mnt/var/log/hadoop/steps/s-1234ABCDEFGH/ e la seconda in /mnt/var/log/hadoop/steps/s-5678IJKLMNOP/ . I log di fase scritti da Amazon EMR sono i seguenti. • controller: informazioni sull'elaborazione della fase. Se la fase ha esito negativo durante il caricamento, puoi trovare la traccia di stack in questo log.

Ubicazione	Descrizione
	• syslog: descrive l'esecuzione dei processi Hadoop nella fase. • stderr: il canale di errore standard di Hadoop durante l'elaborazione della fase. • stdout: il canale di uscita standard di Hadoop durante l'elaborazione della fase.
//container *cluster-id*	Log del container applicazioni. I log per ogni applicazione YARN vengono salvati in queste posizioni.
//hadoop-mapreduce/ *cluster-id*	I registri che contengono informazioni sui dettagli di configurazione e sulla cronologia dei lavori. MapReduce

Visualizzazione dei file di log archiviati in Amazon S3 con la console Amazon S3

1. Accedi a AWS Management Console e apri la console Amazon S3 all'indirizzo. https://console.aws.amazon.com/s3/
2. Apri il bucket S3 specificato quando hai configurato il cluster per archiviare i file di log in Amazon S3.
3. Passare al file di log contenente le informazioni da visualizzare. La tabella precedente fornisce un elenco dei tipi di file di log che sono disponibili e dove è possibile trovarli.
4. Scarica l'oggetto file di log per visualizzarlo. Per istruzioni, consulta Download di un oggetto.

Visualizza le istanze di cluster in Amazon EC2

Per aiutarti a gestire le tue risorse, Amazon ti EC2 consente di assegnare metadati alle risorse sotto forma di tag. Ogni EC2 tag Amazon è composto da una chiave e un valore. I tag ti consentono di classificare le tue EC2 risorse Amazon in diversi modi: ad esempio, per scopo, proprietario o ambiente.

È possibile cercare e filtrare le risorse in base ai tag che vengono aggiunti. I tag che assegni alle risorse tramite il tuo AWS account sono disponibili solo per te. Altri account che condividono la risorsa non possono visualizzare i tag.

Amazon EMR etichetta automaticamente ogni EC2 istanza che avvia con coppie chiave-valore. Le chiavi identificano il cluster e il gruppo di istanze a cui appartiene l'istanza. Ciò semplifica il filtraggio EC2 delle istanze per mostrare, ad esempio, solo le istanze che appartengono a un particolare cluster o per mostrare tutte le istanze attualmente in esecuzione nel gruppo di istanze relativo all'attività. Ciò è particolarmente utile se esegui più cluster contemporaneamente o gestisci un numero elevato di istanze. EC2

Queste sono le coppie predefinite di valori chiave che Amazon EMR assegna:

Chiave	Valore	Definizione di valore
aws:elasticmapreduce:job-flow-id	*job-flow-identifier*	L'ID del cluster per cui viene eseguito il provisioning dell'istanza. Viene visualizzato nel formato `j-XXXXXXXXXXXXX` e può contenere fino a 256 caratteri.
aws:elasticmapreduce:instance-group-role	*group-role*	Il tipo di gruppo di istanze, immesso come uno dei seguenti valori: `master`, `core` o `task`.

È possibile visualizzare e filtrare i tag che Amazon EMR aggiunge. Per ulteriori informazioni, consulta Using tags in Amazon EC2 User Guide. Poiché i tag impostati da Amazon EMR sono tag di sistema e non possono essere modificati o cancellati, le sezioni sui tag di visualizzazione e filtraggio sono le più rilevanti.

> **Note**
>
> Amazon EMR aggiunge tag all' EC2 istanza quando il suo stato viene aggiornato a Running. Se si verifica una latenza tra il momento in cui viene effettuato il provisioning dell' EC2 istanza e il momento in cui il suo stato è impostato su In esecuzione, i tag impostati da Amazon EMR verranno visualizzati all'avvio dell'istanza. Se non vengono visualizzati i tag, attendere qualche minuto e aggiornare la visualizzazione.

CloudWatch eventi e metriche di Amazon EMR

Utilizza eventi e parametri per monitorare l'attività e l'integrità di un cluster Amazon EMR. Gli eventi sono utili per il monitoraggio di una specifica occorrenza in un cluster; ad esempio, quando un cluster passa dallo stato di avvio a quello di esecuzione. I parametri sono utili per monitorare uno specifico valore; ad esempio, la percentuale di spazio disponibile su disco che HDFS utilizza in un cluster.

Per ulteriori informazioni sugli CloudWatch eventi, consulta la Amazon CloudWatch Events User Guide. Per ulteriori informazioni sui CloudWatch parametri, consulta Using Amazon CloudWatch metrics e Creazione di CloudWatchallarmi Amazon nella Amazon CloudWatch User Guide.

Argomenti

- Monitoraggio dei parametri di Amazon EMR con CloudWatch
- Monitoraggio degli eventi di Amazon EMR con CloudWatch
- Risposta agli CloudWatch eventi di Amazon EMR

Monitoraggio dei parametri di Amazon EMR con CloudWatch

Le metriche vengono aggiornate ogni cinque minuti e raccolte e inviate automaticamente CloudWatch per ogni cluster Amazon EMR. Questo intervallo non è configurabile. Non sono previsti costi per i parametri di Amazon EMR riportati in. CloudWatch I parametri di datapoint di cinque minuti vengono tenuti in archivio per 63 giorni, dopodiché vengono eliminati.

Come si utilizzano i parametri di Amazon EMR?

La tabella seguente mostra gli usi comuni per i parametri segnalati da Amazon EMR. Questi suggerimenti sono solo introduttivi e non costituiscono un elenco completo. Per l'elenco completo dei parametri forniti da Amazon EMR, consulta Metriche riportate da Amazon EMR in CloudWatch.

Come...?	Parametri rilevanti
Monitorare l'avanzamento del cluster	Esaminare i parametri `RunningMapTasks` , `RemainingMapTasks` , `RunningReduceTasks` e `RemainingReduceTasks` .
Rilevare cluster inattivi	Il parametro `IsIdle` verifica se un cluster è attivo anche se non esegue attività. È possibile impostare un allarme di modo che venga

Come...?	Parametri rilevanti
	attivato quando il cluster è rimasto inattivo per un determinato periodo di tempo, ad esempio 30 minuti.
Rilevare quando la capacità di storage di un nodo è esaurita	Il parametro `MRUnhealthyNodes` tiene traccia quando uno o più nodi principali o attività esauriscono la capacità di archiviaz ione su disco locale e la transizione a uno stato YARN `UNHEALTHY` . Ad esempio, i nodi principali o attività stanno esaurendo spazio su disco e non saranno in grado di eseguire attività.
Rileva quando un cluster esaurisce la capacità d'archiviazione	Il parametro `HDFSUtilization` monitora la capacità HDFS combinata del cluster e può richiederne il ridimensionamento del cluster per aggiungere più nodi principali. Ad esempio, l'utilizzo di HDFS è elevato, il che può influire sull'integrità dei processi e sul cluster.
Rileva quando un cluster è in esecuzione a capacità ridotta	Il parametro `MRLostNodes` tiene traccia quando uno o più nodi principali o attività non sono in grado di comunicare con il nodo master. Ad esempio, il nodo principale o attività non è raggiungibile dal nodo master.

Per ulteriori informazioni, consultare Il cluster Amazon EMR termina con NO_SLAVE_LEFT e i nodi principali FAILED_BY_MASTER e AWSSupport-AnalyzeEMRLogs.

Accedi ai CloudWatch parametri per Amazon EMR

Puoi visualizzare i parametri che Amazon EMR riporta utilizzando la console o CloudWatch la console di Amazon EMR. CloudWatch Puoi anche recuperare le metriche utilizzando il comando CloudWatch `mon-get-stats` CLI o l'API. CloudWatch `GetMetricStatistics` Per ulteriori informazioni sulla visualizzazione o il recupero dei parametri per l'utilizzo di Amazon EMR, CloudWatch consulta la Amazon User Guide. CloudWatch

Console

Per visualizzare i parametri con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli il cluster per il quale desideri visualizzare le metriche. Si apre la pagina dei dettagli del cluster.
3. Seleziona la scheda Monitoring (Monitoraggio) nella pagina dei dettagli del cluster. Scegli una qualsiasi delle schede Cluster status (Stato del cluster), Node status (Stato del nodo) o Inputs and outputs (Input e output) per caricare i report sull'avanzamento e sull'integrità del cluster.
4. Dopo avere scelto un parametro da visualizzare, puoi ingrandire ciascun grafico. Per filtrare l'intervallo di tempo del grafico, seleziona un'opzione precompilata o scegli Custom (Personalizzato).

Metriche riportate da Amazon EMR in CloudWatch

Le tabelle seguenti elencano i parametri che Amazon EMR riporta nella console e a cui invia i dati. CloudWatch

Parametri di Amazon EMR

Amazon EMR invia dati per diverse metriche a. CloudWatch Tutti i cluster Amazon EMR inviano automaticamente i parametri a intervalli di cinque minuti. I parametri sono conservati per due settimane; dopo tale periodo, i dati vengono eliminati.

Lo spazio dei nomi `AWS/ElasticMapReduce` include i parametri descritti di seguito.

> **Note**
>
> Amazon EMR estrae i parametri da un cluster. Se un cluster diventa inaccessibile, non viene indicato alcun parametro fino a che il cluster non è di nuovo disponibile.

I seguenti parametri sono disponibili per i cluster sui quali sono in esecuzione le versioni 2.x di Hadoop.

Parametro	Descrizione
Stato del cluster	
IsIdle	Indica che un cluster non è più in esecuzione ma è ancora attivo e genera spese. È impostato su 1 se non vi sono task e processi in esecuzione, altrimenti è impostato su 0. Questo valore viene verificato a intervalli di cinque minuti e un valore 1 indica unicamente l'inattività del cluster al momento della verifica e non durante i cinque minuti. Per evitare falsi positivi, devi attivare un allarme quando questo valore è 1 durante due o più verifiche consecutive di cinque minuti. Ad esempio, puoi attivare un allarme se questo valore è 1 per trenta minuti o più. Caso d'uso: monitorare le prestazioni del cluster Unità: booleane
ContainerAllocated	Il numero di contenitori di risorse allocati da. ResourceM anager Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
ContainerReserved	Il numero di container riservati. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
ContainerPending	Il numero di container nella coda non ancora allocati. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
ContainerPendingRatio	Il rapporto tra contenitori in sospeso e contenitori allocati (ContainerPendingRatio = ContainerPending /). Container

Parametro	Descrizione
	Allocated Se ContainerAllocated = 0, allora Container PendingRatio =. ContainerPending Il valore di Container PendingRatio rappresenta un numero, non una percentua le. Questo valore è utile per il dimensionamento delle risorse del cluster in funzione del comportamento di attribuzione dei container. Unità: numero
AppsCompleted	Il numero di applicazioni inviate a YARN che sono state completate. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
AppsFailed	Il numero di applicazioni inviate a YARN il cui completam ento non è riuscito. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero
AppsKilled	Il numero di applicazioni inviate a YARN che sono state interrotte. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero
AppsPending	Il numero di applicazioni inviate a YARN che sono in attesa. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
AppsRunning	Il numero di applicazioni inviate a YARN che sono in esecuzione. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
AppsSubmitted	Il numero di applicazioni inviate a YARN. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
Stato del nodo	
CoreNodesRunning	Il numero di nodi principali attivi. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero
CoreNodesPending	Il numero di nodi principali in attesa di assegnazione. È possibile che non tutti i nodi principali richiesti siano immediatamente disponibili; questo parametro indica le richieste in attesa. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero
LiveDataNodes	La percentuale di nodi dati che ricevono attività da Hadoop. Caso d'uso: monitorare lo stato del cluster Unità: percentuale

Parametro	Descrizione
MRTotalNodi	Il numero di nodi attualmente disponibili per i MapReduce lavori. Equivalente al parametro YARN `mapred.resourcemanager.TotalNodes` . Caso d'uso: monitorare l'avanzamento del cluster Unità: numero Nota: MRTotal i nodi contano solo i nodi attualmente attivi nel sistema. YARN rimuove automaticamente i nodi terminati da questo conteggio e ne interrompe il tracciamento, quindi non vengono considerati nella metrica MRTotal Nodes.
MRActiveNodi	Il numero di nodi che attualmente eseguono MapReduce attività o lavori. Equivalente al parametro YARN `mapred.resourcemanager.NoOfActiveNodes` . Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
MRLostNodi	Il numero di nodi assegnati a MapReduce tali nodi sono stati contrassegnati in uno stato LOST. Equivalente al parametro YARN `mapred.resourcemanager.NoOfLostNodes` . Caso d'uso: monitorare lo stato del cluster, monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
MRUnhealthyNodi	Il numero di nodi disponibili per i MapReduce lavori contrassegnati in uno stato NON SANO. Equivalente al parametro YARN `mapred.resourcemanager.NoOfUnhealthyNodes` . Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
MRDecommissionedNodi	Il numero di nodi assegnati alle MapReduce applicazioni che sono state contrassegnate come DISATTIVATO. Equivalente al parametro YARN `mapred.resourcemanager.NoOfDecommissionedNodes` . Caso d'uso: monitorare lo stato del cluster, monitorare l'avanzamento del cluster Unità: numero
MRRebootedNodi	Il numero di nodi disponibili MapReduce che sono stati riavviati e contrassegnati come RIAVVIATO. Equivalente al parametro YARN `mapred.resourcemanager.NoOfRebootedNodes` . Caso d'uso: monitorare lo stato del cluster, monitorare l'avanzamento del cluster Unità: numero
MultiMasterInstanceGroupNodesRunning	Il numero di nodi master in esecuzione. Caso d'uso: monitorare l'errore e la sostituzione del nodo master Unità: numero

Parametro	Descrizione
MultiMasterInstanceGroupNodesRunningPercentage	La percentuale di nodi master in esecuzione sul numero dell'istanza del nodo master richiesto. Caso d'uso: monitorare l'errore e la sostituzione del nodo master Unità: percentuale
MultiMasterInstanceGroupNodesRequested	Il numero di nodi master richiesti. Caso d'uso: monitorare l'errore e la sostituzione del nodo master Unità: numero
IO	
S3 BytesWritten	Il numero di byte scritti su Amazon S3. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
S3 BytesRead	Il numero di byte letti da Amazon S3. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
HDFSUtilization	La percentuale di storage HDFS attualmente utilizzato. Caso d'uso: analizzare le prestazioni del cluster Unità: percentuale
HDFSBytesLeggi	Il numero di byte letti da HDFS. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
HDFSBytesScritto	Il numero di byte scritti su HDFS. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
MissingBlocks	Il numero di blocchi in cui HDFS non ha repliche. Possono essere blocchi danneggiati. Caso d'uso: monitorare lo stato del cluster Unità: numero
CorruptBlocks	Il numero di blocchi che HDFS ha indicato come danneggia ti. Caso d'uso: monitorare lo stato del cluster Unità: numero

Parametro	Descrizione
TotalLoad	Il numero totale di trasferimenti di dati simultanei. Caso d'uso: monitorare lo stato del cluster Unità: numero
MemoryTotalMB	La quantità totale di memoria nel cluster. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
MemoryReservedMB	La quantità di memoria riservata. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
MemoryAvailableMB	La quantità di memoria disponibile da allocare. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
YARNMemoryAvailablePercentage	La percentuale di memoria rimanente disponibile per YARN (YARNMemoryAvailablePercentage = MemoryAvailable MB/ MemoryTotal MB). Questo valore è utile per il dimension amento delle risorse del cluster in funzione dell'utilizzo della memoria di YARN. Unità: percentuale
MemoryAllocatedMB	La quantità di memoria allocata al cluster. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
PendingDeletionBlocks	Il numero di blocchi contrassegnati per l'eliminazione. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero
UnderReplicatedBlocks	Il numero di blocchi che devono essere replicati una o più volte. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero
DfsPendingReplicationBlocks	Lo stato della replica dei blocchi: blocchi in corso di replica, età delle richieste di replica e richieste di replica non riuscite. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero
CapacityRemainingGB	La quantità di capacità rimanente del disco HDFS. Caso d'uso: monitorare l'avanzamento del cluster, monitorar e lo stato del cluster Unità: numero

Di seguito sono descritti i parametri Hadoop 1:

Parametro	Descrizione
Stato del cluster	
IsIdle	Indica che un cluster non è più in esecuzione ma è ancora attivo e genera spese. È impostato su 1 se non vi sono

Parametro	Descrizione
	task e processi in esecuzione, altrimenti è impostato su 0. Questo valore viene verificato a intervalli di cinque minuti e un valore 1 indica unicamente l'inattività del cluster al momento della verifica e non durante i cinque minuti. Per evitare falsi positivi, devi attivare un allarme quando questo valore è 1 durante due o più verifiche consecutive di cinque minuti. Ad esempio, puoi attivare un allarme se questo valore è 1 per trenta minuti o più. Caso d'uso: monitorare le prestazioni del cluster Unità: booleane
JobsRunning	Il numero di processi nel cluster attualmente in esecuzione. Caso d'uso: monitorare lo stato del cluster Unità: numero
JobsFailed	Il numero di processi nel cluster non riusciti. Caso d'uso: monitorare lo stato del cluster Unità: numero
Mappatura/Riduzione	
MapTasksRunning	Il numero di task di mappatura in esecuzione per ogni processo. Se hai un pianificatore installato e molteplici processi in esecuzione, vengono generati più grafici. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
MapTasksRemaining	Il numero di task di mappatura rimanenti per ogni processo. Se hai un pianificatore installato e molteplici processi in esecuzione, vengono generati più grafici. Un task di mappatura rimanente è un task il cui stato non è Running, Killed o Completed. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
MapSlotsOpen	La capacità dei task di mappatura non utilizzata. Viene calcolata come numero massimo di task di mappatura per un determinato cluster, meno il numero totale di task di mappatura attualmente in esecuzione in quel cluster. Caso d'uso: analizzare le prestazioni del cluster Unità: numero
RemainingMapTasksPerSlot	La proporzione tra i task di mappatura totali rimanenti e gli slot di mappatura totali disponibili nel cluster. Caso d'uso: analizzare le prestazioni del cluster Unità: proporzione
ReduceTasksRunning	Il numero di task di riduzione in esecuzione per ogni processo. Se hai un pianificatore installato e molteplici processi in esecuzione, vengono generati più grafici. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
ReduceTasksRemaining	Il numero di task di riduzione rimanenti per ogni processo. Se hai un pianificatore installato e molteplici processi in esecuzione, vengono generati più grafici. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
ReduceSlotsOpen	La capacità dei task di riduzione non utilizzata. Viene calcolata come capacità massima dei task di riduzione per un determinato cluster, meno il numero di task di riduzione attualmente in esecuzione in quel cluster. Caso d'uso: analizzare le prestazioni del cluster Unità: numero
Stato del nodo	
CoreNodesRunning	Il numero di nodi principali attivi. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero
CoreNodesPending	Il numero di nodi principali in attesa di assegnazione. È possibile che non tutti i nodi principali richiesti siano immediatamente disponibili; questo parametro indica le richieste in attesa. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero

Parametro	Descrizione
LiveDataNodes	La percentuale di nodi dati che ricevono attività da Hadoop. Caso d'uso: monitorare lo stato del cluster Unità: percentuale
TaskNodesRunning	Il numero di nodi di task attivi. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero
TaskNodesPending	Il numero di nodi di task in attesa di assegnazione. È possibile che non tutti i nodi di task richiesti siano immediata mente disponibili; questo parametro indica le richieste in attesa. I punti dati per questo parametro sono indicati solo se esiste un gruppo di istanze corrispondente. Caso d'uso: monitorare lo stato del cluster Unità: numero
LiveTaskTrackers	La percentuale di tracker di task operativi. Caso d'uso: monitorare lo stato del cluster Unità: percentuale
IO	

Parametro	Descrizione
S3 BytesWritten	Il numero di byte scritti su Amazon S3. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
S3 BytesRead	Il numero di byte letti da Amazon S3. Questa metrica aggrega solo i MapReduce lavori e non si applica ad altri carichi di lavoro su Amazon EMR. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
HDFSUtilization	La percentuale di storage HDFS attualmente utilizzato. Caso d'uso: analizzare le prestazioni del cluster Unità: percentuale
HDFSBytesLeggi	Il numero di byte letti da HDFS. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero
HDFSBytesScritto	Il numero di byte scritti su HDFS. Caso d'uso: analizzare le prestazioni del cluster, monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
MissingBlocks	Il numero di blocchi in cui HDFS non ha repliche. Possono essere blocchi danneggiati. Caso d'uso: monitorare lo stato del cluster Unità: numero
TotalLoad	Il numero totale attuale di lettori e scrittori riportato da tutti DataNodes in un cluster. Caso d'uso: diagnostica il grado in cui l'alto I/O potrebbe contribuire a una cattiva esecuzione del lavoro. I nodi di lavoro che eseguono il DataNode demone devono inoltre eseguire operazioni di mappatura e ridurre le attività. TotalLoad Valori costantemente elevati nel tempo possono indicare che un valore elevato I/O potrebbe contribuire a ridurre le prestazioni. Picchi occasionali di questo valore sono comuni e non indicano un problema. Unità: numero

Parametri della capacità del cluster

i parametri seguenti indicano le capacità correnti o di destinazione di un cluster. Queste metriche sono disponibili solo quando sono abilitati il dimensionamento gestito o la terminazione automatica.

Per i cluster composti da parchi istanze, i parametri della capacità del cluster vengono misurate in `Units`. Per i cluster composti da gruppi di istanze, i parametri della capacità del cluster vengono misurate in `Nodes` o in `VCPU` in base al tipo di unità utilizzato nella policy di dimensionamento gestito. Per ulteriori informazioni, consulta Utilizzo del dimensionamento gestito da EMR nella Guida alla gestione di Amazon EMR.

Parametro	Descrizione
• `TotalUnitsRequested` •	Il numero totale previsto di unità units/nodes/vCPUs in un cluster, determinato dalla scalabilità gestita.

Parametro	Descrizione
• `TotalNodesRequested` • `TotalVCPURequested`	Unità: numero
• `TotalUnitsRunning` • `TotalNodesRunning` • `TotalVCPURunning`	Il numero totale corrente di units/nodes/vCPUs disponibilità in un cluster in esecuzione. Quando viene richiesto il ridimensionamento di un cluster, questo parametro verrà aggiornato dopo l'aggiunta o la rimozione delle nuove istanze dal cluster. Unità: numero
• `CoreUnitsRequested` • `CoreNodesRequested` • `CoreVCPURequested`	Il numero target di CORE units/nodes/vCPUs in un cluster determinato dalla scalabilità gestita. Unità: numero
• `CoreUnitsRunning` • `CoreNodesRunning` • `CoreVCPURunning`	Il numero attuale di CORE in units/nodes/vCPUs esecuzione in un cluster. Unità: numero
• `TaskUnitsRequested` • `TaskNodesRequested` • `TaskVCPURequested`	Il numero target di TASK units/nodes/vCPUs in un cluster determinato dalla scalabilità gestita. Unità: numero

Parametro	Descrizione
• `TaskUnitsRunning` • `TaskNodesRunning` • `TaskVCPURunning`	Il numero corrente di TASK in units/nodes/vCPUs esecuzione in un cluster. Unità: numero

Amazon EMR emette le seguenti metriche con una granularità di un minuto quando abiliti la terminazione automatica utilizzando una policy di terminazione automatica. Alcune metriche sono disponibili solo per Amazon EMR versione 6.4.0 e successive. Per ulteriori informazioni sulla terminazione automatica, consulta Utilizzo di una politica di terminazione automatica per la pulizia dei cluster Amazon EMR.

Parametro	Descrizione
`TotalNotebookKernels`	Il numero totale di kernel notebook in esecuzione e inattivi sul cluster. Questa metrica è disponibile solo in Amazon EMR versione 6.4.0 e successive.
`AutoTerminationIsClusterIdle`	Indica se il cluster è in uso. Un valore di 0 indica che il cluster è in uso attivo da uno dei seguenti componenti: • Un'applicazione YARN • HDFS • Un notebook • Un'interfaccia utente on-cluster, come Spark History Server

Parametro	Descrizione
	Un valore di 1 indica che il cluster è inattivo. Amazon EMR verifica l'inattività continua del cluster (`AutoTerminationIsClusterIdle =1`). Quando il tempo di inattività di un cluster è uguale al valore `IdleTimeout` nella tua policy di terminazione automatica, Amazon EMR termina il cluster.

Dimensioni per i parametri Amazon EMR

I dati di Amazon EMR possono essere filtrati utilizzando una qualsiasi delle dimensioni nella tabella esposta di seguito.

Dimensione	Descrizione
JobFlowId	Uguale all'ID del cluster che è l'identificatore univoco di un cluster nel formato `j-XXXXXXXXXXXX` . Puoi trovare questo valore facendo clic sul cluster nella console di Amazon EMR.

Monitoraggio degli eventi di Amazon EMR con CloudWatch

Amazon EMR monitora gli eventi e conserva le informazioni sugli stessi fino a sette giorni nella console di Amazon AMR. Amazon EMR registra gli eventi in caso di modifica dello stato di cluster, gruppi di istanze, parchi istanze, policy di scalabilità automatica o fasi. Gli eventi registrano la data e l'ora in cui si sono verificati, i dettagli sugli elementi interessati e altri dati critici.

La tabella seguente elenca gli eventi di Amazon EMR, insieme allo stato o alla modifica dello stato che l'evento indica, la gravità dell'evento, il tipo di evento, il codice dell'evento e messaggi relativi agli eventi. Amazon EMR rappresenta gli eventi come oggetti JSON e li invia automaticamente a un flusso di eventi. L'oggetto JSON è importante quando configuri le regole per l'elaborazione CloudWatch degli eventi utilizzando Events perché le regole cercano di corrispondere ai modelli dell'oggetto JSON. Per ulteriori informazioni, consulta Events and event patterns e Amazon EMR events nella Amazon CloudWatch Events User Guide.

> **Note**
>
> EMR emette periodicamente eventi con il codice evento EC2 provisioning - Insufficient Instance Capacity. Questi eventi si verificano quando il tuo cluster Amazon EMR rileva un errore di capacità insufficiente da Amazon EMR per la tua flotta o il tuo gruppo di istanze durante la creazione o l'operazione di ridimensionamento del cluster. Un evento potrebbe non includere tutti i tipi di istanze AZs che hai fornito, poiché EMR include solo i tipi di istanza e ha tentato di fornire capacità dall'ultima volta che è stato AZs emesso l'evento Capacità insufficiente. Per informazioni su come rispondere a questi eventi, consulta Risposta agli eventi di capacità insufficiente delle istanze del cluster Amazon EMR.

Eventi di avvio del cluster

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Provisioning della flotta di istanze EMR	EC2 approvvigionamento: capacità insufficiente delle istanze	Non siamo in grado di creare il tuo cluster Amazon EMR per la flotta di istanze `InstanceFleetID` Amazon EC2 ha una capacità Spot insufficiente `ClusterId (ClusterName)` per il tipo di istanza e una capacità `[Instancetype1, Instancetype2]` On-

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				Demand insufficiente per il tipo di istanza `[Instancetype3, Instancetype4]` nella zona di disponibilità. `[AvailabilityZone1, AvaliabilityZone2]` Consulta qui la documentazione per ulteriori informazioni su come rispondere a questo evento.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
CREATING	WARN	Fornitura di gruppi di istanze EMR	EC2 approvvigionamento: capacità insufficiente dell'istanza	Non siamo in grado di creare il tuo cluster Amazon EMR per il gruppo di istanze InstanceGroupID Amazon EC2 ha una capacità Spot insufficiente ClusterId (ClusterName) per il tipo di istanza e una capacità [Instancetype1, Instancetype2] On-Demand insufficiente per il tipo di istanza [Instancetype3, Instancetype4] nella zona di disponibilità. [AvailabilityZone1, AvaliabilityZone

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				2] Consulta qui la documentazione per ulteriori informazioni su come rispondere a questo evento.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
CREATING	WARN	Provisioning della flotta di istanze EMR	EC2 provisioning - Indirizzi gratuiti insufficienti nella sottorete	Non possiamo creare il cluster Amazon EMR ClusterId (ClusterName) che hai richiesto, ad esempio la flotta, InstanceFleetID perché la sottorete specificata [Subnet1, Subnet2] non contiene abbastanza indirizzi IP privati gratuiti per soddisfare la tua richiesta. Usa l'DescribeSubnets operazione per vedere quanti indirizzi IP sono disponibili (non utilizzati) nella tua sottorete. Per informazioni su come rispondere a questo evento, consulta Codici di errore

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				per l' EC2 API Amazon

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
CREATING	WARN	Fornitura di gruppi di istanze EMR	EC2 provisioning - Indirizzi gratuiti insufficienti nella sottorete	Non possiamo creare il cluster Amazon EMR ClusterId (ClusterName) che hai richiesto per esempio il gruppo InstanceGroupID perché la sottorete specificata [Subnet1, Subnet2] non contiene abbastanza indirizzi IP privati gratuiti per soddisfare la tua richiesta. Usa l'DescribeSubnets operazione per vedere quanti indirizzi IP sono disponibili (non utilizzati) nella tua sottorete. Per informazioni su come rispondere a questo evento, consulta

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				Codici di errore per l' EC2 API Amazon
CREATING	WARN	Provisioning della flotta di istanze EMR	EC2 Provisioning: limite vCPU superato	La fornitura di `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di v CPUs (unità di elaborazione virtuali) assegnate alle istanze in esecuzione nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Fornitura di gruppi di istanze EMR	EC2 Provisioning: limite vCPU superato	La fornitura del gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId` viene ritardata perché hai raggiunto il limite del numero di v CPUs (unità di elaborazione virtuali) assegnate alle istanze in esecuzione nel tuo account. `(accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Provisioning della flotta di istanze EMR	EC2 Provisioning: il limite di numero di istanze Spot è stato superato	La fornitura della flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze Spot che puoi avviare nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Fornitura di gruppi di istanze EMR	EC2 Provisioning: il limite di numero di istanze Spot è stato superato	La fornitura del gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze Spot che puoi avviare nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Provisioning della flotta di istanze EMR	EC2 Provisioning: limite di istanze superato	La fornitura della flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze che puoi eseguire contemporaneamente nel tuo. `account (accountID)` Per ulteriori informazioni sui limiti dei EC2 servizi Amazon, consulta Codici di errore per l' EC2 API Amazon.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Fornitura di gruppi di istanze EMR	EC2 Provisioning: limite di istanze superato	La fornitura del gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze che puoi eseguire contemporaneamente nel tuo. `account (accountID)` Per ulteriori informazioni sui limiti dei EC2 servizi Amazon, consulta Codici di errore per l' EC2 API Amazon.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`CREATING`	`WARN`	Fornitura di gruppi di istanze EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` è stato creato alle `Time` ed è pronto per l'uso. oppure Il cluster Amazon EMR `ClusterId (ClusterName)` ha completato l'esecuzione di tutte le fasi in sospeso alle `Time`. **Note** Un cluster il cui stato è `WAITING` può tuttavia elaborare

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				dei processi.
STARTING	INFO	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR ClusterId (ClusterName) è stato richiesto alle Time ed è in fase di creazione.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`STARTING`	`INFO`	Modifica dello stato del cluster EMR	nessuno	**Note** Si applica solo ai cluster con la configurazione delle flotte di istanze e più zone di disponibilità selezionate all'interno di Amazon. EC2 Il cluster Amazon EMR `ClusterId (ClusterName)` è in fase di creazione nella zona (`AvailabilityZoneID`

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				), scelta tra le opzioni di zona di disponibilità specificate.
`STARTING`	`INFO`	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` ha iniziato a eseguire le fasi alle `Time`.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`WAITING`	`INFO`	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` è stato creato alle `Time` ed è pronto per l'uso. oppure Il cluster Amazon EMR `ClusterId (ClusterName)` ha completato l'esecuzione di tutte le fasi in sospeso alle `Time`. **Note** Un cluster il cui stato è `WAITING` può tuttavia elaborare

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				dei processi.

> **Note**
>
> Gli eventi con codice evento `EC2 provisioning - Insufficient Instance Capacity` vengono emessi periodicamente quando il tuo cluster EMR rileva un errore di capacità insufficiente da parte di EC2 Amazon per la tua flotta o il tuo gruppo di istanze durante la creazione o l'operazione di ridimensionamento del cluster. Per ulteriori informazioni su come rispondere a tali eventi, consulta Risposta a eventi di capacità insufficiente dell'istanza del cluster Amazon EMR.

Eventi di chiusura di un cluster

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`TERMINATED`	La gravità dipende dal motivo della modifica dello stato, come descritto di seguito: • **`CRITICAL`** se il cluster è stato terminato con uno dei seguenti motivi di modifica	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` è stato terminato alle `Time` con un motivo di `StateChangeReason: Code` .

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
	dello stato: `INTERNAL_ERROR`, `VALIDATION_ERROR`, `INSTANCE_FAILURE`, `BOOTSTRAP_FAILURE` o `STEP_FAILURE`. • **`INFO`** se il cluster è stato terminato con uno dei seguenti motivi di modifica dello stato: `USER_REQUEST` o `ALL_STEPS_COMPLETED`.			

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`TERMINATED_WITH_ERRORS`	`CRITICAL`	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` è stato terminato con errori alle `Time` con un motivo di `StateChangeReason: Code` .
`TERMINATED_WITH_ERRORS`	`CRITICAL`	Modifica dello stato del cluster EMR	nessuno	Il cluster Amazon EMR `ClusterId (ClusterName)` è stato terminato con errori alle `Time` con un motivo di `StateChangeReason: Code` .

Eventi di modifica dello stato del parco istanze

> **Note**
>
> La configurazione dei parchi istanze è disponibile solo in Amazon EMR rilasci 4.8.0 e successivi, esclusi i rilasci 5.0.0 e 5.0.3.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
Da `PROVISIONING` a `WAITING`	`INFO`		nessuno	Il provisioning per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId` (`ClusterName`) è completo. Il provisioning è stato avviato alle `Time` ed è durato `Num` minuti. Il parco istanze ha ora una capacità On-Demand di `Num` e una capacità Spot di `Num`. La capacità On-Demand di destinazione era `Num` e la capacità Spot di destinazione era `Num`.
Da `WAITING` a `RESIZING`	`INFO`		nessuno	Un ridimensionamento per il parco istanze `InstanceFleetID`

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				nel cluster Amazon EMR `ClusterId (ClusterName)` è iniziato alle `Time`. Il parco istanze è in fase di ridimensionamento da una capacità On-Demand di `Num` a una destinazione di `Num` e da una capacità Spot di `Num` a una destinazione di `Num`.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
Da `RESIZING` a `WAITING`	`INFO`		nessuno	L'operazione di ridimensionamento per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` è stata completata. Il ridimensionamento è stato avviato alle `Time` ed è durato `Num` minuti. Il parco istanze ha ora una capacità On-Demand di `Num` e una capacità Spot di `Num`. La capacità On-Demand di destinazione era `Num` e la capacità Spot di destinazione era `Num`.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
Da RESIZING a WAITING	INFO		nessuno	L'operazione di ridimensionamento per il parco istanze InstanceFleetID nel cluster Amazon EMR ClusterId (ClusterName) ha raggiunto il valore di timeout ed è stata interrotta. Il ridimensionamento è stato avviato alle Time ed è stato interrotto dopo Num minuti. Il parco istanze ha ora una capacità On-Demand di Num e una capacità Spot di Num. La capacità On-Demand di destinazione era Num e la capacità Spot di destinazione era Num.

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
SUSPENDED	ERROR		nessuno	Il parco istanze InstanceFleetID nel cluster Amazon EMR ClusterId (ClusterName) è stato arrestato alle Time per il seguente motivo: ReasonDesc .
RESIZING	WARNING		nessuno	L'operazione di ridimensionamento per il parco istanze InstanceFleetID nel cluster Amazon EMR ClusterId (ClusterName) è bloccata per il seguente motivo: ReasonDesc .

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
`WAITING` o `Running`	`INFO`		nessuno	Non è stato possibile completare l'operazione di dimensionamento per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` poiché Amazon EMR ha aggiunto la capacità Spot nella zona di disponibilità `AvailabilityZone` . Abbiamo annullato la richiesta di fornire capacità spot aggiuntiva. Per informazioni sulle operazioni consigliate, verifica Flessibilità della zona di disponibilità per un cluster

Stato o modifica dello stato	Gravità	Tipo di evento	Codice dell'evento	Messaggio
				Amazon EMR e riprova.
`WAITING` o `Running`	`INFO`		nessuno	Un'operazione di ridimensionamento per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId` `(ClusterName)` è stata avviata da `Entity` alle `Time`.

Eventi di riconfigurazione della flotta di istanze

Stato o modifica dello stato	Gravità	Messaggio
Richiesta di riconfigurazione del parco istanze	`INFO`	Un utente ha richiesto di riconfigurare la flotta di istanze `InstanceFleetID` nel `ClusterId` cluster `ClusterName` Amazon EMR ().
Avvio della riconfigurazione del parco istanze	`INFO`	Amazon EMR ha avviato una riconfigurazione della flotta di istanze nel cluster Amazon EMR () `InstanceFleetID`

Stato o modifica dello stato	Gravità	Messaggio
		all'indirizzo. `ClusterId` `ClusterName` `Time`
Riconfigurazione del parco istanze completata	`INFO`	Amazon EMR ha terminato la riconfigurazione della flotta di istanze nel cluster `InstanceFleetID` Amazon EMR (). `ClusterId` `ClusterName`
Riconfigurazione del parco istanze non riuscita	`WARNING`	Amazon EMR non è riuscito a riconfigurare la flotta di istanze `InstanceFleetID` nel cluster Amazon EMR () all'indirizzo. `ClusterId` `ClusterName` `Time` La riconfigurazione non è riuscita perché. `Reason`
Avvio della riconfigurazione della flotta di istanze	`INFO`	Amazon EMR sta ripristinando la flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId` (`ClusterName`) alla precedente configurazione corretta.
Reversione della riconfigurazione della flotta di istanze completata	`INFO`	Amazon EMR ha completato il ripristino della flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId` (`ClusterName`) alla precedente configurazione corretta.

Stato o modifica dello stato	Gravità	Messaggio
Reversione della riconfigurazione della flotta di istanze non riuscita	`CRITICAL`	Amazon EMR non è riuscito a ripristinare la flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName )` alla configurazione precedentemente riuscita in. `Time` L'inversione della riconfigurazione non è riuscita a causa di. `Reason`
Reversione della riconfigurazione del parco istanze bloccata	`INFO`	Amazon EMR ha temporaneamente bloccato la flotta di istanze nel cluster `InstanceFleetID` Amazon EMR `ClusterId (ClusterName )` in `Time` quanto la flotta di istanze si trova nello stato. `State`

Eventi di ridimensionamento del parco istanze

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`ERROR`	Timeout per il provisioning Spot	L'operazione di ridimensionamento per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` non è stata completata durante l'acquisizione della

Tipo di evento	Gravità	Codice dell'evento	Messaggio
			capacità Spot nella zona di disponibilità `AvailabilityZone` . Abbiamo annullato la tua richiesta e non proveremo più a eseguire il provisioning della capacità Spot aggiuntiva. Il parco istanze ha eseguito il provisioning della capacità Spot di `num`. La capacità Spot di destinazione era `num`. Per ulteriori informazioni e le azioni consigliate, consulta la pagina della documentazione qui e riprova.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	ERROR	Timeout del provisioning On-Demand	L'operazione di ridimensionamento per il parco istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` non è stata completata durante l'acquisizione della capacità On-Demand nella zona di disponibilità `AvailabilityZone` . Abbiamo annullato la tua richiesta e non proveremo più a eseguire il provisioning della capacità On-Demand aggiuntiva. Il parco istanze ha eseguito il provisioning della capacità On-Demand di `num`. La capacità On-Demand di destinazione era `num`. Per ulteriori informazioni e le azioni consigliate, consulta la pagina della documentazione qui e riprova.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`WARNING`	EC2 provisioning: capacità insufficiente dell'istanza	Non siamo in grado di completare l'operazione di ridimensionamento per Instance Fleet `InstanceFleetID` nel cluster EMR `ClusterId (ClusterName)` poiché EC2 Amazon ha una capacità Spot insufficiente per i tipi di istanze e una capacità On-Demand insufficiente per i `[Instancetype1, Instancetype2]` tipi di istanze nella zona `[Instancetype3, Instancetype4]` di disponibilità. `[AvailabilityZone1]` Finora, il parco istanze ha eseguito il provisioning della capacità On-Demand di `num` e la capacità On-Demand di destinazione era `num`. La capacità Spot sottoposta a provisioning è `num`

Tipo di evento	Gravità	Codice dell'evento	Messaggio
			e la capacità Spot di destinazione era num. Consulta qui la documentazione per ulteriori informazioni su come rispondere a questo evento.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`WARNING`	Timeout del provisioning Spot: continuazione del ridimensionamento	Stiamo ancora eseguendo il provisioning della capacità Spot per l'operazione di ridimensionamento del parco istanze avviata alle `time` per il parco istanze con ID `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` per `[Instancetype1, Instancetype2]` nella zona di disponibilità `AvailabilityZone` . Per la precedente operazione di ridimensionamento avviata alle `time`, il periodo di timeout è scaduto, quindi Amazon EMR ha interrotto il provisioning della capacità Spot dopo l'aggiunta di `num` delle `num` istanze richieste al parco istanze. Per ulteriori informazioni, consulta la pagina di documentazione qui.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`WARNING`	Timeout del provisioning On-Demand: continuazione del ridimensionamento	Stiamo ancora eseguendo il provisioning della capacità On-Demand per l'operazione di ridimensionamento del parco istanze avviata alle `time` per il parco istanze con ID `InstanceFleetID` nel cluster Amazon EMR `ClusterId (ClusterName)` per `[Instancetype1, Instancetype2]` nella zona di disponibilità `AvailabilityZone` . Per la precedente operazione di ridimensionamento avviata alle `time`, il periodo di timeout è scaduto, quindi Amazon EMR ha interrotto il provisioning della capacità On-Demand dopo l'aggiunta di `num` delle `num` istanze richieste al parco istanze. Per ulteriori informazioni,

Tipo di evento	Gravità	Codice dell'evento	Messaggio
			consulta la pagina di documentazione qui.
Ridimensionamento della flotta di istanze EMR	`WARNING`	EC2 Provisioning: indirizzo libero insufficiente nella sottorete	Non possiamo completare l'operazione di ridimensionamento, ad esempio la flotta `InstanceFleetID` nel `ClusterId (ClusterName)` cluster Amazon EMR, perché la sottorete specificata [Subnet1, Subnet2] non contiene abbastanza a indirizzi IP privati gratuiti per soddisfare la tua richiesta. Usa l'`DescribeSubnets` operazione per visualizzare quanti indirizzi IP sono disponibili (non utilizzati) nella tua sottorete. Per informazioni su come rispondere a questo evento, consulta Codici di errore per l'EC2 API Amazon.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`WARNING`	EC2 Provisioning - Limite vCPU superato	Il ridimensionamento della flotta di istanze `InstanceFleetID` nel `ClusterName` cluster Amazon EMR viene ritardato perché hai raggiunto il limite del numero di CPUs v (unità di elaborazione virtuali) assegnate alle istanze in esecuzione nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.
Ridimensionamento della flotta di istanze EMR	`WARNING`	EC2 Provisioning: il limite di numero di istanze Spot è stato superato	La fornitura della flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze Spot che puoi avviare nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento della flotta di istanze EMR	`WARNING`	EC2 Provisioning: limite di istanze superato	La fornitura della flotta di istanze `InstanceFleetID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze on-demand che puoi eseguire nel tuo. `account (accountId)` Per ulteriori informazioni sui codici di errore per l' EC2 API Amazon.

> **Note**
>
> Gli eventi di timeout di provisioning vengono emessi quando Amazon EMR interrompe il provisioning della capacità Spot o On-Demand per il parco istanze dopo la scadenza del timeout. Per ulteriori informazioni su come rispondere a tali eventi, consulta Risposta agli eventi di timeout del ridimensionamento del parco istanze del cluster Amazon EMR.

Eventi del gruppo di istanze

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Da `RESIZING` a `Running`	`INFO`	nessuno	L'operazione di ridimensionamento per il gruppo di istanze `InstanceGroupID` nel cluster

Tipo di evento	Gravità	Codice dell'evento	Messaggio
			Amazon EMR `ClusterId (ClusterName)` è stata completata. Il numero di istanze è ora di `Num`. Il ridimensionamento è stato avviato alle `Time` ed è stato completato in `Num` minuti.
Da `RUNNING` a `RESIZING`	`INFO`	nessuno	Un ridimensionamento per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (ClusterName)` è iniziato alle `Time`. In seguito al ridimensionamento il numero di istanze passa da `Num` a `Num`.
`SUSPENDED`	`ERROR`	nessuno	Il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (ClusterName)` è stato arrestato alle `Time` per il seguente motivo: `ReasonDesc` .

Tipo di evento	Gravità	Codice dell'evento	Messaggio
`RESIZING`	`WARNING`	nessuno	L'operazione di ridimensionamento per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (ClusterName)` è bloccata per il seguente motivo: `ReasonDesc` .

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento del gruppo di istanze EMR	`WARNING`	EC2 provisioning: capacità insufficiente dell'istanza	Non siamo in grado di completare l'operazione di ridimensionamento iniziata da `time` for Instance Group `InstanceGroupID` nel cluster EMR `ClusterId (ClusterName)` poiché EC2 Amazon ha una capacità `Spot/On Demand` insufficiente per il `[Instancetype]` tipo di istanza nella zona di disponibilità. `[AvailabilityZone1]` Finora, il gruppo di istanze ha un numero di istanze in esecuzione di `num` e il numero di istanze richiesto era `num`. Consulta qui la documentazione per ulteriori informazioni su come rispondere a questo evento.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento del gruppo di istanze EMR	`WARNING`	EC2 Fornitura : indirizzo libero insufficiente nella sottorete	Non possiamo completare l'operazione di ridimensionamento, ad esempio il gruppo `InstanceGroupID` nel `ClusterId (ClusterName)` cluster Amazon EMR, perché la sottorete specificata [Subnet1, Subnet2] non contiene abbastanza a indirizzi IP privati gratuiti per soddisfare la tua richiesta. Usa l'`DescribeSubnets` operazione per visualizzare quanti indirizzi IP sono disponibili (non utilizzati) nella tua sottorete. Per informazioni su come rispondere a questo evento, consulta Codici di errore per l' EC2 API Amazon.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento del gruppo di istanze EMR	`WARNING`	EC2 Provisioning - Limite vCPU superato	Il ridimensionamento del gruppo di istanze `InstanceGroupID` nel `ClusterName` cluster Amazon EMR viene ritardato perché hai raggiunto il limite del numero di CPUs v (unità di elaborazione virtuali) assegnate alle istanze in esecuzion e nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.
Ridimensionamento del gruppo di istanze EMR	`WARNING`	EC2 Provisioning: il limite di conteggio delle istanze Spot è stato superato	La fornitura del gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze Spot che puoi avviare nel tuo. `account (accountId)` Per ulteriori informazioni, consulta Codici di errore per l' EC2 API Amazon.

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Ridimensionamento del gruppo di istanze EMR	`WARNING`	EC2 Provisioning: limite di istanze superato	La fornitura del gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterID (ClusterName)` viene ritardata perché hai raggiunto il limite del numero di istanze on-demand che puoi eseguire nel tuo. `account (accountId)` Per ulteriori informazioni sui codici di errore per l' EC2 API Amazon.
Da `RUNNING` a `RESIZING`	`INFO`	nessuno	Un ridimensionamento per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (ClusterName)` è stato avviato da `Entity` alle `Time`.

> **Note**
>
> Con Amazon EMR versione 5.21.0 e successive, puoi sovrascrivere le configurazioni del cluster e specificare classificazioni di configurazione aggiuntive per ogni gruppo di istanze in un cluster in esecuzione. A tale scopo, puoi utilizzare la console Amazon EMR, AWS

Command Line Interface (AWS CLI) o l' AWS SDK. Per ulteriori informazioni, consulta Specifica di una configurazione per un gruppo di istanze in un cluster in esecuzione.

La tabella seguente elenca gli eventi di Amazon EMR per l'operazione di riconfigurazione, insieme allo stato o alla modifica dello stato indicato dall'evento, la gravità dell'evento e messaggi relativi agli eventi.

Stato o modifica dello stato	Gravità	Messaggio
RUNNING	INFO	Una riconfigurazione per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) è stata avviata dall'utente alle Time. La versione di configurazione richiesta è Num.
Da RECONFIGURING a Running	INFO	La riconfigurazione per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) è stata completata. La riconfigurazione è stata avviata alle Time e ha richiesto Num minuti per il completamento. La versione di configurazione corrente è Num.
Da RUNNING a RECONFIGURING in	INFO	Un riconfigurazione per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) è iniziata alle Time. Viene eseguita la

Stato o modifica dello stato	Gravità	Messaggio
		configurazione dal numero di versione Num al numero di versione Num.
RESIZING	INFO	L'operazione di riconfigurazione verso la versione di configurazione Num per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) viene temporaneamente bloccata alle Time perché il gruppo di istanze si trova in State.
RECONFIGURING	INFO	L'operazione di ridimensionamento verso il numero di istanze Num per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) viene temporaneamente bloccata alle Time perché il gruppo di istanze si trova in State.

Stato o modifica dello stato	Gravità	Messaggio
RECONFIGURING	WARNING	L'operazione di riconfigu razione per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterN ame) non è riuscita alle Time e ha richiesto Num minuti per la mancata esecuzione. La versione di configurazione non riuscita è Num.
RECONFIGURING	INFO	Viene eseguito il ripristino delle configurazioni al numero di versione riuscita precedent e Num per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterN ame) alle Time. La nuova versione di configurazione è Num.
Da RECONFIGURING a Running	INFO	È stata ripristinata la versione riuscita precedente Num delle configurazioni per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterN ame) alle Time. La nuova versione di configurazione è Num.

Stato o modifica dello stato	Gravità	Messaggio
Da RECONFIGURING a SUSPENDED	CRITICAL	Impossibile ripristinare la versione riuscita precedente Num per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) alle Time.

Eventi di policy di Auto Scaling (scalabilità automatica)

Stato o modifica dello stato	Gravità	Messaggio
PENDING	INFO	Una policy di dimensionamento automatico è stata aggiunta al gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) alle Time. La policy è in attesa di essere collegata. oppure La policy di dimensionamento automatico per il gruppo di istanze InstanceGroupID nel cluster Amazon EMR ClusterId (ClusterName) è stata aggiornata alle Time. La policy è in attesa di essere collegata.
ATTACHED	INFO	La policy di dimensionamento automatico per il gruppo di istanze InstanceGroupID

Stato o modifica dello stato	Gravità	Messaggio
		nel cluster Amazon EMR `ClusterId (Cluster Name)` è stata collegata alle `Time`.
`DETACHED`	`INFO`	La policy di dimensionamento automatico per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (Cluster Name)` è stata scollegata alle `Time`.
`FAILED`	`ERROR`	Non è stato possibile collegare la policy di dimensionamento automatico per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (Cluster Name)` e la policy ha dato esito negativo alle `Time`. oppure Non è stato possibile scollegare la policy di dimensionamento automatico per il gruppo di istanze `InstanceGroupID` nel cluster Amazon EMR `ClusterId (Cluster Name)` e la policy ha dato esito negativo alle `Time`.

Eventi di fase

Stato o modifica dello stato	Gravità	Messaggio
PENDING	INFO	La fase StepID (StepName) è stata aggiunta al cluster Amazon EMR ClusterId (ClusterName) alle Time ed è in attesa di esecuzione.
CANCEL_PENDING	WARN	La fase StepID (StepName) nel cluster Amazon EMR ClusterId (ClusterName) è stata annullata alle Time ed è in attesa di annullamento.
RUNNING	INFO	La fase StepID (StepName) nel cluster Amazon EMR ClusterId (ClusterName) ha iniziato l'esecuzione alle Time.
COMPLETED	INFO	La fase StepID (StepName) nel cluster Amazon EMR ClusterId (ClusterName) ha completato l'esecuzione alle Time. L'esecuzione della fase è stata avviata alle Time ed il completamento ha richiesto Num minuti.
CANCELLED	WARN	La richiesta di annullamento è riuscita per la fase del cluster StepID (StepName) nel cluster Amazon EMR ClusterId (Cluster

Stato o modifica dello stato	Gravità	Messaggio
		`Name)` alle `Time` e la fase è ora annullata.
`FAILED`	`ERROR`	La fase `StepID (StepName )` nel cluster Amazon EMR `ClusterId (ClusterN ame)` non è riuscita alle `Time`.

Eventi di sostituzione dei nodi non funzionanti

Tipo di evento	Gravità	Codice dell'evento	Messaggio	
Sostituzione di nodi non integri con Amazon EMR	`INFO`	È stato rilevato un nodo centrale non integro	Amazon EMR ha identificato che l'istanza principale `[instanceID (Instance Name)]` `InstanceG roup/Flee t` nel cluster Amazon EMR è. `clusterID (ClusterN ame)` `UNHEALTHY` Amazon EMR tenterà di ripristin are o sostituir e correttam	

Tipo di evento	Gravità	Codice dell'evento	Messaggio	
			ente l'istanza. UNHEALTHY	
Sostituzione di nodi non integri con Amazon EMR	INFO	Nodo principale non integro: sostituzione disattivata	Amazon EMR ha identificato che l'istanza principale [instanceID (Instance Name)] InstanceGroup/Fleet nel cluster Amazon EMR è. (clusterID) (ClusterName) UNHEALTHY Attiva la sostituzione graziosa dei nodi core non integri nel tuo cluster per consentire ad Amazon EMR di sostituire UNHEALTHY le istanze nel caso in cui non possano essere ripristinate.	

Tipo di evento	Gravità	Codice dell'evento	Messaggio
Sostituzione di nodi non integri con Amazon EMR	`WARN`	Nodo principale non funzionante non sostituito	Amazon EMR non può sostituire l'istanza *`UNHEALTHY`* principale *`[instanceID (Instance Name)]`* *`InstanceGroup/Fleet`* nel cluster *`clusterID (ClusterName)`* Amazon EMR per un motivo. **Note** Il motivo per cui Amazon EMR non può sostituire il nodo principale varia a seconda dello scenario. Ad esempio, uno dei

Tipo di evento	Gravità	Codice dell'evento	Messaggio
			motivi per cui Amazon EMR non può eliminare un nodo è perché un cluster non avrebbe nodi principali rimanenti .
Sostituzione di nodi non integri con Amazon EMR	INFO	Nodo principale non funzionante recuperato	Amazon EMR ha ripristinato le tue istanze UNHEALTHY principali [instanceID (InstanceName)] InstanceGroup/Fleet nel cluster Amazon EMR clusterID (ClusterName)

Per ulteriori informazioni sulla sostituzione di nodi non integri, consulta Sostituzione di nodi non integri.

Visualizzazione degli eventi con la console Amazon EMR

Per ogni cluster, puoi visualizzare un semplice elenco di eventi nel riquadro dei dettagli, che enumera gli eventi in ordine di occorrenza decrescente. Puoi inoltre visualizzare tutti gli eventi di tutti i cluster in una regione in ordine di occorrenza decrescente.

Se non desideri che un utente visualizzi tutti gli eventi di cluster per una regione, aggiungi un'istruzione che nega l'autorizzazione (`"Effect": "Deny"`) per l'operazione `elasticmapreduce:ViewEventsFromAllClustersInConsole` a una policy collegata all'utente.

Per visualizzare gli eventi per tutti i cluster in una regione con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Eventi.

Per visualizzare gli eventi per un particolare cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli un cluster.
3. Per visualizzare tutti gli eventi, seleziona la scheda Events (Eventi) nella pagina dei dettagli del cluster.

Risposta agli CloudWatch eventi di Amazon EMR

Questa sezione descrive vari modi in cui puoi rispondere agli eventi utilizzabili emessi da Amazon EMR come messaggi di evento. CloudWatch I modi in cui puoi rispondere agli eventi includono la creazione di regole, l'impostazione di allarmi e altre risposte. Le sezioni che seguono includono collegamenti alle procedure e alle risposte consigliate agli eventi comuni.

Argomenti

- Creazione di regole per gli eventi di Amazon EMR con CloudWatch

- Impostazione degli allarmi sulle CloudWatch metriche di Amazon EMR
- Risposta a eventi di capacità insufficiente dell'istanza del cluster Amazon EMR
- Risposta agli eventi di timeout del ridimensionamento del parco istanze del cluster Amazon EMR

Creazione di regole per gli eventi di Amazon EMR con CloudWatch

Amazon EMR invia automaticamente gli eventi a un CloudWatch flusso di eventi. Puoi creare regole che corrispondono a eventi in base a un modello specificato e instradare gli eventi verso target allo scopo di intraprendere delle azioni, ad esempio inviare un'e-mail di notifica. I modelli sono confrontati con l'oggetto JSON dell'evento. Per ulteriori informazioni sui dettagli degli eventi Amazon EMR, consulta gli eventi di Amazon EMR nella Amazon CloudWatch Events User Guide.

Per informazioni sulla configurazione delle regole CloudWatch degli eventi, consulta Creazione di una CloudWatch regola che si attiva in base a un evento.

Impostazione degli allarmi sulle CloudWatch metriche di Amazon EMR

Amazon EMR invia i parametri ad Amazon. CloudWatch In risposta, puoi utilizzare CloudWatch per impostare allarmi sui parametri di Amazon EMR. Ad esempio, puoi configurare un allarme per CloudWatch inviarti un'e-mail ogni volta che l'utilizzo di HDFS supera l'80%. Per istruzioni dettagliate, consulta Creare o modificare un CloudWatch allarme nella Amazon CloudWatch User Guide.

Risposta a eventi di capacità insufficiente dell'istanza del cluster Amazon EMR

Panoramica

I cluster Amazon EMR restituiscono il codice evento `EC2 provisioning - Insufficient Instance Capacity` quando la zona di disponibilità selezionata non ha una capacità sufficiente per soddisfare la richiesta di avvio o ridimensionamento del cluster. L'evento viene emesso periodicamente sia con i gruppi di istanze che con i parchi istanze se Amazon EMR incontra ripetutamente eccezioni di capacità insufficiente e non è in grado di soddisfare la richiesta di provisioning per un'operazione di avvio o ridimensionamento del cluster.

Questa pagina descrive come rispondere al meglio a questo tipo di evento quando si verifica per il cluster EMR.

Risposta consigliata a un evento di capacità insufficiente

Ti consigliamo di rispondere a un evento di capacità insufficiente in uno dei seguenti modi:

- Attendi il ripristino della capacità. La capacità cambia frequentemente, quindi un'eccezione di capacità insufficiente può essere ripristinata autonomamente. Il ridimensionamento dei cluster inizierà o terminerà non appena la EC2 capacità di Amazon sarà disponibile.
- In alternativa, puoi terminare il cluster, modificare le configurazioni del tipo di istanza e creare un nuovo cluster con la richiesta di configurazione del cluster aggiornata. Per ulteriori informazioni, consulta Flessibilità della zona di disponibilità per un cluster Amazon EMR.

Puoi impostare regole o risposte automatiche a un evento di capacità insufficiente, come descritto nella sezione successiva.

Ripristino automatico da un evento di capacità insufficiente

Puoi creare automazione in risposta a eventi di Amazon EMR come quelli con codice evento `EC2 provisioning - Insufficient Instance Capacity`. Ad esempio, la seguente AWS Lambda funzione termina un cluster EMR con un gruppo di istanze che utilizza istanze On-Demand, quindi crea un nuovo cluster EMR con un gruppo di istanze che contiene tipi di istanze diversi rispetto alla richiesta originale.

Le seguenti condizioni attivano il processo automatizzato:

- L'evento di capacità insufficiente viene emesso per i nodi primari o principali da più di 20 minuti.
- Il cluster non è in uno stato READY o WAITING. Per ulteriori informazioni sugli stati del cluster EMR, consulta Comprensione del ciclo di vita del cluster.

> **Note**
>
> Quando crei un processo automatizzato per un'eccezione di capacità insufficiente, devi considerare che l'evento di capacità insufficiente è recuperabile. La capacità cambia spesso e i cluster riprenderanno il ridimensionamento o inizieranno a funzionare non appena la capacità di Amazon EC2 sarà disponibile.

Example funzione per rispondere a un evento di capacità insufficiente

```
// 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")
```

Risposta agli eventi di timeout del ridimensionamento del parco istanze del cluster Amazon EMR

Panoramica

I cluster Amazon EMR emettono eventi durante l'esecuzione dell'operazione di ridimensionamento per i cluster del parco istanze. Gli eventi di timeout di provisioning vengono emessi quando Amazon EMR interrompe il provisioning della capacità Spot o On-Demand per il parco istanze dopo la scadenza del timeout. La durata del timeout può essere configurata dall'utente come parte delle specifiche di ridimensionamento per i parchi istanze. In scenari di ridimensionamento consecutivo per lo stesso parco istanze, Amazon EMR emette gli eventi `Spot provisioning timeout - continuing resize` o `On-Demand provisioning timeout - continuing resize` alla scadenza del timeout per l'operazione di ridimensionamento corrente. Quindi inizia il provisioning della capacità per la successiva operazione di ridimensionamento del parco istanze.

Risposta agli eventi di timeout del ridimensionamento del parco istanze

Ti consigliamo di rispondere a un evento di timeout del provisioning in uno dei seguenti modi:

- Rivedi le specifiche di ridimensionamento e riprova a eseguire l'operazione di ridimensionamento. Poiché la capacità cambia frequentemente, i cluster verranno ridimensionati correttamente non appena la EC2 capacità di Amazon sarà disponibile. Consigliamo ai clienti di configurare valori inferiori per la durata del timeout per i lavori che richiedono requisiti più rigorosi. SLAs
- In alternativa, puoi:
  - Avviare un nuovo cluster con tipi di istanze diversificati in base alle best practice per la flessibilità dell'istanza e della zona di disponibilità oppure
  - Avviare un cluster con la capacità On-demand
- Per il timeout del provisioning, l'evento di ridimensionamento continua, puoi attendere anche l'elaborazione delle operazioni di ridimensionamento. Amazon EMR continuerà a elaborare in sequenza le operazioni di ridimensionamento attivate per il parco istanze, rispettando le specifiche di ridimensionamento configurate.

Puoi anche impostare regole o risposte automatiche a questo evento come descritto nella sezione successiva.

Ripristino automatico da un evento di timeout del provisioning

Puoi creare automazione in risposta agli eventi di Amazon EMR con il codice evento `Spot Provisioning timeout`. Ad esempio, la seguente funzione AWS Lambda arresta un cluster EMR con un parco istanze che utilizza istanze Spot per i nodi Attività, quindi crea un nuovo cluster EMR con un parco istanze che contiene più tipi di istanze diversificati rispetto alla richiesta originale. In questo esempio, l'evento `Spot Provisioning timeout` emesso per i nodi attività attiverà l'esecuzione della funzione Lambda.

Example Funzione di esempio per rispondere all'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")
```

Visualizza i parametri delle applicazioni cluster utilizzando Ganglia con Amazon EMR

Ganglia è disponibile con le versioni di Amazon EMR tra 4.2 e 6.15. Ganglia è un progetto open source per un sistema scalabile e distribuito, progettato per monitorare i cluster e le griglie riducendo al minimo l'impatto sulle loro prestazioni. Quando abiliti Ganglia sul cluster, puoi generare report e visualizzare le prestazioni del cluster nel suo insieme, nonché esaminare le prestazioni di singole istanze nodi. Ganglia è anche configurato per acquisire e visualizzare i parametri Hadoop e Spark. Per ulteriori informazioni, consulta la sezione Ganglia nella Guida ai rilasci di Amazon EMR.

Registrazione delle chiamate AWS API EMR utilizzando AWS CloudTrail

AWS EMR è integrato con AWS CloudTrail, un servizio che fornisce una registrazione delle azioni intraprese da un utente, ruolo o un. Servizio AWS CloudTrail acquisisce tutte le chiamate API per AWS EMR come eventi. Le chiamate acquisite includono chiamate dalla console AWS EMR e chiamate in codice alle operazioni dell'API AWS EMR. Utilizzando le informazioni raccolte da CloudTrail, è possibile determinare la richiesta effettuata a AWS EMR, l'indirizzo IP da cui è stata effettuata la richiesta, quando è stata effettuata e dettagli aggiuntivi.

Ogni evento o voce di log contiene informazioni sull'utente che ha generato la richiesta. Le informazioni di identità consentono di determinare quanto segue:

- Se la richiesta è stata effettuata con le credenziali utente root o utente.
- Se la richiesta è stata effettuata per conto di un utente del Centro identità IAM.
- Se la richiesta è stata effettuata con le credenziali di sicurezza temporanee per un ruolo o un utente federato.
- Se la richiesta è stata effettuata da un altro Servizio AWS.

CloudTrail è attivo nel tuo account Account AWS quando crei l'account e hai automaticamente accesso alla cronologia degli CloudTrail eventi. La cronologia CloudTrail degli eventi fornisce un record visualizzabile, ricercabile, scaricabile e immutabile degli ultimi 90 giorni di eventi di gestione registrati in un. Regione AWSPer ulteriori informazioni, consulta Lavorare con la cronologia degli CloudTrail eventi nella Guida per l'utente.AWS CloudTrail Non sono CloudTrail previsti costi per la visualizzazione della cronologia degli eventi.

Per una registrazione continua degli eventi degli Account AWS ultimi 90 giorni, crea un trail o un data store di eventi CloudTrailLake.

CloudTrail sentieri

> Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Tutti i percorsi creati utilizzando il AWS Management Console sono multiregionali. È possibile creare un trail per una singola Regione o per più Regioni tramite AWS CLI. La creazione di un percorso multiregionale è consigliata in quanto consente di registrare l'intera attività del proprio Regioni AWS account. Se si crea un trail per una singola Regione, è possibile visualizzare solo gli eventi registrati nella Regione AWS del trail. Per ulteriori informazioni sui trail, consulta Creating a trail for your Account AWS e Creating a trail for an organization nella Guida per l'utente di AWS CloudTrail .

Puoi inviare gratuitamente una copia dei tuoi eventi di gestione in corso al tuo bucket Amazon S3 CloudTrail creando un percorso, tuttavia ci sono costi di storage di Amazon S3. Per ulteriori informazioni sui CloudTrail prezzi, consulta la pagina Prezzi.AWS CloudTrail Per informazioni sui prezzi di Amazon S3, consulta Prezzi di Amazon S3.

CloudTrail Lake Event Data Store

CloudTrail Lake ti consente di eseguire query basate su SQL sui tuoi eventi. CloudTrail Lake converte gli eventi esistenti in formato JSON basato su righe in formato Apache ORC. ORC è un formato di archiviazione a colonne ottimizzato per il recupero rapido dei dati. Gli eventi vengono aggregati in archivi di dati degli eventi, che sono raccolte di eventi immutabili basate sui criteri selezionati applicando i selettori di eventi avanzati. I selettori applicati a un archivio di dati degli eventi controllano quali eventi persistono e sono disponibili per l'esecuzione della query. Per ulteriori informazioni su CloudTrail Lake, consulta Working with AWS CloudTrail Lake nella Guida per l'utente.AWS CloudTrail

CloudTrail Gli archivi e le richieste di dati sugli eventi di Lake comportano dei costi. Quando crei un datastore di eventi, scegli l'opzione di prezzo da utilizzare per tale datastore. L'opzione di prezzo determina il costo per l'importazione e l'archiviazione degli eventi, nonché il periodo di conservazione predefinito e quello massimo per il datastore di eventi. Per ulteriori informazioni sui CloudTrail prezzi, consulta Prezzi.AWS CloudTrail

AWS Eventi relativi ai dati EMR in CloudTrail

Gli eventi di dati forniscono informazioni sulle operazioni delle risorse eseguite su o in una risorsa (ad esempio, lettura o scrittura su un oggetto Amazon S3). Queste operazioni sono definite anche operazioni del piano dei dati. Gli eventi di dati sono spesso attività che interessano volumi elevati di dati. Per impostazione predefinita, CloudTrail non registra gli eventi relativi ai dati. La cronologia CloudTrail degli eventi non registra gli eventi relativi ai dati.

Per gli eventi di dati sono previsti costi aggiuntivi. Per ulteriori informazioni sui CloudTrail prezzi, consulta la sezione AWS CloudTrail Prezzi.

È possibile registrare gli eventi relativi ai dati per i tipi di risorse AWS EMR utilizzando la CloudTrail console o AWS CLI le operazioni CloudTrail API. Per ulteriori informazioni su come registrare gli eventi di dati, consulta Registrazione di eventi di dati con AWS Management Console e Registrazione di eventi di dati con AWS Command Line Interface nella Guida all'utente AWS CloudTrail .

La tabella seguente elenca i tipi di risorse AWS EMR per i quali è possibile registrare gli eventi relativi ai dati. La colonna Data event type (console) mostra il valore da scegliere dall'elenco Data event type

(console) sulla CloudTrail console. La colonna del valore resources.type mostra il `resources.type` valore da specificare durante la configurazione dei selettori di eventi avanzati utilizzando o. AWS CLI CloudTrail APIs La CloudTrail colonna Dati APIs registrati mostra le chiamate API registrate per il tipo di risorsa. CloudTrail

Per ulteriori informazioni su queste operazioni API, consulta il riferimento alla CLI di Amazon EMR WAL (EMRWAL). Amazon EMR registra alcune operazioni Data API in CloudTrail cui non si HBase effettuano chiamate direttamente. Queste operazioni non sono incluse nel riferimento alla CLI EMRWAL.

Tipo di evento di dati (console)	valore resources.type	Dati registrati su APIs CloudTrail
Spazio di lavoro di registrazione write-ahead di Amazon EMR	`AWS::EMRWAL::Workspace`	• GetCurrentWALTime • ListTagsForResource • Elenco WALs • ListWorkspaces • TrimWal • Completo WALFlush

È possibile configurare selettori di eventi avanzati per filtrare i campi `eventName`, `readOnly` e `resources.ARN` per registrare solo gli eventi importanti per l'utente. Per ulteriori informazioni su questi campi, consulta AdvancedFieldSelector in Riferimento API AWS CloudTrail .

AWS Eventi di gestione EMR in CloudTrail

Gli eventi di gestione forniscono informazioni sulle operazioni di gestione eseguite sulle risorse dell'azienda. Account AWS Queste operazioni sono definite anche operazioni del piano di controllo (control-plane). Per impostazione predefinita, CloudTrail registra gli eventi di gestione.

AWS EMR registra tutte le operazioni del piano di controllo AWS EMR come eventi di gestione. Per un elenco delle operazioni del piano di controllo AWS EMR a cui AWS EMR accede, CloudTrail vedere il riferimento all'API EMR.AWS

AWS Esempi di eventi EMR

Un evento rappresenta una singola richiesta proveniente da qualsiasi fonte e include informazioni sull'operazione API richiesta, la data e l'ora dell'operazione, i parametri della richiesta e così via. CloudTrail i file di registro non sono una traccia stack ordinata delle chiamate API pubbliche, quindi gli eventi non vengono visualizzati in un ordine specifico.

L'esempio seguente mostra una voce di CloudTrail registro che illustra l'RunJobFlowazione.

```
{
 "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
   ]
}
```


Per informazioni sul contenuto dei CloudTrail record, consultate il contenuto dei CloudTrail record nella Guida per l'AWS CloudTrail utente.

Best practice per l'osservabilità dell'EMR

L'osservabilità dell'EMR comprende un approccio completo di monitoraggio e gestione per i cluster EMR. AWS La base si basa su Amazon CloudWatch come servizio di monitoraggio principale, integrato da EMR Studio e da strumenti di terze parti come Prometheus e Grafana per una maggiore visibilità. In questo documento, esploriamo aspetti specifici dell'osservabilità dei cluster:

1. Spark observability (GitHub): per quanto riguarda l'interfaccia utente Spark, in Amazon EMR sono disponibili tre opzioni.
2. Spark troubleshooting (GitHub) — Risoluzioni per errori.
3. Monitoraggio del cluster EMR (GitHub): monitoraggio delle prestazioni del cluster.
4. Risoluzione dei problemi EMR (GitHub): identifica, diagnostica e risolvi i problemi più comuni del cluster EMR.

5. Ottimizzazione dei costi (GitHub): questa sezione descrive le migliori pratiche per l'esecuzione di carichi di lavoro convenienti.

Strumento di ottimizzazione delle prestazioni per le applicazioni Apache Spark

1. AWS Lo strumento EMR Advisor analizza i registri degli eventi di Spark per fornire consigli personalizzati per ottimizzare le configurazioni dei cluster EMR, migliorare le prestazioni e ridurre i costi. Sfruttando i dati storici, suggerisce le dimensioni e le impostazioni dell'infrastruttura ideali degli esecutori, consentendo un utilizzo più efficiente delle risorse e migliorando le prestazioni complessive del cluster.
2. Lo strumento Amazon CodeGuru Profiler aiuta gli sviluppatori a identificare i colli di bottiglia e le inefficienze nelle loro applicazioni Spark raccogliendo e analizzando i dati di runtime. Lo strumento si integra perfettamente con le applicazioni Spark esistenti, richiede una configurazione minima e fornisce informazioni dettagliate tramite la AWS console sull'utilizzo della CPU, sui modelli di memoria e sugli hotspot prestazionali.

Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione

Puoi regolare il numero di EC2 istanze Amazon disponibili per un cluster Amazon EMR automaticamente o manualmente in risposta a carichi di lavoro con esigenze diverse. Per utilizzare il dimensionamento automatico, sono disponibili due opzioni. Puoi abilitare il dimensionamento gestito da Amazon EMR o creare una policy di dimensionamento automatico personalizzato. La tabella seguente descrive le differenze tra le due opzioni.

	Dimensionamento gestito da Amazon EMR	Dimensionamento automatico personalizzato
Policy e regole di dimension amento	Non è richiesta alcuna policy. Amazon EMR gestisce l'attività di dimensionamento automatic o valutando continuam ente i parametri del cluster e prendendo decisioni di dimensionamento ottimizzate.	È necessario definire e gestire le policy e le regole di dimensionamento automatic o, ad esempio le condizion i specifiche che attivano le attività di dimensionamento, i

	Dimensionamento gestito da Amazon EMR	Dimensionamento automatico personalizzato
		periodi di valutazione, i periodi di attesa e così via.
Versioni di Amazon EMR supportate	Amazon EMR versione 5.30.0 e successive (tranne Amazon EMR versione 6.0.0)	Amazon EMR 4.0.0 e versioni successive
Composizione cluster supportata	Gruppi di istanze o parchi istanze	Solo gruppi di istanze
Configurazione dei limiti di dimensionamento	I limiti di dimensionamento sono configurati per l'intero cluster.	I limiti di dimensionamento possono essere configura ti solo per ogni gruppo di istanze.
Frequenza di valutazione dei parametri	Ogni 5-10 secondi Una valutazione più frequente dei parametri consente ad Amazon EMR di prendere decisioni di dimensionamento più precise.	È possibile definire i periodi di valutazione solo in incrementi di cinque minuti.
Applicazioni supportate	Sono supportate solo le applicazioni YARN, come Spark, Hadoop, Hive, Flink. La scalabilità gestita di Amazon EMR non supporta applicazi oni non basate su YARN, come Presto o. HBase	È possibile scegliere quali applicazioni sono supportat e quando si definiscono le regole di dimensionamento automatico.

Considerazioni

- Un cluster Amazon EMR è sempre costituito da uno o tre nodi primari. Una volta configurato inizialmente il cluster, è possibile dimensionare solo i nodi principali e i nodi attività. Non è possibile dimensionare il numero di nodi primari per il cluster.
- Per i gruppi di istanze, le operazioni di riconfigurazione e le operazioni di ridimensionamento avvengono consecutivamente e non contemporaneamente. Se si avvia una riconfigurazione durante il ridimensionamento di un gruppo di istanze, la riconfigurazione inizia quando il gruppo di istanze completa il ridimensionamento in corso. Al contrario, se si avvia un'operazione di ridimensionamento mentre un gruppo di istanze ne esegue la riconfigurazione.

Utilizzo del dimensionamento gestito in Amazon EMR

> **Important**
>
> Ti consigliamo vivamente di utilizzare l'ultima versione di Amazon EMR (Amazon EMR 7.10.0) per la scalabilità gestita. In alcune versioni preliminari di Amazon EMR, potresti riscontrare esiti negativi intermittenti delle applicazioni o ritardi nel dimensionamento. Amazon EMR. ha risolto questo problema con le versioni 5.x 5.30.2, 5.31.1, 5.32.1, 5.33.1 e successive e con le versioni 6.x 6.1.1, 6.2.1, 6.3.1 e successive. Per ulteriori informazioni sulla regione e sulla disponibilità del rilascio, consulta Disponibilità di dimensionamento gestito.

Panoramica

Con Amazon EMR versioni 5.30.0 e successive (tranne Amazon EMR 6.0.0), puoi abilitare il dimensionamento gestito da Amazon EMR. Il dimensionamento gestito ti permette di aumentare o diminuire automaticamente il numero di istanze o unità nel cluster in base al carico di lavoro. Amazon EMR valuta continuamente i parametri dei cluster per prendere decisioni di dimensionamento che ottimizzano i cluster in termini di costi e velocità. Il dimensionamento gestito è disponibile per i cluster composti da gruppi di istanze o parchi istanze.

Disponibilità di dimensionamento gestito

- Di seguito Regioni AWS, la scalabilità gestita di Amazon EMR è disponibile con Amazon EMR 6.14.0 e versioni successive:

  - Europa (Spagna) (eu-south-2)
- Di seguito Regioni AWS, la scalabilità gestita di Amazon EMR è disponibile con Amazon EMR 5.30.0 e 6.1.0 e versioni successive:
  - Stati Uniti orientali (Virginia settentrionale) (us-east-1)
  - Stati Uniti orientali (Ohio) (us-east-2)
  - Stati Uniti occidentali (Oregon) (us-west-2)
  - Stati Uniti occidentali (California settentrionale) (us-west-1)
  - Africa (Città del Capo) (af-south-1)
  - Asia Pacifico (Hong Kong) (ap-east-1)
  - Asia Pacifico (Mumbai) (ap-south-1)
  - Asia Pacifico (Hyderabad) (ap-south-2)
  - Asia Pacifico (Seoul) (ap-northeast-2)
  - Asia Pacifico (Singapore) (ap-southeast-1)
  - Asia Pacifico (Sydney) (ap-southeast-2)
  - Asia Pacific (Giacarta) (ap-southeast-3)
  - Asia Pacifico (Tokyo) (ap-northeast-1)
  - Asia Pacifico (Osaka-Locale) (ap-northeast-3)
  - Canada (Centrale) (ca-central-1)
  - Sud America (San Paolo) (sa-east-1)
  - Europa (Francoforte) (eu-central-1)
  - Europa (Zurigo) (eu-central-2)
  - Europa (Irlanda) (eu-west-1)
  - Europa (Londra) (eu-west-2)
  - Europa (Milano) (eu-south-1)
  - Europe (Parigi) (eu-west-3)
  - Europa (Stoccolma) (eu-north-1)
  - Israele (Tel Aviv) (il-central-1)
  - Medio Oriente (EAU) (me-central-1)
  - Cina (Pechino) cn-north-1

  - Cina (Ningxia) cn-nordovest-1

  - AWS GovCloud (Stati Uniti orientali) (-1) us-gov-east
  - AWS GovCloud (Stati Uniti occidentali) (us-gov-west-1)
- Il dimensionamento gestito da Amazon EMR funziona solo con applicazioni YARN, come Spark, Hadoop, Hive e Flink. Non supporta applicazioni che non sono basate su YARN, come Presto e. HBase

Parametri del dimensionamento gestito

È necessario configurare i seguenti parametri per il dimensionamento gestito. Il limite si applica solo ai nodi core e task. Il nodo primario non può essere dimensionato dopo la configurazione iniziale.

- Minimum (`MinimumCapacityUnits`): il limite inferiore della EC2 capacità consentita in un cluster. Si misura in core vCPU (Virtual Central Processing Unit) o istanze per gruppi di istanze. Si misura in unità per parchi istanze.
- Maximum (`MaximumCapacityUnits`): il limite superiore della EC2 capacità consentita in un cluster. Si misura in core vCPU (Virtual Central Processing Unit) o istanze per gruppi di istanze. Si misura in unità per parchi istanze.
- Limite On-Demand (`MaximumOnDemandCapacityUnits`) (Facoltativo): il limite superiore della EC2 capacità consentita per il tipo di mercato On-Demand in un cluster. Se questo parametro non è specificato, il valore predefinito viene impostato su `MaximumCapacityUnits`.
  - Questo parametro viene utilizzato per dividere l'allocazione della capacità tra istanze on demand e Spot. Ad esempio, se imposti 2 istanze come parametro minimo, il parametro massimo sarà 100 istanze e il limite on demand sarà 10 istanze, pertanto il dimensionamento gestito da Amazon EMR aumenta fino a 10 istanze on demand e assegna la capacità rimanente alle istanze spot. Per ulteriori informazioni, consulta Scenari di assegnazione dei nodi.
- Numero massimo di nodi principali (`MaximumCoreCapacityUnits`) (opzionale): il limite superiore della EC2 capacità consentita per il tipo di nodo principale in un cluster. Se questo parametro non è specificato, il valore predefinito viene impostato su `MaximumCapacityUnits`.
  - Questo parametro viene utilizzato per dividere l'allocazione della capacità tra nodi principali e attività. Ad esempio, se imposti 2 istanze come parametro minimo, il parametro massimo sarà 100 istanze e il nodo core massimo sarà 17 istanze, pertanto il dimensionamento gestito da Amazon EMR aumenta fino a 17 nodi core e assegna le 83 istanze rimanenti ai nodi attività. Per ulteriori informazioni, consulta Scenari di assegnazione dei nodi.

Per ulteriori informazioni sui parametri del dimensionamento gestito, consulta `ComputeLimits`.

Considerazioni sul dimensionamento gestito di Amazon EMR

- La scalabilità gestita è supportata in versioni limitate Regioni AWS e di Amazon EMR. Per ulteriori informazioni, consulta Disponibilità di dimensionamento gestito.
- È necessario configurare i parametri richiesti per il dimensionamento gestito da Amazon EMR. Per ulteriori informazioni, consulta Parametri del dimensionamento gestito.
- Per utilizzare il dimensionamento gestito in API Gateway, il processo di raccolta dei parametri deve essere in grado di connettersi all'endpoint API pubblico. Se utilizzi un nome DNS privato con Amazon Virtual Private Cloud, la scalabilità gestita non funzionerà correttamente. Per garantire che il dimensionamento gestito funzioni, consigliamo di eseguire una delle seguenti operazioni:
  - Rimuovi l'endpoint VPC dell'interfaccia API Gateway dal tuo Amazon VPC.
  - Segui le istruzioni in Perché ricevo un errore HTTP 403 Forbidden quando mi connetto al mio API Gateway APIs da un VPC? per disabilitare l'impostazione del nome DNS privato.
  - In alternativa, avvia il cluster in una sottorete privata. Per ulteriori informazioni, consulta l'argomento in Sottoreti private.
- Se i processi YARN sono rallentati in modo intermittente durante la riduzione e i log di YARN Resource Manager mostrano che la maggior parte dei nodi sono stati elencati come negati durante quel periodo, puoi regolare la soglia di timeout di disattivazione.

  Riduci il `spark.blacklist.decommissioning.timeout` da un'ora a un minuto per rendere disponibile il nodo per altri container in sospeso per continuare l'elaborazione del processo.

  Inoltre, è necessario impostare `YARN.resourcemanager.nodemanager-graceful-decommission-timeout-secs` su un valore maggiore per garantire che Amazon EMR non forzi la terminazione del nodo mentre il "Processo Spark" più lungo è ancora in esecuzione sul nodo. Il valore predefinito attuale è 60 minuti, il che significa che la YARN forza la terminazione del container dopo 60 minuti una volta che il nodo entra nello stato di disattivazione.

  Il seguente esempio di riga del log Resource Manager di YARN mostra i nodi aggiunti allo stato di disattivazione:

```
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: []
```

Scopri maggiori dettagli su come Amazon EMR si integra con l'elenco rifiutati di YARN durante la disattivazione dei nodi, i casi in cui i nodi in Amazon EMR possono essere aggiunti all'elenco rifiutati e la configurazione del comportamento di disattivazione dei nodi di Spark.

- Per i carichi di lavoro Spark, la disabilitazione di Spark Dynamic Resource Allocator (DRA) modificando la proprietà Spark spark.dynamicAllocation.Enabled in `FALSE` può causare problemi di Managed Scaling, in cui i cluster possono essere scalati più del necessario per i carichi di lavoro (fino al massimo di calcolo). Quando utilizzi Managed Scaling per questi carichi di lavoro, ti consigliamo di mantenere Spark DRA abilitato, che è lo stato predefinito di questa proprietà.
- L'utilizzo eccessivo dei volumi EBS può causare problemi di dimensionamento gestito. Si consiglia di mantenere il volume EBS inferiore al 90% di utilizzo. Per ulteriori informazioni, consulta Opzioni e comportamento di storage delle istanze in Amazon EMR.
- CloudWatch I parametri di Amazon sono fondamentali per il funzionamento della scalabilità gestita di Amazon EMR. Ti consigliamo di monitorare attentamente i CloudWatch parametri di Amazon per assicurarti che non manchino dati. Per ulteriori informazioni su come configurare gli CloudWatch allarmi per rilevare le metriche mancanti, consulta Using Amazon CloudWatch alarms.
- Le operazioni di dimensionamento gestito su cluster 5.30.0 e 5.30.1 senza Presto installato possono causare errori delle applicazioni o far sì che un gruppo di istanze o un parco istanze uniforme mantenga lo stato `ARRESTED`, in particolare quando un'operazione di riduzione è seguita rapidamente da un'operazione di aumento.

  Come soluzione alternativa, scegli Presto come applicazione da installare quando crei un cluster con Amazon EMR rilasci 5.30.0 e 5.30.1, anche se il tuo processo non richiede Presto.
- Quando imposti il nodo core massimo e il limite on demand per il dimensionamento gestito da Amazon EMR, considera le differenze tra gruppi di istanze e parchi istanze. Ogni gruppo di istanze è costituito dallo stesso tipo di istanza e dalla stessa opzione di acquisto per istanze: on demand o Spot. Per ogni parco istanze, puoi specificare fino a cinque tipi di istanze, che possono essere assegnate come istanze on demand e Spot. Per ulteriori informazioni, consulta Creazione di un cluster con parchi istanze o gruppi di istanze uniformi, Opzioni dei parchi istanze e Scenari di assegnazione dei nodi.
- Con Amazon EMR rilascio 5.30.0 e successivi, se si rimuove la regola predefinita Allow All (Consenti tutto) in uscita su 0.0.0.0/ nel gruppo di sicurezza principale è necessario aggiungere una regola per consentire la connettività TCP in uscita al gruppo di sicurezza di accesso al servizio sulla porta 9443. Inoltre, il gruppo di sicurezza di accesso al servizio deve consentire il traffico

TCP in ingresso sulla porta 9443 dal gruppo di sicurezza principale. Per ulteriori informazioni sulla configurazione dei gruppi di sicurezza, consulta la sezione Amazon EMR-managed security group for the primary instance (private subnets) (Gruppo di sicurezza gestito da Amazon EMR per l'istanza primaria [sottoreti private]).

- Puoi usarlo AWS CloudFormation per configurare la scalabilità gestita di Amazon EMR. Per ulteriori informazioni, consulta AWS::EMR::Cluster nella Guida per l'utente di AWS CloudFormation .
- Se utilizzi nodi Spot, valuta la possibilità di utilizzare le etichette dei nodi per impedire ad Amazon EMR di rimuovere i processi applicativi quando Amazon EMR rimuove i nodi Spot. Per ulteriori informazioni sulle etichette dei nodi, consulta Task nodes.
- L'etichettatura dei nodi non è supportata per impostazione predefinita nelle versioni 6.15 o precedenti di Amazon EMR. Per ulteriori informazioni, consulta Comprendere i tipi di nodi: nodi primari, principali e nodi di attività.
- Se utilizzi le versioni 6.15 o precedenti di Amazon EMR, puoi assegnare etichette ai nodi solo per tipo di nodo, ad esempio nodi core e task. Tuttavia, se utilizzi Amazon EMR versione 7.0 o successiva, puoi configurare le etichette dei nodi per tipo di nodo e tipo di mercato, ad esempio On-Demand e Spot.
- Se la domanda dei processi applicativi aumenta e la domanda degli esecutori diminuisce quando si limita il processo applicativo ai nodi principali, è possibile aggiungere nuovamente i nodi principali e rimuovere i nodi di attività nella stessa operazione di ridimensionamento. Per ulteriori informazioni, vedere Comprensione della strategia e degli scenari di allocazione dei nodi.
- Amazon EMR non etichetta i task node, quindi non puoi impostare le proprietà YARN per limitare i processi applicativi solo per i task node. Tuttavia, se desideri utilizzare i tipi di mercato come etichette dei nodi, puoi utilizzare le `SPOT` etichette `ON_DEMAND` o per il posizionamento dei processi applicativi. Non è consigliabile utilizzare i nodi Spot per i processi primari delle applicazioni.
- Quando utilizzi le etichette dei nodi, il totale delle unità in esecuzione nel cluster può temporaneamente superare la capacità di calcolo massima impostata nella tua politica di scalabilità gestita, mentre Amazon EMR disattiva alcune delle tue istanze. Le unità totali richieste rimarranno sempre pari o inferiori alla capacità di calcolo massima prevista dalla tua policy.
- La scalabilità gestita supporta solo le etichette dei nodi `SPOT` e/o `ON_DEMAND` `CORE` e. `TASK` Le etichette personalizzate dei nodi non sono supportate.
- Amazon EMR crea etichette per i nodi durante la creazione del cluster e il provisioning delle risorse. Amazon EMR non supporta l'aggiunta di etichette dei nodi durante la riconfigurazione del cluster. Inoltre, non è possibile modificare le etichette dei nodi durante la configurazione della scalabilità gestita dopo l'avvio del cluster.

- La scalabilità gestita ridimensiona i nodi principali e i nodi di attività in modo indipendente in base al processo applicativo e alla richiesta dell'esecutore. Per evitare problemi di perdita di dati HDFS durante il core scaling down, segui la procedura standard per i nodi principali. Per ulteriori informazioni sulle best practice relative ai nodi principali e alla replica HDFS, consulta Considerazioni e best practice.
- Non è possibile posizionare sia il processo applicativo che gli esecutori solo sul nodo o. `core` `ON_DEMAND` Se desideri aggiungere sia il processo di applicazione che gli esecutori su uno dei nodi, non utilizzare la `yarn.node-labels.am.default-node-label-expression` configurazione.

  Ad esempio, per posizionare sia il processo dell'applicazione che gli esecutori nei `ON_DEMAND` nodi, impostate max computute sullo stesso valore del valore massimo nel nodo. `ON_DEMAND` Rimuovi anche la configurazione. `yarn.node-labels.am.default-node-label-expression`

  Per aggiungere sia il processo di applicazione che gli esecutori sui `core` nodi, rimuovi la `yarn.node-labels.am.default-node-label-expression` configurazione.
- Quando utilizzi la scalabilità gestita con le etichette dei nodi, imposta la proprietà `yarn.scheduler.capacity.maximum-am-resource-percent: 1` se prevedi di eseguire più applicazioni in parallelo. In questo modo si garantisce che i processi applicativi utilizzino appieno i nodi `CORE` o `ON_DEMAND` disponibili.
- Se utilizzi la scalabilità gestita con etichette dei nodi, imposta la proprietà `yarn.resourcemanager.decommissioning.timeout` su un valore più lungo dell'applicazione in esecuzione più a lungo del cluster. In questo modo si riduce la possibilità che lo scaling gestito di Amazon EMR debba riprogrammare le applicazioni per la rimessa in servizio o i nodi. `CORE` `ON_DEMAND`
- Per ridurre il rischio di errori delle applicazioni dovuti alla perdita casuale dei dati, Amazon EMR raccoglie i parametri dal cluster per determinare i nodi che dispongono di dati di shuffle transitori esistenti della fase attuale e precedente. In rari casi, i parametri possono continuare a riportare dati obsoleti per applicazioni già completate o terminate. Ciò può influire sulla tempestiva riduzione delle istanze nel cluster. Per i cluster con una grande quantità di dati shuffle, prendi in considerazione l'utilizzo delle versioni EMR 6.13 e successive.

Cronologia delle funzionalità

Questa tabella elenca gli aggiornamenti della funzionalità di dimensionamento gestito da Amazon EMR.

Data di rilascio	Funzionalità	Versioni di Amazon EMR
20 novembre 2024	La scalabilità gestita è disponibile nelle regioni di `il-central-1` Israele (Tel Aviv), `me-central-1` Medio Oriente (Emirati Arabi Uniti) e `ap-northeast-3` Asia Pacifico (Osaka).	5.30.0 e 6.1.0 e versioni successive
15 novembre 2024	La scalabilità gestita è disponibile nella regione `eu-central-2` Europa (Zurigo).	5.30.0 e 6.1.0 e versioni successive
20 agosto 2024	Le etichette dei nodi sono ora disponibili nella versione Managed Scaling, quindi puoi etichettare le istanze in base al tipo di mercato o al tipo di nodo per migliorare la scalabilità automatica.	7.2.0 e versioni successive
31 marzo 2024	La scalabilità gestita è disponibile nella regione `ap-south-2` Asia Pacifico (Hyderabad).	6.14.0 e versioni successive
13 febbraio 2024	La scalabilità gestita è disponibile nella `eu-south-2` regione Europa (Spagna).	6.14.0 e versioni successive
10 ottobre 2023	Il dimensionamento gestito è disponibile nella regione Asia	6.14.0 e versioni successive

Data di rilascio	Funzionalità	Versioni di Amazon EMR
	Pacific (Giacarta) `ap-southeast-3` .	
28 luglio 2023	Dimensionamento gestito migliorato per passare a gruppi di istanze di attività diversi su scala verticale quando Amazon EMR subisce un ritardo nel dimension amento verticale rispetto al gruppo di istanze corrente.	5.34.0 e versioni successive, 6.4.0 e versioni successive
16 giugno 2023	Dimensionamento gestito migliorato per rilevare i nodi che eseguono il master dell'applicazione in modo che tali nodi non vengano dimensionati ridotti. Per ulteriori informazioni, consulta Comprendere la strategia e gli scenari di allocazione dei nodi di Amazon EMR.	5.34.0 e versioni successive, 6.4.0 e versioni successive

Data di rilascio	Funzionalità	Versioni di Amazon EMR
21 marzo 2022	Aggiunta la consapevolezza dei dati di shuffle di Spark utilizzata durante la riduzione verticale dei cluster. Per i cluster Amazon EMR con Apache Spark e la caratteristica di dimensionamento gestito abilitato, Amazon EMR monitora continuamente gli esecutori Spark e le posizioni intermedie dei dati di shuffle. Utilizzando queste informazioni, Amazon EMR riduce verticalmente solo le istanze sottoutilizzate che non contengono dati di shuffle utilizzati attivamente. Ciò impedisce il ricalcolo dei dati di shuffle persi, contribuendo a ridurre i costi e migliorare le prestazioni dei processi. Per ulteriori informazioni, consulta la Guida di programmazione Spark.	5.34.0 e versioni successive, 6.4.0 e versioni successive

Configurazione della scalabilità gestita per Amazon EMR

Nelle sezioni seguenti viene illustrato come avviare un cluster EMR che utilizza la scalabilità gestita con AWS Management Console AWS SDK per Java, il o. AWS Command Line Interface

Argomenti

- Utilizzare il per configurare la AWS Management Console scalabilità gestita
- Usa il per configurare la AWS CLI scalabilità gestita
- AWS SDK per Java Da utilizzare per configurare la scalabilità gestita

Utilizzare il per configurare la AWS Management Console scalabilità gestita

Puoi utilizzare la console Amazon EMR per configurare il dimensionamento gestito quando crei un cluster o per modificare una policy di dimensionamento gestita per un cluster in esecuzione.

Console

Per configurare la scalabilità gestita quando crei un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Scegli una versione di Amazon EMR emr-5.30.0 o successiva, tranne quella emr-6.0.0.
4. In Cluster scaling and provisioning option (Opzione di dimensionamento e provisioning del cluster), scegli Use EMR-managed scaling (Utilizza dimensionamento gestito da EMR). Specifica il numero Minimum (Minimo) e Maximum (Massimo) di istanze, il numero di istanze Maximum core node (Massimo del nodo core) e il numero di istanze Maximum On-Demand (Massimo on demand).
5. Scegli qualsiasi altra opzione applicabile al cluster.
6. Per avviare il cluster, scegli Create cluster (Crea cluster).

Per configurare la scalabilità gestita su un cluster esistente con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare.
3. Nella scheda Instances (Istanze) della pagina dei dettagli del cluster, cerca la sezione Instance group settings (Impostazioni del gruppo di istanze). Seleziona Edit cluster scaling (Modifica dimensionamento del cluster) per specificare nuovi valori per il numero Minimum (Minimo) e Maximum (Massimo) di istanze e per il limite On-Demand (On demand).

Usa il per configurare la AWS CLI scalabilità gestita

Puoi usare AWS CLI i comandi per Amazon EMR per configurare la scalabilità gestita quando crei un cluster. È possibile utilizzare una sintassi abbreviata, specificando la configurazione JSON inline

all'interno dei relativi comandi, oppure si può fare riferimento a un file contenente la configurazione JSON. Inoltre puoi applicare una policy di dimensionamento gestito a un cluster esistente e rimuovere una policy di dimensionamento gestito applicata in precedenza. Inoltre, è possibile recuperare i dettagli di una configurazione di policy di dimensionamento da un cluster in esecuzione.

Abilitazione del dimensionamento gestito durante l'avvio del cluster

Puoi abilitare il dimensionamento gestito durante l'avvio del cluster, come illustrato nell'esempio seguente.

```
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}'
```

È inoltre possibile specificare una configurazione di policy gestita utilizzando l'managed-scaling-policy opzione -- quando si utilizza. `create-cluster`

Applicazione di una policy di dimensionamento gestito a un cluster esistente

Puoi applicare una policy di dimensionamento gestito a un cluster esistente come illustrato nell'esempio seguente.

```
aws emr put-managed-scaling-policy
--cluster-id j-123456
--managed-scaling-policy ComputeLimits='{MinimumCapacityUnits=1,
MaximumCapacityUnits=10,  MaximumOnDemandCapacityUnits=10, UnitType=Instances}'
```

Inoltre puoi applicare una policy di dimensionamento gestito a un cluster esistente utilizzando il comando `aws emr put-managed-scaling-policy`. L'esempio seguente utilizza un riferimento a un file JSON, `managedscaleconfig.json`, che specifica la configurazione della policy di dimensionamento gestito.

```
aws emr put-managed-scaling-policy --cluster-id j-123456 --managed-scaling-policy
 file://./managedscaleconfig.json
```

Nell'esempio seguente viene illustrato il contenuto del file `managedscaleconfig.json` che definisce la policy di dimensionamento gestito.

```
{
    "ComputeLimits": {
        "UnitType": "Instances",
        "MinimumCapacityUnits": 1,
        "MaximumCapacityUnits": 10,
        "MaximumOnDemandCapacityUnits": 10
    }
}
```

Recupero di una configurazione di policy di dimensionamento gestito

Il comando `GetManagedScalingPolicy` recupera la configurazione della policy. Ad esempio, il seguente comando recupera la configurazione per il cluster con un ID cluster pari a `j-123456`.

```
aws emr get-managed-scaling-policy --cluster-id j-123456
```

Il comando produce il seguente esempio di output.

```
{
   "ManagedScalingPolicy": {
      "ComputeLimits": {
         "MinimumCapacityUnits": 1,
         "MaximumOnDemandCapacityUnits": 10,
         "MaximumCapacityUnits": 10,
         "UnitType": "Instances"
      }
   }
}
```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Rimozione della policy di dimensionamento gestito

Il comando `RemoveManagedScalingPolicy` rimuove la configurazione della policy. Ad esempio, il seguente comando recupera la configurazione per il cluster con ID cluster `j-123456`.

```
aws emr remove-managed-scaling-policy --cluster-id j-123456
```

AWS SDK per Java Da utilizzare per configurare la scalabilità gestita

Il seguente estratto di programma mostra come configurare il dimensionamento gestito utilizzando AWS SDK per 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() { }
}
```

Scalabilità avanzata per Amazon EMR

A partire da Amazon EMR nella EC2 versione 7.0, puoi sfruttare Advanced Scaling per controllare l'utilizzo delle risorse del cluster. Advanced Scaling introduce una scala di utilizzo/prestazioni per ottimizzare l'utilizzo delle risorse e il livello di prestazioni in base alle esigenze aziendali. Il valore impostato determina se il cluster è maggiormente incentrato sulla conservazione delle risorse o sulla scalabilità per gestire carichi di lavoro sensibili service-level-agreement (SLA), in cui il completamento rapido è fondamentale. Quando il valore di scalabilità viene modificato, la scalabilità gestita interpreta l'intento dell'utente e si ridimensiona in modo intelligente per ottimizzare le risorse. Per ulteriori informazioni sulla scalabilità gestita, consulta Configurare la scalabilità gestita per Amazon EMR.

Impostazioni di ridimensionamento avanzate

Il valore impostato per Advanced Scaling ottimizza il cluster in base alle tue esigenze. I valori vanno da 1 a 100. I valori possibili sono 1, 25, 50, 75 e 100. Se si imposta l'indice su valori diversi da questi, si verifica un errore di convalida.

I valori di scalabilità si riferiscono alle strategie di utilizzo delle risorse. L'elenco seguente definisce alcune di queste:

- Utilizzo ottimizzato [1]: questa impostazione impedisce il sovradimensionamento delle risorse. Utilizzate un valore basso quando desiderate mantenere bassi i costi e dare priorità all'utilizzo efficiente delle risorse. Fa sì che il cluster si espanda in modo meno aggressivo. Questa soluzione è ideale per i casi d'uso in cui si verificano regolarmente picchi di carico di lavoro e non si desidera che le risorse aumentino troppo rapidamente.
- Balanced [50]: bilancia l'utilizzo delle risorse e le prestazioni lavorative. Questa impostazione è adatta per carichi di lavoro stabili in cui la maggior parte delle fasi ha un'autonomia stabile. È adatta anche per carichi di lavoro con un mix di fasi di breve e lunga durata. Ti consigliamo di iniziare con questa impostazione se non sei sicuro di quale scegliere.
- Ottimizzazione delle prestazioni [100]: questa strategia dà priorità alle prestazioni. Il cluster si espande in modo aggressivo per garantire che i lavori vengano completati rapidamente e soddisfino gli obiettivi prestazionali. L'ottimizzazione delle prestazioni è adatta per carichi di lavoro sensibili service-level-agreement (SLA) in cui i tempi di esecuzione rapidi sono fondamentali.

> **Note**
>
> I valori intermedi disponibili forniscono una via di mezzo tra le strategie per ottimizzare il comportamento di Advanced Scaling del cluster.

Vantaggi della scalabilità avanzata

Data la variabilità dell'ambiente e dei requisiti, ad esempio la modifica dei volumi di dati, l'adeguamento degli obiettivi di costo e le implementazioni degli SLA, la scalabilità del cluster può aiutarvi a modificare la configurazione del cluster per raggiungere i vostri obiettivi. I vantaggi principali includono:

- Controllo granulare avanzato: l'introduzione dell'impostazione delle prestazioni di utilizzo consente di regolare facilmente il comportamento di scalabilità del cluster in base alle proprie esigenze. Puoi scalare verso l'alto per soddisfare la domanda di risorse di elaborazione o verso il basso per risparmiare risorse, in base ai tuoi modelli di utilizzo.
- Migliore ottimizzazione dei costi: puoi scegliere un valore di utilizzo basso, in base ai requisiti richiesti, per raggiungere più facilmente i tuoi obiettivi di costo.

Iniziare con l'ottimizzazione

Installazione e configurazione

Utilizza questi passaggi per impostare l'indice delle prestazioni e ottimizzare la tua strategia di scalabilità.

1. Il comando seguente aggiorna un cluster esistente con la strategia di scalabilità ottimizzata per l'utilizzo: `[1]`

```
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"
```

Gli attributi `ScalingStrategy` e `UtilizationPerformanceIndex` sono nuovi e pertinenti per l'ottimizzazione della scalabilità. È possibile selezionare diverse strategie di scalabilità impostando i valori corrispondenti (1, 25, 50, 75 e 100) per l'`UtilizationPerformanceIndex`attributo nella politica di scalabilità gestita.

2. Per ripristinare la strategia di scalabilità gestita predefinita, esegui il comando senza includere gli attributi and. `put-managed-scaling-policy` `ScalingStrategy` `UtilizationPerformanceIndex` (Questo è facoltativo). Questo esempio mostra come eseguire questa operazione:

```
aws emr put-managed-scaling-policy \
--cluster-id 'cluster-id' \
--managed-scaling-policy '{"ComputeLimits":
{"UnitType":"Instances","MinimumCapacityUnits":1,"MaximumCapacityUnits":2,"MaximumOnDemandC
 \
--region "region-name"
```

Utilizzo delle metriche di monitoraggio per tenere traccia dell'utilizzo del cluster

A partire dalla versione 7.3.0 di EMR, Amazon EMR pubblica quattro nuove metriche relative alla memoria e alla CPU virtuale. È possibile utilizzarli per misurare l'utilizzo del cluster attraverso strategie di scalabilità. Queste metriche sono disponibili per qualsiasi caso d'uso, ma puoi utilizzare i dettagli forniti qui per monitorare Advanced Scaling.

Le metriche utili disponibili includono quanto segue:

- YarnContainersUsedMemoryGBSeconds— Quantità di memoria consumata dalle applicazioni gestite da YARN.
- YarnContainersTotalMemoryGBSeconds— Capacità di memoria totale allocata a YARN all'interno del cluster.
- YarnNodesUsedVCPUSeconds— Secondi VCPU totali per ogni applicazione gestita da YARN.
- YarnNodesTotalVCPUSeconds— Secondi VCPU totali aggregati per la memoria consumata, inclusa la finestra temporale in cui yarn non è pronto.

È possibile analizzare le metriche delle risorse utilizzando Logs Insights. Amazon CloudWatch Le funzionalità includono un linguaggio di interrogazione appositamente progettato che consente di estrarre metriche specifiche per l'uso e la scalabilità delle risorse.

La seguente query, che puoi eseguire nella Amazon CloudWatch console, utilizza la matematica metrica per calcolare l'utilizzo medio della memoria (e1) dividendo la somma corrente della memoria consumata (e2) per la somma corrente della memoria totale (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"
}
```

Per interrogare i log, puoi selezionare nella console. CloudWatch AWS Per ulteriori informazioni sulla scrittura di query per CloudWatch, consulta Analyzing log data with CloudWatch Logs Insights nella Amazon CloudWatch Logs User Guide.

L'immagine seguente mostra queste metriche per un cluster di esempio:

Considerazioni e limitazioni

- L'efficacia delle strategie di scalabilità può variare a seconda delle caratteristiche uniche del carico di lavoro e della configurazione del cluster. Ti invitiamo a sperimentare l'impostazione di scalabilità per determinare un valore di indice ottimale per il tuo caso d'uso.

- Amazon EMR Advanced Scaling è particolarmente adatto per carichi di lavoro in batch. Per i carichi di lavoro SQL/data-warehousing e streaming, consigliamo di utilizzare la strategia di scalabilità gestita predefinita per prestazioni ottimali.
- La strategia di scalabilità ottimizzata per le prestazioni consente un'esecuzione dei lavori più rapida mantenendo elevate risorse di elaborazione per un periodo più lungo rispetto alla strategia di scalabilità gestita predefinita. Questa modalità dà priorità alla scalabilità rapida per soddisfare la domanda di risorse, con conseguente completamento più rapido del lavoro. Ciò potrebbe comportare costi più elevati rispetto alla strategia predefinita.
- Nei casi in cui il cluster è già ottimizzato e completamente utilizzato, l'attivazione di Advanced Scaling potrebbe non offrire vantaggi aggiuntivi. In alcune situazioni, l'abilitazione di Advanced Scaling potrebbe comportare un aumento dei costi in quanto i carichi di lavoro potrebbero durare più a lungo. In questi casi, consigliamo di utilizzare la strategia di scalabilità gestita predefinita per garantire l'allocazione ottimale delle risorse e l'efficienza dei costi.
- Nel contesto della scalabilità gestita, l'enfasi si sposta sull'utilizzo delle risorse rispetto al tempo di esecuzione poiché l'impostazione viene modificata da ottimizzata per le prestazioni [100] a ottimizzata per l'utilizzo [1]. Tuttavia, è importante notare che i risultati potrebbero variare in base alla natura del carico di lavoro e alla topologia del cluster. Per garantire risultati ottimali per il tuo caso d'uso, ti consigliamo vivamente di testare le strategie di scalabilità con i tuoi carichi di lavoro per determinare l'impostazione più adatta.
- PerformanceUtilizationIndexAccetta solo i seguenti valori:
  - 1
  - 25
  - 50
  - 75
  - 100

  Qualsiasi altro valore inviato genera un errore di convalida.

Comprendere la strategia e gli scenari di allocazione dei nodi di Amazon EMR

In questa sezione viene fornita una panoramica della strategia di allocazione dei nodi e degli scenari di dimensionamento comuni che è possibile utilizzare con il dimensionamento gestito da Amazon EMR.

Strategia di assegnazione dei nodi

Il dimensionamento gestito da Amazon EMR assegna nodi core e attività in base alle seguenti strategie di aumento e riduzione:

Strategia di aumento

- Per le versioni 7.2 e successive di Amazon EMR, la scalabilità gestita aggiunge innanzitutto i nodi in base alle etichette dei nodi e alla proprietà YARN di restrizione del processo applicativo.
- Per le release 7.2 e successive di Amazon EMR, se hai abilitato le etichette dei nodi e i processi applicativi limitati ai nodi`CORE`, la scalabilità gestita di Amazon EMR aumenta la scalabilità verso i nodi core e i nodi task se la domanda dei processi applicativi aumenta e la domanda degli esecutori aumenta. Allo stesso modo, se sono state abilitate le etichette dei nodi e i processi applicativi sono limitati ai `ON_DEMAND` nodi, la scalabilità gestita aumenta i nodi su richiesta se la domanda del processo applicativo aumenta e aumenta i nodi spot se aumenta la domanda degli esecutori.
- Se le etichette dei nodi non sono abilitate, il posizionamento dei processi applicativi non è limitato a nessun tipo di nodo o mercato.
- Utilizzando le etichette dei nodi, la scalabilità gestita può aumentare e ridimensionare diversi gruppi di istanze e flotte di istanze nella stessa operazione di ridimensionamento. Ad esempio, in uno scenario in cui `instance_group1` ha un `ON_DEMAND` nodo e `instance_group2` ha un `SPOT` nodo e le etichette dei nodi sono abilitate e i processi applicativi sono limitati ai nodi con l'etichetta. `ON_DEMAND` La scalabilità gestita si ridurrà `instance_group1` e aumenterà `instance_group2` se la domanda dei processi applicativi diminuisce e la domanda degli esecutori aumenta.
- Quando Amazon EMR subisce un ritardo nel dimensionamento verticale rispetto al gruppo di istanze corrente, i cluster che utilizzano il dimensionamento gestito passano automaticamente a un gruppo di istanze di attività diverso.
- Se hai impostato il parametro `MaximumCoreCapacityUnits`, Amazon EMR ridimensiona i nodi principali fino a quando le unità principali non raggiungono il limite massimo consentito. La capacità rimanente viene aggiunta ai nodi attività.
- Se hai impostato il parametro `MaximumOnDemandCapacityUnits`, Amazon EMR ridimensiona il cluster utilizzando le istanze on demand fino a quando le unità on demand non raggiungono il limite massimo consentito. La capacità rimanente viene aggiunta utilizzando istanze Spot.
- Se hai impostato entrambi i parametri `MaximumCoreCapacityUnits` e `MaximumOnDemandCapacityUnits`, Amazon EMR considera entrambi i limiti durante il dimensionamento.

Ad esempio, se l'opzione `MaximumCoreCapacityUnits` è minore di `MaximumOnDemandCapacityUnits`, Amazon EMR ridimensiona innanzitutto i nodi principali fino al raggiungimento del limite di capacità principale. Per la capacità rimanente, Amazon EMR utilizza innanzitutto le istanze on demand per ridimensionare i nodi attività fino al raggiungimento del limite on demand e, in seguito, utilizza le istanze Spot per i nodi attività.

Strategia di riduzione verticale

- Analogamente alla strategia di scalabilità verticale, Amazon EMR rimuove i nodi in base alle etichette dei nodi. Per ulteriori informazioni sulle etichette dei nodi, consulta Comprendere i tipi di nodi: nodi primari, principali e nodi di attività.
- Se non hai abilitato le etichette dei nodi, la scalabilità gestita rimuove i nodi di attività e quindi rimuove i nodi principali fino a raggiungere la capacità target di scalabilità verso il basso desiderata. La scalabilità gestita non riduce mai il cluster al di sotto dei vincoli minimi specificati nella politica di scalabilità gestita.
- Le versioni 5.34.0 e successive di Amazon EMR e le versioni 6.4.0 e successive di Amazon EMR supportano il riconoscimento dei dati shuffle di Spark, che impedisce la scalabilità verso il basso di un'istanza mentre Managed Scaling è a conoscenza dei dati shuffle esistenti. Per ulteriori informazioni sulle operazioni di shuffle, consulta la Guida di programmazione Spark. Managed Scaling fa del suo meglio per evitare che i nodi vengano ridimensionati in modo casuale, con dati provenienti dalla fase corrente e precedente di qualsiasi applicazione Spark attiva, fino a un massimo di 30 minuti. Questo aiuta a ridurre al minimo la perdita involontaria dei dati provenienti dallo shuffle, evitando la necessità di ripetere i tentativi di lavoro e il ricalcolo dei dati intermedi. Tuttavia, la prevenzione della perdita di dati in modalità shuffle non è garantita. Per una protezione garantita, consigliamo la consapevolezza dello shuffle su cluster con etichetta di release 7.4 o successiva. Di seguito viene illustrato come impostare la protezione shuffle garantita.
- La scalabilità gestita rimuove prima i nodi delle attività e poi i nodi principali fino a raggiungere la capacità target di scalabilità verso il basso desiderata. Il cluster non scende mai al di sotto dei vincoli minimi specificati nella politica di scalabilità gestita.
- Per i cluster avviati con Amazon EMR 5.x rilasci 5.34.0 e successivi e 6.x rilasci 6.4.0 e successivi, il dimensionamento gestito da Amazon EMR non riduce i nodi su cui è in esecuzione `ApplicationMaster` per Apache Spark. Ciò riduce al minimo gli errori e i nuovi tentativi del processo, il che aiuta a migliorare le prestazioni lavorative e a ridurre i costi. Per confermare su quali nodi del cluster è in esecuzione `ApplicationMaster`, visita Spark History Server e filtra i driver nella scheda Esecutori dell'ID applicazione Spark.

- Sebbene la scalabilità intelligente con EMR Managed Scaling riduca al minimo la perdita di dati shuffle per Spark, in alcuni casi i dati shuffle transitori potrebbero non essere protetti durante uno scale-down. Per fornire una maggiore resilienza dei dati shuffle durante lo scale-down, consigliamo di abilitare Graceful Decommissioning for Shuffle Data in YARN. Quando Graceful Decommissioning for Shuffle Data è abilitato in YARN, i nodi selezionati per lo scale-down con dati shuffle entreranno nello stato Decommissioning e continueranno a servire i file shuffle. YARN attende che i nodi segnalino l'assenza di file ResourceManager shuffle prima di rimuovere i nodi dal cluster.
  - La versione 6.11.0 e successive di Amazon EMR supportano lo smantellamento graduale basato su Yarn per i dati Hive shuffle per i gestori Tez e Shuffle. MapReduce
    - Abilita `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` Graceful `true` Decommissioning for Shuffle Data impostando su.
  - La versione 7.4.0 e successive di Amazon EMR supportano lo smantellamento graduale basato su Yarn per i dati Spark shuffle quando il servizio shuffle esterno è abilitato (abilitato per impostazione predefinita in EMR attivo). EC2
    - Il comportamento predefinito del servizio shuffle esterno Spark, quando si esegue Spark su Yarn, prevede che Yarn rimuova i file shuffle delle applicazioni al momento della chiusura dell'applicazione. NodeManager Ciò può avere un impatto sulla velocità di disattivazione dei nodi e sull'utilizzo del calcolo. Per le applicazioni con esecuzione prolungata, valuta la possibilità di `spark.shuffle.service.removeShuffle` `true` impostare la rimozione dei file shuffle non più in uso per consentire una più rapida disattivazione dei nodi senza dati shuffle attivi.
    - Se il `yarn.nodemanager.shuffledata-monitor.interval-ms` flag o il sono `spark.dynamicAllocation.executorIdleTimeout` stati modificati rispetto ai valori predefiniti, assicurati che la condizione `spark.dynamicAllocation.executorIdleTimeout > yarn.nodemanager.shuffledata-monitor.interval-ms` `true` rimanga aggiornata aggiornando il flag necessario.

Se il cluster non ha alcun carico, Amazon EMR annulla l'aggiunta di nuove istanze da una valutazione precedente ed esegue operazioni di dimensionamento verso il basso. Se il cluster presenta un carico pesante, Amazon EMR annulla la rimozione delle istanze ed esegue operazioni di aumento.

Considerazioni sull'assegnazione

È consigliabile utilizzare l'opzione di acquisto on demand per i nodi principali al fine di evitare la perdita di dati HDFS in caso di recupero Spot. È possibile utilizzare l'opzione di acquisto Spot per i nodi attività al fine di ridurre i costi e ottenere un'esecuzione più rapida dei processi quando vengono aggiunte più istanze Spot ai nodi attività.

Scenari di assegnazione dei nodi

È possibile creare vari scenari di dimensionamento in base alle proprie esigenze impostando i parametri massimo, minimo, limite on demand e nodo principale massimo in diverse combinazioni.

Scenario 1: Dimensionare i soli nodi principali

Per ridimensionare solo i nodi principali, i parametri di dimensionamento gestiti devono soddisfare i seguenti requisiti:

- Il limite on demand è uguale al limite massimo.
- Il nodo principale massimo è uguale al limite massimo.

Quando non vengono specificati i parametri limite on demand e nodo principale massimo, entrambi i parametri impostano il limite massimo.

Questo scenario non è applicabile se si utilizza la scalabilità gestita con etichette dei nodi e si limita l'esecuzione dei processi applicativi solo sui `CORE` nodi, poiché la scalabilità gestita ridimensiona i nodi delle attività per soddisfare la domanda degli esecutori.

I seguenti esempi dimostrano lo scenario di dimensionamento dei nodi principali.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimension amento
Gruppi di istanze Principale: 1 on demand Attività: 1 on demand e 1 Spot	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20	Dimension amento da 1 a 20 istanze o unità di parchi istanze sui

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
	`MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 20	nodi principali utilizzando il tipo on demand. Nessun dimensionamento sui nodi attività.
Parchi istanze Principale: 1 on demand Attività: 1 on demand e 1 Spot	UnitType: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 20	Quando si utilizza la scalabilità gestita con etichette di nodi e si limitano i processi applicativi ai `ON_DEMAND` nodi, il cluster scalerà da 1 a 20 istanze o unità del parco istanze sui `CORE` nodi utilizzando il tipo o, a seconda del `On-Demand` `Spot` tipo di domanda.

Scenario 2: Dimensionamento dei soli nodi attività

Per ridimensionare solo i nodi attività, i parametri di dimensionamento gestiti devono soddisfare il seguente requisito:

- Il nodo principale massimo è uguale al limite minimo.

I seguenti esempi dimostrano lo scenario di dimensionamento dei nodi attività.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimension amento
Gruppi di istanze Principale: 2 on demand Attività: 1 Spot	`UnitType`: Istanze `MinimumCapacityUnits` 2: `MaximumCapacityUnits` : 20 `MaximumCoreCapacityUnits` 2:	Mantenimento dei nodi principal i a 2 e dimension amento dei soli nodi attività da 0 a 18 istanze o unità di parchi istanze. La capacità tra i limiti minimo e massimo viene aggiunta solo ai nodi attività. Quando si utilizza la scalabilità gestita con etichette dei nodi e si limitano i processi applicativi ai nodi ON_DEMAND, il cluster manterrà i nodi principal i fissi su 2 e ridimensionerà solo i nodi di attività tra 0 e 18 istanze o unità del parco
Parchi istanze Principale: 2 on demand Attività: 1 Spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` 2: `MaximumCapacityUnits` : 20 `MaximumCoreCapacityUnits` 2:	

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
		di istanze che utilizzano il `Spot` tipo `On-demand` o, a seconda del tipo di domanda.

Scenario 3: Solo istanze On Demand nel cluster

Per avere solo istanze on demand, il cluster e i parametri di dimensionamento gestito devono soddisfare i requisiti seguenti:

- Il limite on demand è uguale al limite massimo.

  Quando il limite on demand non è specificato, il valore del parametro viene impostato di default sul limite massimo. Il valore di default indica che Amazon EMR ridimensiona solo le istanze on demand.

Se il nodo principale massimo è inferiore al limite massimo, è possibile utilizzare il parametro nodo principale massimo per dividere l'assegnazione della capacità tra nodi principali e nodi attività.

Per abilitare questo scenario in un cluster composto da gruppi di istanze, tutti i gruppi di nodi nel cluster devono utilizzare il tipo di mercato on demand durante la configurazione iniziale.

Questo scenario non è applicabile se si utilizza la scalabilità gestita con etichette dei nodi e si limita l'esecuzione dei processi applicativi solo sui `ON_DEMAND` nodi, poiché la scalabilità gestita ridimensiona i `Spot` nodi per soddisfare la domanda degli esecutori.

Negli esempi seguenti viene illustrato lo scenario che presenta istanze on demand nell'intero cluster.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
Gruppi di istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 12	Dimensionamento da 1 a 12 istanze o unità di parchi istanze sui nodi principali utilizzando il tipo on demand. Dimensionamento della capacità rimanente utilizzando l'opzione on demand sui nodi attività. Nessun dimensionamento utilizzando istanze Spot. Quando si utilizza la scalabilità gestita con etichette di nodi e si limitano i processi applicativi ai `CORE` nodi, il cluster viene scalato da 1 a 20 istanze o unità del parco istanze su `CORE` nodi o
Parchi istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 12	

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
		`task` nodi che utilizzano il tipo, a seconda del `ON_DEMAND` tipo di domanda. La scalabilità sui nodi principali non supererà le 12 istanze o unità del parco istanze.

Scenario 4: Solo istanze Spot nel cluster

Per avere solo istanze Spot, i parametri di dimensionamento gestito devono soddisfare i requisiti seguenti:

- Il limite on demand è impostato su 0.

Se il nodo principale massimo è inferiore al limite massimo, è possibile utilizzare il parametro nodo principale massimo per dividere l'assegnazione della capacità tra nodi principali e nodi attività.

Per abilitare questo scenario in un cluster composto da gruppi di istanze, il gruppo di istanze principale deve utilizzare l'opzione acquisto Spot durante la configurazione iniziale. Se non è presente alcuna istanza spot nel gruppo di istanze attività, il dimensionamento gestito da Amazon EMR crea un gruppo di attività utilizzando le istanze spot quando necessario.

Questo scenario non è applicabile se si utilizza la scalabilità gestita con etichette dei nodi e si limita l'esecuzione dei processi applicativi solo sui `ON_DEMAND` nodi, poiché la scalabilità gestita ridimensiona i nodi per soddisfare la domanda `ON_DEMAND` dei processi applicativi.

Negli esempi seguenti viene illustrato lo scenario che presenta istanze Spot nell'intero cluster.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
Gruppi di istanze Principale: 1 Spot Attività: 1 Spot	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 0	Dimensionamento da 1 a 20 istanze o unità di parchi istanze sui nodi principali utilizzando il tipo Spot. Nessun dimensionamento utilizzando il tipo on demand. Quando si utilizza la scalabilità gestita con etichette di nodi e si limitano i processi applicativi ai `CORE` nodi, il cluster viene scalato da 1 a 20 istanze o unità del parco istanze su `CORE` o `TASK` nodi che utilizzano Spot, a seconda del tipo di domanda. Amazon EMR non è scalabile
Parchi istanze Principale: 1 Spot Attività: 1 Spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 0	

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
		utilizzando questo tipo. `ON_DEMAND`

Scenario 5: Dimensionamento delle istanze On Demand sui nodi principali e delle istanze Spot sui nodi attività

Per dimensionare istanze On Demand sui nodi principali e istanze Spot sui nodi attività, i parametri di ridimensionamento gestiti devono soddisfare i seguenti requisiti:

- Il limite on demand deve essere uguale al nodo principale massimo.
- Sia il limite on demand che il nodo principale massimo devono essere inferiori al limite massimo.

Per abilitare questo scenario in un cluster composto da gruppi di istanze, il gruppo di nodi principale deve utilizzare l'opzione di acquisto on demand.

Questo scenario non è applicabile se utilizzi la scalabilità gestita con etichette di nodi e limiti i processi applicativi all'esecuzione solo su `ON_DEMAND` nodi o `CORE` nodi.

Negli esempi seguenti viene illustrato lo scenario di dimensionamento delle istanze on demand sui nodi principali e delle istanze Spot sui nodi attività.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
Gruppi di istanze Principale: 1 on demand Attività: 1 on demand e 1 Spot	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 7	Aumento a 6 unità on demand sul nodo principale poiché nel nodo attività è già presente 1 unità

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimension amento
	`MaximumCoreCapacityUnits` : 7	on demand e il limite massimo per il tipo on demand è 7. In seguito, aumento a 13 unità Spot sui nodi attività.
Parchi istanze Principale: 1 on demand Attività: 1 on demand e 1 Spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 7 `MaximumCoreCapacityUnits` : 7	

Scenario 6: scalare **`CORE`** le istanze per la domanda dei processi applicativi e **`TASK`** le istanze per la domanda degli esecutori.

Questo scenario è applicabile solo se si utilizza la scalabilità gestita con etichette dei nodi e si limita l'esecuzione dei processi applicativi solo sui nodi. `CORE`

Per scalare `CORE` i nodi in base alla richiesta del processo applicativo e `TASK` i nodi in base alla domanda degli esecutori, è necessario impostare le seguenti configurazioni all'avvio del cluster:

- `yarn.node-labels.enabled:true`
- `yarn.node-labels.am.default-node-label-expression: 'CORE'`

Se non si specificano il `ON_DEMAND` limite e i parametri massimi del `CORE` nodo, entrambi i parametri utilizzano per impostazione predefinita il limite massimo.

Se il `ON_DEMAND` nodo massimo è inferiore al limite massimo, la scalabilità gestita utilizza il parametro del `ON_DEMAND` nodo massimo per suddividere l'allocazione della capacità tra i nodi. `ON_DEMAND` `SPOT` Se si imposta il parametro del `CORE` nodo massimo su un valore inferiore o uguale al parametro di capacità minima, `CORE` i nodi rimangono statici alla capacità core massima.

Gli esempi seguenti illustrano lo scenario di scalabilità delle istanze CORE in base alla domanda dei processi applicativi e delle istanze TASK in base alla domanda dell'esecutore.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimension amento
Gruppi di istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 10 `MaximumCoreCapacityUnits` : 20	Scala i `CORE` nodi tra 1 e 20 nodi in base alla domanda del processo applicativo del cluster utilizzan do il tipo di mercato On-Demand o Spot. Ridimensi ona i `TASK` nodi in base alla domanda dell'esecutore e alla capacità disponibile residua dopo l'allocazione dei nodi da parte di Amazon EMR. `CORE` La somma delle richieste `CORE` e dei `TASK` nodi non supererà il valore di 20. `maximumCa pacity` La somma dei nodi principali su
Parchi istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 10 `MaximumCoreCapacityUnits` : 20	

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
		richiesta e dei nodi di attività su richiesta non supererà il numero `maximumOnDemandCapacity` di 10. I nodi core o task aggiuntivi utilizzano il tipo di mercato Spot.

Scenario 7: scalare **`ON_DEMAND`** le istanze per la domanda dei processi applicativi e **`SPOT`** le istanze per la domanda degli esecutori.

Questo scenario è applicabile solo se si utilizza la scalabilità gestita con etichette dei nodi e si limita l'esecuzione dei processi applicativi solo sui nodi. `ON_DEMAND`

Per scalare `ON_DEMAND` i nodi in base alla richiesta del processo applicativo e `SPOT` i nodi in base alla domanda degli esecutori, è necessario impostare le seguenti configurazioni all'avvio del cluster:

- `yarn.node-labels.enabled:true`
- `yarn.node-labels.am.default-node-label-expression: 'ON_DEMAND'`

Se non si specificano il `ON_DEMAND` limite e i parametri massimi del `CORE` nodo, entrambi i parametri utilizzano per impostazione predefinita il limite massimo.

Se il `CORE` nodo massimo è inferiore al limite massimo, la scalabilità gestita utilizza il parametro del `CORE` nodo massimo per suddividere l'allocazione della capacità tra i nodi. `CORE` `TASK` Se si imposta il parametro del `CORE` nodo massimo su un valore inferiore o uguale al parametro di capacità minima, `CORE` i nodi rimangono statici alla capacità core massima.

Gli esempi seguenti illustrano lo scenario di scalabilità delle istanze On-Demand in base alla domanda del processo applicativo e delle istanze Spot in base alla domanda dell'esecutore.

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
Gruppi di istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: Istanze `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 10	Ridimensiona i `ON_DEMAND` nodi tra 1 e 20 nodi in base alla richiesta del processo applicativo del cluster utilizzando il tipo di nodo or. `CORE` `TASK` Ridimensiona i `SPOT` nodi in base alla domanda dell'esecutore e alla capacità disponibile residua dopo l'allocazione dei nodi da parte di Amazon EMR. `ON_DEMAND` La somma delle richieste `ON_DEMAND` e dei `SPOT` nodi non supererà il valore di 20.
Parchi istanze Principale: 1 on demand Attività: 1 on demand	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 10	

Stato iniziale del cluster	Parametri di dimensionamento	Comportamento del dimensionamento
		`maximumCapacity` La somma dei nodi core on-demand richiesti e dei nodi core spot non supererà il valore `maximumCoreCapacity` di 10. I nodi on-demand o spot aggiuntivi utilizzano il tipo di `TASK` nodo.

Comprensione delle metriche di scalabilità gestita in Amazon EMR

Amazon EMR pubblica i parametri di alta risoluzione con dati a una granularità di un minuto quando il dimensionamento gestito è abilitato per un cluster. Puoi visualizzare gli eventi relativi a ogni avvio e completamento del ridimensionamento controllati dalla scalabilità gestita con la console Amazon EMR o la console Amazon. CloudWatch CloudWatch le metriche sono fondamentali per il funzionamento della scalabilità gestita di Amazon EMR. Ti consigliamo di monitorare attentamente i CloudWatch parametri per assicurarti che non manchino dati. Per ulteriori informazioni su come configurare gli CloudWatch allarmi per rilevare le metriche mancanti, consulta Using Amazon CloudWatch alarms. Per ulteriori informazioni sull'utilizzo CloudWatch degli eventi con Amazon EMR, consulta Monitoraggio CloudWatch degli eventi.

i parametri seguenti indicano le capacità correnti o di destinazione di un cluster. Questi parametri sono disponibili solo quando è abilitata il dimensionamento gestito. Per i cluster composti da parchi istanze, i parametri della capacità del cluster vengono misurate in `Units`. Per i cluster composti da gruppi di istanze, i parametri della capacità del cluster vengono misurate in `Nodes` o in `vCPU` in base al tipo di unità utilizzato nella policy di dimensionamento gestito.

Parametro	Descrizione
• `TotalUnitsRequested` • `TotalNodesRequested` • `TotalVCPURequested`	Il numero totale previsto di unità units/nodes/vCPUs in un cluster, determinato dalla scalabilità gestita. Unità: numero
• `TotalUnitsRunning` • `TotalNodesRunning` • `TotalVCPURunning`	Il numero totale corrente di units/nodes/vCPUs disponibi lità in un cluster in esecuzione. Quando viene richiesto il ridimensionamento di un cluster, questo parametro verrà aggiornato dopo l'aggiunta o la rimozione delle nuove istanze dal cluster. Unità: numero
• `CoreUnitsRequested` • `CoreNodesRequested` • `CoreVCPURequested`	Il numero target di CORE units/nodes/vCPUs in un cluster determinato dalla scalabilità gestita. Unità: numero
• `CoreUnitsRunning` • `CoreNodesRunning` • `CoreVCPURunning`	Il numero attuale di CORE in units/nodes/vCPUs esecuzione in un cluster. Unità: numero
• `TaskUnitsRequested` • `TaskNodesRequested` •	Il numero target di TASK units/nodes/vCPUs in un cluster determinato dalla scalabilità gestita. Unità: numero

Parametro	Descrizione
`TaskVCPURequested`	
• `TaskUnitsRunning` • `TaskNodesRunning` • `TaskVCPURunning`	Il numero corrente di TASK in units/nodes/vCPUs esecuzione in un cluster. Unità: numero

I parametri seguenti indicano lo stato di utilizzo del cluster e delle applicazioni. Questi parametri sono disponibili per tutte le caratteristiche Amazon EMR, ma vengono pubblicati a una risoluzione più elevata con dati a una granularità di un minuto quando il dimensionamento gestito è abilitato per un cluster. È possibile correlare i parametri seguenti con i parametri della capacità del cluster nella tabella precedente per comprendere le decisioni relative al dimensionamento gestito.

Parametro	Descrizione
`AppsCompleted`	Il numero di applicazioni inviate a YARN che sono state completate. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
`AppsPending`	Il numero di applicazioni inviate a YARN che sono in attesa. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
`AppsRunning`	Il numero di applicazioni inviate a YARN che sono in esecuzione. Caso d'uso: monitorare l'avanzamento del cluster

Parametro	Descrizione
	Unità: numero
`ContainerAllocated`	Il numero di contenitori di risorse allocati da. ResourceM anager Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
`ContainerPending`	Il numero di container nella coda non ancora allocati. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
`ContainerPendingRatio`	Il rapporto tra contenitori in sospeso e contenitori allocati (ContainerPendingRatio = ContainerPending /). Container Allocated Se ContainerAllocated = 0, allora Container PendingRatio =. ContainerPending Il valore di Container PendingRatio rappresenta un numero, non una percentua le. Questo valore è utile per il dimensionamento delle risorse del cluster in funzione del comportamento di attribuzione dei container. Unità: numero
`HDFSUtilization`	La percentuale di storage HDFS attualmente utilizzato. Caso d'uso: analizzare le prestazioni del cluster Unità: percentuale

Parametro	Descrizione
`IsIdle`	Indica che un cluster non è più in esecuzione ma è ancora attivo e genera spese. È impostato su 1 se non vi sono task e processi in esecuzione, altrimenti è impostato su 0. Questo valore viene verificato a intervalli di cinque minuti e un valore 1 indica unicamente l'inattività del cluster al momento della verifica e non durante i cinque minuti. Per evitare falsi positivi, devi attivare un allarme quando questo valore è 1 durante due o più verifiche consecutive di cinque minuti. Ad esempio, puoi attivare un allarme se questo valore è 1 per trenta minuti o più. Caso d'uso: monitorare le prestazioni del cluster Unità: booleane
`MemoryAvailableMB`	La quantità di memoria disponibile da allocare. Caso d'uso: monitorare l'avanzamento del cluster Unità: numero
`MRActiveNodes`	Il numero di nodi che attualmente eseguono MapReduce attività o lavori. Equivalente al parametro YARN `mapred.resourcemanager.NoOfActiveNodes` . Caso d'uso: monitorare l'avanzamento del cluster Unità: numero

Parametro	Descrizione
YARNMemoryAvailablePercentage	La percentuale di memoria rimanente disponibile per YARN (YARNMemoryAvailablePercentage = MemoryAvailable MB/ MemoryTotal MB). Questo valore è utile per il dimensionamento delle risorse del cluster in funzione dell'utilizzo della memoria di YARN. Unità: percentuale

Le seguenti metriche forniscono informazioni sulle risorse utilizzate dai contenitori e dai nodi YARN. Queste metriche del gestore delle risorse YARN offrono approfondimenti sulle risorse utilizzate dai contenitori e dai nodi in esecuzione nel cluster. Il confronto di queste metriche con le metriche della capacità del cluster della tabella precedente fornisce un quadro più chiaro dell'impatto della scalabilità gestita:

Parametro	Versioni associate	Descrizione
YarnContainersUsedMemoryGBSeconds	Disponibile con l'etichetta di rilascio 7.3.0 e versioni successive	Memoria del contenitore consumata * secondi per il periodo di pubblicazione. Unità: GB* secondi
YarnContainersTotalMemoryGBSeconds	Disponibile con l'etichetta di rilascio 7.3.0 e versioni successive	Il totale del contenitore di filato* secondi per il periodo di pubblicazione. Unità: GB* secondi
YarnContainersUsedVCPUSeconds	Disponibile con l'etichetta di rilascio 7.5.0 e versioni successive	I secondi VCPU del container consumati per il periodo di pubblicazione. Unità: VCPU * secondi

Parametro	Versioni associate	Descrizione
YarnContainersTotalVCPUSeconds	Disponibile con l'etichetta di rilascio 7.5.0 e versioni successive	Il numero totale di secondi VCPU del container per il periodo di pubblicazione. Unità: VCPU * secondi
YarnNodesUsedMemoryGBSeconds	Disponibile con l'etichetta di rilascio 7.5.0 e versioni successive	Memoria del nodo consumata* secondi per il periodo di pubblicazione. Unità: GB* secondi
YarnNodesTotalMemoryGBSeconds	Disponibile con l'etichetta di rilascio 7.5.0 e versioni successive	Memoria totale del nodo* secondi per il periodo di pubblicazione. Unità: GB* secondi
YarnNodesUsedVCPUSeconds	Disponibile con l'etichetta di rilascio 7.3.0 e versioni successive	I secondi VCPU del nodo utilizzati per il periodo di pubblicazione. Unità: VCPU * secondi
YarnNodesTotalVCPUSeconds	Disponibile con l'etichetta di rilascio 7.3.0 e versioni successive	Il numero totale di secondi VCPU del nodo per il periodo di pubblicazione. Unità: VCPU * secondi

Grafici dei parametri di dimensionamento gestito

Puoi visualizzare in grafico i parametri per vedere i modelli di carico di lavoro del cluster e le corrispondenti decisioni di dimensionamento adottate dal dimensionamento gestito da Amazon EMR come illustrato nella procedura riportata di seguito.

Per rappresentare graficamente le metriche di scalabilità gestita nella console CloudWatch

1. Apri la CloudWatch console.
2. Nel riquadro di navigazione, seleziona Amazon EMR. È possibile cercare il cluster da monitorare in base al relativo identificatore.
3. Scorrere fino al parametro da rappresentare graficamente. Aprire un parametro per visualizzare il grafico.
4. Per rappresentare graficamente uno o più parametri, seleziona la casella di controllo accanto a ciascun parametro.

Nell'esempio seguente viene illustrata l'attività di dimensionamento gestito da Amazon EMR di un cluster. Il grafico mostra tre periodi di dimensionamento automatico, che consentono di risparmiare sui costi quando è presente un carico di lavoro meno attivo.

Tutti i parametri relativi alla capacità e all'utilizzo del cluster vengono pubblicati a intervalli di un minuto. Anche altre informazioni statistiche sono associate a ogni dato di un minuto, permettendo il monitoraggio di varie funzioni come `Percentiles`, `Min`, `Max`, `Sum`, `Average`, `SampleCount`.

Ad esempio, il grafico seguente traccia lo stesso parametro `YARNMemoryAvailablePercentage` in percentili diversi, P10, P50, P90, P99, insieme a `Sum`, `Average`, `Min`, `SampleCount`.

Utilizzo del ridimensionamento automatico con una politica personalizzata, ad esempio gruppi in Amazon EMR

La scalabilità automatica con una policy personalizzata nelle versioni 4.0 e successive di Amazon EMR consente di scalare orizzontalmente e scalare in modo programmatico nei nodi principali e nei nodi task in base a CloudWatch una metrica e ad altri parametri specificati in una politica di scalabilità. Il dimensionamento automatico con una policy personalizzata è disponibile con la configurazione dei gruppi di istanze e non è disponibile quando si utilizzano i parchi istanze. Per ulteriori informazioni sui gruppi di istanze e i parchi di istanze, vedere Crea un cluster Amazon EMR con flotte di istanze o gruppi di istanze uniformi.

> **Note**
>
> Per utilizzare la scalabilità automatica con una caratteristica di policy personalizzata in Amazon EMR, è necessario impostare `true` per il parametro `VisibleToAllUsers` quando si crea un cluster. Per ulteriori informazioni, consulta SetVisibleToAllUsers.

La policy di dimensionamento è parte di un gruppo di istanze di configurazione. È possibile specificare una policy durante la configurazione iniziale di un gruppo di istanze o modificando un gruppo di istanze in un cluster esistente, anche quando tale gruppo di istanze è attivo. Ogni gruppo di istanze in un cluster, eccetto il gruppo di istanze primarie, può avere la propria policy di dimensionamento, che consiste in regole di aumento e riduzione orizzontali. Le regole di scalabilità verticale e orizzontale possono essere configurate in modo indipendente, con parametri diversi per ciascuna regola.

Puoi configurare le politiche di scalabilità con AWS Management Console AWS CLI, the o l'API Amazon EMR. Quando utilizzi l'API AWS CLI o Amazon EMR, specifichi la politica di scalabilità in formato JSON. Inoltre, se utilizzi l' AWS CLI API Amazon EMR, puoi specificare parametri personalizzati CloudWatch . I parametri personalizzati non sono selezionabili con la AWS Management Console. Quando crei inizialmente una policy di dimensionamento con la console, viene preconfigurata una policy predefinita adatta a molte applicazioni per facilitare l'avvio. È possibile eliminare o modificare le regole predefinite.

Anche se la scalabilità automatica consente di regolare la on-the-fly capacità del cluster EMR, è comunque necessario considerare i requisiti di base del carico di lavoro e pianificare le configurazioni dei nodi e dei gruppi di istanze. Per ulteriori informazioni, consulta Linee guida di configurazione del cluster.

> **Note**
>
> Per la maggior parte dei carichi di lavoro, è auspicabile impostare regole di scalabilità orizzontale e verticale per ottimizzare l'utilizzo delle risorse. L'impostazione di una regola senza l'altra significa che è necessario ridimensionare manualmente il conteggio delle istanze dopo un'attività di dimensionamento. In altre parole, questo imposta una policy di scalabilità orizzontale o verticale automatica "unidirezionale" o con un reset manuale.

Creazione del ruolo IAM per la scalabilità automatica

La scalabilità automatica in Amazon EMR richiede un ruolo IAM con autorizzazioni per aggiungere e terminare le istanze quando vengono attivate le attività di dimensionamento. A questo scopo è disponibile un ruolo predefinito configurato con la policy di ruolo e di attendibilità più appropriata, ovvero `EMR_AutoScaling_DefaultRole`. Quando crei un cluster con una politica di scalabilità per la prima volta con AWS Management Console, Amazon EMR crea il ruolo predefinito e allega la policy gestita predefinita per le autorizzazioni,. `AmazonElasticMapReduceforAutoScalingRole`

Quando crei un cluster con una politica di scalabilità automatica con AWS CLI, devi innanzitutto assicurarti che esista il ruolo IAM predefinito o di disporre di un ruolo IAM personalizzato con una policy allegata che fornisca le autorizzazioni appropriate. Per creare il ruolo predefinito, è possibile eseguire il comando `create-default-roles` prima di creare un cluster. È quindi possibile specificare l'opzione `--auto-scaling-role EMR_AutoScaling_DefaultRole` al momento della creazione di un cluster. In alternativa, è possibile creare un ruolo di scalabilità automatica personalizzato, quindi specificarlo quando si crea un cluster, ad esempio `--auto-scaling-role` *`MyEMRAutoScalingRole`*. Se si crea un ruolo di scalabilità automatica personalizzato per Amazon EMR, si consiglia di basare le policy di autorizzazione per il proprio ruolo personalizzato sulla base della policy gestita. Per ulteriori informazioni, consulta Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS.

Comprensione delle regole di scalabilità automatica

Quando una regola di scalabilità orizzontale attiva un'attività di scalabilità per un gruppo di istanze, le istanze Amazon vengono aggiunte al gruppo di EC2 istanze in base alle tue regole. I nuovi nodi possono essere utilizzati da applicazioni come Apache Spark, Apache Hive e Presto non appena l' EC2 istanza Amazon entra nello stato. `InService` È anche possibile impostare una regola di scalabilità verticale che chiude le istanze e rimuove i nodi. Per ulteriori informazioni sul ciclo di vita delle EC2 istanze Amazon con scalabilità automatica, consulta Auto Scaling lifecycle nella Amazon Auto Scaling User Guide. EC2

Puoi configurare il modo in cui un cluster termina le EC2 istanze Amazon. Puoi scegliere di terminare l'operazione entro il limite EC2 orario-istanza di Amazon per la fatturazione o al completamento dell'attività. Questa impostazione si applica sia alla scalabilità automatica che alle operazioni di ridimensionamento manuale. Per ulteriori informazioni su questa configurazione, consulta Opzioni di scalabilità verso il basso per i cluster Amazon EMR.

I seguenti parametri per ciascuna regola di una policy determinano il comportamento di scalabilità automatica.

> **Note**
>
> I parametri elencati qui si basano AWS Management Console su Amazon EMR. Quando utilizzi l'API AWS CLI o Amazon EMR, sono disponibili opzioni di configurazione avanzate aggiuntive. Per ulteriori informazioni sulle opzioni avanzate, consulta SimpleScalingPolicyConfigurationAmazon EMR API Reference.

- Il numero massimo di istanze e istanze minime. Il vincolo Maximum instances specifica il numero massimo di istanze Amazon che possono essere incluse nel gruppo di EC2 istanze e si applica a tutte le regole di scalabilità orizzontale. Analogamente, il vincolo Minimum instances specifica il numero minimo di EC2 istanze Amazon e si applica a tutte le regole di scalabilità.
- Il Rule name (Nome della regola), che deve essere univoco all'interno della policy.
- La regolazione della scalabilità, che determina il numero di EC2 istanze da aggiungere (per le regole di scalabilità orizzontale) o terminare (per le regole di scalabilità orizzontale) durante l'attività di scalabilità attivata dalla regola.
- La CloudWatch metrica, che viene controllata per rilevare una condizione di allarme.
- Un operatore di confronto, utilizzato per confrontare la CloudWatch metrica con il valore Threshold e determinare una condizione di attivazione.
- Un periodo di valutazione, con incrementi di cinque minuti, per il quale la CloudWatch metrica deve trovarsi in una condizione di attivazione prima che venga attivata l'attività di scalabilità.
- Un Cooldown period (Periodo di attesa), in secondi, che determina la quantità di tempo che deve intercorrere tra un'attività di dimensionamento avviato da una regola e l'inizio dell'attività di dimensionamento successiva, indipendentemente dalla regola che la attiva. Quando un gruppo di istanze ha terminato un'attività di scalabilità e ha raggiunto lo stato successivo alla scalabilità, il periodo di cooldown offre l'opportunità di stabilizzare le CloudWatch metriche che potrebbero innescare le successive attività di scalabilità. Per ulteriori informazioni, consulta Auto Scaling cooldown nella Amazon Auto Scaling EC2 User Guide.

Considerazioni e limitazioni

- CloudWatch I parametri di Amazon sono fondamentali per il funzionamento della scalabilità automatica di Amazon EMR. Ti consigliamo di monitorare attentamente i CloudWatch parametri di Amazon per assicurarti che non manchino dati. Per ulteriori informazioni su come configurare gli

CloudWatch allarmi Amazon per rilevare i parametri mancanti, consulta Using Amazon CloudWatch alarms.

- L'utilizzo eccessivo dei volumi EBS può causare problemi di dimensionamento gestito. Si consiglia di monitorare attentamente l'utilizzo del volume EBS per assicurarsi che il volume EBS sia inferiore al 90% di utilizzo. Consulta Archiviazione dell'istanza per informazioni su come specificare volumi EBS aggiuntivi.
- Il ridimensionamento automatico con una politica personalizzata nelle versioni di Amazon EMR da 5.18 a 5.28 può comportare errori di scalabilità causati dalla mancanza intermittente di dati nei parametri di Amazon. CloudWatch Si consiglia di utilizzare le versioni più recenti di Amazon EMR per migliorare la scalabilità automatica. Se desideri utilizzare un rilascio di Amazon EMR compreso tra 5.18 e 5.28, puoi anche contattare il Supporto AWS per richiedere una patch.

AWS Management Console Utilizzo di per configurare il ridimensionamento automatico

Quando crei un cluster, configuri una policy di dimensionamento per i gruppi di istanze mediante le opzioni di configurazione avanzate del cluster. È anche possibile creare o modificare una policy di dimensionamento per un gruppo di istanze attive modificando i gruppi di istanze nelle impostazioni Hardware di un cluster esistente.

1. Passa alla nuova console Amazon EMR e seleziona Passa alla vecchia console dalla barra di navigazione laterale. Per ulteriori informazioni su cosa aspettarti quando passi alla vecchia console, consulta Utilizzo della vecchia console.
2. Se stai creando un cluster, nella console di Amazon EMR seleziona Create Cluster (Crea cluster), quindi Go to advanced options (Vai alle opzioni avanzate), scegli l'opzione per Step 1: Software and Steps (Fase 1: software e fasi) e procedi a Step 2: Hardware Configuration (Fase 2: configurazione hardware).

   - o -

   Se si modifica un gruppo di istanze in un cluster in esecuzione, selezionare il cluster dall'elenco del cluster, quindi espandere la sezione Hardware.
3. In Cluster scaling and provisioning option (Opzione di dimensionamento e provisioning del cluster), seleziona Enable cluster scaling (Abilita dimensionamento del cluster). Quindi selezionare Create a custom automatic scaling policy (Crea policy di dimensionamento automatico personalizzato).

Nella tabella Custom automatic scaling policies (Policy di dimensionamento automatico personalizzato), fare clic sull'icona a forma di matita visualizzata nella riga del gruppo di istanze che si desidera configurare. Si apre la schermata delle regole di Auto Scaling.

4. Digitare il Maximum instances (Numero massimo di istanze) che si desidera che il gruppo di istanze contenga dopo che è stato dimensionato orizzontalmente e digitare il Minimum instances (Numero di istanze minimo) che si desidera che il gruppo di istanze contenga dopo che è stato dimensionato verticalmente.
5. Fare clic sulla matita per modificare i parametri della regola, fare clic sulla X per rimuovere una regola dalla policy e fare clic su Add rule (Aggiungi regola) per aggiungere ulteriori regole.
6. Scegliere i parametri delle regole come descritto in precedenza in questo argomento. Per le descrizioni dei parametri disponibili per Amazon EMR, consulta le CloudWatch metriche e le dimensioni di Amazon EMR nella Amazon User Guide. CloudWatch

Utilizzo di per configurare il AWS CLI ridimensionamento automatico

Puoi usare AWS CLI i comandi per Amazon EMR per configurare la scalabilità automatica quando crei un cluster e quando crei un gruppo di istanze. È possibile utilizzare una sintassi abbreviata, specificando la configurazione JSON inline all'interno dei relativi comandi, oppure si può fare riferimento a un file contenente la configurazione JSON. È inoltre possibile applicare una policy di scalabilità automatica a un gruppo di istanze esistente e rimuovere una policy di scalabilità automatica precedentemente applicata. Inoltre, è possibile recuperare i dettagli di una configurazione di policy di dimensionamento da un cluster in esecuzione.

> **Important**
>
> Quando si crea un cluster basato su una policy di scalabilità automatica, è necessario utilizzare il comando `--auto-scaling-role` *`MyAutoScalingRole`* per specificare il ruolo IAM per la scalabilità automatica. Il ruolo predefinito è *`EMR_AutoScaling_DefaultRole`* e può essere creato con il comando `create-default-roles`. Il ruolo può essere aggiunto solo quando il cluster viene creato e non può essere aggiunto a un cluster esistente.

Per una descrizione dettagliata dei parametri disponibili durante la configurazione di una politica di scalabilità automatica, consulta PutAutoScalingPolicyAmazon EMR API Reference.

Creazione di un cluster con una policy di scalabilità automatica applicata a un gruppo di istanze

È possibile specificare una configurazione di scalabilità automatica all'interno dell'opzione `--instance-groups` del comando `aws emr create-cluster`. L'esempio seguente mostra un comando di creazione di cluster in cui viene fornita una policy di scalabilità automatica per il gruppo di istanze principale. Il comando crea una configurazione di scalabilità equivalente alla politica di scalabilità orizzontale predefinita che appare quando si crea una politica di scalabilità automatica con Amazon AWS Management Console EMR. Per brevità, non viene mostrata una policy di scalabilità verticale. Non è consigliabile creare una regola di scalabilità orizzontale senza una regola di scalabilità verticale.

```
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
```

Il comando seguente illustra come utilizzare la riga di comando per fornire la definizione della policy di dimensionamento automatico come parte di un file di configurazione del gruppo di istanze denominato *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 il contenuto del file di configurazione come segue:

```
[
{
  "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}"
             }
          ]
        }
       }
      }
     ]
   }
```

```
}
]
```

Aggiunta di un gruppo di istanze con policy di scalabilità automatica a un cluster

Puoi specificare una configurazione della policy di dimensionamento utilizzando l'opzione `--instance-groups` con il comando `add-instance-groups` nello stesso modo in cui puoi farlo quando utilizzi `create-cluster`. L'esempio seguente utilizza un riferimento a un file JSON, *`instancegroupconfig.json`*, con la configurazione di un gruppo di istanze.

```
aws emr add-instance-groups --cluster-id j-1EKZ3TYEVF1S2 --instance-groups file://your/
path/to/instancegroupconfig.json
```

Applicazione di una policy di scalabilità automatica a un gruppo di istanze esistenti o modifica di una policy applicata

Utilizzare il comando `aws emr put-auto-scaling-policy` per applicare una policy di scalabilità automatica a un gruppo di istanze esistente. Il gruppo di istanze deve essere parte di un cluster che utilizza il ruolo IAM di scalabilità automatica. L'esempio seguente utilizza un riferimento a un file JSON, *`autoscaleconfig.json`*, che specifica la configurazione automatica della policy di scalabilità automatica.

```
aws emr put-auto-scaling-policy --cluster-id j-1EKZ3TYEVF1S2 --instance-group-
id ig-3PLUZBA6WLS07 --auto-scaling-policy file://your/path/to/autoscaleconfig.json
```

Il contenuto del file `autoscaleconfig.json`, che definisce la stessa regola di scalabilità orizzontale mostrata nell'esempio precedente, è mostrato di seguito.

```
{
          "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"
                        }
                }
        }]
 }
```

Rimozione di una policy di scalabilità automatica da un gruppo di istanze

```
aws emr remove-auto-scaling-policy --cluster-id j-1EKZ3TYEVF1S2 --instance-group-
id ig-3PLUZBA6WLS07
```

Recupero di una configurazione di policy di scalabilità automatica

Il `describe-cluster` comando recupera la configurazione della politica nel blocco. InstanceGroup Ad esempio, il seguente comando recupera la configurazione per il cluster con un ID cluster pari a `j-1CWOHP4PI30VJ`.

```
aws emr describe-cluster --cluster-id j-1CWOHP4PI30VJ
```

Il comando produce il seguente esempio di output.

```
{
    "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."
            }
        }
    }
}
```

Ridimensiona manualmente un cluster Amazon EMR in esecuzione

Puoi aggiungere e rimuovere istanze da gruppi di istanze core e task e flotte di istanze in un cluster in esecuzione con o l' AWS Management Console AWS CLI API Amazon EMR. Se un cluster utilizza gruppi di istanze, deve essere cambiato esplicitamente il conteggio delle istanze. Se il cluster utilizza parchi di istanze, è possibile modificare le unità di destinazione per le istanze on demand e le istanze Spot. Il parco di istanze aggiunge e rimuove le istanze per soddisfare la nuova destinazione. Per ulteriori informazioni, consulta Opzioni del parco istanze. Le applicazioni possono utilizzare EC2 istanze Amazon di nuova generazione per ospitare i nodi non appena le istanze sono disponibili. Quando le istanze vengono rimosse, Amazon EMR chiude le attività in modo da non interrompere i lavori e proteggerli dalla perdita di dati. Per ulteriori informazioni, consulta Terminazione al completamento dell'attività.

Ridimensionamento di un cluster con la console

È possibile utilizzare la console di Amazon EMR per ridimensionare un cluster in esecuzione.

Console

Modifica del conteggio delle istanze di un cluster esistente con la nuova console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster e seleziona il cluster che desideri aggiornare. Il cluster deve essere in esecuzione; non è possibile ridimensionare un cluster di provisioning o terminato.
3. Nella scheda Instances (Istanze) della pagina dei dettagli del cluster, visualizza il pannello Instance groups (Gruppi di istanze).
4. Per ridimensionare un gruppo di istanze esistente, utilizza il pulsante di opzione accanto al gruppo di istanze core o attività che desideri ridimensionare, quindi scegli Resize instance group (Ridimensiona gruppo di istanze). Specifica il nuovo numero di istanze per il gruppo di istanze, quindi seleziona Resize (Ridimensiona).

   > **Note**
   >
   > Se scegli di ridurre le dimensioni di un gruppo di istanze in esecuzione, Amazon EMR selezionerà in modo intelligente le istanze da rimuovere dal gruppo per ridurre al minimo la perdita di dati. Per un controllo più granulare dell'operazione di ridimensionamento, puoi selezionare l'ID del gruppo di istanze, scegliere le istanze che desideri rimuovere e quindi utilizzare l'opzione Terminate (Termina). Per ulteriori informazioni sul comportamento di riduzione intelligente, consulta la sezione Opzioni di scalabilità verso il basso per i cluster Amazon EMR.

5. Se desideri annullare l'operazione di ridimensionamento, puoi utilizzare il pulsante di opzione per un gruppo di istanze con lo stato Resizing (In fase di ridimensionamento) e quindi scegliere Stop resize (Interrompi il ridimensionamento) dall'elenco delle operazioni.
6. Per aggiungere uno o più gruppi di istanze attività al cluster in risposta all'aumento del carico di lavoro, scegli Add task instance group (Aggiungi gruppo di istanze attività) dall'elenco delle operazioni. Scegli il tipo di EC2 istanza Amazon, inserisci il numero di istanze per il gruppo di

attività, quindi seleziona Aggiungi gruppo di istanze di attività per tornare al pannello Gruppi di istanze per il tuo cluster.

Quando si modifica il numero di nodi, Status (Stato) del gruppo di istanze viene aggiornato. Quando la modifica richiesta è completata, lo Status (Stato) è Running (In esecuzione).

Ridimensiona un cluster con AWS CLI

È possibile utilizzare il AWS CLI per ridimensionare un cluster in esecuzione. Si può aumentare o diminuire il numero di nodi di attività e si può aumentare il numero di nodi principali in un cluster in esecuzione. È anche possibile chiudere un'istanza nel gruppo di istanze principale con AWS CLI o l'API. Questo deve essere fatto con cautela. La chiusura di un'istanza nel gruppo di istanze core comporta la perdita di dati e l'istanza non viene sostituita automaticamente.

Oltre a ridimensionare i gruppi core e attività, è inoltre possibile aggiungere uno o più gruppi di istanze attività a un cluster in esecuzione con il comando AWS CLI.

Per ridimensionare un cluster modificando il conteggio delle istanze con AWS CLI

È possibile aggiungere istanze al gruppo principale o al gruppo di attività e rimuovere istanze dal gruppo di attività con il AWS CLI `modify-instance-groups` sottocomando con il parametro. `InstanceCount` Per aggiungere istanze ai gruppi principali o di attività, aumentare il `InstanceCount`. Per ridurre il numero di istanze nel gruppo di attività, diminuire `InstanceCount`. La modifica del conteggio delle istanze del gruppo di attività a 0 rimuove tutte le istanze ma non il gruppo di istanze.

- Per aumentare il numero di istanze nel gruppo di istanze di attività da 3 a 4, digita il comando seguente e sostituiscilo *`ig-31JXXXXXXBT0`* con l'ID del gruppo di istanze.

  ```
  aws emr modify-instance-groups --instance-groups
   InstanceGroupId=ig-31JXXXXXXBT0,InstanceCount=4
  ```

  Per recuperare `InstanceGroupId`, utilizzare il sottocomando `describe-cluster`. L'output è un oggetto JSON chiamato `Cluster` contenente l'ID dell'istanza di ciascun gruppo. Per utilizzare questo comando, è necessario l'ID del cluster (che è possibile recuperare con il comando `aws emr list-clusters` o la console). Per recuperare l'ID del gruppo di istanze, digitate il comando seguente e sostituitelo *`j-2AXXXXXXGAPLF`* con l'ID del cluster.

```
aws emr describe-cluster --cluster-id j-2AXXXXXXGAPLF
```

Con AWS CLI, potete anche terminare un'istanza nel gruppo di istanze principale con il `--modify-instance-groups` sottocomando.

> **Warning**
>
> La specifica di `EC2InstanceIdsToTerminate` deve essere eseguita con cautela. Le istanze vengono terminate immediatamente, indipendentemente dallo stato delle applicazioni in esecuzione su di esse, e l'istanza non viene sostituita automaticamente. Questo è vero indipendentemente dalla configurazione del Scale down behavior (Comportamento di ridimensionamento) del cluster. La cessazione di un'istanza in questo modo comporta il rischio di perdita di dati e di un comportamento imprevedibile del cluster.

Per terminare un'istanza specifica sono necessari l'ID del gruppo di istanze (restituito dal `aws emr describe-cluster --cluster-id` sottocomando) e l'ID dell'istanza (restituito dal `aws emr list-instances --cluster-id` sottocomando). Digitate il comando seguente, sostituitelo con l'ID del gruppo di istanze e *i-f9XXXXf2* sostituitelo *ig-6RXXXXXX07SA* con l'ID dell'istanza.

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-6RXXXXXX07SA,EC2InstanceIdsToTerminate=i-f9XXXXf2
```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Per ridimensionare un cluster aggiungendo gruppi di istanze di attività con AWS CLI

Con AWS CLI, è possibile aggiungere da 1 a 48 gruppi di istanze di attività a un cluster con il sottocomando. `--add-instance-groups` I gruppi di istanze attività possono essere aggiunti solo a un cluster contenente un gruppo di istanze primarie e un gruppo di istanze core. Quando si utilizza il AWS CLI, è possibile aggiungere fino a cinque gruppi di istanze di task ogni volta che si utilizza il sottocomando. `--add-instance-groups`

1. Per aggiungere un singolo gruppo di istanze di task a un cluster, digita il comando seguente e sostituiscilo *j-JXBXXXXXX37R* con l'ID del cluster.

```
aws emr add-instance-groups --cluster-id j-JXBXXXXXX37R --instance-groups
 InstanceCount=6,InstanceGroupType=task,InstanceType=m5.xlarge
```

2. Per aggiungere più gruppi di istanze di attività a un cluster, digita il comando seguente e sostituiscilo *j-JXBXXXXXX37R* con l'ID del cluster. È possibile aggiungere fino a cinque gruppi di istanze di attività in un unico 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
```

   Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Interruzione di un ridimensionamento

Utilizzando Amazon EMR versione 4.1.0 o successive, è possibile emettere un ridimensionamento nel corso di un'operazione di ridimensionamento. Inoltre, è possibile interrompere una richiesta di ridimensionamento precedentemente inviata o inviare una nuova richiesta per sovrascrivere una richiesta precedente senza attendere che termini. È possibile arrestare un ridimensionamento esistente anche dalla console o con la chiamata API `ModifyInstanceGroups` con il conteggio corrente come conteggio target del cluster.

Il seguente screenshot mostra un gruppo di istanze di attività di ridimensionamento, ma possono essere arrestate scegliendo Stop (Interrompi).

Per interrompere un ridimensionamento con AWS CLI

È possibile utilizzare il AWS CLI per interrompere un ridimensionamento con il sottocomando. `modify-instance-groups` Supponiamo di avere sei istanze nel proprio gruppo di istanze e di volerle portare a 10. In seguito si decide di annullare la richiesta:

- La richiesta iniziale:

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-myInstanceGroupId,InstanceCount=10
```

La seconda richiesta di bloccare la prima richiesta:

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-myInstanceGroupId,InstanceCount=6
```

> **Note**
>
> Poiché questo processo è asincrono, è possibile vedere i conteggi delle istanze cambiare rispetto alle richieste API precedenti prima che le richieste successive vengano onorate. In caso di riduzione, è possibile che, se si ha un lavoro in corso sui nodi, il gruppo di istanze non si riduca fino a quando i nodi non hanno completato il loro lavoro.

Stato sospeso

Un gruppo di istanze entra in uno stato di sospensione se incontra troppi errori durante il tentativo di avviare i nuovi nodi del cluster. Ad esempio, se i nuovi nodi hanno esito negativo durante l'esecuzione delle operazioni di bootstrap, il gruppo di istanze passa allo stato SUSPENDED (SOSPESO) piuttosto che al provisioning continuo dei nuovi nodi. Dopo aver risolto il problema sottostante, reimpostare il numero desiderato di nodi sul gruppo di istanze del cluster, quindi il gruppo di istanze riprende l'allocazione dei nodi. La modifica di un gruppo di istanze istruisce Amazon EMR a tentare di fornire nuovamente i nodi. Nessun nodo in esecuzione viene riavviato o terminato.

In AWS CLI, il `list-instances` sottocomando restituisce tutte le istanze e i relativi stati, così come il sottocomando. `describe-cluster` Se Amazon EMR rileva un guasto in un gruppo di istanze, modifica lo stato del gruppo su `SUSPENDED`.

Per ripristinare un cluster in uno stato SUSPENDED con AWS CLI

Digitare il sottocomando `describe-cluster` con il parametro `--cluster-id` per visualizzare lo stato delle istanze nel cluster.

- Per visualizzare le informazioni su tutte le istanze e i gruppi di istanze in un cluster, digita il comando seguente e sostituiscilo *`j-3KVXXXXXXY7UG`* con l'ID del cluster.

```
aws emr describe-cluster --cluster-id j-3KVXXXXXXY7UG
```

L'output mostra le informazioni sui gruppi di istanze e lo stato delle istanze:

```
{
    "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-3SUXXXXXQ9ZM",
                "Market": "ON_DEMAND",
                "RunningInstanceCount": 1
            }
...
}
```

Per visualizzare le informazioni su un particolare gruppo di istanze, digitare il sottocomando `list-instances` con i parametri `--cluster-id` e `--instance-group-types`. È possibile visualizzare informazioni per i gruppi primari, core o attività.

```
aws emr list-instances --cluster-id j-3KVXXXXXXY7UG --instance-group-types "CORE"
```

Usare il sottocomando `modify-instance-groups` con il parametro `--instance-groups` per resettare un cluster nello stato SUSPENDED. Il sottocomando `describe-cluster` restituisce l'id del gruppo di istanze.

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-3SUXXXXXQ9ZM,InstanceCount=3
```

Considerazioni sulla riduzione delle dimensioni del cluster

Se scegli di ridurre le dimensioni di un cluster in esecuzione, considera il comportamento di Amazon EMR e le best practice seguenti:

- Per ridurre l'impatto sui processi in corso, Amazon EMR seleziona in modo intelligente le istanze da rimuovere. Per maggiori informazioni sul comportamento di riduzione del cluster, consulta Terminazione al completamento dell'attività nella Guida alla gestione di Amazon EMR.
- Quando riduci le dimensioni di un cluster, Amazon EMR copia i dati dalle istanze rimosse alle istanze rimanenti. Verifica che vi sia una capacità di archiviazione sufficiente per questi dati nelle istanze che rimangono nel gruppo.
- Amazon EMR tenta di rimuovere le autorizzazioni HDFS sulle istanze del gruppo. Prima di ridurre le dimensioni di un cluster, ti consigliamo di ridurre al minimo le operazioni I/O di scrittura HDFS.
- Per un controllo più granulare quando si riducono le dimensioni di un cluster, puoi visualizzare il cluster nella console e accedere alla scheda Instances (Istanze). Seleziona l'ID per il gruppo di istanze che desideri ridimensionare. Quindi utilizza l'opzione Terminate (Termina) per le istanze specifiche che desideri rimuovere.

Configurazione dei timeout di provisioning per controllare la capacità in Amazon EMR

Quando si utilizzano parchi istanze, è possibile configurare i timeout di provisioning. Un timeout di provisioning indica ad Amazon EMR di interrompere il provisioning della capacità dell'istanza se il cluster supera una soglia temporale specifica durante l'avvio o le operazioni di dimensionamento del cluster . Gli argomenti seguenti illustrano come configurare un timeout di provisioning per l'avvio del cluster e per le operazioni di dimensionamento del cluster.

Argomenti

- Configurare i timeout di provisioning per l'avvio del cluster in Amazon EMR
- Personalizzare un periodo di timeout di provisioning per il ridimensionamento del cluster in Amazon EMR

Configurare i timeout di provisioning per l'avvio del cluster in Amazon EMR

Puoi definire un periodo di timeout per il provisioning delle istanze Spot per ogni parco istanze del tuo cluster. Se Amazon EMR non è in grado di eseguire il provisioning della capacità Spot, puoi scegliere

di terminare il cluster o di effettuare il provisioning della capacità on demand. Se il periodo di timeout termina durante il processo di ridimensionamento del cluster, Amazon EMR annulla le richieste Spot non sottoposte a provisioning. Le istanze Spot non sottoposte a provisioning non vengono trasferite alla capacità on demand.

Esegui i seguenti passaggi per personalizzare un periodo di timeout di provisioning per l'avvio del cluster con la console Amazon EMR.

Console

Per configurare il timeout di provisioning quando crei un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Nella pagina Crea cluster, vai alla Configurazione del cluster e seleziona Parchi istanze.
4. In Opzione di dimensionamento e provisioning del cluster, specifica la dimensione Spot per i parchi istanze principali e di attività.
5. In Configurazione del timeout Spot, seleziona Termina il cluster dopo il timeout Spot o Passa a on demand dopo il timeout Spot. Quindi, specifica il periodo di timeout per il provisioning delle istanze Spot. Il valore predefinito è 1 ora.
6. Scegli qualsiasi altra opzione applicabile al tuo cluster.
7. Per avviare il tuo cluster con il timeout configurato, scegli Crea cluster.

AWS CLI

Per specificare un timeout di provisioning con il 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"}]'
```

Personalizzare un periodo di timeout di provisioning per il ridimensionamento del cluster in Amazon EMR

È possibile definire un periodo di timeout per il provisioning delle istanze spot per ogni parco istanze nel tuo cluster. Se Amazon EMR non è in grado di fornire la capacità Spot, annulla la richiesta di ridimensionamento e interrompe i tentativi di fornire capacità Spot aggiuntiva. Quando crei un cluster, puoi configurare il timeout. Per un cluster in esecuzione, puoi aggiungere o aggiornare un timeout.

Alla scadenza del periodo di timeout, Amazon EMR invia automaticamente gli eventi a uno stream di Amazon Events. CloudWatch Con CloudWatch, puoi creare regole che corrispondano agli eventi in base a uno schema specifico e quindi indirizzare gli eventi agli obiettivi per agire. Ad esempio, è possibile configurare una regola per inviare una notifica via email. Per ulteriori informazioni su come creare regole, consulta Creazione di regole per gli eventi di Amazon EMR con CloudWatch. Per ulteriori informazioni sui dati dei diversi eventi, consulta Eventi di modifica dello stato del parco istanze.

Esempi di timeout di provisioning per il ridimensionamento del cluster

Per specificare un timeout di provisioning per il ridimensionamento con la AWS CLI

L'esempio seguente utilizza il comando `create-cluster` per aggiungere un timeout di provisioning per il ridimensionamento.

```
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"}]'
```

L'esempio seguente utilizza il comando `modify-instance-fleet` per aggiungere un timeout di provisioning per il ridimensionamento.

```
aws emr modify-instance-fleet \
--cluster-id j-XXXXXXXXXXXX \
--instance-fleet '{"InstanceFleetId":"if-XXXXXXXXXXX","ResizeSpecifications":
{"SpotResizeSpecification":{"TimeoutDurationMinutes":30},"OnDemandResizeSpecification":
{"TimeoutDurationMinutes":60}}}' \
--region us-east-1
```

L'esempio seguente utilizza il `add-instance-fleet-command` per aggiungere un timeout di provisioning per il ridimensionamento.

```
aws emr add-instance-fleet \
--cluster-id j-XXXXXXXXXXXX \
--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
```

Specificate un timeout di provisioning per il ridimensionamento e l'avvio con AWS CLI

L'esempio seguente utilizza il comando `create-cluster` per aggiungere un timeout di provisioning per il ridimensionamento e l'avvio.

```
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"}]'
```

Considerazioni sui timeout del provisioning del ridimensionamento

Quando configuri i timeout di provisioning del cluster per i tuoi parchi istanze, considera i seguenti comportamenti.

- È possibile configurare i timeout di provisioning sia per le istanze spot che on demand. Il timeout minimo di provisioning è di 5 minuti. Il timeout massimo di provisioning è di 7 giorni.
- È possibile configurare solo i timeout di provisioning per un cluster EMR che utilizza parchi istanze. È necessario configurare separatamente ciascun parco istanze principale e dell'attività.
- Quando crei un cluster, puoi configurare i timeout di provisioning. È possibile aggiungere un timeout o aggiornare un timeout esistente per un cluster in esecuzione.
- Se invii più operazioni di ridimensionamento, Amazon EMR tiene traccia dei timeout di provisioning per ogni operazione di ridimensionamento. Ad esempio, imposta il timeout di provisioning su un cluster su minuti. *60* Quindi, invia un'operazione *R1* di ridimensionamento alla volta. *T1* Invia una

seconda operazione di ridimensionamento *R2* alla volta. *T2* Il timeout di provisioning per R1 scade alle. *T1 + 60 minutes* Il timeout di provisioning per R2 scade alle. *T2 + 60 minutes*

- Se invii una nuova operazione di ridimensionamento scalabile prima della scadenza del timeout, Amazon EMR continua a tentare di fornire capacità per il tuo cluster EMR.

Opzioni di scalabilità verso il basso per i cluster Amazon EMR

> **Note**
>
> A partire da Amazon EMR rilascio 5.10.0, le opzioni relative al comportamento di riduzione non sono più supportate. Grazie all'introduzione della fatturazione al secondo in Amazon EC2, il comportamento predefinito di scalabilità verso il basso per i cluster Amazon EMR ora termina al completamento dell'attività.

Con le versioni di Amazon EMR da 5.1.0 a 5.9.1, sono disponibili due opzioni per il comportamento di scalabilità: terminare al limite dell'ora di istanza per la fatturazione Amazon o terminare al completamento dell'attività. EC2 A partire dalla versione 5.10.0 di Amazon EMR, l'impostazione per la terminazione al limite dell'ora di istanza è obsoleta a causa dell'introduzione della fatturazione al secondo in Amazon. EC2 Si sconsiglia di specificare la terminazione allo scadere dell'ora dell'istanza nelle versioni in cui l'opzione è disponibile.

> **Warning**
>
> Se utilizzi il AWS CLI per emettere un `modify-instance-groups` account`EC2InstanceIdsToTerminate`, queste istanze vengono terminate immediatamente, senza considerare queste impostazioni e indipendentemente dallo stato delle applicazioni in esecuzione su di esse. La cessazione di un'istanza in questo modo comporta il rischio di perdita di dati e di un comportamento imprevedibile del cluster.

Quando viene specificata la terminazione al completamento dell'attività, Amazon EMR nega le elenca e scarica le attività dai nodi prima di terminare le istanze Amazon. EC2 Con uno dei due comportamenti specificati, Amazon EMR non interrompe le istanze EC2 Amazon nei gruppi di istanze principali se ciò potrebbe causare il danneggiamento di HDFS.

Terminazione al completamento dell'attività

Amazon EMR consente di ridimensionare il cluster senza influenzare il carico di lavoro. Amazon EMR tenta di disattivare senza problemi YARN, HDFS e altri daemon sui nodi core e task durante un'operazione di ridimensionamento senza perdere dati o interrompere i lavori. Amazon EMR riduce le dimensioni dei gruppi di istanze solo se il lavoro assegnato ai gruppi è stato completato e questi sono inattivi. Per YARN NodeManager Graceful Decommission, puoi regolare manualmente il tempo di attesa di disattivazione di un nodo.

> **Note**
>
> Quando si verifica una disattivazione regolare, può verificarsi una perdita di dati. Assicurati di eseguire il backup dei dati.

> **Important**
>
> È possibile che i dati HDFS vadano persi definitivamente durante la sostituzione graduale di un'istanza core non integra. Ti consigliamo di eseguire sempre il backup dei dati.

In questo momento è impostato utilizzando una proprietà nella classificazione di configurazione `YARN-site`. Se utilizzi il rilascio 5.12.0 e successivi di Amazon EMR, specifica la proprietà `YARN.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Se utilizzi rilasci precedenti di Amazon EMR, specifica la proprietà `YARN.resourcemanager.decommissioning.timeout`.

Se ci sono ancora contenitori o applicazioni YARN in esecuzione al superamento del tempo di disattivazione, il nodo è costretto a essere disattivato e YARN ricondiziona i contenitori interessati su altri nodi. Il valore predefinito è 3600 secondi (1 ora). Puoi impostare questo timeout con un valore arbitrariamente alto per forzare la riduzione graduale ad attendere più a lungo. Per ulteriori informazioni, consulta la sezione Graceful Decommission of YARN nodes (Disattivazione normale dei nodi YARN) nella documentazione di Apache Hadoop.

Gruppi di nodi attività

Amazon EMR seleziona in modo intelligente le istanze che non presentano attività in esecuzione su qualsiasi fase o applicazione e rimuove prima tali istanze da un cluster. Se tutte le istanze del

cluster sono in uso, Amazon EMR attende il completamento delle attività su un'istanza prima di rimuoverla dal cluster. Il tempo di attesa predefinito è 1 ora. Questo valore può essere modificato con l'impostazione `YARN.resourcemanager.decommissioning.timeout`. Amazon EMR utilizza dinamicamente la nuova impostazione. Puoi impostarlo su un numero arbitrariamente elevato per assicurarti che Amazon EMR non termini alcuna attività riducendo al contempo le dimensioni del cluster.

Gruppi di nodi principali

Sui nodi principali, i DataNode daemon YARN NodeManager e HDFS devono essere disattivati per ridurre il gruppo di istanze. Per quanto riguarda YARN, la riduzione graduale garantisce che un nodo contrassegnato per la disattivazione passi allo stato `DECOMMISSIONED` solo se non vi sono contenitori o applicazioni in attesa o incompleti. La disattivazione termina immediatamente se non vi sono contenitori in funzione sul nodo all'inizio della disattivazione.

Per gli HDFS, le riduzioni graduali assicurano che la capacità nominale degli HDFS sia sufficientemente grande da adattarsi a tutti i blocchi esistenti. Se la capacità target non è sufficientemente grande, solo un numero parziale di istanze principali viene disattivato in modo che i nodi rimanenti possano gestire i dati correnti che risiedono negli HDFS. È necessario garantire una capacità HDFS aggiuntiva per consentire un'ulteriore disattivazione. Dovresti anche cercare di ridurre al minimo la scrittura I/O prima di tentare di ridurre i gruppi di istanze. Una scrittura eccessiva I/O potrebbe ritardare il completamento dell'operazione di ridimensionamento.

Un altro limite è il fattore di replica predefinito, `dfs.replication` all'interno di `/etc/hadoop/conf/hdfs-site`. Durante la creazione di un cluster, Amazon EMR configura il valore in base al numero di istanze del cluster: 1 con 1-3 istanze, 2 per i cluster con 4-9 istanze e 3 per i cluster con oltre 10 istanze.

> **Warning**
>
> 1. L'impostazione di `dfs.replication` su 1 per i cluster con meno di quattro nodi può causare la perdita di dati HDFS in caso di disattivazione anche di un singolo nodo. Ti consigliamo di utilizzare un cluster con almeno quattro nodi principali per i carichi di lavoro di produzione.
> 2. Amazon EMR non consente ai cluster di dimensionare i nodi principali al di sotto di `dfs.replication`. Ad esempio, se `dfs.replication = 2`, il numero minimo di nodi principali è 2.

3. Quando utilizzi il dimensionamento gestito, il dimensionamento automatico o scegli di dimensionare manualmente il cluster, ti consigliamo di impostare `dfs.replication` su 2 o su un valore superiore.

Una riduzione graduale non consente di ridurre i nodi principali al di sotto del fattore di replica HDFS. Questo per consentire a HDFS di chiudere i file a causa di repliche insufficienti. Per aggirare questo limite, abbassate il fattore di replica e riavviate il demone. NameNode

Configura il comportamento di dimensionamento verso il basso per Amazon EMR.

> **Note**
>
> L'opzione di comportamento di riduzione con terminazione allo scadere del limite di ore-istanze non è più supportata per Amazon EMR rilascio 5.10.0 e successivi. Le seguenti opzioni relative alla riduzione vengono visualizzate nella console Amazon EMR solo per i rilasci da 5.1.0 a 5.9.1.

Puoi utilizzare l' AWS Management Console, the o l' AWS CLI API Amazon EMR per configurare il comportamento di scalabilità verso il basso durante la creazione di un cluster.

Console

Per configurare il comportamento di scalabilità verso il basso con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster, quindi scegli Crea cluster.
3. Nella sezione Opzioni di scalabilità e provisioning del cluster, scegli Usa scalabilità automatica personalizzata. In Politiche di scalabilità automatica personalizzate, scegli il pulsante di azione Plus per aggiungere Scala nelle politiche. Ti consigliamo di aggiungere le politiche Scale in e Scale out. L'aggiunta di un solo set di politiche significa che Amazon EMR eseguirà solo la scalabilità unidirezionale e dovrai eseguire manualmente le altre azioni.
4. Scegli qualsiasi altra opzione applicabile al cluster.
5. Per avviare il cluster, scegli Create cluster (Crea cluster).

AWS CLI

Per configurare il comportamento di scalabilità verso il basso con AWS CLI

- Per l'opzione `--scale-down-behavior` puoi specificare `TERMINATE_AT_INSTANCE_HOUR` o `TERMINATE_AT_TASK_COMPLETION`.

Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa

Questa sezione descrive i metodi per terminare un cluster. Per informazioni sull'abilitazione della protezione da cessazione e sulla cessazione automatica dei cluster, consulta Controlla la terminazione dei cluster Amazon EMR. Puoi terminare i cluster il cui stato è `STARTING`, `RUNNING` o `WAITING`. Un cluster il cui stato è `WAITING` deve essere terminato altrimenti viene eseguito a tempo indeterminato generando spese sul tuo account. Puoi terminare un cluster che non riesce a passare dallo stato `STARTING` a un altro stato o che non è in grado di completare una fase.

Se desideri terminare un cluster per il quale la protezione da cessazione è attivata, devi disattivare tale protezione per poter terminare il cluster. I cluster possono essere terminati utilizzando la console, il o utilizzando l' AWS CLI API a livello di codice. `TerminateJobFlows`

A seconda della configurazione del cluster, potrebbero essere necessari dai 5 ai 20 minuti prima che il cluster termini completamente e rilasci le risorse allocate, ad esempio le istanze. EC2

> **Note**
>
> Non è possibile riavviare un cluster terminato, ma è possibile clonare un cluster terminato per riutilizzarne la configurazione per un nuovo cluster. Per ulteriori informazioni, consulta Clonare un cluster Amazon EMR utilizzando la console.

> **Important**
>
> Amazon EMR utilizza il ruolo di servizio Amazon EMR e il `AWSServiceRoleForEMRCleanup` ruolo per pulire le risorse del cluster nel tuo account che non usi più, come le istanze Amazon. EC2 È necessario includere azioni relative alle policy del ruolo per eliminare o terminare le risorse. Altrimenti, Amazon EMR non può eseguire

queste azioni di pulizia e potresti incorrere in costi per le risorse inutilizzate che rimangono nel cluster.

Terminazione di un cluster con la console

Puoi terminare uno o più cluster utilizzando la console di Amazon EMR. La procedura per terminare un cluster nella console varia a seconda dell'attivazione o meno della protezione da cessazione. Per terminare un cluster protetto, devi dapprima disattivare la protezione da cessazione.

Console

Per terminare un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. Seleziona Clusters (Cluster), quindi scegli il cluster da terminare.
3. Nel menu a discesa Actions (Azioni), scegli Terminate cluster (Termina cluster) per aprire il prompt Terminate cluster (Termina cluster).
4. Al prompt, scegli Terminate (Termina). A seconda della configurazione del cluster, la terminazione potrebbe richiedere da 5 a 10 minuti. Per ulteriori informazioni sui cluster Amazon EMR, consulta la sezione Termina un cluster Amazon EMR nello stato di avvio, in esecuzione o in attesa.

Terminazione di un cluster mediante la AWS CLI

Per terminare un cluster non protetto utilizzando il AWS CLI

Per terminare un cluster non protetto utilizzando il AWS CLI, utilizzate il `terminate-clusters` sottocomando con il parametro --cluster-ids.

- Digita il comando seguente per terminare un singolo cluster e sostituirlo con il tuo ID del cluster. *j-3KVXXXXXXX7UG*

  ```
  aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG
  ```

  Per terminare più cluster, digita il comando seguente e sostituisci «e» *j-WJ2XXXXXX8EU* con il *j-3KVXXXXXXX7UG* tuo cluster. IDs

```
aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG j-WJ2XXXXXX8EU
```

Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Per terminare un cluster protetto utilizzando il AWS CLI

Per terminare un cluster protetto utilizzando il AWS CLI, disattivate innanzitutto la protezione da terminazione utilizzando il `modify-cluster-attributes` sottocomando con il parametro. `--no-termination-protected` Utilizzare quindi il sottocomando `terminate-clusters` con il parametro `--cluster-ids` per terminare il cluster.

1. Digitare il comando seguente per disabilitare la protezione dalla terminazione e sostituirla *j-3KVTXXXXXX7UG* con l'ID del cluster.

   ```
   aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-protected
   ```

2. Per terminare il cluster, digita il comando seguente e sostituiscilo *j-3KVXXXXXXX7UG* con il tuo ID del cluster.

   ```
   aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG
   ```

   Per terminare più cluster, digita il seguente comando e sostituisci «e» *j-WJ2XXXXXX8EU* con il *j-3KVXXXXXXX7UG* tuo cluster. IDs

   ```
   aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG j-WJ2XXXXXX8EU
   ```

   Per ulteriori informazioni sull'utilizzo dei comandi Amazon EMR in AWS CLI, consulta. https://docs.aws.amazon.com/cli/latest/reference/emr

Terminazione di un cluster mediante l'API

L'`TerminateJobFlows`operazione termina l'elaborazione delle fasi, carica tutti i dati di registro da Amazon EC2 ad Amazon S3 (se configurato) e termina il cluster Hadoop. Un cluster viene terminato

automaticamente anche se `KeepJobAliveWhenNoSteps` è impostato su `False` in una richiesta `RunJobFlows`.

Puoi utilizzare questa azione per terminare un singolo cluster o un elenco di cluster in base al relativo cluster. IDs

Per ulteriori informazioni sui parametri di input esclusivi di`TerminateJobFlows`, vedere. TerminateJobFlows Per ulteriori informazioni sui parametri generici nella richiesta, consulta la sezione relativa ai Parametri di richiesta comuni.

Clonare un cluster Amazon EMR utilizzando la console

La console di Amazon EMR consente di clonare un cluster, ossia di creare una copia della configurazione del cluster originale da utilizzare come base per un nuovo cluster.

Console

Per clonare un cluster con la console

1. Accedi a e apri AWS Management Console la console Amazon EMR su https://console.aws.amazon.com /emr.
2. In EMR attivo EC2 nel riquadro di navigazione a sinistra, scegli Cluster.
3. Clonazione di un cluster dall'elenco dei cluster
   a. Utilizza le opzioni di ricerca e filtro per trovare il cluster che desideri clonare nella visualizzazione a elenco.
   b. Seleziona la casella di controllo a sinistra della riga relativa al cluster da clonare.
   c. L'opzione Clone (Clona) sarà ora disponibile nella parte superiore della visualizzazione a elenco. Seleziona Clone (Clona) per avviare il processo di clonazione. Se il cluster ha delle fasi configurate, scegli Include steps (Includi fasi) e Continue (Continua) se desideri clonare le fasi insieme alle altre configurazioni del cluster.
   d. Esamina le impostazioni del nuovo cluster, che sono state copiate dal cluster clonato. Se necessario, modifica le impostazioni. Quando reputi la configurazione del nuovo cluster soddisfacente, seleziona Create cluster (Crea cluster) per avviare il nuovo cluster.
4. Clonazione di un cluster da una pagina di dettaglio del cluster
   a. Per accedere alla pagina dei dettagli del cluster che desideri clonare, seleziona il relativo Cluster ID (ID cluster dalla visualizzazione a elenco dei cluster.

b. Nella parte superiore della pagina dei dettagli del cluster, seleziona Clone cluster (Clona cluster) dal menu Actions (Operazioni) per avviare il processo di clonazione. Se il cluster ha delle fasi configurate, scegli Include steps (Includi fasi) e Continue (Continua) se desideri clonare le fasi insieme alle altre configurazioni del cluster.

c. Esamina le impostazioni del nuovo cluster, che sono state copiate dal cluster clonato. Se necessario, modifica le impostazioni. Quando reputi la configurazione del nuovo cluster soddisfacente, seleziona Create cluster (Crea cluster) per avviare il nuovo cluster.

Automatizza i cluster Amazon EMR ricorrenti con AWS Data Pipeline

AWS Data Pipeline è un servizio che automatizza lo spostamento e la trasformazione dei dati. Puoi utilizzare questo metodo per programmare il trasferimento di dati di input a Amazon S3 e l'avvio di cluster per elaborare quei dati. Ad esempio, immaginiamo che disponi di un server Web che registra log di traffico. Se desideri eseguire un cluster settimanale per analizzare i dati sul traffico, puoi utilizzarlo AWS Data Pipeline per pianificare tali cluster. AWS Data Pipeline è un flusso di lavoro basato sui dati, in modo che un'attività (avvio del cluster) possa dipendere da un'altra attività (spostamento dei dati di input su Amazon S3). Dispone inoltre di una potente caratteristica di ripetizione tentativi.

Per ulteriori informazioni su AWS Data Pipeline, consulta la AWS Data Pipeline Developer Guide, in particolare i tutorial su Amazon EMR:

- Tutorial: Avvio di un flusso di lavoro di Amazon EMR
- Guida introduttiva: elabora i log Web con AWS Data Pipeline Amazon EMR e Hive
- Tutorial: importazione ed esportazione di Amazon DynamoDB con AWS Data Pipeline

Tutorial Amazon EMR

Un cluster EMR funziona in un ecosistema complesso. Il seguente tutorial tratta casi d'uso importanti.

Argomenti

- Amazon EMR attivo EC2 : monitoraggio avanzato con CloudWatch l'utilizzo di parametri e log personalizzati
- Monitora le applicazioni Apache Spark su Amazon EMR con Amazon CloudWatch
- Monitora lo stato delle applicazioni Amazon EMR con l'integrazione CloudWatch

Amazon EMR attivo EC2 : monitoraggio avanzato con CloudWatch l'utilizzo di parametri e log personalizzati

Panoramica

Amazon EMR offre funzionalità di elaborazione di big data potenti e convenienti. Per massimizzare le prestazioni e l'utilizzo delle risorse, è essenziale un monitoraggio efficace. Amazon CloudWatch offre un'osservabilità completa per i cluster EMR, consentendoti di tracciare metriche e log in tempo reale. Questo documento descrive come:

1. Configurare l' CloudWatch agente a cui inviare EMR nei log EC2 CloudWatch
2. Aggiungi Hadoop, YARN e metriche personalizzate tramite classificazioni HBase
3. Monitora le metriche tramite dashboard integrate
4. Tieni traccia dei log del cluster tramite gruppi di log CloudWatch

Prerequisiti e contesto

Per impostazione predefinita, Amazon EMR invia i parametri di base CloudWatch ogni cinque minuti senza costi aggiuntivi. Con EMR Release 7.0+, è possibile distribuire l'agente per: CloudWatch

- Raccogli 34 metriche dettagliate aggiuntive a intervalli di un minuto (si applicano costi aggiuntivi)
- Raccogli le metriche da tutti i nodi del cluster
- Aggrega i dati sul nodo principale prima di inviarli a CloudWatch
- Accedi alle metriche tramite la scheda Monitoraggio o la console della console EMR CloudWatch

EMR 7.1 estende queste funzionalità, consentendoti di configurare l'agente per acquisire metriche specializzate da Hadoop, YARN e componenti. HBase Per gli ambienti che utilizzano Prometheus, i parametri possono essere inoltrati ad Amazon Managed Service for Prometheus.

CloudWatch Configurazione dell'agente per i log

Per acquisire i log in EMR CloudWatch, crea un file cloudwatch-config.json che definisca quali file di registro raccogliere:

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 Bootstrap per la configurazione dell'agente CloudWatch

Per applicare la tua CloudWatch configurazione personalizzata ai nodi EMR, crea uno script di bootstrap che riavvierà l' CloudWatch agente con le tue impostazioni. Questo script garantisce che l'agente venga eseguito con i parametri specifici di raccolta dei log dopo il provisioning del cluster.

Creazione dello script Bootstrap

Crea un file denominato cloudwatch-agent-bootstrap.sh con il seguente contenuto:

```
#!/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 importante sulla configurazione

> **Important**
>
> Prima di caricare lo script, sostituiscilo **<amzn-s3-demo-bucket1>** con il nome effettivo del tuo bucket S3 in cui hai archiviato il file cloudwatch-config.json del passaggio precedente. Ciò garantisce che lo script bootstrap possa recuperare il file di configurazione durante l'inizializzazione del cluster.

Questo script di bootstrap consentirà di:

- Attendi il completamento del provisioning del nodo
- Scarica la tua configurazione personalizzata CloudWatch
- Arresta qualsiasi CloudWatch agente in esecuzione
- Riavvia l'agente con la tua configurazione specifica
- Registra lo stato dell'agente per la risoluzione dei problemi

Classificazioni metriche personalizzate per Hadoop, YARN e HBase

Oltre alle CloudWatch metriche predefinite, è possibile migliorare le capacità di monitoraggio configurando metriche personalizzate specifiche dell'applicazione per i componenti del cluster EMR. L'API di configurazione di Amazon EMR offre un modo flessibile per definire esattamente quali parametri desideri raccogliere.

Configurazione di metriche personalizzate

È possibile implementare la raccolta di metriche personalizzate in due modi:

- Durante la creazione di cluster per nuovi cluster
- Come riconfigurazione per i cluster esistenti tramite la console EMR

Creazione di un file di classificazione

Il file di classificazione definisce quali metriche specifiche dei componenti devono essere raccolte dal cluster. Di seguito è riportata una struttura di esempio per la raccolta di metriche Hadoop personalizzate:

```
[
  {
    "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": {}
  }
]
```

Passaggi dell'implementazione

1. Crea un file JSON con le classificazioni metriche desiderate.
2. Personalizza le metriche in base ai tuoi requisiti di monitoraggio.
3. Salva il file e caricalo nel tuo bucket S3.
4. Fai riferimento a questo file quando crei un nuovo cluster o ne riconfiguri uno esistente.

Best practice

- Raccogli solo metriche che forniscono informazioni significative per i tuoi carichi di lavoro.
- Considera l'intervallo di raccolta delle metriche in base alle tue esigenze di monitoraggio.
- AWS Consulta la documentazione per l'elenco completo delle metriche disponibili per ciascun componente.
- Raggruppa le metriche correlate all'interno della stessa classificazione per una migliore organizzazione.

Questo approccio consente di concentrare il monitoraggio sulle metriche più critiche per le applicazioni EMR specifiche, offrendo una visibilità più approfondita sulle prestazioni del cluster.

Implementazione di un cluster EMR con integrazione CloudWatch

Segui questi passaggi per creare un cluster Amazon EMR che invii automaticamente log e parametri personalizzati a: CloudWatch

Passaggio 1: abilitare l'agente CloudWatch

Quando si crea un cluster EMR tramite la console di AWS gestione:

1. Passa alla sezione Applicazioni durante la creazione del cluster.
2. Seleziona le caselle di controllo per le tue applicazioni principali (Hadoop, Spark, ecc.).
3. Scorri per trovare e selezionare l'opzione Amazon CloudWatch Agent.
4. Ciò abilita l'agente sul tuo cluster, il che è essenziale per raccogliere metriche e log avanzati.

L' CloudWatch agente verrà installato su tutti i nodi del cluster, consentendogli di raccogliere i parametri di sistema e delle applicazioni agli intervalli configurati.

Creazione di un cluster e visualizzazione dei pacchetti disponibili.

> **Note**
>
> L' CloudWatch agente è disponibile nella versione EMR 7.0 e successive. L'abilitazione di questo componente è necessaria per la raccolta di metriche personalizzate e l'inoltro dei log descritti in questa guida.

Passaggio 2: aggiungere l'azione Bootstrap per Log Collection

Per configurare l' CloudWatch agente per raccogliere e inoltrare file di registro specifici a CloudWatch:

1. Nella procedura guidata per la creazione del cluster EMR, vai alla sezione Azioni Bootstrap
2. Fai clic su Aggiungi azione bootstrap
3. Seleziona Azione personalizzata dal menu a discesa
4. Fornisci un nome per l'azione di bootstrap (ad esempio, Configure CloudWatch Agent)
5. Nel campo Posizione dello script, inserisci il percorso S3 dello cloudwatch-agent-bootstrap script.sh (ad esempio, s3:///.sh) your-bucket-name cloudwatch-agent-bootstrap
6. Fai clic su Aggiungi per salvare l'azione bootstrap

Questa azione di bootstrap verrà eseguita durante l'avvio del cluster, assicurando che CloudWatchagent sia configurata correttamente con le impostazioni personalizzate per raccogliere e inoltrare i file di registro specificati nel file di configurazione.

L'agente inizierà automaticamente a raccogliere i log una volta effettuato il provisioning dei nodi, fornendo una visibilità quasi in tempo reale delle operazioni del cluster tramite Logs. CloudWatch

Utilizzo delle azioni bootstrap.

Passaggio 3: configurare la raccolta di metriche personalizzate

Per abilitare la raccolta di Hadoop, YARN o HBase metriche personalizzate oltre al set predefinito:

1. Nella procedura guidata per la creazione del cluster EMR, vai alla sezione Configurazioni.
2. Fai clic sul pulsante Modifica configurazioni per espandere le opzioni di configurazione.
3. Seleziona l'opzione Carica JSON da Amazon S3 dal menu a discesa del metodo di configurazione.
4. Inserisci il percorso URI S3 del tuo file di classificazione delle metriche personalizzate (ad esempio, s3://amzn-s3-demo-bucket1/ .json). emr-metrics-classification
5. Fai clic su Carica per analizzare la configurazione.
6. Verifica che la configurazione appaia correttamente nell'interfaccia della console.
7. Fai clic su Salva modifiche per applicare queste configurazioni di metriche al cluster.

Questo passaggio indica all' CloudWatch agente di raccogliere le metriche dei componenti specifiche definite nel file di classificazione. Le metriche verranno raccolte agli intervalli specificati nella configurazione e pubblicate su CloudWatch, dove potranno essere visualizzate e analizzate.

Le metriche personalizzate forniscono informazioni più approfondite sulle caratteristiche prestazionali del cluster, consentendo un monitoraggio e una risoluzione dei problemi più precisi delle applicazioni EMR.

Sostituisci le configurazioni predefinite.

Aggiornamento della configurazione delle metriche per Running Clusters

È possibile modificare le impostazioni di raccolta delle metriche per un cluster EMR esistente senza interrompere le operazioni seguendo questi passaggi:

1. Accedere al cluster EMR attivo nella console di AWS gestione.
2. Seleziona la scheda Configurazioni nella visualizzazione dei dettagli del cluster.
3. Trova la sezione Configurazioni dei gruppi di istanze.
4. Fai clic sul pulsante Riconfigura per modificare le impostazioni.
5. Scegli Load JSON da Amazon S3 o modifica direttamente la configurazione.
6. Inserisci la posizione del file di classificazione delle metriche aggiornato o apporta le modifiche nell'editor.
7. Applica le modifiche per aggiornare il comportamento di raccolta delle metriche.

Questa funzionalità di riconfigurazione consente di ottimizzare l'approccio di monitoraggio man mano che i requisiti del carico di lavoro si evolvono. L' CloudWatch agente si adatterà automaticamente alla nuova configurazione, raccogliendo il set aggiornato di metriche senza richiedere riavvii o tempi di inattività del cluster.

> **Important**
>
> La propagazione delle modifiche alla configurazione può richiedere diversi minuti su tutti i nodi del cluster. Continua a monitorare i CloudWatch dashboard per confermare che le nuove metriche vengano visualizzate come previsto.

Configurazioni dei gruppi di istanze.

Convalida dell'integrazione CloudWatch

Dopo aver completato i passaggi di configurazione, è il momento di verificare che la configurazione di monitoraggio funzioni correttamente:

Fase 1: Implementazione del cluster EMR

1. Controlla tutte le impostazioni di configurazione per verificarne la precisione.
2. Assicurati che le azioni di bootstrap e i file di classificazione siano referenziati correttamente.
3. Fai clic su Crea cluster per avviare il tuo ambiente EMR.
4. Attendi che il cluster raggiunga lo stato Running (in genere 5-15 minuti).

Fase 2: Eseguire le applicazioni di test

Invia diverse applicazioni Spark di prova per generare metriche significative:

- Esegui un semplice job Spark che elabora dati di esempio.
- Esegui un'attività di analisi di lunga durata per osservare l'utilizzo delle risorse.
- Prova diverse configurazioni di applicazioni per confrontare le metriche delle prestazioni.

Una volta completate le applicazioni (o mentre sono in esecuzione):

- Vai alla CloudWatch console.
- Controlla i gruppi di log configurati per i log delle applicazioni.
- Esamina i dashboard delle metriche per osservare le metriche specifiche di CPU, memoria e applicazioni.
- Verifica che le metriche personalizzate definite nel tuo file di classificazione vengano visualizzate in. CloudWatch

Questo processo di convalida conferma che l' CloudWatch integrazione sta acquisendo correttamente sia i log che le metriche, offrendoti una visibilità completa sulle prestazioni e sul comportamento delle applicazioni del cluster EMR.

Accesso ai log EMR nei gruppi di log CloudWatch

Dopo l'esecuzione del cluster EMR e la corretta configurazione dell' CloudWatch agente, i registri dell'applicazione e del sistema saranno disponibili nella sezione Registri. CloudWatch Segui questi passaggi per accedervi e analizzarli:

Visualizzazione dei gruppi di log

1. Accedere alla CloudWatch console in AWS Management Console.
2. Seleziona Log groups dal riquadro di navigazione a sinistra.
3. Cerca i gruppi di log creati dalla tua configurazione, ad esempio:
   - /emr/yarn/resourcemngerper i ResourceManager log YARN.
   - /emr/hdfs/namenodeper i log HDFS NameNode .
   - Eventuali gruppi di log aggiuntivi specificati nel file di configurazione.

Ogni gruppo di log contiene flussi di log organizzati per ID di istanza, che consentono di tracciare i log su nodi specifici del cluster.

Lavorare con i dati di registro

- Cerca nei dati di registro: utilizza CloudWatch Logs Insights per eseguire query strutturate tra i tuoi gruppi di log.
- Crea metriche: estrai le metriche dai modelli di registro per creare metriche personalizzate. CloudWatch
- Imposta avvisi: configura gli allarmi in base a modelli di errore specifici o frequenze di registro.
- Esporta registri: scarica i registri per l'analisi o l'archiviazione offline.

Retention dei log

> **Note**
>
> Per impostazione predefinita, i log vengono conservati per 30 giorni. È possibile modificare la politica di conservazione per ogni gruppo di log per conservare i log per periodi più lunghi, se necessario per scopi di conformità o analisi.

CloudWatch Logs fornisce una posizione centralizzata per tutti i dati di registro EMR, eliminando la necessità di utilizzare SSH nei singoli nodi del cluster per risolvere problemi o analizzare il comportamento delle applicazioni.

Visualizzazione delle metriche personalizzate nel dashboard di monitoraggio EMR

Dopo l'esecuzione del cluster EMR con la configurazione dell' CloudWatch agente e delle metriche personalizzate, puoi monitorare facilmente queste metriche direttamente nella console EMR:

Accesso alle metriche personalizzate

1. Accedi al tuo cluster EMR nella console di AWS gestione.
2. Seleziona la scheda Monitoraggio nella pagina dei dettagli del cluster.
3. Individua il menu a discesa di classificazione delle metriche Filter nella parte superiore dei dashboard di monitoraggio.
4. Utilizza questo filtro per selezionare categorie di metriche specifiche:
   - Scegli HDFS per visualizzare NameNode e DataNode metriche.
   - Seleziona YARN per visualizzare ResourceManager e contenere le metriche.
   - Scegli HBaseper dati HBase prestazionali specifici.
   - Seleziona le classificazioni metriche personalizzate che hai definito.

La dashboard si aggiornerà dinamicamente per visualizzare i grafici relativi alle metriche selezionate, mostrando le tendenze delle prestazioni nel tempo.

Utilizzo delle visualizzazioni metriche

- Modifica gli intervalli di tempo: modifica la finestra temporale per visualizzare le attività recenti o le tendenze storiche.
- Confronta le metriche: visualizza più metriche correlate side-by-side per l'analisi della correlazione.
- Funzioni di zoom: concentrati su periodi di tempo specifici in cui compaiono anomalie o schemi.
- Aggiorna i dati: aggiorna le visualizzazioni con i dati delle metriche più recenti quasi in tempo reale.

Questo approccio di monitoraggio integrato consente di tenere traccia sia delle metriche EMR standard che delle metriche personalizzate in un dashboard unificato, semplificando l'identificazione di problemi di prestazioni, vincoli di risorse o colli di bottiglia delle applicazioni senza uscire dalla console EMR.

Classificazione delle metriche di filtraggio.

Monitora le applicazioni Apache Spark su Amazon EMR con Amazon CloudWatch

Per migliorare l'efficienza di un'applicazione Spark, è essenziale monitorarne le prestazioni e il comportamento. Nel seguente post del blog, un tutorial mostra in qualche step-by-step modo come pubblicare metriche Spark dettagliate da Amazon EMR a. CloudWatch Questo ti dà la possibilità di identificare i colli di bottiglia ottimizzando al contempo l'utilizzo delle risorse: monitora le applicazioni Apache Spark su Amazon EMR con Amazon CloudWatch

Per ulteriori informazioni su CloudWatch, consulta What is Amazon CloudWatch? .

Monitora lo stato delle applicazioni Amazon EMR con l'integrazione CloudWatch

Quando effettui l'integrazione CloudWatch con Amazon EMR, puoi tenere traccia degli stati critici di applicazioni come 2. HiveServer Puoi pubblicare lo stato CloudWatch su parametri personalizzati e configurare avvisi di indisponibilità del servizio.

In particolare, puoi creare uno script per monitorare applicazioni Amazon EMR come YARN ResourceManager e HiveServer 2 su un nodo primario. Consulta Pubblicare e monitorare lo stato di un'applicazione Amazon EMR con CloudWatch integrazione nel Knowledge Center re:POST per dettagli su come configurare questo caso d'uso.

Risoluzione dei problemi relativi ai cluster Amazon EMR

Un cluster EMR funziona in un ecosistema complesso che comprende software open source, codice applicativo personalizzato e. Servizi AWS Quando si verifica un problema con una di queste parti, l'esecuzione del cluster potrebbe non riuscire o impiegare più tempo del previsto per il completamento. Gli argomenti seguenti possono aiutarti a identificare e a risolvere i problemi del cluster.

Argomenti

- Quali strumenti sono disponibili per la risoluzione dei problemi di un cluster Amazon EMR?
- Visualizzazione e riavvio di Amazon EMR e dei processi applicativi (daemon)
- Raccolte di errori comuni in Amazon EMR
- Risolvi un cluster Amazon EMR che non funziona con un codice di errore
- Risolvi i problemi di un cluster Amazon EMR lento
- Risolvi i problemi più comuni durante l'utilizzo di Amazon EMR con Lake Formation AWS

Quando si sviluppa una nuova applicazione Hadoop, si consiglia di abilitare il debug ed elaborare un sottoinsieme piccolo ma rappresentativo dei dati per testare l'applicazione. È inoltre possibile eseguire l'applicazione step-by-step per testare ogni passaggio separatamente. Per ulteriori informazioni, consultare Configurazione del logging e del debug dei cluster Amazon EMR e Fase 5: testare il cluster Amazon EMR passo dopo passo.

Quali strumenti sono disponibili per la risoluzione dei problemi di un cluster Amazon EMR?

Per identificare e correggere gli errori del cluster, puoi utilizzare gli strumenti descritti in questa pagina. Potrebbe essere necessario inizializzare alcuni strumenti all'avvio del cluster. Per impostazione predefinita, sono disponibili altri strumenti per ogni cluster.

Argomenti

- Visualizzazione dei dettagli del cluster EMR
- Visualizzazione dei dettagli degli errori del cluster EMR
- Esecuzione di script e configurazione di processi Amazon EMR

- Visualizzare file di log di
- Monitoraggio delle prestazioni del cluster EMR

Visualizzazione dei dettagli del cluster EMR

È possibile utilizzare l'API AWS Management Console AWS CLI, o EMR per recuperare informazioni dettagliate su un cluster EMR e sull'esecuzione del processo. Per ulteriori informazioni sull'utilizzo di AWS Management Console and AWS CLI, vedere. Visualizza lo stato e i dettagli del cluster Amazon EMR

Riquadro dei dettagli della console di Amazon EMR

Nell'elenco Cluster sulla console di Amazon EMR puoi visualizzare le informazioni di alto livello sullo stato di ogni cluster nel tuo account e Regione AWS. L'elenco mostra tutti i cluster attivi e terminati che hai avviato negli ultimi due mesi. Dall'elenco Clusters (Cluster) è possibile selezionare un Name (Nome) di cluster per visualizzare i dettagli del cluster. Queste informazioni sono organizzate in diverse categorie per facilitarne la navigazione.

La funzionalità Interfacce utente dell'applicazione, disponibile nella pagina dei dettagli del cluster, può essere utile per risolvere i problemi dei cluster. Fornisce lo stato delle applicazioni YARN e, per alcune di esse, come ad esempio le applicazioni Spark, puoi esplorare diversi parametri e sfaccettature, come ad esempio processi, fasi ed esecutori. Per ulteriori informazioni, consulta Visualizza la cronologia delle applicazioni Amazon EMR. Questa funzionalità è disponibile solo per Amazon EMR versione 5.8.0 e successive.

Interfaccia a riga di comando di Amazon EMR

È possibile individuare i dettagli su un cluster utilizzando l'argomento AWS CLI con l'`--describe`argomento.

API Amazon EMR

È possibile individuare i dettagli su un cluster dall'API utilizzando l'azione `DescribeJobFlows`.

Visualizzazione dei dettagli degli errori del cluster EMR

Quando un cluster EMR termina con un errore, `DescribeCluster` e `ListClusters` APIs restituisce un codice di errore e un messaggio di errore. Per alcuni errori del cluster, l'array di dati `ErrorDetail` può aiutarti a risolvere l'errore.

Per un elenco dei codici di errore che includono dati `ErrorDetail`, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR.

> **Note**
>
> Per assicurarci di fornirti le informazioni più recenti e pertinenti, miglioriamo continuamente i nostri messaggi di errore. Non è consigliabile analizzare il testo di `ErrorMessage` perché è soggetto a modifiche.

Esecuzione di script e configurazione di processi Amazon EMR

Come parte del processo di risoluzione dei problemi, potrebbe essere utile eseguire script personalizzati sul cluster o visualizzare e configurare i processi del cluster.

Visualizzare e riavviare i processi di applicazione

Può essere utile visualizzare i processi in esecuzione sul cluster per diagnosticare potenziali problemi. È possibile arrestare e riavviare i processi del cluster effettuando la connessione al nodo principale del cluster. Per ulteriori informazioni, consulta Visualizzazione e riavvio di Amazon EMR e dei processi applicativi (daemon).

Esegui comandi e script senza connessione SSH

Per eseguire un comando o uno script sul cluster come fase, è possibile utilizzare gli strumenti `command-runner.jar` o `script-runner.jar` senza stabilire una connessione SSH al nodo principale. Per ulteriori informazioni, consulta Esegui comandi e script su un cluster Amazon EMR.

Visualizzare file di log di

Amazon EMR e Hadoop generano entrambi file di log quando il cluster è in esecuzione. Puoi accedere a questi file di log da diversi strumenti, in base alla configurazione specificata quando hai avviato il cluster. Per ulteriori informazioni, consulta Configurazione del logging e del debug dei cluster Amazon EMR.

File di log sul nodo master

Ogni cluster pubblica i file di registro nella directory/mnt/var/log/sul nodo master. Questi file di log sono disponibili solo quando il cluster è in esecuzione.

File di log archiviati in Amazon S3

Se avvii il cluster e specifichi un percorso di log di Amazon S3, il cluster copia i file di log archiviati in/mnt/var/log/sul nodo master su Amazon S3 a intervalli di 5 minuti. Questo assicura l'accesso ai file di log anche dopo la chiusura del cluster. Poiché i file vengono archiviati a intervalli di 5 minuti, gli ultimi minuti di un cluster terminato improvvisamente potrebbero non essere disponibili.

Monitoraggio delle prestazioni del cluster EMR

Amazon EMR fornisce diversi strumenti per controllare le prestazioni del cluster.

Interfacce Web Hadoop

Ogni cluster pubblica una serie di interfacce web sul nodo master che contengono informazioni sul cluster. È possibile accedere a queste pagine Web utilizzando un tunnel SSH per collegarle al nodo master. Per ulteriori informazioni, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

CloudWatch metriche

Ogni cluster riporta le metriche a. CloudWatch CloudWatch è un servizio web che tiene traccia delle metriche e che puoi utilizzare per impostare allarmi su tali metriche. Per ulteriori informazioni, consulta Monitoraggio dei parametri di Amazon EMR con CloudWatch.

Visualizzazione e riavvio di Amazon EMR e dei processi applicativi (daemon)

Durante la risoluzione dei problemi relativi a un cluster, è possibile che tu voglia elencare i processi in esecuzione. Puoi anche interrompere o riavviare i processi. Ad esempio, puoi riavviare il processo dopo aver modificato una configurazione o notato un problema con un determinato processo dopo l'analisi di file di log e messaggi di errore.

Esistono due tipi di processi che vengono eseguiti su un cluster: i processi Amazon EMR (ad esempio, instance-controller e Log Pusher) e i processi associati alle applicazioni installate nel cluster (ad esempio, e). hadoop-hdfs-namenode hadoop-yarn-resourcemanager

Per utilizzare i processi direttamente su un cluster, è necessario innanzitutto collegare al nodo principale. Per ulteriori informazioni, consulta Connettiti a un cluster Amazon EMR.

Visualizzazione dei processi in esecuzione

Il metodo utilizzato per visualizzare i processi in esecuzione su un cluster differisce a seconda della versione di Amazon EMR utilizzata.

EMR 5.30.0 and 6.0.0 and later

Example : elenca tutti i processi in esecuzione

Nell'esempio seguente viene utilizzato `systemctl` e specifica `--type` per visualizzare tutti i processi.

```
systemctl --type=service
```

Example : elenca processi specifici

Nell'esempio seguente sono elencati tutti i processi con nomi che contengono `hadoop`.

```
systemctl --type=service | grep -i hadoop
```

Output di esempio:

```
 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 : consulta un report dettagliato sullo stato per un processo specifico

Nell'esempio seguente viene visualizzato un report dettagliato sullo stato del servizio `hadoop-hdfs-namenode`.

```
sudo systemctl status hadoop-hdfs-namenode
```

Output di esempio:

```
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 : elenca tutti i processi in esecuzione

Nell'esempio seguente sono elencati tutti i processi in esecuzione.

```
initctl list
```

EMR 2.x - 3.x

Example : elenca tutti i processi in esecuzione

Nell'esempio seguente sono elencati tutti i processi in esecuzione.

```
ls /etc/init.d/
```

Arresto e riavvio di processi

Dopo aver determinato quali processi sono in esecuzione, puoi interromperli e riavviarli se necessario.

EMR 5.30.0 and 6.0.0 and later

Example : arresta un processo

L'esempio seguente arresta il processo `hadoop-hdfs-namenode`.

```
sudo systemctl stop hadoop-hdfs-namenode
```

Puoi eseguire una query sul `status` per verificare che il processo sia stato arrestato.

```
sudo systemctl status hadoop-hdfs-namenode
```

Output di esempio:

```
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 : avvia un processo

Nell'esempio seguente viene avviato il processo `hadoop-hdfs-namenode`.

```
sudo systemctl start hadoop-hdfs-namenode
```

È possibile eseguire una query sullo stato per verificare che il processo sia in esecuzione.

```
sudo systemctl status hadoop-hdfs-namenode
```

Output di esempio:

```
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 : arresta un processo in esecuzione

L'esempio seguente arresta il servizio `hadoop-hdfs-namenode`.

```
sudo stop hadoop-hdfs-namenode
```

Example : riavvia un processo arrestato

L'esempio seguente riavvia il servizio `hadoop-hdfs-namenode`. È necessario utilizzare il comando `start` e non `restart`.

```
sudo start hadoop-hdfs-namenode
```

Example : verifica lo stato del processo

Quanto segue recupera lo stato di `hadoop-hdfs-namenode`. Puoi utilizzare il comando `status` per verificare che il processo sia stato arrestato o avviato.

```
sudo status hadoop-hdfs-namenode
```

EMR 2.x - 3.x

Example : arresta un processo di applicazione

L'esempio seguente arresta il servizio `hadoop-hdfs-namenode`, associato alla versione di Amazon EMR installata sul cluster.

```
sudo /etc/init.d/hadoop-hdfs-namenode stop
```

Example : riavvia un processo di applicazione

Il seguente comando di esempio riavvia il processo `hadoop-hdfs-namenode`:

```
sudo /etc/init.d/hadoop-hdfs-namenode start
```

Example : arresta un processo Amazon EMR

L'esempio seguente arresta un processo, come il controller di istanza, che non è associato alla versione di Amazon EMR sul cluster.

```
sudo /sbin/stop instance-controller
```

Example : riavvia un processo Amazon EMR

L'esempio seguente riavvia un processo, come il controller di istanza, che non è associato alla versione di Amazon EMR sul cluster.

```
sudo /sbin/start instance-controller
```

> **Note**
>
> I comandi `/sbin/start, stop` e `restart` sono symlinks a `/sbin/intictl`. Per ulteriori informazioni su `initctl`, consulta la pagina initctl man digitando `man initctl` al prompt dei comandi.

Raccolte di errori comuni in Amazon EMR

A volte, l'elaborazione dei dati da parte dei cluster non riesce oppure è lenta. Le seguenti sezioni elencano i problemi più comuni relativi ai cluster. Gli errori includono errori di bootstrap ed errori di convalida, con suggerimenti su come risolverli.

Argomenti

- Codici di errore con ErrorDetail informazioni in Amazon EMR
- Errori di risorse durante le operazioni del cluster Amazon EMR
- Errori di input e output del cluster durante le operazioni di Amazon EMR
- Errori di autorizzazione durante le operazioni del cluster Amazon EMR
- Errori del cluster Hive
- Errori VPC durante le operazioni del cluster Amazon EMR
- Errori di streaming del cluster Amazon EMR
- Amazon EMR: errori di cluster JAR personalizzati

- Errori di Amazon EMR AWS GovCloud (Stati Uniti occidentali)
- Ricerca di un cluster mancante

Codici di errore con ErrorDetail informazioni in Amazon EMR

Quando un cluster EMR termina con un errore, `DescribeCluster` e `ListClusters` APIs restituisce un codice di errore e un messaggio di errore. Per alcuni errori del cluster, l'array di dati `ErrorDetail` può aiutarti a risolvere l'errore.

Gli errori che includono un array `ErrorDetail` forniscono i seguenti dettagli:

`ErrorCode`

Un codice di errore univoco che puoi utilizzare per l'accesso programmatico.

`ErrorData`

Un elenco di identificatori in coppie chiave-valore che puoi utilizzare per la programmazione o la ricerca manuale. Per le descrizioni dei valori `ErrorData` inclusi in un codice di errore, consulta la pagina di risoluzione dei problemi relativa al codice di errore.

`ErrorMessage`

Descrizione dell'errore con un collegamento a ulteriori informazioni nella documentazione di Amazon EMR.

> **Note**
>
> Non è consigliabile analizzare il testo di `ErrorMessage` perché è soggetto a modifiche.

Codici di errore per categoria

- Codici di errore di errore di bootstrap in Amazon EMR
- Codici di errore interni per Amazon EMR
- Codici di errore di convalida non riuscita in Amazon EMR

Codici di errore di errore di bootstrap in Amazon EMR

Le seguenti sezioni forniscono informazioni sulla risoluzione dei problemi relativi ai codici di errore per gli errori di bootstrap.

Argomenti

- 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

Panoramica

Quando un cluster termina con un errore `BOOTSTRAP_FAILURE_PRIMARY_WITH_NON_ZERO_CODE`, un'azione di bootstrap non è riuscita nell'istanza primaria. Per ulteriori informazioni sulle operazioni di bootstrap, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR.

Risoluzione

Per risolvere questo errore, esamina i dettagli restituiti nell'errore dell'API, modifica lo script dell'operazione di bootstrap e crea un nuovo cluster con l'operazione di bootstrap aggiornata.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID dell'istanza primaria in cui l'azione di bootstrap non è riuscita.

`bootstrap-action`

Il numero ordinale dell'operazione di bootstrap che non è riuscita. Uno script con un valore `bootstrap-action` di `1` è la prima azione di bootstrap da eseguire sull'istanza.

`return-code`

Il codice restituito per l'operazione di bootstrap non riuscita.

`amazon-s3-path`

La posizione Amazon S3 dell'azione di bootstrap non riuscita.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere la causa principale dell'errore dell'operazione di bootstrap. Quindi avvia un nuovo cluster.

1. Esamina i file di log dell'operazione di bootstrap in Amazon S3 per identificare la causa principale dell'errore. Per ulteriori informazioni su come visualizzare i log di Amazon EMR, consulta Visualizza i file di log di Amazon EMR.
2. Se hai attivato i log del cluster quando hai creato l'istanza, consulta il log `stdout` per ulteriori informazioni. Puoi trovare il log `stdout` per l'operazione di bootstrap in questa posizione Amazon S3:

   ```
   s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-actions/Failed_Bootstrap_Action_Number/stdout.gz
   ```

   Per ulteriori informazioni sui log del cluster, consulta Configurazione del logging e del debug dei cluster Amazon EMR.
3. Per determinare l'errore dell'operazione di bootstrap, esamina le eccezioni nei log `stdout` e il valore in `return-code` in `ErrorData`.
4. Usa i risultati del passaggio precedente per rivedere l'operazione di bootstrap in modo da evitare le eccezioni o gestirle correttamente quando si verificano.
5. Avvia un nuovo cluster con l'operazione di bootstrap aggiornata.

BOOTSTRAP_FAILURE_BA_DOWNLOAD_FAILED_PRIMARY

Panoramica

Un cluster termina con l'errore `BOOTSTRAP_FAILURE_BA_DOWNLOAD_FAILED_PRIMARY` quando l'istanza primaria non è in grado di scaricare lo script di un'operazione di bootstrap dalla posizione Amazon S3 specificata. Le cause potenziali includono:

- Il file script dell'operazione di bootstrap non si trova nella posizione Amazon S3 specificata.
- Il ruolo di servizio per EC2 le istanze Amazon sul cluster (chiamato anche profilo di EC2 istanza per Amazon EMR) non dispone delle autorizzazioni per accedere al bucket Amazon S3 in cui risiede lo script di azione bootstrap. Per ulteriori informazioni sui ruoli di servizio, consulta Ruolo di servizio per le istanze del cluster (profilo EC2 dell'istanza) EC2.

Per ulteriori informazioni sulle operazioni di bootstrap, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR.

Risoluzione

Per risolvere questo errore, assicurati che l'istanza primaria disponga dell'accesso appropriato allo script dell'operazione di bootstrap.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID dell'istanza primaria in cui l'azione di bootstrap non è riuscita.

`bootstrap-action`

Il numero ordinale dell'operazione di bootstrap che non è riuscita. Uno script con un valore `bootstrap-action` di `1` è la prima azione di bootstrap da eseguire sull'istanza.

`amazon-s3-path`

La posizione Amazon S3 dell'azione di bootstrap non riuscita.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere la causa principale dell'errore dell'operazione di bootstrap. Quindi avvia un nuovo cluster.

Fasi per la risoluzione dei problemi

1. Utilizza il valore `amazon-s3-path` dell'array `ErrorData` per trovare lo script dell'operazione di bootstrap pertinente in Amazon S3.
2. Se hai attivato i log del cluster quando hai creato l'istanza, consulta il log `stdout` per ulteriori informazioni. Puoi trovare il log `stdout` per l'operazione di bootstrap in questa posizione Amazon S3:

   ```
   s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-actions/Failed_Bootstrap_Action_Number/stdout.gz
   ```

   Per ulteriori informazioni sui log del cluster, consulta Configurazione del logging e del debug dei cluster Amazon EMR.
3. Per determinare l'errore dell'operazione di bootstrap, esamina le eccezioni nei log `stdout` e il valore in `return-code` in `ErrorData`.
4. Usa i risultati del passaggio precedente per rivedere l'operazione di bootstrap in modo da evitare le eccezioni o gestirle correttamente quando si verificano.
5. Avvia un nuovo cluster con l'operazione di bootstrap aggiornata.

BOOTSTRAP_FAILURE_FILE_NOT_FOUND_PRIMARY

Panoramica

L'errore `BOOTSTRAP_FAILURE_FILE_NOT_FOUND_PRIMARY` indica che l'istanza primaria non riesce a trovare lo script dell'operazione di bootstrap che l'istanza ha appena scaricato dal bucket Amazon S3 specificato.

Risoluzione

Per risolvere questo errore, verifica che l'istanza primaria disponga dell'accesso appropriato allo script dell'operazione di bootstrap.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID dell'istanza primaria in cui l'azione di bootstrap non è riuscita.

`bootstrap-action`

Il numero ordinale dell'operazione di bootstrap che non è riuscita. Uno script con un valore `bootstrap-action` di `1` è la prima azione di bootstrap da eseguire sull'istanza.

`amazon-s3-path`

La posizione Amazon S3 dell'azione di bootstrap non riuscita.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere la causa principale dell'errore dell'operazione di bootstrap. Quindi avvia un nuovo cluster.

1. Per trovare lo script dell'operazione di bootstrap pertinente in Amazon S3, utilizza il valore `amazon-s3-path` dell'array `ErrorData`.
2. Esamina i file di log dell'operazione di bootstrap in Amazon S3 per identificare la causa principale dell'errore. Per ulteriori informazioni su come visualizzare i log di Amazon EMR, consulta Visualizza i file di log di Amazon EMR.

   > **Note**
   >
   > Se non hai attivato i log per il tuo cluster, devi creare un nuovo cluster con le stesse configurazioni e operazioni di bootstrap. Per garantire che i log del cluster siano stati attivati, consulta Configurazione del logging e del debug dei cluster Amazon EMR.

3. Esamina il log `stdout` per le operazioni di bootstrap e conferma che non vi siano processi personalizzati che eliminano i file nella cartella `/emr/instance-controller/lib/bootstrap-actions` sulle istanze primarie. Puoi trovare il log `stdout` per l'operazione di bootstrap in questa posizione Amazon S3:

   ```
   s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-actions/Failed_Bootstrap_Action_Number/stdout.gz
   ```

4. Avvia un nuovo cluster con l'operazione di bootstrap aggiornata.

BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_PRIMARY

Panoramica

L'`BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_PRIMARY`errore indica che l'istanza principale non dispone di spazio su disco sufficiente per l'installazione del software necessario.

Risoluzione

Per risolvere questo errore, verifica che l'istanza principale disponga di spazio su disco sufficiente sul volume principale.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID dell'istanza principale con spazio su disco insufficiente.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

1. Consulta le best practice per il volume del dispositivo root EBS del tuo cluster. Consulta Personalizzazione del volume del dispositivo root Amazon EBS nella Guida alla gestione di Amazon EMR.
2. Avvia un nuovo cluster con un volume del dispositivo root EBS più grande.

BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_WORKER

Panoramica

L'`BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_WORKER`errore indica che una o più istanze di lavoro non dispongono di spazio su disco sufficiente per l'installazione del software necessario.

Risoluzione

Per risolvere questo errore, verifica che le istanze di lavoro abbiano spazio su disco sufficiente sul volume principale.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`worker-instance-ids`

Le istanze IDs di lavoro con spazio su disco insufficiente.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

1. Consulta le best practice per il volume del dispositivo root EBS del tuo cluster. Consulta Personalizzazione del volume del dispositivo root Amazon EBS nella Guida alla gestione di Amazon EMR.
2. Avvia un nuovo cluster con un volume del dispositivo root EBS più grande.

BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_PRIMARY

Panoramica

L'`BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_PRIMARY`errore indica che l'istanza principale non è in grado di stabilire una connessione all'Hive Metastore esterno configurato.

Risoluzione

Per risolvere questo errore, conferma che il tuo Hive Metastore esterno sia configurato correttamente e che l'istanza principale sia autorizzata a connettersi ad esso.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID dell'istanza principale non è in grado di stabilire una connessione all'Hive Metastore esterno configurato.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

1. Consulta le migliori pratiche per configurare un metastore esterno per Hive. Vedi Configurazione di un metastore esterno per Hive.
2. Avvia un nuovo cluster con la configurazione aggiornata.

BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_WORKER

Panoramica

L'`BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_WORKER`errore indica che una o più istanze di lavoro non sono in grado di stabilire una connessione all'Hive Metastore esterno configurato.

Risoluzione

Per risolvere questo errore, verifica che l'Hive Metastore esterno sia configurato correttamente e che le istanze di lavoro siano autorizzate a connettersi ad esso.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`worker-instance-ids`

Le istanze IDs di lavoro che non sono in grado di stabilire una connessione all'Hive Metastore esterno configurato.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

1. Consulta le best practice per configurare un metastore esterno per Hive. Vedi Configurazione di un metastore esterno per Hive.
2. Avvia un nuovo cluster con la configurazione aggiornata.

Codici di errore interni per Amazon EMR

Le seguenti sezioni forniscono informazioni sulla risoluzione dei problemi relativi ai codici di errore interni, inclusi i codici relativi alla capacità insufficiente o inesistente.

Argomenti

- ERRORE_INTERNO_ EC2 _INSUFFICIENT_CAPACITY_AZ
- INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY
- INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY

ERRORE_INTERNO_ EC2 _INSUFFICIENT_CAPACITY_AZ

Panoramica

Un cluster termina con un `INTERNAL_ERROR_EC2_INSUFFICIENT_CAPACITY_AZ` errore quando la zona di disponibilità selezionata non ha una capacità sufficiente per soddisfare la richiesta del tipo di EC2 istanza Amazon. La sottorete selezionata per un cluster determina la zona di disponibilità. Per ulteriori informazioni sulle sottoreti per Amazon EMR, consulta Configurazione della rete in un VPC per Amazon EMR.

Risoluzione

Per risolvere questo errore, modifica le configurazioni del tipo di istanza e crea un nuovo cluster con la richiesta aggiornata.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`instance-type`

Il tipo di istanza che ha esaurito la capacità.

`availability-zone`

La zona di disponibilità in cui si risolve la sottorete.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere la causa principale dell'errore di configurazione del cluster:

- Rivedi le best practice per la configurazione dei cluster. Consulta Configurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot nella Guida alla gestione di Amazon EMR.
- Risolvi i problemi di avvio e rivedi la configurazione. Consulta Risolvere i problemi di avvio delle istanze nella Amazon EC2 User Guide.
- Avvia un nuovo cluster con la configurazione aggiornata.

INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY

Panoramica

Un cluster termina con un errore `INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY` quando Amazon EMR non è in grado di soddisfare la richiesta di istanza Spot per il nodo primario perché le istanze non sono disponibili al prezzo Spot massimo o inferiore. Per ulteriori informazioni, consulta le istanze Spot nella Amazon EC2 User Guide.

Risoluzione

Per risolvere questo errore, specifica i tipi di istanza per il cluster che rientrano nel tuo target di prezzo o aumenta il limite di prezzo per lo stesso tipo di istanza.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID per l'istanza primaria del cluster che ha avuto esito negativo.

`instance-type`

Il tipo di istanza che ha esaurito la capacità.

`availability-zone`

La zona di disponibilità in cui risiede la sottorete.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per risolvere i problemi relativi alla tua strategia di configurazione del cluster, quindi avvia un nuovo cluster:

1. Consulta le best practice per le istanze Amazon EC2 Spot e rivedi la tua strategia di configurazione del cluster. Per ulteriori informazioni, consulta la sezione Best practice for EC2 Spot nella Amazon EC2 User Guide eConfigurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot.
2. Modifica le configurazioni del tipo di istanza o la zona di disponibilità e crea un nuovo cluster con la richiesta aggiornata.
3. Se il problema persiste, utilizza la capacità On-Demand per l'istanza primaria.

INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY

Panoramica

Un cluster termina con un errore `INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY` quando la capacità non è sufficiente per soddisfare una richiesta di istanza Spot per il nodo primario. Per ulteriori informazioni, consulta le istanze Spot nella Amazon EC2 User Guide.

Risoluzione

Per risolvere questo errore, specifica i tipi di istanza per il cluster che rientrano nel tuo target di prezzo o aumenta il limite di prezzo per lo stesso tipo di istanza.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`primary-instance-id`

L'ID per l'istanza primaria del cluster che ha avuto esito negativo.

`instance-type`

Il tipo di istanza che ha esaurito la capacità.

`availability-zone`

La zona di disponibilità in cui si risolve la sottorete.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per risolvere i problemi relativi alla tua strategia di configurazione del cluster, quindi avvia un nuovo cluster:

1. Consulta le best practice per le istanze Amazon EC2 Spot e rivedi la tua strategia di configurazione del cluster. Per ulteriori informazioni, consulta la sezione Best practice for EC2 Spot nella Amazon EC2 User Guide eConfigurazione dei tipi di istanze del cluster Amazon EMR e delle best practice per le istanze Spot.
2. Modifica le configurazioni dei tipi di istanza e crea un nuovo cluster con la richiesta aggiornata.
3. Se il problema persiste, utilizza la capacità On-Demand per l'istanza primaria.

Codici di errore di convalida non riuscita in Amazon EMR

Le seguenti sezioni forniscono informazioni sulla risoluzione dei problemi relativi ai codici di errore per gli errori di convalida.

Argomenti

- 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

Panoramica

Quando il cluster e le sottoreti a cui fai riferimento per il cluster appartengono a diversi cloud privati virtuali (VPCs), il cluster termina con un errore. `VALIDATION_ERROR_SUBNET_NOT_FROM_ONE_VPC` Puoi avviare i cluster con Amazon EMR con la configurazione dei parchi istanze tra le sottoreti in un VPC. Per ulteriori informazioni sui parchi istanze, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR nella Guida alla gestione di Amazon EMR.

Risoluzione

Per risolvere questo errore, utilizza sottoreti che appartengono allo stesso VPC del cluster.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`vpc`

Per ogni coppia subnet:VPC, l'ID del VPC a cui appartiene la sottorete.

`subnet`

Per ogni coppia subnet:VPC, l'ID della sottorete.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere l'errore:

1. Esamina le IDs sottoreti elencate nell'`ErrorData`array e verifica che appartengano al VPC in cui desideri avviare il cluster EMR.

2. Modifica le configurazioni delle sottoreti. Puoi utilizzare uno dei metodi seguenti per trovare tutte le sottoreti pubbliche e private disponibili in un VPC.

   - Passa alla console di Amazon VPC. Scegli Subnet ed elenca tutte le sottoreti che risiedono all'interno del tuo cluster. Regione AWS Per trovare solo sottoreti pubbliche o private, applica il filtro Assegna automaticamente gli indirizzi pubblici. IPv4 Per trovare e selezionare le sottoreti nel VPC utilizzato dal cluster, utilizza l'opzione Filtra per VPC. Per ulteriori informazioni su come creare sottoreti, consulta Creazione di una sottorete nella Guida per l'utente di Amazon Virtual Private Cloud.
   - Utilizza il AWS CLI per trovare tutte le sottoreti pubbliche e private disponibili nel VPC utilizzato dal cluster. Per ulteriori informazioni, consulta l'API describe-subnets. Per creare nuove sottoreti in un VPC, consulta l'API create-subnet.
3. Avvia un nuovo cluster con sottoreti dallo stesso VPC del cluster.

VALIDATION_ERROR_SECURITY_GROUP_NOT_FROM_ONE_VPC

Panoramica

Quando il cluster e i gruppi di sicurezza assegnati al cluster appartengono a diversi cloud privati virtuali (VPCs), il cluster termina con un errore. `VALIDATION_ERROR_SECURITY_GROUP_NOT_FROM_ONE_VPC` Per ulteriori informazioni sui gruppi di sicurezza, consulta Specifica dei gruppi di sicurezza gestiti da Amazon EMR e aggiuntivi e Controlla il traffico di rete con gruppi di sicurezza per il tuo cluster Amazon EMR.

Risoluzione

Per risolvere questo errore, utilizza gruppi di sicurezza che appartengono allo stesso VPC del cluster.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`vpc`

Per ogni coppia security-group:VPC, l'ID del VPC a cui appartiene il gruppo di sicurezza.

`security-group`

Per ogni coppia security-group:VPC, l'ID del gruppo di sicurezza.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere l'errore:

1. Esamina i gruppi IDs di sicurezza elencati nell'`ErrorData`array e conferma che appartengano al VPC in cui desideri avviare il cluster EMR.
2. Passa alla console di Amazon VPC. Scegli Gruppi di sicurezza per elencare tutti i gruppi di sicurezza all'interno della regione selezionata. Trova i gruppi di sicurezza provenienti dallo stesso VPC del cluster, quindi modifica la configurazione del gruppo di sicurezza.
3. Avvia un nuovo cluster con gruppi di sicurezza provenienti dallo stesso VPC del cluster.

VALIDATION_ERROR_INVALID_SSH_KEY_NAME

Panoramica

Un cluster termina con un `VALIDATION_ERROR_INVALID_SSH_KEY_NAME` errore quando utilizzi una coppia di EC2 chiavi Amazon non valida per SSH nell'istanza primaria. Il nome della coppia di chiavi potrebbe non essere corretto o la coppia di chiavi potrebbe non esistere nella richiesta Regione AWS. Per ulteriori informazioni sulle coppie di chiavi, consulta le coppie di EC2 chiavi Amazon e le istanze Linux nella Amazon EC2 User Guide.

Risoluzione

Per risolvere questo errore, crea un nuovo cluster con un nome valido per la coppia di chiavi SSH.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`ssh-key`

Il nome della coppia di chiavi SSH che hai fornito quando hai creato il cluster.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere l'errore:

1. Controlla il tuo *keypair* file.pem e conferma che corrisponda al nome della chiave SSH che vedi nella console Amazon EMR.
2. Accedi alla EC2 console Amazon. Verifica che il nome della chiave SSH che hai usato sia disponibile nella Regione AWS chiave utilizzata dal cluster. Puoi trovare il tuo ID Regione AWS accanto al tuo account nella parte superiore di AWS Management Console.
3. Avvia un nuovo cluster con un nome valido per la chiave SSH.

VALIDATION_ERROR_INSTANCE_TYPE_NOT_SUPPORTED

Panoramica

Un cluster termina con un errore `VALIDATION_ERROR_INSTANCE_TYPE_NOT_SUPPORTED` quando la Regione AWS e le zone di disponibilità del cluster non supportano il tipo di istanza specificato per uno o più gruppi di istanze. Amazon EMR è in grado di supportare un tipo di istanza in una zona di disponibilità all'interno di una Regione ma non in un'altra. La sottorete selezionata per un cluster determina la zona di disponibilità all'interno della regione. Per un elenco dei tipi di istanza e delle regioni supportati da Amazon EMR, consulta Tipi di istanze supportati con Amazon EMR.

Risoluzione

Per risolvere questo errore, specifica i tipi di istanza per il tuo cluster supportati da Amazon EMR nella regione e nella zona di disponibilità in cui richiedi il cluster.

Per risolvere il problema relativo al cluster EMR guasto, fare riferimento alle `ErrorDetail` informazioni restituite da e. `DescribeCluster` `ListClusters` APIs Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR. L'array `ErrorData` all'interno di `ErrorDetail` restituisce le seguenti informazioni per questo codice di errore:

`instance-types`

L'elenco dei tipi di istanza non supportati.

`availability-zones`

L'elenco delle zone di disponibilità in cui viene risolta la sottorete.

`public-doc`

L'URL pubblico della documentazione per il codice di errore.

Passaggi da completare

Esegui i passaggi seguenti per identificare e correggere l'errore:

1. Utilizzare il AWS CLI per recuperare i tipi di istanza disponibili in una zona di disponibilità. A tale scopo, è possibile utilizzare il `ec2 describe-instance-type-offerings` comando per filtrare i tipi di istanze disponibili in base alla posizione (Regione AWS o alla zona di disponibilità). Ad esempio, il comando seguente restituisce i tipi di istanza offerti nella zona di disponibilità specificata, *`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
   ```

   Per ulteriori informazioni su come scoprire i tipi di istanza disponibili, consulta Trova un tipo di EC2 istanza Amazon.
2. Dopo aver determinato i tipi di istanza disponibili nella stessa regione e zona di disponibilità del cluster, scegli una delle seguenti risoluzioni per continuare:

   a. Crea un nuovo cluster e scegli una sottorete per il cluster in una zona di disponibilità in cui il tipo di istanza è disponibile e supportato da Amazon EMR.

   b. Crea un nuovo cluster nella stessa regione e nella stessa EC2 sottorete Amazon del cluster in cui si è verificato l'errore, ma con un tipo di istanza supportato in quella posizione da Amazon EMR.

Per un elenco dei tipi di istanza e delle regioni supportati da Amazon EMR, consulta Tipi di istanze supportati con Amazon EMR. Per confrontare le funzionalità dei tipi di istanze, consulta la sezione Tipi di EC2 istanze Amazon.

Errori di risorse durante le operazioni del cluster Amazon EMR

I seguenti errori sono comunemente causati da risorse vincolate sul cluster. La guida descrive ogni errore e fornisce assistenza per la risoluzione dei problemi.

Argomenti

- Il cluster Amazon EMR termina con NO_SLAVE_LEFT e i nodi principali FAILED_BY_MASTER
- Errore del cluster Amazon EMR: impossibile replicare il blocco, è riuscito a replicare solo su zero nodi.
- Errore del cluster Amazon EMR: EC2 QUOTA SUPERATA
- Errore del cluster Amazon EMR: troppi errori di fetch
- Errore del cluster Amazon EMR: il file può essere replicato solo su 0 nodi anziché su 1
- Errore del cluster Amazon EMR: nodi con elenco negato
- Errori di limitazione con un cluster Amazon EMR
- Errore del cluster Amazon EMR: tipo di istanza non supportato
- Errore del cluster Amazon EMR: capacità EC2 insufficiente
- Errore del cluster Amazon EMR: errore del fattore di replica HDFS
- Errore del cluster Amazon EMR: errore di spazio HDFS insufficiente

Il cluster Amazon EMR termina con NO_SLAVE_LEFT e i nodi principali FAILED_BY_MASTER

In genere, ciò accade perché la protezione da cessazione è disabilitata e tutti i nodi principali superano la capacità di storage su disco come specificato da una soglia di utilizzo massimo nella classificazione di configurazione `yarn-site`, che corrisponde al file `yarn-site.xml`. Per impostazione predefinita, questo valore è 90%. Quando l'utilizzo del disco per un nodo principale supera la soglia di utilizzo, il servizio sanitario YARN segnala il nodo come. NodeManager `UNHEALTHY` Mentre è in questo stato, Amazon EMR elenca come negato il nodo e non assegna i container YARN a tale nodo. Se il nodo non è integro per 45 minuti, Amazon EMR contrassegna l'istanza EC2 Amazon associata per la terminazione come. `FAILED_BY_MASTER` Quando tutte le EC2 istanze Amazon associate ai nodi principali sono contrassegnate per la chiusura, il cluster termina con lo stato `NO_SLAVE_LEFT` perché non ci sono risorse per eseguire i job.

Il superamento dell'utilizzo del disco su un nodo principali potrebbe causare una reazione a catena. Se un singolo nodo supera la soglia di utilizzo del disco a causa di HDFS, è probabile che anche altri nodi siano vicini alla soglia. Il primo nodo supera la soglia di utilizzo del disco, quindi Amazon EMR lo elenca come negato. Ciò aumenta il carico di utilizzo del disco per i nodi rimanenti perché questi iniziano a replicare tra loro i dati HDFS che hanno perso sul nodo elencato come negato. Di conseguenza, ogni nodo diventa `UNHEALTHY` nello stesso modo e infine il cluster viene terminato.

Best practice e raccomandazioni

Configurazione di hardware cluster con archiviazione adeguata

Quando crei un cluster, accertati che vi sia un numero sufficiente di nodi principali e che ogni nodo disponga di volumi adeguati di storage EBS e instance store per HDFS. Per ulteriori informazioni, consulta Calcolo della capacità HDFS richiesta di un cluster. Puoi inoltre aggiungere istanze principali ai gruppi di istanze esistenti manualmente o utilizzando il dimensionamento automatico. Le nuove istanze hanno la stessa configurazione di storage delle altre istanze nel gruppo di istanze. Per ulteriori informazioni, consulta Usa la scalabilità dei cluster Amazon EMR per adattarti ai carichi di lavoro in continua evoluzione.

Abilitare la protezione da cessazione

Abilita la protezione da cessazione. In questo modo, se un nodo principale è in elenco negato, puoi connetterti all' EC2 istanza Amazon associata utilizzando SSH per risolvere i problemi e ripristinare i dati. Se abiliti la protezione dalla terminazione, tieni presente che Amazon EMR non sostituisce l'istanza EC2 Amazon con una nuova istanza. Per ulteriori informazioni, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

Crea un allarme per la metrica Nodes MRUnhealthy CloudWatch

Questo parametro indica il numero di nodi con stato `UNHEALTHY`. È equivalente al parametro YARN `mapred.resourcemanager.NoOfUnhealthyNodes`. Puoi impostare una notifica per questo allarme in modo che ti vengano segnalati i nodi non integri prima che venga raggiunto il timeout di 45 minuti. Per ulteriori informazioni, consulta Monitoraggio dei parametri di Amazon EMR con CloudWatch.

Impostazioni Tweak mediante yarn-site

Le impostazioni riportate di seguito possono essere regolate in base ai requisiti dell'applicazione. Ad esempio, potresti voler aumentare la soglia di utilizzo del disco in base alla quale un nodo segnala lo stato `UNHEALTHY` aumentando il valore di `yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage`.

Puoi impostare questi valori quando crei un cluster utilizzando la classificazione di configurazione `yarn-site`. Per ulteriori informazioni, consulta Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR. Puoi anche connetterti alle EC2 istanze Amazon associate ai nodi principali tramite SSH e quindi aggiungere i valori `/etc/hadoop/conf.empty/yarn-site.xml` utilizzando

un editor di testo. Dopo aver apportato la modifica, è necessario riavviare il sistema hadoop-yarn-nodemanager come illustrato di seguito.

> **Important**
>
> Quando si riavvia il NodeManager servizio, i contenitori YARN attivi vengono eliminati a meno che non `yarn.nodemanager.recovery.enabled` sia impostato l'`true`utilizzo della classificazione di `yarn-site` configurazione al momento della creazione del cluster. Devi inoltre specificare la directory in cui memorizzare lo stato del container utilizzando la proprietà `yarn.nodemanager.recovery.dir`.

```
sudo /sbin/stop hadoop-yarn-nodemanager
sudo /sbin/start hadoop-yarn-nodemanager
```

Per ulteriori informazioni sulle proprietà `yarn-site` correnti e i valori predefiniti, consulta la pagina delle impostazioni predefinite di YARN nella documentazione di Apache Hadoop.

Proprietà	Valore predefinito	Descrizione
yarn.nodemanager. disk-health-checker.intervallo-ms	120000	La frequenza (in secondi) con cui viene eseguito il controllo dello stato del disco.
yarn.nodemanager. disk-health-checker. min-healthy-disks	0.25	La frazione minima del numero di dischi che devono essere integri per NodeManager poter avviare nuovi contenitori. Ciò corrisponde sia a yarn.nodemanager.local-dirs (per impostazione predefinita `/mnt/yarn` in Amazon EMR) che a yarn.nodemanager.log-dirs (per impostazione predefinita `/var/log/hadoop-yarn/containers` , del quale

Proprietà	Valore predefinito	Descrizione
		è eseguito il symlink a `mnt/var/log/hadoop-yarn/containers` in Amazon EMR).
`yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage`	90,0	La percentuale massima di utilizzo dello spazio su disco consentita dopo che un disco viene contrassegnato come difettoso. I valori possono andare da 0,0 a 100,0. Se il valore è maggiore o uguale a 100, NodeManager verifica la presenza di un disco completo. Ciò è valido per `yarn-nodemanager.local-dirs` e `yarn.nodemanager.log-dirs`.
`yarn.nodemanager.disk-health-checker.min-free-space-per-disk-mb`	0	Lo spazio minimo che deve essere disponibile su un disco affinché possa essere utilizzato. Ciò è valido per `yarn-nodemanager.local-dirs` e `yarn.nodemanager.log-dirs`.

Errore del cluster Amazon EMR: impossibile replicare il blocco, è riuscito a replicare solo su zero nodi.

Generalmente, l'errore "Cannot replicate block, only managed to replicate to zero nodes (Impossibile replicare il blocco, gestito solo per la replica su zero nodi)." si verifica quando un cluster non dispone di sufficiente spazio di archiviazione HDFS. Questo errore si verifica quando la quantità di dati generata nel cluster è superiore alla capacità di archiviazione di HDFS. Questo errore viene

visualizzato solo durante l'esecuzione del cluster in quanto quando termina il processo rilascia lo spazio HDFS che stava utilizzando.

La quantità di spazio HDFS disponibile per un cluster dipende dal numero e dal tipo di EC2 istanze Amazon utilizzate come nodi principali. I nodi di task non vengono utilizzati per lo storage HDFS. Tutto lo spazio su disco di ogni EC2 istanza Amazon, inclusi i volumi di storage EBS collegati, è disponibile per HDFS. Per ulteriori informazioni sulla quantità di storage locale per ogni tipo di EC2 istanza, consulta Tipi e famiglie di istanze nella Amazon EC2 User Guide.

L'altro fattore che può influenzare la quantità di spazio HDFS disponibile è il fattore di replica, ovvero il numero di copie di ciascun blocco di dati che viene archiviato in HDFS per la ridondanza. Il fattore di replica aumenta con il numero di nodi nel cluster: sono disponibili 3 copie di ogni blocco di dati per un cluster con 10 o più nodi, 2 copie di ogni blocco per un cluster con un numero di nodi da 4 a 9 e 1 copia (nessuna ridondanza) per i cluster con al massimo 3 nodi. Lo spazio HDFS totale disponibile è diviso per il fattore di replica. In alcuni casi, ad esempio incrementando il numero di nodi da 9 a 10, l'aumento del fattore di replica può causare effettivamente la diminuzione della quantità di spazio HDFS disponibile.

Ad esempio, per un cluster con 10 nodi principali di tipo m1.xlarge sarebbero disponibili 2833 GB di spazio in HDFS ((10 nodi x 850 GB per nodo)/fattore di replica 3).

Se le dimensioni del cluster superano la quantità di spazio disponibile per HDFS, puoi aggiungere altri nodi principali nel cluster o utilizzare la compressione dati per creare più spazio HDFS. Se il tuo cluster può essere interrotto e riavviato, potresti prendere in considerazione l'utilizzo di nodi core di un tipo di EC2 istanza Amazon più grande. Puoi anche valutare la modifica del fattore di replica. Tiene presente, tuttavia, che la riduzione del fattore di replica riduce la ridondanza dei dati HDFS e la capacità del cluster di recuperare da blocchi HDFS persi o danneggiati.

Errore del cluster Amazon EMR: EC2 QUOTA SUPERATA

Se viene visualizzato un messaggio `EC2 QUOTA EXCEEDED`, le cause potrebbero essere diverse. A seconda della differenze di configurazione, potrebbero essere necessari tra i 5 e i 20 minuti affinché i cluster precedenti vengano terminati e le risorse allocate rilasciate. Se visualizzi un errore `EC2 QUOTA EXCEEDED` al tentativo di avvio di un cluster, potrebbe essere dovuto al rilascio non ancora avvenuto delle risorse di un terminato recentemente. Questo messaggio può anche essere causato dal ridimensionamento di un gruppo di istanze o di un parco istanze in una dimensione di destinazione maggiore della quota di istanza corrente per l'account. Questa operazione può essere eseguita manualmente o automaticamente attraverso il dimensionamento automatico.

Considera le seguenti opzioni per risolvere il problema:

- Segui le istruzioni in Service Quotas di AWS in Riferimenti generali di Amazon Web Services per richiedere un aumento dei limiti del servizio. Per alcuni APIs, l'impostazione di un CloudWatch evento potrebbe essere un'opzione migliore rispetto all'aumento dei limiti. Per ulteriori dettagli, consulta Quando configurare gli eventi EMR in CloudWatch.
- Se uno o più cluster in esecuzione non hanno capacità, ridimensiona i gruppi di istanze o riduci le capacità di destinazione sui parchi istanze per l'esecuzione di cluster.
- Crea cluster con un minor numero di EC2 istanze o una capacità target ridotta.

Errore del cluster Amazon EMR: troppi errori di fetch

La presenza di messaggi di errore "Too many fetch-failures (Troppi errori fetch)" o "Error reading task output (Errore di lettura output attività)" nella fase o nei log del tentativo di attività indica che l'attività in esecuzione dipende dall'output di un'altra attività. Ciò si verifica spesso quando un'attività di riduzione viene accodata per l'esecuzione e richiede l'output di una o più attività di mappatura e tale output non è ancora disponibile.

L'output potrebbe non essere disponibile per diversi motivi:

- L'attività preliminare è ancora in fase di elaborazione. Questa è spesso un'attività di mappatura.
- I dati potrebbero non essere disponibili a causa di scarsa connettività di rete se i dati si trovano su un'istanza diversa.
- Se si utilizza HDFS per recuperare l'output, potrebbe verificarsi un problema con HDFS.

La causa più comune di questo errore è che l'attività precedente è ancora in corso di elaborazione. Ciò è probabile soprattutto se gli errori si verificano durante il primo tentativo di esecuzione delle attività di riduzione. Puoi verificare se questo è il motivo esaminando il log syslog per la fase di cluster che restituisce l'errore. Se il syslog mostra che entrambe le attività di mappatura e riduzione stanno avanzando, significa che la fase di riduzione è iniziata mentre ci sono attività di mappatura non ancora completate.

Ti consigliamo di cercare nei log una percentuale di avanzamento della mappatura che raggiunge il 100% e quindi scende nuovamente a un valore più basso. Quando la percentuale di mappatura è pari al 100%, questo non significa che tutte le attività di mappatura sono state completate, ma semplicemente che Hadoop sta eseguendo tutte le attività di mappatura. Se questo valore scende nuovamente sotto il 100%, significa che un'attività di mappatura non è riuscita e, a seconda della configurazione, Hadoop potrebbe tentare di pianificare di nuovo l'attività. Se la percentuale della mappa rimane al 100% nei log, esamina le CloudWatch metriche, in particolare`RunningMapTasks`,

per verificare se l'attività della mappa è ancora in fase di elaborazione. Puoi anche trovare queste informazioni utilizzando l'interfaccia Web Hadoop sul nodo master.

Se si verifica questo problema, puoi provare a eseguire una serie di operazioni:

- Configura la fase di riduzione per aspettare più a lungo prima dell'avvio. A questo scopo, puoi modificare l'impostazione di configurazione Hadoop mapred.reduce.slowstart.completed.maps su un tempo più lungo. Per ulteriori informazioni, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR.
- Associa il conteggio del riduttore alla funzionalità del riduttore totale del cluster. A questo scopo, modifica l'impostazione di configurazione Hadoop mapred.reduce.tasks per il processo.
- Utilizza un codice classe combiner per ridurre al minimo la quantità di output da recuperare.
- Verifica che non vi siano problemi con il EC2 servizio Amazon che influiscono sulle prestazioni di rete del cluster. A questo scopo, puoi utilizzare il Service Health Dashboard.
- Esamina le risorse di CPU e della memoria delle istanze nel cluster per verificare che l'elaborazione dei dati non stia sovraccaricando le risorse dei nodi. Per ulteriori informazioni, consulta Configurazione dell'hardware e della rete del cluster Amazon EMR.
- Controlla la versione di Amazon Machine Image (AMI) utilizzata nel cluster Amazon EMR. Se la versione è compresa tra 2.3.0 e 2.4.4, esegui l'aggiornamento a una versione più recente. Le versioni AMI nell'intervallo specificato utilizzano una versione di Jetty che potrebbe non essere in grado di fornire l'output dalla fase di mappatura. L'errore di recupero si verifica quando i riduttori non sono in grado di ottenere l'output dalla fase di mappatura.

  Jetty è un server HTTP open source utilizzato per le comunicazioni da macchina a macchina all'interno di un cluster Hadoop.

Errore del cluster Amazon EMR: il file può essere replicato solo su 0 nodi anziché su 1

Un file scritto in HDFS viene replicato in più nodi principali. Quando viene visualizzato questo errore, significa che il NameNode demone non dispone di DataNode istanze disponibili su cui scrivere dati in HDFS. In altre parole, la replica dei blocchi non viene eseguita. Questo errore può essere causato da diversi problemi:

- Possibile mancanza di spazio libero per il filesystem HDFS. Questa è la causa più probabile.
- DataNode le istanze potrebbero non essere disponibili al momento dell'esecuzione del processo.
- DataNode le istanze potrebbero essere state bloccate dalla comunicazione con il nodo master.

- Istanze non disponibili nel gruppo di istanze principale.
- Possibile mancanza di autorizzazioni. Ad esempio, il JobTracker demone potrebbe non disporre delle autorizzazioni per creare informazioni sul job tracker.
- L'impostazione dello spazio riservato per un' DataNode istanza potrebbe essere insufficiente. Controlla se questo è il caso verificando l'impostazione di configurazione dfs.datanode.du.reserved.

Per verificare se questo problema è causato dall'esaurimento dello spazio su disco di HDFS, guarda la `HDFSUtilization` metrica in. CloudWatch Se questo valore è troppo elevato, puoi aggiungere altri nodi principali al cluster. Se hai un cluster che pensi possa esaurire lo spazio su disco HDFS, puoi impostare un allarme CloudWatch per avvisarti quando il valore di `HDFSUtilization` supera un certo livello. Per ulteriori informazioni, consultare Ridimensiona manualmente un cluster Amazon EMR in esecuzione e Monitoraggio dei parametri di Amazon EMR con CloudWatch.

Se l'esaurimento dello spazio su HDFS non era il problema, controllate i log, DataNode i NameNode log e la connettività di rete per verificare la presenza di altri problemi che avrebbero potuto impedire a HDFS di replicare i dati. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Errore del cluster Amazon EMR: nodi con elenco negato

Il NodeManager daemon è responsabile del lancio e della gestione dei container sui nodi core e task. I contenitori vengono assegnati al NodeManager demone dal demone in esecuzione sul nodo ResourceManager master. ResourceManager Monitora il nodo tramite un battito cardiaco. NodeManager

Ci sono un paio di situazioni in cui il ResourceManager daemon deny elenca a NodeManager, rimuovendolo dal pool di nodi disponibili per elaborare le attività:

- Se non NodeManager ha inviato un heartbeat al ResourceManager demone negli ultimi 10 minuti (600.000 millisecondi). Questo periodo di tempo può essere configurato utilizzando l'impostazione di configurazione `yarn.nm.liveness-monitor.expiry-interval-ms`. Per ulteriori informazioni sulla modifica delle classificazioni di configurazione Yarn, consulta la sezione relativa alla Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR.
- NodeManager controlla lo stato dei dischi determinato da e. `yarn.nodemanager.local-dirs` `yarn.nodemanager.log-dirs` I controlli includono le autorizzazioni e lo spazio libero su disco (< 90%). Se un disco non supera il controllo, NodeManager smette di utilizzare quel particolare disco ma riporta comunque lo stato del nodo come integro. Se più dischi non superano il controllo, il nodo viene segnalato come non integro ResourceManager e i nuovi contenitori non vengono assegnati al nodo.

Il master dell'applicazione può anche negare l'elenco di un NodeManager nodo se presenta più di tre attività non riuscite. Puoi modificare questa impostazione su un valore più alto utilizzando il parametro di configurazione `mapreduce.job.maxtaskfailures.per.tracker`. Altre impostazioni di configurazione che è possibile modificare sono il numero di tentativi di esecuzione di una task prima che venga contrassegnata come non riuscita: `mapreduce.map.max.attempts` per task di mappatura e `mapreduce.reduce.maxattempts` per task di riduzione. Per ulteriori informazioni sulla modifica delle classificazioni di configurazione, consulta la sezione relativa alla Configurazione delle applicazioni nella Guida ai rilasci di Amazon EMR.

Errori di limitazione con un cluster Amazon EMR

Gli errori «Throttled from *Amazon EC2* while launching cluster» e «Failed to provisioning delle istanze a causa della limitazione da» si verificano *Amazon EC2* quando Amazon EMR non è in grado di completare una richiesta perché un altro servizio ha limitato l'attività. Amazon EC2 è la fonte più comune di errori di limitazione, ma altri servizi possono essere la causa di errori di limitazione. AWS i limiti di servizio si applicano su base regionale per migliorare le prestazioni e un errore di limitazione indica che hai superato il limite di servizio per il tuo account in quella regione.

Possibili cause

La fonte più comune di errori di EC2 throttling di Amazon è l'avvio di un gran numero di istanze di cluster, in modo da superare il limite di servizio per EC2 le istanze. Le istanze cluster possono essere avviate per i seguenti motivi:

- Vengono creati nuovi cluster.
- I cluster vengono ridimensionati manualmente. Per ulteriori informazioni, consulta Ridimensiona manualmente un cluster Amazon EMR in esecuzione.
- I gruppi di istanze in un cluster aggiungono istanze (ridimensionamento) come risultato di una regola di ridimensionamento automatico. Per ulteriori informazioni, consulta Comprensione delle regole di scalabilità automatica.
- I parchi istanze in un cluster aggiungono le istanze per soddisfare una maggiore capacità di destinazione. Per ulteriori informazioni, consulta Pianificazione e configurazione di flotte di istanze per il tuo cluster Amazon EMR.

È anche possibile che la frequenza o il tipo di richiesta API effettuata ad Amazon EC2 causi errori di throttling. Per ulteriori informazioni su come Amazon limita EC2 le richieste API, consulta Query API request rate nell'Amazon EC2 API Reference.

Soluzioni

Prendi in considerazione le soluzioni seguenti:

- Segui le istruzioni in Service Quotas di AWS in Riferimenti generali di Amazon Web Services per richiedere un aumento dei limiti del servizio. Per alcuni APIs, organizzare un CloudWatch evento potrebbe essere un'opzione migliore rispetto all'aumento dei limiti. Per ulteriori dettagli, consulta Quando configurare gli eventi EMR in CloudWatch.
- Se disponi di cluster che si avviano con la stessa pianificazione, ad esempio all'inizio di ogni ora, puoi aspettarti tempi di avvio straordinari.
- Se disponi di cluster dimensionati per la domanda di picco e hai periodicamente la capacità di istanza, considera la possibilità di specificare il ridimensionamento automatico per aggiungere e rimuovere le istanze on demand. In questo modo, le istanze vengono utilizzate in modo più efficiente e, in base al profilo della domanda, è possibile richiedere un numero inferiore di istanze in un dato momento in un account. Per ulteriori informazioni, consulta Utilizzo del ridimensionamento automatico con una politica personalizzata, ad esempio gruppi in Amazon EMR.

Errore del cluster Amazon EMR: tipo di istanza non supportato

Se crei un cluster e non riesci a visualizzare il messaggio di errore «Il tipo di istanza richiesto non *InstanceType* è supportato nella zona di disponibilità richiesta», significa che hai creato il cluster e specificato un tipo di istanza per uno o più gruppi di istanze che non è supportato da Amazon EMR nella regione e nella zona di disponibilità in cui è stato creato il cluster. Amazon EMR è in grado di supportare un tipo di istanza in una zona di disponibilità all'interno di una Regione ma non in un'altra. La sottorete selezionata per un cluster determina la zona di disponibilità all'interno della regione.

Soluzione

Determina i tipi di istanze disponibili in una zona di disponibilità utilizzando il AWS CLI

- Utilizzare il comando `ec2 run-instances` con l'opzione `--dry-run`. Nell'esempio seguente, sostituisci *m5.xlarge* con il tipo di istanza che desideri utilizzare, *ami-035be7bafff33b6b6* con l'AMI associato a quel tipo di istanza e *subnet-12ab3c45* con una sottorete nella zona di disponibilità che desideri interrogare.

```
aws ec2 run-instances --instance-type m5.xlarge --dry-run --image-
id ami-035be7bafff33b6b6 --subnet-id subnet-12ab3c45
```

Per istruzioni su come trovare un ID AMI, consulta Ricerca di un'AMI Linux. Per trovare un ID di sottorete, puoi utilizzare il comando describe-subnets.

Per ulteriori informazioni su come scoprire i tipi di istanza disponibili, consulta Trova un tipo di EC2 istanza Amazon.

Dopo aver determinato i tipi di istanza disponibili, è possibile eseguire una delle operazioni seguenti:

- Crea il cluster nella stessa regione e EC2 sottorete e scegli un tipo di istanza diverso con funzionalità simili a quelle della tua scelta iniziale. Per una lista di tipi di istanze supportate, consulta Tipi di istanze supportati con Amazon EMR. Per confrontare le funzionalità dei tipi di EC2 istanze, consulta i tipi di EC2 istanze Amazon.
- Scegli una sottorete per il cluster in una zona di disponibilità in cui il tipo di istanza è disponibile e supportato da Amazon EMR.

Riduci gli errori di avvio del cluster della flotta di istanze a causa di tipi di istanze primarie non supportati in Amazon EMR

I nodi primari sono essenziali nei cluster Amazon EMR. L'avvio di un cluster EMR potrebbe non riuscire con un `instance type not supported` errore in cui Amazon EMR tenta di avviare il cluster in una zona di disponibilità se il tipo di istanza principale non è supportato. La selezione avanzata della zona di disponibilità, ad esempio i cluster di flotte in Amazon EMR, filtra automaticamente le istanze AZs non supportate per i tipi di istanze principali specificati nella configurazione del cluster. Ciò significa che Amazon EMR non sceglierà una zona di disponibilità in cui i tipi di istanza primarie configurati non siano supportati, il che impedisce errori di avvio del cluster a causa di tipi di istanze non supportati.

Per consentire questo miglioramento, aggiungi l'autorizzazione richiesta al ruolo o alla policy di servizio per il tuo cluster. L'ultima versione di `AmazonEMRServicePolicy_v2` include questa autorizzazione, quindi se si utilizza tale politica, il miglioramento è già disponibile. Se utilizzi un ruolo o una policy di servizio personalizzati, aggiungi l'autorizzazione all'`ec2:DescribeInstanceTypeOfferings`avvio del cluster.

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "ec2:DescribeInstanceTypeOfferings"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describeinstancetypeofferings"
    }
  ]
}
```

Errore del cluster Amazon EMR: capacità EC2 insufficiente

Un EC2 *`InstanceType`* errore di esaurimento della capacità si verifica quando si tenta di creare un cluster o di aggiungere istanze a un cluster in una zona di disponibilità che non ha più il tipo di EC2 istanza specificato. La sottorete selezionata per un cluster determina la zona di disponibilità.

Per creare un cluster, effettua una delle operazioni indicate di seguito:

- Specifica un diverso tipo di istanza con funzionalità simili
- Crea il cluster in una Regione diversa
- Seleziona una sottorete in una zona di disponibilità in cui potrebbe essere disponibile il tipo di istanza che desideri.

Per aggiungere istanze a un cluster in esecuzione, effettua una delle seguenti operazioni:

- Modifica le configurazioni dei gruppi di istanze o le configurazioni dei parchi istanze per aggiungere tipi di istanze disponibili con capacità simili. Per una lista di tipi di istanze supportate, consulta Tipi di istanze supportati con Amazon EMR. Per confrontare le funzionalità dei tipi di EC2 istanze, consulta i tipi di EC2 istanze Amazon.
- Termina il cluster e ricrealo in una Regione e in una zona di disponibilità in cui è disponibile il tipo di istanza.

Errore del cluster Amazon EMR: errore del fattore di replica HDFS

Quando rimuovi un nodo principale da un gruppo di istanze principali o da una flotta di istanze, Amazon EMR potrebbe riscontrare un errore di replica HDFS. Questo errore si verifica quando rimuovi i nodi principali e il numero di nodi principali scende al di sotto del fattore dfs.replication configurato per l'Hadoop Distributed File System (HDFS). Pertanto, Amazon EMR non è in grado di eseguire l'operazione in sicurezza. Per determinare il valore predefinito della `dfs.replication` configurazione, configurazione HDFS.

Possibili cause

Di seguito sono riportate le possibili cause dell'errore del fattore di replica HDFS:

- Se si ridimensiona manualmente un gruppo di istanze principale o un parco di istanze al di sotto del fattore configurato. `dfs.replication`
- Le tue politiche per la scalabilità gestita o la scalabilità automatica potrebbero consentire la scalabilità per ridurre il numero di nodi principali al di sotto della soglia di. `dfs.replication`
- Questo errore può verificarsi anche se Amazon EMR tenta di sostituire un nodo centrale non integro quando un cluster ha il numero minimo di nodi core definito da. `dfs.replication`

Soluzioni e best practice

Per le soluzioni e le best practice, consulta quanto segue:

- Quando ridimensionate manualmente un cluster Amazon EMR, non ridimensionatelo al di sotto, `dfs.replication` poiché Amazon EMR non può completare il ridimensionamento in modo sicuro.
- Quando utilizzi la scalabilità gestita o la scalabilità automatica, assicurati che la capacità minima del cluster non sia inferiore al fattore. `dfs.replication`
- Il numero di istanze principali deve essere almeno più uno. `dfs.replication` Ciò garantisce che Amazon EMR possa sostituire con successo un nodo principale non integro se hai abilitato la sostituzione del core non integro.

> **Important**
>
> Il guasto di un singolo nodo core può portare alla perdita di dati HDFS se impostato su 1. `dfs.replication` Se il tuo cluster dispone di storage HDFS, ti consigliamo di configurare

il cluster con almeno quattro nodi principali per i carichi di lavoro di produzione per evitare la perdita di dati e di impostare il `dfs.replication` fattore su almeno 2.

Errore del cluster Amazon EMR: errore di spazio HDFS insufficiente

Se si tenta di rimuovere un nodo principale, può verificarsi un errore di spazio insufficiente nell'Hadoop Distributed File System (HDFS), ma Amazon EMR non può completare l'operazione in modo sicuro a causa dello spazio insufficiente rimasto nell'HDFS. Prima che Amazon EMR rimuova un nodo principale, tutti i dati HDFS sul nodo devono essere trasferiti su altri nodi principali per garantire la ridondanza dei dati. Tuttavia, se non c'è abbastanza spazio sugli altri nodi principali per la replica, Amazon EMR non può disattivare correttamente il nodo.

Possibili cause

Di seguito è riportato un elenco delle possibili cause dell'errore di spazio insufficiente HDFS:

- Se si ridimensiona manualmente un gruppo di istanze principale o un parco di istanze quando non c'è abbastanza spazio HDFS sui nodi rimanenti per la replica dei dati prima del ridimensionamento.
- La scalabilità gestita o la scalabilità automatica riducono un gruppo di istanze principale o un parco di istanze quando non c'è abbastanza spazio HDFS per la replica dei dati.
- Amazon EMR tenta di sostituire un nodo principale non integro, ma non è in grado di sostituirlo in modo sicuro a causa dello spazio HDFS insufficiente.

Soluzioni e best practice

Per le soluzioni e le best practice, consulta quanto segue:

- Aumenta il numero di nodi principali nel tuo cluster Amazon EMR. Se utilizzi la scalabilità gestita o la scalabilità automatica, aumenta la capacità minima dei tuoi nodi principali.
- Usa volumi EBS più grandi per i tuoi nodi principali quando crei il tuo cluster EMR.
- Elimina i dati HDFS non necessari nel tuo cluster EMR. Ti consigliamo di impostare CloudWatch allarmi per monitorare la `HDFSUtilization` metrica nel cluster per sapere se lo spazio sul cluster EMR è insufficiente.

Errori di input e output del cluster durante le operazioni di Amazon EMR

I seguenti errori sono comuni nelle operazioni di input e output del cluster. Utilizza le indicazioni contenute in questo argomento per risolvere i problemi e verificare la configurazione.

Argomenti

- Il percorso per Amazon Simple Storage Service (Amazon S3) presenta almeno tre barre?
- Stai cercando di attraversare le directory di input in modo ricorsivo?
- La directory di output esiste già?
- Si sta tentando di specificare una risorsa utilizzando un URL HTTP?
- Si sta facendo riferimento a un bucket Amazon S3 che utilizza un formato di nome non valido?
- Si stanno verificando problemi di caricamento dei dati in e da Amazon S3?

Il percorso per Amazon Simple Storage Service (Amazon S3) presenta almeno tre barre?

Quando si specifica un bucket Amazon S3, è necessario includere una barra di terminazione alla fine dell'URL. Ad esempio, invece di fare riferimento a un bucket come «s3n: //amzn-s3-demo-bucket1", dovreste usare «s3n: //amzn-s3-demo-bucket1/», altrimenti Hadoop danneggia il cluster nella maggior parte dei casi.

Stai cercando di attraversare le directory di input in modo ricorsivo?

Hadoop non cerca i file nelle directory di input in modo ricorsivo. Se si dispone di una struttura di directory come /corpus/01/01.txt, /corpus/01/02.txt, /corpus/02/01.txt, ecc. e si specifica /corpus/ come parametro di input per il cluster, Hadoop non trova alcun file di input perché la directory /corpus/ è vuota e Hadoop non controlla il contenuto delle sottodirectory. Allo stesso modo, Hadoop non controlla ricorsivamente le sottodirectory bucket Amazon S3.

I file di input devono trovarsi direttamente nella directory di input o nel bucket Amazon S3 specificato, non nelle sottocartelle.

La directory di output esiste già?

Se si specifica un percorso di output che esiste già, Hadoop farà fallire il cluster nella maggior parte dei casi. Questo significa che se si esegue un cluster una volta e poi lo si esegue di nuovo con esattamente gli stessi parametri, probabilmente funzionerà la prima volta e poi mai più; dopo la prima esecuzione, il percorso di output esiste e quindi causa il fallimento di tutte le esecuzioni successive.

Si sta tentando di specificare una risorsa utilizzando un URL HTTP?

Hadoop non accetta le posizioni di risorsa specificate utilizzando il prefisso http:// Non è possibile fare riferimento a una risorsa utilizzando un URL HTTP. Ad esempio, se si passa http://mysite/myjar.jar come parametro JAR il cluster fallisce.

Si sta facendo riferimento a un bucket Amazon S3 che utilizza un formato di nome non valido?

Se tenti di utilizzare un nome di bucket come «amzn-s3-demo-bucket1.1" con Amazon EMR, il cluster avrà esito negativo perché Amazon EMR richiede che i nomi dei bucket siano nomi host RFC 2396 validi; il nome non può terminare con un numero. Inoltre, a causa dei requisiti di Hadoop, i nomi dei bucket Amazon S3 utilizzati con Amazon EMR devono contenere solo lettere minuscole, numeri, punti (.) e trattini alti (-). Per ulteriori informazioni su come formattare i nomi dei bucket Amazon S3, consulta Restrizioni e limitazioni dei bucket nella Guida per l'utente di Amazon Simple Storage Service.

Si stanno verificando problemi di caricamento dei dati in e da Amazon S3?

Amazon S3 è la più popolare fonte di input e output per Amazon EMR. Un errore comune è trattare Amazon S3 come un normale file system. Occorre tener conto delle differenze tra Amazon S3 e un file system quando si esegue il cluster.

- Se si verifica un errore interno in Amazon S3, l'applicazione deve gestirlo in modo semplice e ripetere l'operazione.
- Se le chiamate ad Amazon S3 impiegano troppo tempo per tornare indietro, l'applicazione potrebbe dover ridurre la frequenza con cui chiama Amazon S3.
- L'elenco di tutti gli oggetti in un bucket Amazon S3 costituisce una chiamata costosa. L'applicazione dovrebbe ridurre al minimo il numero di volte che ciò accade.

Ci sono diversi modi per migliorare il modo in cui il cluster interagisce con Amazon S3.

- Avviare il cluster utilizzando la versione più recente di Amazon EMR.
- Usa S3 DistCp per spostare oggetti dentro e fuori da Amazon S3. S3 DistCp implementa la gestione degli errori, i nuovi tentativi e i back-off per soddisfare i requisiti di Amazon S3. Per ulteriori informazioni, consulta Copia distribuita con S3. DistCp

- Progettare l'applicazione con la consistenza finale in mente. Utilizza HDFS per la memorizzazione dei dati intermedi mentre il cluster è in esecuzione e Amazon S3 solo per inserire i dati iniziali e produrre i risultati finali.
- Se i cluster impegneranno 200 o più transazioni al secondo in Amazon S3, contatta il supporto per preparare il bucket per più transazioni al secondo e prendI in considerazione l'utilizzo delle strategie di partizione chiave descritte in Consigli per le prestazioni in Amazon S3.
- Impostare l'impostazione di configurazione Hadoop io.file.buffer.size su 65536. In questo modo Hadoop trascorre meno tempo a cercare tra gli oggetti Amazon S3.
- Considera la possibilità di disabilitare la funzione di esecuzione speculativa di Hadoop se il cluster riscontra problemi di concomitanza con Amazon S3. Questo è utile anche quando si esegue la risoluzione dei problemi di un cluster lento. È possibile eseguire questa operazione impostando le proprietà `mapreduce.map.speculative` e `mapreduce.reduce.speculative` su `false`. Quando si avvia un cluster, è possibile impostare questi valori utilizzando la classificazione di configurazione `mapred-env`. Per ulteriori informazioni, consulta Configurazione delle applicazioni nella Guida alle versioni di Amazon EMR.
- Se è in esecuzione un cluster Hive, consulta Si stanno verificando problemi di caricamento dei dati in e da Amazon S3 in Hive?.

Per ulteriori informazioni, consulta Best practice per gli errori di Amazon S3 nella Guida per l'utente di Amazon Simple Storage Service.

Errori di autorizzazione durante le operazioni del cluster Amazon EMR

Di seguito sono descritti alcuni degli errori comuni relativi all'utilizzo di autorizzazioni o credenziali.

Argomenti

- Le credenziali specificate per SSH sono corrette?
- Se utilizzi IAM, disponi delle EC2 politiche Amazon appropriate?

Le credenziali specificate per SSH sono corrette?

Se non sei in grado di utilizzare SSH per eseguire la connessione al nodo master, il problema è probabilmente legato alle tue credenziali di sicurezza.

In primo luogo, verifica che il file .pem contenente la chiave SSH disponga delle autorizzazioni appropriate. Puoi utilizzare chmod per modificare le autorizzazioni per il file .pem come illustrato nell'esempio seguente, dove devi sostituire mykey.pem con il nome del tuo file .pem.

```
chmod og-rwx mykey.pem
```

La seconda possibilità è che non stai utilizzando la coppia di chiavi che hai specificato alla creazione del cluster. Si tratta di un errore comune se hai creato più coppie di chiavi. Verifica i dettagli del cluster nella console di Amazon EMR (o utilizza l'opzione `--describe` nella CLI) per determinare il nome della coppia di chiavi specificato alla creazione del cluster.

Dopo aver verificato di utilizzare la coppia di chiavi corretta e che le autorizzazioni siano impostate correttamente nel file .pem, puoi utilizzare il comando seguente per utilizzare SSH per la connessione al nodo master, sostituendo mykey.pem con il nome del file .pem e `hadoop@ec2-01-001-001-1.compute-1.amazonaws.com` con il nome DNS pubblico del nodo master (disponibile mediante l'opzione `--describe` nella CLI o tramite la console di Amazon EMR).

> **Important**
>
> Dovrai utilizzare il nome di login `hadoop` quando esegui la connessione a un nodo di cluster Amazon EMR, altrimenti potrebbe verificarsi un errore simile a `Server refused our key`.

```
ssh -i mykey.pem hadoop@ec2-01-001-001-1.compute-1.amazonaws.com
```

Per ulteriori informazioni, consulta Connect al nodo primario del cluster Amazon EMR tramite SSH.

Se utilizzi IAM, disponi delle EC2 politiche Amazon appropriate?

Poiché Amazon EMR utilizza EC2 le istanze come nodi, gli utenti di Amazon EMR devono inoltre disporre di determinate EC2 politiche Amazon per consentire ad Amazon EMR di gestire tali istanze per conto dell'utente. Se non disponi delle autorizzazioni richieste, Amazon EMR restituisce l'errore: «l'account non è autorizzato a EC2 chiamare».

Per ulteriori informazioni sulle EC2 politiche Amazon che il tuo account IAM deve impostare per eseguire Amazon EMR, consulta. Funzionamento di Amazon EMR con IAM

Errori del cluster Hive

In genere, puoi trovare la causa di un errore Hive nel file `syslog`, a cui ti colleghi dal riquadro Steps (Fasi). Se non riesci a determinare il problema, controlla nel messaggio di errore del tentativo di attività Hadoop tramite il riquadro Task Attempts (Tentativi attività).

I seguenti errori sono comuni ai cluster Hive.

Argomenti

- Stai usando la versione più recente di Hive?
- Hai rilevato un errore di sintassi nello script Hive?
- Un processo non è riuscito durante l'esecuzione in modalità interattiva?
- Si stanno verificando problemi di caricamento dei dati in e da Amazon S3 in Hive?

Stai usando la versione più recente di Hive?

La versione più recente di Hive contiene tutte le patch e le correzioni dei bug correnti e può risolvere il problema.

Hai rilevato un errore di sintassi nello script Hive?

Se una fase non riesce, cerca nel file `stdout` dei log la fase che ha eseguito lo script Hive. Se l'errore non esiste, cerca nel file `syslog` dei log dei tentativi di attività il tentativo di attività non riuscito. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Un processo non è riuscito durante l'esecuzione in modalità interattiva?

Se stai eseguendo Hive in maniera interattiva sul nodo master e il cluster non è riuscito, nel log dei tentativi di attività cerca le voci `syslog` relative al tentativo di attività non riuscito. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Si stanno verificando problemi di caricamento dei dati in e da Amazon S3 in Hive?

Se si verificano problemi di accesso ai dati in Amazon S3, controlla innanzitutto le possibili cause elencate in Si stanno verificando problemi di caricamento dei dati in e da Amazon S3?. Se nessuno di questi problemi è la causa, considera le seguenti opzioni specifiche di Hive.

- Assicurati di utilizzare la versione più recente di Hive che contiene tutte le patch correnti e le correzioni dei bug che possono risolvere il problema. Per ulteriori informazioni, consulta Apache Hive.
- L'uso di `INSERT OVERWRITE` richiede l'elenco dei contenuti della cartella o del bucket Amazon S3. Questa è un'operazione costosa. Se possibile, elimina manualmente il percorso anziché lasciare che sia Hive a elencare ed eliminare gli oggetti esistenti.
- Con le versioni di Amazon EMR precedenti alla 5.0, è possibile utilizzare il comando seguente in HiveQL per memorizzare anticipatamente nella cache i risultati di un'operazione di elenco Amazon S3 localmente sul cluster:

```
set hive.optimize.s3.query=true;
```

- Ove possibile, utilizza partizioni statiche.
- In alcune versioni di Hive e Amazon EMR, potrebbe non essere possibile utilizzare ALTER TABLE (ALTERA TABELLA) perché la tabella è archiviata in un percorso diverso da quello previsto da Hive. La soluzione è aggiungere o aggiornare seguendo in `/home/hadoop/conf/core-site.xml`:

```
<property>
    <name>fs.s3n.endpoint</name>
    <value>s3.amazonaws.com</value>
</property>
```

Errori VPC durante le operazioni del cluster Amazon EMR

Gli errori indicati di seguito sono comuni nella configurazione VPC in Amazon EMR.

Argomenti

- Configurazione sottorete non valida
- Set opzioni DHCP mancante
- Errori di autorizzazioni
- Errori che generano START_FAILED
- Cluster e non si avvia Terminated with errors NameNode

Configurazione sottorete non valida

Nella pagina Cluster Details (Dettagli cluster), nel campo Status (Stato), verrà visualizzato un errore simile al seguente:

```
The subnet configuration was invalid: Cannot find route to InternetGateway
in main RouteTable rtb-id for vpc vpc-id.
```

Per risolvere questo problema, devi creare un gateway Internet e collegarlo al tuo VPC. Per ulteriori informazioni, consulta la pagina relativa all'Aggiunta di un gateway Internet al VPC.

In alternativa, verifica di aver configurato il tuo VPC con Enable DNS resolution (Abilita risoluzione DNS) e Enable DNS hostname support (Abilita supporto nome host DNS). Per ulteriori informazioni, consulta Utilizzo del DNS con il tuo VPC.

Set opzioni DHCP mancante

Riscontri un errore di fase nel log di sistema del cluster (syslog) simile al seguente:

```
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.
```

oppure

```
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.
```

Per risolvere il problema, devi configurare un VPC che include un set di opzioni DHCP con parametri impostati sui seguenti valori:

> **Note**
>
> Se utilizzi la regione AWS GovCloud (Stati Uniti occidentali), imposta domain-name su **`us-gov-west-1.compute.internal`** anziché sul valore utilizzato nell'esempio seguente.

- domain-name = **`ec2.internal`**

Utilizza **`ec2.internal`** se la regione è Stati Uniti orientali (Virginia settentrionale). Per altre regioni, usa. *`region-name`* **`.compute.internal`** Ad esempio, in us-west-2, utilizza domain-name=**`us-west-2.compute.internal`**.

- domain-name-servers = **`AmazonProvidedDNS`**

Per ulteriori informazioni, consulta la pagina relativa ai Set di opzioni DHCP.

Errori di autorizzazioni

Un errore nel log `stderr` per una fase indica che una risorsa Amazon S3 non dispone delle autorizzazioni appropriate. Si tratta di un errore 403 simile a questo:

```
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
```

Se ActionOnFailure è impostato su`TERMINATE_JOB_FLOW`, ciò comporterebbe la terminazione del cluster con lo stato,`SHUTDOWN_COMPLETED_WITH_ERRORS`.

Alcuni modi per risolvere questo problema sono:

- Se utilizzi una policy di bucket Amazon S3 all'interno di un VPC, assicurati di concedere l'accesso a tutti i bucket creando un endpoint VPC e selezionando Allow all (Consenti tutti) in corrispondenza dell'opzione Policy durante la creazione dell'endpoint.
- Assicurati che le policy associate alle risorse S3 includano il VPC in cui avvii il cluster.
- Prova a eseguire il seguente comando dal cluster per verificare che sia possibile accedere al bucket

  ```
  hadoop fs -copyToLocal s3://path-to-bucket /tmp/
  ```

- Per ottenere informazioni di debug più specifiche, imposta il parametro `log4j.logger.org.apache.http.wire` su `DEBUG` nel file `/home/hadoop/conf/log4j.properties` nel cluster. Puoi controllare il file di log `stderr` dopo aver provato ad accedere al bucket dal cluster. Il file di log fornirò informazioni più dettagliate:

  ```
  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]"
```

Errori che generano **START_FAILED**

Prima dell'AMI 3.7.0, per VPCs cui veniva specificato un nome host, Amazon EMR mappa i nomi host interni della sottorete con indirizzi di dominio personalizzati come segue:. `ip-`*`X.X.X.X.customdomain.com`*`.tld` Ad esempio, se il nome host fosse `ip-10.0.0.10` e il VPC avesse l'opzione del nome di dominio impostata su customdomain.com, il nome host risultante mappato da Amazon EMR sarebbe `ip-10.0.1.0.customdomain.com`. Viene aggiunta una voce in `/etc/hosts` per risolvere il nome host in 10.0.0.10. Questo comportamento è stato modificato con l'AMI 3.7.0 e ora Amazon EMR mantiene la configurazione DHCP del VPC in modo completo. In passato, i clienti potevano anche specificare la mappatura del nome host tramite un'operazione di bootstrap.

Se preferisci mantenere questo comportamento, devi fornire la configurazione di risoluzione per DNS e inoltro necessaria per il dominio personalizzato.

Cluster e non si avvia **Terminated with errors** NameNode

Quando avvii un cluster EMR in un VPC che utilizza un nome di dominio DNS personalizzato, potrebbero verificarsi errori sul cluster con il seguente messaggio nella console:

```
Terminated with errors  On the master instance(instance-id), bootstrap action 1
 returned a  non-zero return code
```

L'errore è dovuto all' NameNode impossibilità di avviarsi. Ciò comporterà il seguente errore rilevato nei NameNode log, il cui URI Amazon S3 ha il seguente 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)
```

Ciò è dovuto a un potenziale problema per cui un' EC2 istanza può avere più set di nomi di dominio completi all'avvio di cluster EMR in un VPC, che utilizza sia un server DNS fornito sia AWS un server DNS personalizzato fornito dall'utente. Se il server DNS fornito dall'utente non prevede alcun record puntatore (PTR) per eventuali record A utilizzati per designare i nodi in un cluster EMR, il cluster non potrà avviarsi se configurato in questo modo. La soluzione consiste nell'aggiungere 1 record PTR per ogni record A creato all'avvio di un' EC2 istanza in una delle sottoreti del VPC.

Errori di streaming del cluster Amazon EMR

In genere, puoi individuare la causa di un errore di streaming in un file `syslog`. Un collegamento a questo file è visualizzato nel riquadro Steps (Fasi).

Gli errori esposti di seguito sono comuni ai cluster di streaming.

Argomenti

- I dati sono inviati al mappatore in un formato errato?
- Si è verificato il timeout dello script?
- Stai passando argomenti di streaming non validi?

- Lo script è terminato con un errore?

I dati sono inviati al mappatore in un formato errato?

Per determinare se è il caso, cerca un messaggio di errore nel file `syslog` di un tentativo di attività nei log dei tentativi di attività. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Si è verificato il timeout dello script?

Il timeout predefinito per uno script mappatore o riduttore è di 600 secondi. Se lo script richiede più tempo, il tentativo di attività non riuscirà. Per determinare se si è verificato tale problema, controlla il file `syslog` di un tentativo di attività non riuscito nei log dei tentativi di attività. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Puoi modificare il limite di tempo impostando un nuovo valore per l'impostazione di configurazione `mapred.task.timeout`. Questa impostazione specifica il numero di millisecondi dopo i quali Amazon EMR terminerà un'attività che non ha letto l'input, scritto l'output o aggiornato la stringa di stato. Puoi aggiornare questo valore passando un ulteriore argomento di streaming `-jobconf mapred.task.timeout=800000`.

Stai passando argomenti di streaming non validi?

Lo streaming di Hadoop supporta solo gli argomenti elencati di seguito. Se passi degli argomenti che non sono tra questi, si verificherà un errore nel cluster.

```
-blockAutoGenerateCacheFiles
-cacheArchive
-cacheFile
-cmdenv
-combiner
-debug
-input
-inputformat
-inputreader
-jobconf
-mapper
-numReduceTasks
-output
-outputformat
```

```
-partitioner
-reducer
-verbose
```

Inoltre, lo streaming Hadoop riconosce solo gli argomenti passati utilizzando la sintassi Java, ovvero preceduti da un solo trattino. Se passi argomenti preceduti da un doppio trattino, si verificherà un errore nel cluster.

Lo script è terminato con un errore?

Se lo script mappatore o riduttore termina con un errore, puoi individuare l'errore nel file `stderr` relativo al tentativo di attività non riuscito nei log dei tentativi di attività. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Amazon EMR: errori di cluster JAR personalizzati

I seguenti errori sono comuni ai cluster JAR personalizzati.

Argomenti

- Il cluster JAR genera un'eccezione prima di creare un processo?
- Il cluster JAR genera un errore all'interno di un'attività di mappatura?

Il cluster JAR genera un'eccezione prima di creare un processo?

Se il programma principale del cluster JAR personalizzato genera un'eccezione durante la creazione del processo Hadoop, si consiglia di cercare nel file `syslog` del log delle fasi. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Il cluster JAR genera un errore all'interno di un'attività di mappatura?

Se il cluster JAR personalizzato e il mappatore generano un'eccezione durante l'elaborazione dei dati di input, si consiglia di cercare nel file `syslog` dei log del tentativo di attività. Per ulteriori informazioni, consulta Visualizza i file di log di Amazon EMR.

Errori di Amazon EMR AWS GovCloud (Stati Uniti occidentali)

La regione AWS GovCloud (Stati Uniti occidentali) si differenzia dalle altre aree per sicurezza, configurazione e impostazioni predefinite. Di conseguenza, utilizza la seguente lista di controllo per

risolvere gli errori di Amazon EMR specifici della regione (Stati Uniti occidentali) prima di utilizzare consigli più generali per AWS GovCloud la risoluzione dei problemi.

- Verifica che i ruoli IAM siano configurati correttamente. Per ulteriori informazioni, consulta Configurazione dei ruoli di servizio IAM per le autorizzazioni di Amazon EMR per i servizi e risorse AWS.
- Assicurati che la configurazione del VPC abbia configurato correttamente il resolution/hostname supporto DNS, l'Internet Gateway e i parametri del set di opzioni DHCP. Per ulteriori informazioni, consulta Errori VPC durante le operazioni del cluster Amazon EMR.

Se questa procedura non risolve il problema, continua con la procedura per la risoluzione degli errori più comuni in Amazon EMR. Per ulteriori informazioni, consulta Raccolte di errori comuni in Amazon EMR.

Ricerca di un cluster mancante

Se il cluster non è presente nell'elenco delle console o nell'API `ListClusters`, controlla quanto segue:

- Verifica che l'età del cluster dal momento del completamento sia inferiore a due mesi. Amazon EMR conserva le informazioni sui metadati relativi ai cluster completati per due mesi senza alcun costo aggiuntivo. Non puoi eliminare i cluster completati dalla console, ma Amazon EMR elimina automaticamente i cluster completati dopo due mesi.
- Conferma che disponi delle autorizzazioni di ruolo per visualizzare il cluster.
- Conferma che stai visualizzando lo stesso Regione AWS luogo in cui si trova il cluster.

Risolvi un cluster Amazon EMR che non funziona con un codice di errore

In questa sezione viene descritto il processo di risoluzione dei problemi di un cluster in cui si sono verificati errori. Ciò significa che il cluster è stato terminato con un codice di errore.

> **Note**
>
> Quando un cluster EMR termina con un errore, `DescribeCluster` e `ListClusters` APIs restituisce un codice di errore e un messaggio di errore. Per alcuni errori del cluster, l'array di

dati `ErrorDetail` può anche aiutarti a risolvere l'errore. Per ulteriori informazioni, consulta Codici di errore con ErrorDetail informazioni in Amazon EMR.

Se il cluster viene eseguito ma impiega molto tempo per restituire risultati, consulta Risolvi i problemi di un cluster Amazon EMR lento.

Argomenti

- Fase 1: raccogliere dati sul problema con il cluster Amazon EMR
- Fase 2: Controllo dell'ambiente
- Fase 3: Esame dell'ultima modifica dello stato
- Fase 4: Esamina i file di log di Amazon EMR
- Fase 5: testare il cluster Amazon EMR passo dopo passo

Fase 1: raccogliere dati sul problema con il cluster Amazon EMR

Il primo passaggio per la risoluzione dei problemi di un cluster consiste nel raccogliere informazioni su ciò che non ha funzionato, nonché sullo stato corrente e sulla configurazione del cluster. Queste informazioni verranno utilizzate nei passaggi seguenti per confermare o escludere possibili cause del problema.

Definizione del problema

Una chiara definizione del problema è il primo punto da cui iniziare. Alcune domande da porsi:

- Cosa mi aspettavo che succedesse? Che cos'è successo invece?
- Quando si è verificato questo problema per la prima volta? Quante volte è successo da allora?
- È cambiato qualcosa nel modo in cui configuro o eseguo il cluster?

Dettagli del cluster

I seguenti dettagli del cluster sono utili per individuare i problemi. Per ulteriori informazioni su come raccogliere tali informazioni, consulta Visualizza lo stato e i dettagli del cluster Amazon EMR.

- Identificatore del cluster (chiamato anche identificatore del flusso di lavoro).

- Regione AWS e nella zona di disponibilità in cui è stato avviato il cluster.
- Stato del cluster, inclusi i dettagli dell'ultima modifica dello stato.
- Tipo e numero di EC2 istanze specificati per i nodi master, core e task.

Fase 2: Controllo dell'ambiente

Amazon EMR funziona come parte di un ecosistema di servizi Web e software open source. Fattori che influenzano tali dipendenze possono influire sulle prestazioni di Amazon EMR.

Argomenti

- Verifica della presenza di interruzioni del servizio
- Controllo dei limiti di utilizzo
- Controllo della versione di rilascio
- Controllo della configurazione della sottorete Amazon VPC

Verifica della presenza di interruzioni del servizio

Amazon EMR utilizza diversi servizi Amazon Web Services al suo interno. Esegue server virtuali su Amazon EC2, archivia dati e script su Amazon S3 e riporta i parametri a. CloudWatch Gli eventi che disturbano questi servizi sono rari, ma quando si verificano possono causare problemi in Amazon EMR.

Prima di proseguire, controllare il Service Health Dashboard. Controlla la Regione in cui è stato avviato il cluster per verificare se esistono eventi di interruzione in uno di questi servizi.

Controllo dei limiti di utilizzo

Se stai avviando un cluster di grandi dimensioni, hai avviato più cluster contemporaneamente o sei un utente che condivide un cluster Account AWS con altri utenti, il cluster potrebbe aver avuto esito negativo perché hai superato un limite di servizio. AWS

Amazon EC2 limita il numero di istanze di server virtuali in esecuzione su una singola AWS regione a 20 istanze on-demand o riservate. Se avvii un cluster con più di 20 nodi o avvii un cluster che fa sì che il numero totale di EC2 istanze attive sul tuo Account AWS computer superi le 20, il cluster non sarà in grado di avviare tutte le EC2 istanze necessarie e potrebbe fallire. Quando ciò si verifica, Amazon EMR restituisce un errore `EC2 QUOTA EXCEEDED`. Puoi richiedere l' AWS aumento del

numero di EC2 istanze che puoi eseguire sul tuo account inviando un'applicazione Request to Increase Amazon EC2 Instance Limit.

Un altro fattore che può causare il superamento dei limiti di utilizzo è il ritardo tra il momento in cui un cluster viene terminato e quello in cui rilascia tutte le sue risorse. A seconda della configurazione, potrebbero essere necessari tra i 5 e i 20 minuti affinché un cluster venga terminato e le risorse allocate rilasciate. Se visualizzi un errore `EC2 QUOTA EXCEEDED` al tentativo di avvio di un cluster, potrebbe essere dovuto al rilascio non ancora avvenuto delle risorse di un cluster terminato recentemente. In questo caso, puoi richiedere che la tua EC2 quota Amazon venga aumentata oppure puoi attendere venti minuti e riavviare il cluster.

Amazon S3 limita il numero di bucket creati su un account a 100. Se il cluster crea un nuovo bucket che supera questo limite, la creazione del bucket avrà esito negativo e potrebbe causare l'esito negativo del cluster.

Controllo della versione di rilascio

Confronta l'etichetta di rilascio utilizzata per avviare il cluster con il rilascio di Amazon EMR più recente. Ogni rilascio di Amazon EMR include miglioramenti quali nuove applicazioni, caratteristiche, patch e correzioni di bug. Il problema che si presenta nel cluster potrebbe essere già stato corretto nella versione del rilascio più recente. Se possibile, riavvia il cluster utilizzando la versione più recente.

Controllo della configurazione della sottorete Amazon VPC

Se il cluster è stato avviato in una sottorete Amazon VPC, la sottorete deve essere configurata come descritto in Configurazione della rete in un VPC per Amazon EMR. Occorre inoltre verificare che la sottorete in cui si avvia il cluster disponga di indirizzi IP elastici liberi sufficienti per assegnarne uno a ciascun nodo del cluster.

Fase 3: Esame dell'ultima modifica dello stato

L'ultima modifica dello stato fornisce informazioni su cosa si è verificato l'ultima volta che lo stato del cluster è cambiato: Questo spesso può fornire un'indicazione su cosa non ha funzionato quando lo stato del cluster cambia in `FAILED`. Ad esempio, se si avvia un cluster di streaming e si specifica un percorso di output che esiste già in Amazon S3, il cluster restituisce un errore con un'ultima modifica di stato di "Streaming output directory already exists (La directory di output dello streaming esiste già)".

Puoi individuare il valore dell'ultima modifica dello stato dalla console visualizzando il riquadro dei dettagli per il cluster, dalla CLI utilizzando gli argomenti `list-steps` o `describe-cluster`, oppure dall'API utilizzando le azioni `DescribeCluster` e `ListSteps`. Per ulteriori informazioni, consulta Visualizza lo stato e i dettagli del cluster Amazon EMR.

Fase 4: Esamina i file di log di Amazon EMR

Il passaggio successivo consiste nell'esaminare i file di log per individuare un codice di errore o altra indicazione relativa al problema riscontrato dal cluster. Per informazioni sui file di log disponibili, su dove trovarli e su come visualizzarli, consulta Visualizza i file di log di Amazon EMR.

Potrebbe essere necessaria qualche indagine per determinare l'accaduto. Hadoop esegue il lavoro dei processi nei tentativi di attività su vari nodi del cluster. Amazon EMR può avviare tentativi di attività speculative, terminando gli altri tentativi di attività che non vengono completati per primi. Questo genera un'attività significativa che viene registrata nei file di log del controller, stderr e syslog mentre si verifica. Inoltre, anche se più tentativi di attività vengono eseguiti contemporaneamente, un file di log può visualizzare i risultati solo in modo lineare.

Inizia controllando i log delle operazioni di bootstrap per individuare errori o modifiche impreviste alla configurazione durante l'avvio del cluster. Successivamente, cerca nei log di fase per identificare i processi Hadoop avviati come parte di una fase con errori. Esamina i log dei processi Hadoop per identificare i tentativi di attività non riusciti. Il log dei tentativi di attività conterrà dettagli su ciò che ha causato l'esito negativo di un tentativo di attività.

Le sezioni seguenti descrivono come utilizzare i vari file di log per identificare gli errori nel cluster.

Controllo dei log delle operazioni di bootstrap

Le operazioni di bootstrap eseguono script sul cluster mentre quest'ultimo viene avviato. Vengono comunemente utilizzate per installare software aggiuntivi nel cluster o per modificare le impostazioni di configurazione rispetto ai valori predefiniti. Il controllo di questi log può fornire informazioni dettagliate sugli errori verificatisi durante la configurazione del cluster e sulle modifiche alle impostazioni di configurazione che potrebbero influire sulle prestazioni.

Controllo dei log della fase

Esistono quattro tipi di log di fase.

- controller: contiene i file generati da Amazon EMR che derivano da errori riscontrati durante il tentativo di eseguire la fase. Se la fase ha esito negativo durante il caricamento, è possibile

trovare la traccia di stack in questo log. Gli errori che si verificano durante il caricamento o l'accesso all'applicazione sono spesso descritti nel log, mentre mancano gli errori relativi ai file del mappatore.

- stderr: contiene messaggi di errore che si sono presentati durante l'elaborazione della fase. Gli errori di caricamento delle applicazioni sono spesso descritti in questo log. Talvolta, questo log contiene una traccia di stack.
- stdout: contiene lo stato generato dagli eseguibili del mappatore e del riduttore. Gli errori di caricamento delle applicazioni sono spesso descritti in questo log. Talvolta, questo log contiene messaggi di errore dell'applicazione.
- syslog: contiene log generati da software non Amazon, come Apache e Hadoop. Gli errori di streaming sono spesso descritti in questo log.

Controlla stderr per rilevare errori evidenti. Se stderr visualizza un breve elenco di errori, la fase è giunta a un rapido arresto generando un errore. Questo è spesso causato da un errore nelle applicazioni di mappatore e riduttore eseguite nel cluster.

Esamina le ultime righe di controller e syslog per cercare avvisi di errore o guasti. Segui tutte le notifiche relative alle attività non riuscite, in particolare "Job Failed (Processo non riuscito)".

Controllo dei log del tentativo di attività

Se l'analisi precedente dei log di fase ha rilevato una o più attività non riuscite, esamina i log dei tentativi di attività corrispondenti per ottenere informazioni più dettagliate sugli errori.

Fase 5: testare il cluster Amazon EMR passo dopo passo

Una tecnica utile quando cerchi di rintracciare l'origine di un errore è riavviare il cluster e inviare le fasi una alla volta. Ciò consente di controllare i risultati di ciascuna fase prima di elaborare quella successiva e offre l'opportunità di correggere ed eseguire nuovamente una fase non riuscita. Inoltre ha il vantaggio di poter caricare i dati di input una sola volta.

Per eseguire il test dettagliato di un cluster

1. Avviare un nuovo cluster con keep-alive e la protezione da cessazione abilitati. Keep-alive consente di mantenere il cluster in esecuzione dopo che tutte le fasi in sospeso sono state elaborate. La protezione da cessazione impedisce l'arresto di un cluster in caso di errore. Per ulteriori informazioni, consultare Configurazione di un cluster Amazon EMR per continuare

o terminare dopo l'esecuzione delle fasi e Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

2. Inviare una fase al cluster. Per ulteriori informazioni, consulta Invia il lavoro a un cluster Amazon EMR.
3. Al termine dell'elaborazione della fase, verificare la presenza di errori nei file di log della fase. Per ulteriori informazioni, consulta Fase 4: Esamina i file di log di Amazon EMR. Per individuare in modo veloce questi file di log, collegati al nodo master e visualizza qui i file di log. I file di log vengono visualizzati solo se la fase rimane in esecuzione per qualche tempo, termina o non riesce.
4. Se la fase viene completata senza errori, eseguire la fase successiva. In caso di errori, analizzare l'errore nei file di log. Se si è verificato un errore nel codice, apportare la correzione ed eseguire nuovamente la fase. Continuare finché tutte le fasi non vengono eseguite senza errore.
5. Al termine del debug del cluster, se desideri terminarlo, devi farlo manualmente. Questo è necessario perché il cluster è stato avviato con la protezione da cessazione abilitata. Per ulteriori informazioni, consulta Utilizzo della protezione dalle terminazioni per proteggere i cluster Amazon EMR da arresti accidentali.

Risolvi i problemi di un cluster Amazon EMR lento

Questa sezione descrive la procedura di risoluzione dei problemi relativi a un cluster in esecuzione che impiega troppo tempo per restituire i risultati. Per ulteriori informazioni sulle soluzioni da adottare nel caso in cui il cluster venga terminato con un codice di errore, consulta Risolvi un cluster Amazon EMR che non funziona con un codice di errore

Amazon EMR ti consente di specificare il numero e il tipo di istanze nel cluster. Queste specifiche sono quelle che influiscono maggiormente sulla velocità di elaborazione dei dati. Una cosa da prendere in considerazione è rieseguire il cluster, questa volta specificando EC2 istanze con maggiori risorse o specificando un numero maggiore di istanze nel cluster. Per ulteriori informazioni, consulta Configurazione dell'hardware e della rete del cluster Amazon EMR.

I seguenti argomenti illustrano la procedura di identificazione delle altre cause che possono rallentare un cluster.

Argomenti

- Fase 1: raccogliere dati sul problema con il cluster Amazon EMR
- Fase 2: Verificare l'ambiente del cluster EMR

- Fase 3: Esamina i file di log per il cluster Amazon EMR
- Fase 4: verifica dello stato del cluster e dell'istanza Amazon EMR
- Fase 5: Verifica della presenza di gruppi sospesi
- Fase 6: Rivedere le impostazioni di configurazione per il cluster Amazon EMR
- Fase 7: esaminare i dati di input per il cluster Amazon EMR

Fase 1: raccogliere dati sul problema con il cluster Amazon EMR

Il primo passaggio per la risoluzione dei problemi di un cluster consiste nel raccogliere informazioni su ciò che non ha funzionato, nonché sullo stato corrente e sulla configurazione del cluster. Queste informazioni verranno utilizzate nei passaggi seguenti per confermare o escludere possibili cause del problema.

Definizione del problema

Una chiara definizione del problema è il primo punto da cui iniziare. Alcune domande da porsi:

- Cosa mi aspettavo che succedesse? Che cos'è successo invece?
- Quando si è verificato questo problema per la prima volta? Quante volte è successo da allora?
- È cambiato qualcosa nel modo in cui configuro o eseguo il cluster?

Dettagli del cluster

I seguenti dettagli del cluster sono utili per individuare i problemi. Per ulteriori informazioni su come raccogliere tali informazioni, consulta Visualizza lo stato e i dettagli del cluster Amazon EMR.

- Identificatore del cluster (chiamato anche identificatore del flusso di lavoro).
- Regione AWS e nella zona di disponibilità in cui è stato avviato il cluster.
- Stato del cluster, inclusi i dettagli dell'ultima modifica dello stato.
- Tipo e numero di EC2 istanze specificati per i nodi master, core e task.

Fase 2: Verificare l'ambiente del cluster EMR

Controlla il tuo ambiente per vedere se ci sono interruzioni del servizio o se hai superato un limite di servizio. AWS

Argomenti

- Verifica della presenza di interruzioni del servizio
- Controllo dei limiti di utilizzo
- Controllo della configurazione della sottorete Amazon VPC
- Riavvio del cluster

Verifica della presenza di interruzioni del servizio

Amazon EMR utilizza diversi servizi Amazon Web Services al suo interno. Esegue server virtuali su Amazon EC2, archivia dati e script su Amazon S3 e riporta i parametri a. CloudWatch Gli eventi che disturbano questi servizi sono rari, ma quando si verificano possono causare problemi in Amazon EMR.

Prima di proseguire, controllare il Service Health Dashboard. Controlla la Regione in cui è stato avviato il cluster per verificare se esistono eventi di interruzione in uno di questi servizi.

Controllo dei limiti di utilizzo

Se stai avviando un cluster di grandi dimensioni, hai avviato più cluster contemporaneamente o sei un utente che condivide un cluster Account AWS con altri utenti, il cluster potrebbe aver avuto esito negativo perché hai superato un limite di servizio. AWS

Amazon EC2 limita il numero di istanze di server virtuali in esecuzione su una singola AWS regione a 20 istanze on-demand o riservate. Se avvii un cluster con più di 20 nodi o avvii un cluster che fa sì che il numero totale di EC2 istanze attive sul tuo Account AWS computer superi le 20, il cluster non sarà in grado di avviare tutte le EC2 istanze necessarie e potrebbe fallire. Quando ciò si verifica, Amazon EMR restituisce un errore `EC2 QUOTA EXCEEDED`. Puoi richiedere l' AWS aumento del numero di EC2 istanze che puoi eseguire sul tuo account inviando un'applicazione Request to Increase Amazon EC2 Instance Limit.

Un altro fattore che può causare il superamento dei limiti di utilizzo è il ritardo tra il momento in cui un cluster viene terminato e quello in cui rilascia tutte le sue risorse. A seconda della configurazione, potrebbero essere necessari tra i 5 e i 20 minuti affinché un cluster venga terminato e le risorse allocate rilasciate. Se visualizzi un errore `EC2 QUOTA EXCEEDED` al tentativo di avvio di un cluster, potrebbe essere dovuto al rilascio non ancora avvenuto delle risorse di un cluster terminato recentemente. In questo caso, puoi richiedere che la tua EC2 quota Amazon venga aumentata oppure puoi attendere venti minuti e riavviare il cluster.

Amazon S3 limita il numero di bucket creati su un account a 100. Se il cluster crea un nuovo bucket che supera questo limite, la creazione del bucket avrà esito negativo e potrebbe causare l'esito negativo del cluster.

Controllo della configurazione della sottorete Amazon VPC

Se il cluster è stato avviato in una sottorete Amazon VPC, la sottorete deve essere configurata come descritto in Configurazione della rete in un VPC per Amazon EMR. Occorre inoltre verificare che la sottorete in cui si avvia il cluster disponga di indirizzi IP elastici liberi sufficienti per assegnarne uno a ciascun nodo del cluster.

Riavvio del cluster

Il rallentamento dell'elaborazione può essere causato da una condizione transitoria. Valuta la possibilità di terminare e riavviare il cluster per verificare se tale operazione migliora le prestazioni.

Fase 3: Esamina i file di log per il cluster Amazon EMR

Il passaggio successivo consiste nell'esaminare i file di log per individuare un codice di errore o altra indicazione relativa al problema riscontrato dal cluster. Per informazioni sui file di log disponibili, su dove trovarli e su come visualizzarli, consulta Visualizza i file di log di Amazon EMR.

Potrebbe essere necessaria qualche indagine per determinare l'accaduto. Hadoop esegue il lavoro dei processi nei tentativi di attività su vari nodi del cluster. Amazon EMR può avviare tentativi di attività speculative, terminando gli altri tentativi di attività che non vengono completati per primi. Questo genera un'attività significativa che viene registrata nei file di log del controller, stderr e syslog mentre si verifica. Inoltre, anche se più tentativi di attività vengono eseguiti contemporaneamente, un file di log può visualizzare i risultati solo in modo lineare.

Inizia controllando i log delle operazioni di bootstrap per individuare errori o modifiche impreviste alla configurazione durante l'avvio del cluster. Successivamente, cerca nei log di fase per identificare i processi Hadoop avviati come parte di una fase con errori. Esamina i log dei processi Hadoop per identificare i tentativi di attività non riusciti. Il log dei tentativi di attività conterrà dettagli su ciò che ha causato l'esito negativo di un tentativo di attività.

Le sezioni seguenti descrivono come utilizzare i vari file di log per identificare gli errori nel cluster.

Controllo dei log delle operazioni di bootstrap

Le operazioni di bootstrap eseguono script sul cluster mentre quest'ultimo viene avviato. Vengono comunemente utilizzate per installare software aggiuntivi nel cluster o per modificare le impostazioni

di configurazione rispetto ai valori predefiniti. Il controllo di questi log può fornire informazioni dettagliate sugli errori verificatisi durante la configurazione del cluster e sulle modifiche alle impostazioni di configurazione che potrebbero influire sulle prestazioni.

Controllo dei log della fase

Esistono quattro tipi di log di fase.

- controller: contiene i file generati da Amazon EMR che derivano da errori riscontrati durante il tentativo di eseguire la fase. Se la fase ha esito negativo durante il caricamento, è possibile trovare la traccia di stack in questo log. Gli errori che si verificano durante il caricamento o l'accesso all'applicazione sono spesso descritti nel log, mentre mancano gli errori relativi ai file del mappatore.
- stderr: contiene messaggi di errore che si sono presentati durante l'elaborazione della fase. Gli errori di caricamento delle applicazioni sono spesso descritti in questo log. Talvolta, questo log contiene una traccia di stack.
- stdout: contiene lo stato generato dagli eseguibili del mappatore e del riduttore. Gli errori di caricamento delle applicazioni sono spesso descritti in questo log. Talvolta, questo log contiene messaggi di errore dell'applicazione.
- syslog: contiene log generati da software non Amazon, come Apache e Hadoop. Gli errori di streaming sono spesso descritti in questo log.

Controlla stderr per rilevare errori evidenti. Se stderr visualizza un breve elenco di errori, la fase è giunta a un rapido arresto generando un errore. Questo è spesso causato da un errore nelle applicazioni di mappatore e riduttore eseguite nel cluster.

Esamina le ultime righe di controller e syslog per cercare avvisi di errore o guasti. Segui tutte le notifiche relative alle attività non riuscite, in particolare "Job Failed (Processo non riuscito)".

Controllo dei log del tentativo di attività

Se l'analisi precedente dei log di fase ha rilevato una o più attività non riuscite, esamina i log dei tentativi di attività corrispondenti per ottenere informazioni più dettagliate sugli errori.

Verifica dei log dei daemon Hadoop

In rari casi, Hadoop stesso potrebbe non funzionare. Per comprendere la natura del problema, è necessario esaminare i log di Hadoop. Questi log si trovano all'indirizzo `/var/log/hadoop/` su ogni nodo.

Puoi utilizzare JobTracker i log per mappare un tentativo di operazione fallito al nodo su cui è stato eseguito. Una volta che conosci il nodo associato al tentativo di operazione, puoi controllare lo stato dell' EC2 istanza che ospita quel nodo per vedere se si sono verificati problemi come l'esaurimento della CPU o della memoria.

Fase 4: verifica dello stato del cluster e dell'istanza Amazon EMR

Un cluster Amazon EMR è composto da nodi in esecuzione su istanze Amazon EC2 . Se tali istanze sono limitate dalle risorse (ad esempio, CPU o memoria esaurita), presentano problemi di connettività di rete o vengono terminate, la velocità di elaborazione del cluster diminuisce.

In un cluster sono presenti fino a tre tipi di nodi:

- nodo master: gestisce il cluster. Eventuali problemi di prestazioni nel nodo si ripercuotono all'intero cluster.
- nodi principali: elaborano le attività map-reduce e gestiscono il file system distribuito Hadoop (HDFS). Eventuali problemi di prestazioni in uno di questi nodi possono provocare il rallentamento delle operazioni HDFS nonché dell'elaborazione map-reduce. Puoi aggiungere dei nodi principali a un cluster per migliorare le prestazioni, ma non puoi rimuoverne. Per ulteriori informazioni, consulta Ridimensiona manualmente un cluster Amazon EMR in esecuzione.
- nodi attività: elaborano le attività map-reduce. Sono esclusivamente risorse di calcolo che non archiviano dati. Puoi aggiungere nodi di task a un cluster per migliorare le prestazioni oppure rimuovere quelli non necessari. Per ulteriori informazioni, consulta Ridimensiona manualmente un cluster Amazon EMR in esecuzione.

Quando esamini lo stato di un cluster, devi considerare le prestazioni globali del cluster nonché quelle delle singole istanze. Sono disponibili vari strumenti che puoi utilizzare:

Verifica lo stato del cluster con CloudWatch

Ogni cluster Amazon EMR riporta i parametri a. CloudWatch Questi parametri forniscono informazioni di riepilogo sulle prestazioni del cluster, ad esempio il carico totale, l'utilizzo di HDFS, le attività in esecuzione, le attività rimanenti, i blocchi danneggiati e altro ancora. L'analisi delle CloudWatch metriche offre un quadro generale di ciò che sta accadendo al cluster e può fornire informazioni sulla causa del rallentamento dell'elaborazione. Oltre CloudWatch a utilizzare per analizzare un problema di prestazioni esistente, puoi impostare allarmi che avvisino CloudWatch se si verificano problemi di prestazioni futuri. Per ulteriori informazioni, consulta Monitoraggio dei parametri di Amazon EMR con CloudWatch.

Verifica dell'integrità del processo e di HDFS

Utilizzare Cronologia applicazione nella pagina dei dettagli del cluster per visualizzare informazioni sull'applicazione YARN. Per alcune applicazioni, puoi esaminare ulteriori dettagli e accedere direttamente ai log. Ciò è particolarmente utile per le applicazioni Spark. Per ulteriori informazioni, consulta Visualizza la cronologia delle applicazioni Amazon EMR.

Hadoop fornisce una serie di interfacce Web che puoi utilizzare per visualizzare le informazioni. Per ulteriori informazioni su come accedere a queste interfacce Web, consulta Visualizzazione di interfacce Web ospitate su cluster Amazon EMR.

- JobTracker — fornisce informazioni sullo stato di avanzamento del lavoro elaborato dal cluster. Puoi utilizzare questa interfaccia per identificare quando un processo si blocca.
- HDFS NameNode : fornisce informazioni sulla percentuale di utilizzo di HDFS e sullo spazio disponibile su ciascun nodo. Puoi utilizzare questa interfaccia per identificare quando HDFS è limitato dalle risorse e richiede capacità supplementare.
- TaskTracker — fornisce informazioni sulle attività del lavoro elaborato dal cluster. Puoi utilizzare questa interfaccia per identificare quando un'attività si blocca.

Verifica lo stato delle istanze con Amazon EC2

Un altro modo per cercare informazioni sullo stato delle istanze nel cluster consiste nell'utilizzare la EC2 console Amazon. Poiché ogni nodo del cluster viene eseguito su un' EC2 istanza, puoi utilizzare gli strumenti forniti da Amazon EC2 per verificarne lo stato. Per ulteriori informazioni, consulta Visualizza le istanze di cluster in Amazon EC2.

Fase 5: Verifica della presenza di gruppi sospesi

Un gruppo di istanze è considerato sospeso quando si verificano troppi errori durante il tentativo di avvio di nodi. Ad esempio, in caso di errori ripetuti in nuovi nodi durante l'esecuzione di operazioni di bootstrap, il gruppo di istanze passerà, dopo un determinato periodo di tempo, allo stato `SUSPENDED` anziché tentare reiteratamente di effettuare il provisioning di nuovi nodi.

L'avvio di un nodo potrebbe non riuscire se:

- Hadoop o il cluster è in qualche modo danneggiato e non accetta un nuovo nodo nel cluster.
- Un'operazione di bootstrap non riesce sul nuovo nodo.

- Il nodo non funziona correttamente e la sessione Hadoop non riesce.

Se lo stato di un gruppo di istanze è `SUSPENDED` e quello del cluster è `WAITING`, puoi aggiungere una fase di cluster per ripristinare il numero desiderato di nodi di task e principali. Con l'aggiunta di una fase l'elaborazione del cluster riprende e il gruppo di istanze ritorna allo stato `RUNNING`.

Per ulteriori informazioni su come ripristinare un cluster con stato sospeso, consulta Stato sospeso.

Fase 6: Rivedere le impostazioni di configurazione per il cluster Amazon EMR

Le impostazioni di configurazione specificano dettagli relativi all'esecuzione di un cluster, come il numero di nuovi tentativi per un'attività e la quantità di memoria disponibile per l'ordinamento. Quando avvii un cluster utilizzando Amazon EMR, oltre alle impostazioni di configurazione di Hadoop standard sono disponibili anche impostazioni specifiche di Amazon EMR. Le impostazioni di configurazione sono memorizzate nel nodo master del cluster. Puoi verificare le impostazioni di configurazione per assicurarti che il cluster disponga delle risorse necessarie per una corretta esecuzione.

Amazon EMR definisce le impostazioni di configurazione di Hadoop di default che utilizza per avviare un cluster. I valori sono basati sull'AMI e sul tipo di istanza specificato per il cluster. Puoi modificare i valori di default delle impostazioni di configurazione mediante un'operazione di bootstrap o specificando nuovi valori nei parametri di esecuzione dei processi. Per ulteriori informazioni, consulta Crea azioni di bootstrap per installare software aggiuntivo con un cluster Amazon EMR. Per determinare se un'operazione di bootstrap ha modificato le impostazioni di configurazione, controlla i log delle operazioni di bootstrap.

Amazon EMR registra le impostazioni Hadoop utilizzate per eseguire ogni processo. I dati di log vengono archiviati in un file denominato `job_`*`job-id`*`_conf.xml` nella `/mnt/var/log/hadoop/history/` directory del nodo master, dove *`job-id`* vengono sostituiti dall'identificatore del lavoro. Se hai abilitato l'archiviazione dei log, questi dati vengono copiati in Amazon S3 nella `logs/`*`date`*`/`*`jobflow-id`*`/jobs` cartella, *`date`* dove si trova la data di esecuzione del processo *`jobflow-id`* e l'identificatore del cluster.

Le seguenti impostazioni di configurazione di processi Hadoop sono particolarmente utili per la risoluzione dei problemi di prestazioni. Per ulteriori informazioni sulle impostazioni di configurazione di Hadoop e su come queste influiscono sul comportamento di Hadoop, visita la pagina all'indirizzo http://hadoop.apache.org/docs/.

⚠ Warning

1. L'impostazione di `dfs.replication` su 1 per i cluster con meno di quattro nodi può causare la perdita di dati HDFS in caso di disattivazione anche di un singolo nodo. Ti consigliamo di utilizzare un cluster con almeno quattro nodi principali per i carichi di lavoro di produzione.
2. Amazon EMR non consente ai cluster di dimensionare i nodi principali al di sotto di `dfs.replication`. Ad esempio, se `dfs.replication = 2`, il numero minimo di nodi principali è 2.
3. Quando utilizzi il dimensionamento gestito, il dimensionamento automatico o scegli di dimensionare manualmente il cluster, ti consigliamo di impostare `dfs.replication` su 2 o su un valore superiore.

Impostazione di configurazione	Descrizione
dfs.replication	Il numero di nodi HDFS in cui un singolo blocco (ad esempio un blocco di disco rigido) viene copiato per generare un ambiente di tipo RAID. Determina il numero di nodi HDFS che contengono una copia del blocco.
io.sort.mb	La memoria totale disponibile per l'ordinamento. Il valore deve essere 10x io.sort.factor. Questa impostazione può anche essere utilizzata per calcolare la memoria totale utilizzata dal nodo di task moltiplicando io.sort.mb per mapred.tasktracker.ap.tasks.maximum.
io.sort.spill.percent	Utilizzata durante l'ordinamento. In tal caso il disco viene utilizzato in quanto la memoria di ordinamento assegnata è esaurita.
mapred.child.java.opts	Obsoleta. Utilizzare mapred.map.child.java.opts e mapred.reduce.child.java.opts. Le opzioni Java TaskTracker utilizzate all'avvio di una JVM per eseguire un'attività all'interno. "-Xmx" è un parametro comune per l'impostazione della dimensione massima della memoria.

Impostazione di configurazione	Descrizione
mapred.map.child.java.opts	Le opzioni Java TaskTracker utilizzate all'avvio di una JVM per eseguire un'attività di mappa all'interno. "-Xmx" è un parametro comune per l'impostazione della dimension e massima dell'heap di memoria.
mapred.map.tasks.speculativ e.execution	Determina se i tentativi di attività di mappatura della stessa attività possono essere avviati in parallelo.
mapred.reduce.tasks.specula tive.execution	Determina se i tentativi di attività di riduzione della stessa attività possono essere avviati in parallelo.
mapred.map.max.attempts	Il numero massimo di tentativi per un'attività di mappatura . Se tutti i tentativi non riescono, l'attività di mappatura viene contrassegnata come non riuscita.
mapred.reduce.child.java.opts	Le opzioni Java TaskTracker vengono utilizzate all'avvio di una JVM per un'attività di riduzione da eseguire all'inter no. "-Xmx" è un parametro comune per l'impostazione della dimensione massima dell'heap di memoria.
mapred.reduce.max.attempts	Il numero massimo di tentativi per un'attività di riduzione . Se tutti i tentativi non riescono, l'attività di mappatura viene contrassegnata come non riuscita.
mapred.reduce.slowstart.com pleted.maps	Il numero di attività di mappatura che devono essere completate prima dell'avvio delle attività di riduzione. Se il tempo di attesa è troppo breve, è possibile che vengano generati errori "Too many fetch-failure (Troppi errori di recupero)" nei tentativi.
mapred.reuse.jvm.num.tasks	Un'attività viene eseguita in una singola JVM. Specifica il numero di attività che possono riutilizzare la stessa JVM.
mapred.tasktracker.map.task s.maximum	Il numero massimo di attività che possono essere eseguite in parallelo per nodo di task durante la mappatura.

Impostazione di configurazione	Descrizione
mapred.tasktracker.reduce.tasks.maximum	Il numero massimo di attività che possono essere eseguite in parallelo per nodo di task durante la riduzione.

Se le attività del cluster utilizzano molta memoria, è possibile migliorare le prestazioni utilizzando un minor numero di attività per nodo principale e riducendo la dimensione heap del job tracker.

Fase 7: esaminare i dati di input per il cluster Amazon EMR

Esamina i dati di input. Sono distribuiti in modo uniforme tra i valori chiave? Se i dati sono ripartiti soprattutto su uno o soltanto alcuni valori chiave, il carico di elaborazione può essere mappato a un numero ridotto di nodi, mentre altri sono inattivi. Questa distribuzione squilibrata del lavoro può comportare tempi di elaborazione più elevati.

Un esempio di set di dati squilibrati sarebbe un cluster eseguito per ordinare alfabeticamente le parole, ma con un set di dati che contiene unicamente parole che iniziano con la lettera "a". Dopo la mappatura del lavoro, il nodo che elabora i valori che iniziano con "a" risulterebbe sovraccaricato, mentre i nodi che elaborano le parole che iniziano con altre lettere sarebbero inattivi.

Risolvi i problemi più comuni durante l'utilizzo di Amazon EMR con Lake Formation AWS

Questa sezione descrive il processo di risoluzione dei problemi comuni relativi all'utilizzo di Amazon EMR con AWS Lake Formation.

Accesso al data lake non consentito

Prima di poter analizzare ed elaborare i dati nel data lake, devi attivare esplicitamente il filtro dei dati sui cluster Amazon EMR. Quando l'accesso ai dati ha esito negativo, verrà visualizzato un messaggio generico `Access is not allowed` nell'output delle voci del notebook.

Per istruzioni sull'attivazione del filtro dei dati su Amazon EMR, consulta Attivazione del filtro dei dati su Amazon EMR nella Guida per gli sviluppatori di AWS Lake Formation .

Scadenza della sessione

Il timeout della sessione per gli EMR Notebooks e Zeppelin è controllato dall'impostazione `Maximum CLI/API session duration` del ruolo IAM per Lake Formation. Il valore predefinito per questa impostazione è un'ora. Quando si verifica un timeout di sessione, verrà visualizzato il seguente messaggio nell'output delle voci del notebook quando si tenta di eseguire i comandi SQL Spark.

```
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…
```

Per convalidare la sessione, aggiorna la pagina. Verrà richiesto di eseguire nuovamente l'autenticazione utilizzando il provider di identità e di essere reindirizzati al notebook. Puoi continuare a eseguire query dopo la riautenticazione.

Nessuna autorizzazione per l'utente sulla tabella richiesta

Quando tenti di accedere a una tabella a cui non hai accesso, visualizzerai la seguente eccezione nell'output delle voci del notebook quando tenti di eseguire i comandi SQL Spark.

```
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: …
```

Per accedere alla tabella, devi concedere l'accesso all'utente aggiornando le autorizzazioni associate a questa tabella in Lake Formation.

Interrogazione dei dati tra account condivisi con Lake Formation

Quando utilizzi Amazon EMR per accedere ai dati condivisi con te da un altro account, alcune librerie Spark tenteranno di chiamare l'operazione API `Glue:GetUserDefinedFunctions`. Poiché le versioni 1 e 2 delle autorizzazioni AWS RAM gestite non supportano questa azione, ricevi il seguente messaggio di errore:

```
"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"`

Per risolvere questo errore, l'amministratore del data lake che ha creato la condivisione di risorse deve aggiornare le autorizzazioni AWS RAM gestite allegate alla condivisione di risorse. La versione 3 delle autorizzazioni gestite da AWS RAM consente ai responsabili di eseguire l'operazione `glue:GetUserDefinedFunctions`.

Se crei una nuova condivisione di risorse, Lake Formation applica la versione più recente dell'autorizzazione AWS RAM gestita per impostazione predefinita e non è richiesta alcuna azione da parte tua. Per abilitare l'accesso ai dati tra account diversi per le condivisioni di risorse esistenti, devi aggiornare le autorizzazioni AWS RAM gestite alla versione 3.

Puoi visualizzare le AWS RAM autorizzazioni assegnate alle risorse condivise con te in. AWS RAM Le autorizzazioni seguenti sono incluse nella versione 3:

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
  AWSRAMPermissionGlueAllTablesReadWriteForCatalog
  AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

Per aggiornare la versione delle autorizzazioni AWS RAM gestite delle condivisioni di risorse esistenti

L'amministratore del data lake può aggiornare le autorizzazioni AWS RAM gestite a una versione più recente seguendo le istruzioni contenute nella Guida per l'AWS RAM utente oppure revocare tutte le autorizzazioni esistenti per il tipo di risorsa e concederle nuovamente. Se revochi le autorizzazioni, AWS RAM elimina la condivisione di risorse associata al tipo di risorsa. AWS RAM Quando concedi nuovamente le autorizzazioni, AWS RAM crea nuove condivisioni di risorse allegando la versione più recente delle autorizzazioni gestite. AWS RAM

Inserimento, creazione e modifica di tabelle

L'inserimento, la creazione e la modifica di tabelle nei database protetti da policy Lake Formation non sono supportati. Quando esegui queste operazioni, visualizzerai la seguente eccezione nell'output delle voci del notebook quando tenti di eseguire i comandi SQL Spark:

```
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: …
```

Per ulteriori informazioni, consulta Limitazioni dell'integrazione di Amazon EMR con. AWS Lake Formation

Scrivi applicazioni che avviano e gestiscono i cluster Amazon EMR

Puoi accedere alla funzionalità fornita dall'API Amazon EMR richiamando le funzioni wrapper in una delle. AWS SDKs AWS SDKs Forniscono funzioni specifiche della lingua che racchiudono l'API del servizio Web e semplificano la connessione al servizio Web, gestendo al posto tuo molti dettagli di connessione. Per ulteriori informazioni su come chiamare Amazon EMR utilizzando uno dei seguenti SDKs, consulta. SDKs Usalo per chiamare Amazon EMR APIs

Argomenti

- End-to-end Esempio di codice sorgente Java di Amazon EMR
- Concetti comuni per le chiamate API Amazon EMR
- SDKs Usalo per chiamare Amazon EMR APIs
- Gestione delle Service Quotas di Amazon EMR

> **Important**
>
> Il tasso massimo di richieste per Amazon EMR è una richiesta ogni dieci secondi.

End-to-end Esempio di codice sorgente Java di Amazon EMR

Gli sviluppatori possono chiamare l'API di Amazon EMR utilizzando codice Java personalizzato per eseguire le stesse operazioni possibili con la console Amazon EMR o la CLI. Questa sezione fornisce i end-to-end passaggi necessari per installare AWS Toolkit for Eclipse ed eseguire un esempio di codice sorgente Java completamente funzionale che aggiunge passaggi a un cluster Amazon EMR.

> **Note**
>
> Questo esempio si concentra su Java, ma Amazon EMR supporta anche diversi linguaggi di programmazione con una raccolta di Amazon EMR. SDKs Per ulteriori informazioni, consulta SDKs Usalo per chiamare Amazon EMR APIs.

Questo esempio di codice sorgente Java illustra come eseguire le seguenti attività utilizzando l'API di Amazon EMR:

- Recupera AWS le credenziali e inviale ad Amazon EMR per effettuare chiamate API
- Configurare una nuova fase personalizzata e una nuova fase predefinita
- Aggiunta di nuove fasi a un cluster Amazon EMR esistente
- Recupera la fase IDs del cluster da un cluster in esecuzione

> **Note**
>
> In questo esempio viene illustrato come aggiungere fasi a un cluster esistente e, pertanto, richiede che sia disponibile un cluster attivo sull'account.

Prima di iniziare, installare una versione di Eclipse IDE per Java EE Developers che corrisponde alla piattaforma del computer. Per ulteriori informazioni, vai a Eclipse Downloads (Download di Eclipse).

Quindi, installa il plug-in Database Development per Eclipse.

Installazione del plug-in Database Development per Eclipse

1. Aprire l'IDE Eclipse.
2. Scegliere Help (Guida) e Install New Software (Installa nuovo software).
3. Nel campo Work with: (Utilizza con:), digitare **`http://download.eclipse.org/releases/kepler`** o il percorso che corrisponde al numero di versione dell'IDE Eclipse.
4. Nell'elenco di elementi, scegliere Database Development (Sviluppo di database) e Finish (Fine).
5. Riavviare Eclipse quando richiesto.

Quindi, installa il Toolkit for Eclipse per rendere disponibili i modelli di progetto del codice sorgente preconfigurati.

Installazione del Toolkit for Eclipse

1. Aprire l'IDE Eclipse.
2. Scegliere Help (Guida) e Install New Software (Installa nuovo software).
3. Nel campo Utilizza con:, digitare **`https://aws.amazon.com/eclipse`**.

4. Nell'elenco di elementi, scegli AWS Toolkit for Eclipse e Finish (Fine).
5. Riavviare Eclipse quando richiesto.

Quindi, crea un nuovo progetto AWS Java ed esegui il codice sorgente Java di esempio.

Per creare un nuovo progetto AWS Java

1. Aprire l'IDE Eclipse.
2. Scegliere File, New (Nuovo)e Other (Altro).
3. Nella finestra di dialogo Select a wizard (Seleziona una procedura guidata), scegli AWS Java Project (Progetto Java e Next (Successivo).
4. Nella finestra di dialogo Nuovo progetto AWS Java, nel **`Project name:`** campo, inserisci il nome del tuo nuovo progetto, ad esempio**`EMR-sample-code`**.
5. Scegli Configura AWS account..., inserisci le tue chiavi di accesso pubbliche e private e scegli Fine. Per ulteriori informazioni sulla creazione di chiavi di accesso, consulta la sezione Come ottenere le credenziali di sicurezza in Riferimenti generali di Amazon Web Services.

   > **Note**
   >
   > Si consiglia di non incorporare le chiavi di accesso direttamente nel codice. L'SDK Amazon EMR consente di inserire le chiavi di accesso in percorsi noti in modo da evitare di conservarle nel codice.

6. Nel nuovo progetto Java, fare clic con il pulsante destro del mouse sulla cartella src, quindi scegliere New (Nuovo) e Class (Classe).
7. Nella finestra di dialogo Java Class (Classe Java), nel campo Name (Nome), immettere un nome per la nuova classe, ad esempio **`main`**.
8. Nella sezione Which method stubs would you like to create? (Quali stub del metodo si desidera creare?), scegliere public static void main(String[] args) e Finish (Fine).
9. Immettere il codice sorgente Java all'interno della nuova classe e aggiungere le istruzioni import appropriate per le classi e metodi nell'esempio. Per praticità, il codice sorgente completo è riportato di seguito.

> **Note**
>
> Nel seguente codice di esempio, sostituisci l'esempio cluster ID (JobFlowId)*j-xxxxxxxxxxxx*, con un ID cluster valido nel tuo account che si trova nel AWS Management Console o utilizzando il seguente AWS CLI comando:
>
> ```
> aws emr list-clusters --active | grep "Id"
> ```
>
> Inoltre, sostituisci il percorso Amazon S3 di esempio, *s3://path/to/my/jarfolder*, con un percorso valido per il JAR. Infine, sostituisci il nome classe di esempio, *com.my.Main1*, con il nome corretto della classe nel tuo JAR, se applicabile.

```
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. Scegli Run (Esegui), Run As (Esegui come) e Java Application (Applicazione Java).
11. Se l'esempio viene eseguito correttamente, nella finestra della IDs console IDE di Eclipse viene visualizzato un elenco dei nuovi passaggi. L'output corretto sarà simile al seguente:

```
[s-39BLQZRJB2E5E, s-1L6A4ZU2SAURC]
```

Concetti comuni per le chiamate API Amazon EMR

Quando scrivi un'applicazione che chiama l'API di Amazon EMR, ci sono diversi concetti da tenere a mente quando si chiama una delle funzioni wrapper di un SDK.

Argomenti

- Endpoint per Amazon EMR

- Specifica dei parametri del cluster in Amazon EMR
- Zone di disponibilità in Amazon EMR
- Come utilizzare file e librerie aggiuntivi in cluster Amazon EMR

Endpoint per Amazon EMR

Un endpoint è un URL che rappresenta il punto di partenza per un servizio Web. Ogni richiesta di servizio Web deve contenere un endpoint. L'endpoint specifica la AWS regione in cui i cluster vengono creati, descritti o terminati. Il suo formato è `elasticmapreduce.`*`regionname`*`.amazonaws.com`. Se specifichi l'endpoint generale (`elasticmapreduce.amazonaws.com`), la richiesta viene indirizzata da Amazon EMR a un endpoint nella regione predefinita. Per account creati a partire dall'8 marzo 2013, la regione predefinita è us-west-2; per i vecchi account, la regione predefinita è us-east-1.

Per ulteriori informazioni sugli endpoint per Amazon EMR, consulta la sezione Regioni ed endpoint in Riferimenti generali di Amazon Web Services.

Specifica dei parametri del cluster in Amazon EMR

I `Instances` parametri consentono di configurare il tipo e il numero di EC2 istanze per creare nodi per elaborare i dati. Hadoop distribuisce l'elaborazione dei dati su più nodi del cluster. Il nodo master serve a monitorare l'integrità dei nodi principali e di task ed esegue il polling dei nodi per lo stato del risultato del processo. I nodi principali e di task eseguono l'elaborazione effettiva dei dati. Se si dispone di un cluster a nodo singolo, il nodo svolge la funzione di nodo master e principale.

Il parametro `KeepJobAlive` in una richiesta `RunJobFlow` determina se terminare il cluster quando esaurisce le fasi del cluster da eseguire. Impostare questo valore su `False` quando l'esecuzione del cluster è quella prevista. Durante la risoluzione dei problemi del flusso di elaborazione e l'aggiunta di fasi mentre l'esecuzione del cluster è sospesa, è opportuno impostare il valore su `True`. Questo consente di ridurre il tempo e le spese di caricamento dei risultati in Amazon Simple Storage Service (Amazon S3), solo per ripetere il processo dopo la modifica di una fase per riavviare il cluster.

In caso `KeepJobAlive` `true` affermativo, dopo aver completato con successo il funzionamento del cluster, è necessario inviare una `TerminateJobFlows` richiesta o il cluster continuerà a funzionare e generare AWS addebiti.

Per ulteriori informazioni sui parametri che sono unici per`RunJobFlow`, vedere RunJobFlow. Per ulteriori informazioni sui parametri generici nella richiesta, consulta la sezione relativa ai Parametri di richiesta comuni.

Zone di disponibilità in Amazon EMR

Amazon EMR utilizza EC2 le istanze come nodi per elaborare i cluster. Queste EC2 istanze hanno sedi composte da zone di disponibilità e regioni. Le regioni sono disperse e situate in aree geografiche separate. Le zone di disponibilità sono ubicazioni distinte all'interno di una Regione isolata dai guasti che si verificano in altre zone di disponibilità. Ogni zona di disponibilità offe una connettività di rete economica, a bassa latenza ad altre zone di disponibilità nella stessa Regione. Per un elenco delle regioni e degli endpoint per Amazon EMR, consulta la sezione Regioni ed endpoint in Riferimenti generali di Amazon Web Services.

Il parametro `AvailabilityZone` specifica il percorso generale del cluster. Questo parametro è facoltativo e, in generale, ne sconsigliamo l'utilizzo. Quando `AvailabilityZone` non è specificato, Amazon EMR sceglie automaticamente il valore `AvailabilityZone` ottimale per il cluster. Questo parametro può essere utile se desideri co-individuare le tue istanze con altre istanze in esecuzione esistenti e il cluster deve leggere o scrivere dati di tali istanze. Per ulteriori informazioni, consulta la Amazon EC2 User Guide.

Come utilizzare file e librerie aggiuntivi in cluster Amazon EMR

Talvolta potrebbe essere necessario utilizzare file aggiuntivi o librerie personalizzate con applicazioni mappatore o riduttore. Ad esempio, potrebbe essere necessario utilizzare una libreria che consente di convertire un file PDF in testo normale.

Per memorizzare nella cache un file utilizzato dal mappatore o riduttore durante lo streaming Hadoop

- Nel campo `args` JAR, aggiungere il seguente argomento:

  ```
  -cacheFile s3://bucket/path_to_executable#local_path
  ```

  Il file, `local_path`, si trova nella directory di lavoro del mappatore, che potrebbe fare riferimento al file.

SDKs Usalo per chiamare Amazon EMR APIs

Argomenti

- Utilizzo di AWS SDK per Java per creare un cluster Amazon EMR

AWS SDKs Forniscono funzioni che racchiudono l'API e si occupano di molti dettagli della connessione, come il calcolo delle firme, la gestione dei nuovi tentativi di richiesta e la gestione degli errori. SDKs Inoltre contengono codice di esempio, tutorial e altre risorse per aiutarti a iniziare a scrivere applicazioni che chiamano. AWS La chiamata alle funzioni wrapper in un SDK può semplificare notevolmente il processo di scrittura di un'applicazione. AWS

Per ulteriori informazioni su come scaricare e utilizzare AWS SDKs, consulta la SDKs sezione Tools for Amazon Web Services.

Utilizzo di AWS SDK per Java per creare un cluster Amazon EMR

AWS SDK per Java Fornisce tre pacchetti con funzionalità Amazon EMR:

- com.amazonaws.services.elasticmapreduce
- com.amazonaws.services.elasticmapreduce.model
- com.amazonaws.services.elasticmapreduce.util

Per ulteriori informazioni su questi pacchetti, consulta la Guida di riferimento alle API di AWS SDK per Java.

L'esempio seguente illustra come SDKs possono semplificare la programmazione con Amazon EMR. Il codice di esempio sottostante utilizza l'oggetto `StepFactory`, una classe helper per la creazione di tipi di fase Amazon EMR comuni, per creare un cluster Hive interattivo con il debug abilitato.

```
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());

 }

}
```

Come minimo, è necessario assegnare un ruolo di servizio e un ruolo jobflow corrispondenti rispettivamente a EMR_ DefaultRole e EMR_ _. EC2 DefaultRole È possibile farlo richiamando questo comando per lo stesso account. AWS CLI Innanzitutto, verifica se i ruoli esistono già:

```
aws iam list-roles | grep EMR
```

Se esistono, verranno visualizzati sia il profilo dell'istanza (EC2EMR_ _DefaultRole) che il ruolo di servizio (EMR_DefaultRole):

```
"RoleName": "EMR_DefaultRole",
            "Arn": "arn:aws:iam::AccountID:role/EMR_DefaultRole"
            "RoleName": "EMR_EC2_DefaultRole",
            "Arn": "arn:aws:iam::AccountID:role/EMR_EC2_DefaultRole"
```

Se i ruoli predefiniti non esistono, puoi utilizzare il seguente comando per crearli:

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

Gestione delle Service Quotas di Amazon EMR

Argomenti

- Cosa sono le Service Quotas di Amazon EMR
- Come gestire le Service Quotas di Amazon EMR
- Quando configurare gli eventi EMR in CloudWatch

Gli argomenti di questa sezione descrivono le quote del servizio EMR (precedentemente denominate limiti di servizio), come gestirle e quando è vantaggioso utilizzare CloudWatch gli eventi anziché le quote di servizio per monitorare i cluster e attivare azioni. AWS Management Console

Cosa sono le Service Quotas di Amazon EMR

L' AWS account dispone di quote di servizio predefinite, note anche come limiti, per ogni servizio. AWS Il servizio EMR presenta due tipi di limiti:

- Limiti sulle risorse: è possibile utilizzare EMR per creare EC2 risorse. Tuttavia, queste EC2 risorse sono soggette a quote di servizio. I limiti delle risorse in questa categoria sono:
  - Il numero massimo di cluster attivi che possono essere eseguiti nello stesso momento.
  - Il numero massimo di istanze attive per gruppo di istanze.
- Limiti APIs: quando si utilizza EMR APIs, i due tipi di limitazioni sono:
  - Limite di frammentazione: questo è il numero massimo di chiamate API che puoi effettuare contemporaneamente. Ad esempio, il numero massimo di richieste AddInstanceFleet API che è possibile effettuare al secondo è impostato su 5 calls/second come impostazione predefinita. Ciò implica che il limite di burst dell' AddInstanceFleet API è di 5 chiamate/secondo o che, in qualsiasi momento, è possibile effettuare al massimo 5 AddInstanceFleet chiamate API. Tuttavia, dopo aver utilizzato il limite di frammentazione, le chiamate successive sono limitate dal limite di frequenza.
  - Limite di frequenza: questa è la velocità di rifornimento della capacità di espansione dell'API. Ad esempio, la frequenza di rifornimento delle AddInstanceFleet chiamate è impostata a 0,5 come impostazione predefinita. calls/second Ciò significa che dopo aver raggiunto il limite di frammentazione, è necessario attendere almeno 2 secondi (0,5 chiamate/secondo x 2 secondi = 1 chiamata) per effettuare la chiamata API. Se si effettua una chiamata prima di questo valore temporale, si è limitati dal servizio Web EMR. In qualsiasi momento, è possibile effettuare solo un numero di chiamate pari alla capacità di espansione senza alcuna limitazione. Per

ogni ulteriore secondo di attesa, la capacità di espansione aumenta di 0,5 chiamate fino a raggiungere il limite massimo di 5, ossia il limite di frammentazione.

Come gestire le Service Quotas di Amazon EMR

Service Quotas è una AWS funzionalità che puoi utilizzare per visualizzare e gestire le quote o i limiti del servizio Amazon EMR da una posizione centrale, utilizzando l'API o il AWS Management Console. AWS CLI Per ulteriori informazioni sulla visualizzazione delle quote e sulla richiesta di un aumento, consulta Service Quotas di AWS in Riferimenti generali di Amazon Web Services.

Important

Le quote di servizio possono essere applicate indipendentemente l'una dall'altra e, a volte, i loro effetti possono sovrapporsi. Alcune hanno un ambito specifico APIs, mentre altre impongono limiti a livello di account. In particolare, a livello di account vengono applicate le seguenti quote:

- La velocità massima di rifornimento del bucket per tutte le operazioni EMR
- Il numero massimo di richieste API che puoi effettuare al secondo

Tieni presente che se richiedi con successo un aumento di quota specifico e continui a riscontrare limitazioni, ad esempio per le richieste API, è possibile che una di queste quote a livello di account continui a limitare le frequenze di chiamata API. Per la risoluzione dei problemi, puoi utilizzare la console Service Quotas, disponibile a https://console.aws.amazon.com/servicequotas/casa, per verificare i limiti di quota e richiedere aumenti. Se effettui richieste di aumento e continui a riscontrare limitazioni, contatta l'assistenza. È inoltre possibile trovare i limiti e le descrizioni predefiniti per le quote del servizio EMR nella guida di riferimento AWS generale.

Per alcuni APIs, l'organizzazione di un CloudWatch evento potrebbe essere un'opzione migliore rispetto all'aumento delle quote di servizio. Puoi anche risparmiare tempo impostando allarmi e attivando l' CloudWatch aumento delle richieste in modo proattivo, prima di raggiungere la quota di servizio. Per ulteriori dettagli, consulta Quando configurare gli eventi EMR in CloudWatch.

Quando configurare gli eventi EMR in CloudWatch

Per alcuni sondaggi APIs, ad esempio, e DescribeCluster DescribeStep ListClusters, l'impostazione di un CloudWatch evento può ridurre i tempi di risposta alle modifiche e liberare le quote di servizio. Se hai impostato una funzione Lambda che si attiva al variare dello stato di un cluster, ad esempio quando una fase viene completata o il cluster viene terminato, puoi utilizzare tale attivazione per avviare l'operazione successiva nel flusso di lavoro anziché attendere il polling successivo. Altrimenti, se disponi di EC2 istanze Amazon dedicate o funzioni Lambda che controllano costantemente le modifiche dell'API EMR, non solo sprecheresti risorse di elaborazione, ma potresti anche raggiungere la tua quota di servizio.

Di seguito sono riportati alcuni casi che mostrano come trarre vantaggio dal passaggio a un'architettura basata su eventi.

Caso 1: sondaggio EMR DescribeCluster utilizzando chiamate API per il completamento delle fasi

Example Sondaggio EMR DescribeCluster tramite chiamate API per il completamento delle fasi

Uno schema comune consiste nell'inviare una fase a un cluster in esecuzione e interrogare Amazon EMR per verificare lo stato della fase, in genere utilizzando DescribeCluster o. DescribeStep APIs Questa attività può anche essere eseguita con un ritardo minimo collegandosi all'evento di modifica dello stato della fase Amazon EMR.

Questo evento include le seguenti informazioni nel relativo payload.

```
{
  "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."
  }
}
```

Nella mappa di dettaglio, una funzione Lambda potrebbe analizzare "state", "stepId" o "clusterId" per trovare informazioni pertinenti.

Caso 2: Polling di EMR per i cluster disponibili per l'esecuzione di flussi di lavoro

Example Polling di EMR per i cluster disponibili per l'esecuzione di flussi di lavoro

Un modello utile per i clienti che eseguono più cluster consiste nell'eseguire flussi di lavoro su cluster non appena sono disponibili. Se ci sono molti cluster in esecuzione ed è necessario eseguire un flusso di lavoro su un cluster in attesa, uno schema potrebbe essere quello di interrogare EMR utilizzando DescribeCluster o chiamate ListClusters API per i cluster disponibili. Un altro modo per ridurre il ritardo nel sapere quando un cluster è pronto per una fase è elaborare l'evento di modifica dello stato del cluster Amazon EMR.

Questo evento include le seguenti informazioni nel relativo payload.

```
{
  "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 ..."
  }
}
```

Per questo evento, potresti impostare una funzione Lambda per inviare immediatamente un flusso di lavoro in attesa a un cluster non appena il suo stato cambia in WAITING (IN ATTESA).

Caso 3: Polling di EMR per la terminazione del cluster

Example Polling di EMR per la terminazione del cluster

Un modello comune tra i clienti che eseguono molti cluster EMR è il polling di Amazon EMR per cluster terminati in modo che il lavoro non venga più inviato a esso. Puoi implementare questo modello con le chiamate DescribeCluster e ListClusters API o utilizzando l'evento Amazon EMR Cluster State Change in.

In seguito alla terminazione del cluster, l'evento emesso ha un aspetto simile all'esempio seguente.

```
{
  "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 sezione "Detail (Dettaglio)" del payload include il clusterId e lo stato su cui è possibile agire.

AWS Glossario

Per la AWS terminologia più recente, consultate il AWS glossario nella sezione Reference.Glossario AWS

Cronologia dei documenti

La tabella seguente descrive le principali modifiche alla Amazon EMR Management Guide dall'ultima versione di Amazon EMR. Per ulteriori informazioni su versioni rilasciate specifiche di Amazon EMR, consulta Informazioni sulle versioni di Amazon EMR.

Modifica	Descrizione	Data di rilascio
Aggiornamento della policy gestita	Aggiornamenti Amazon EMR alle policy AWS gestite — Aggiornamenti Amazon EMR alle policy AWS gestite — Autorizzazioni aggiuntive per Amazon Policy_v2. EMRService	4 marzo 2025
Versione iniziale	Versione iniziale della Amazon EMR Managemen t Guide per le versioni 4.x e successive di Amazon EMR. Per ulteriori informazioni sulle versioni precedenti di Amazon EMR, consulta la Amazon EMR Developer Guide.	24 luglio 2015

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.