



Guida per gli sviluppatori

AWS Lake Formation



AWS Lake Formation: Guida per gli sviluppatori

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 dei rispettivi proprietari, che possono o meno essere affiliati, collegati o sponsorizzati da Amazon.

Table of Contents

Che cos'è AWS Lake Formation?	1
Caratteristiche di Lake Formation	2
Inserimento e gestione dei dati	2
Gestione della sicurezza	3
Inserisci i tuoi dati nel Data Catalog	4
Come funziona	6
Flusso di lavoro per la gestione delle autorizzazioni di Lake Formation	6
Autorizzazioni per i metadati	8
Gestione degli accessi allo storage	11
Condivisione dei dati tra account in Lake Formation	13
Componenti di Lake Formation	14
Console Lake Formation	14
API e interfaccia a riga di comando di Lake Formation	14
Altri AWS servizi	14
Terminologia dei Lake Formation	15
Lago di dati	15
Accesso ai dati	15
Modalità di accesso ibrida	15
Piano	16
Flusso di lavoro	16
Catalogo dati	16
Dati sottostanti	17
Principale	17
Amministratore del data lake	17
AWS integrazioni di servizi con Lake Formation	17
Risorse aggiuntive per Lake Formation	19
Blog	19
Conferenze tecniche e webinar	20
Architettura moderna	20
Risorse Data Mesh	20
Guide alle migliori pratiche	20
Guida introduttiva a Lake Formation	21
Nozioni di base	22
Completa le attività AWS di configurazione iniziali	22

Iscriviti per un Account AWS	22
Crea un utente con accesso amministrativo	23
Concessione dell'accesso programmatico	24
Configurare AWS Lake Formation	26
Configura le risorse di Lake Formation utilizzando il CloudFormation modello	27
Crea un amministratore del data lake	28
Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida	33
Assegna le autorizzazioni agli utenti di Lake Formation	35
Configura una posizione Amazon S3 per il tuo data lake	36
(Facoltativo) Impostazioni di filtraggio dei dati esterni	37
(Facoltativo) Concedi l'accesso alla chiave di crittografia Data Catalog	38
(Facoltativo) Crea un ruolo IAM per i flussi di lavoro	38
Aggiornamento delle autorizzazioni per AWS Glue i dati al modello Lake Formation	40
Informazioni sulle autorizzazioni predefinite	40
Elenca le autorizzazioni esistenti	42
Configura i permessi di Lake Formation	44
Concedi agli utenti le autorizzazioni IAM	45
Passa al modello di autorizzazioni Lake Formation	46
Fase 5: Proteggi le nuove risorse del Data Catalog	49
Fase 6: Offri agli utenti una nuova policy IAM	49
Fase 7: Pulisci le policy IAM esistenti	51
Configurazione degli endpoint Amazon VPC (AWS PrivateLink)	51
Considerazioni sugli endpoint VPC di Lake Formation	52
Creazione di un endpoint VPC di interfaccia per Lake Formation	52
Creazione di una policy sugli endpoint VPC per Lake Formation	52
Esercitazioni	54
Creare un data lake da una fonte AWS CloudTrail	55
Destinatari principali	56
Prerequisiti	57
Fase 1: Creare un utente analista di dati	58
Passaggio 2: Aggiungere le autorizzazioni per leggere i AWS CloudTrail registri al ruolo del flusso di lavoro	59
Fase 3: creare un bucket Amazon S3 per il data lake	59
Fase 4: Registrare un percorso Amazon S3	60
Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati	60
Fase 6: Creare un database nel Data Catalog	61

Passaggio 7: concedere le autorizzazioni per i dati	61
Fase 8: Utilizzare un blueprint per creare un flusso di lavoro	62
Passaggio 9: Esegui il flusso di lavoro	63
Fase 10: concedere SELECT sui tavoli	64
Passaggio 11: interrogare il data lake utilizzando Amazon Athena	64
Creazione di un data lake da una sorgente JDBC	65
Destinatari principali	66
Prerequisiti	67
Fase 1: Creare un utente analista di dati	67
Fase 2: Creare una connessione in AWS Glue	68
Fase 3: creare un bucket Amazon S3 per il data lake	69
Fase 4: Registrare un percorso Amazon S3	69
Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati	70
Fase 6: Creare un database nel Data Catalog	70
Passaggio 7: concedere le autorizzazioni per i dati	70
Fase 8: Utilizzare un blueprint per creare un flusso di lavoro	71
Passaggio 9: Esegui il flusso di lavoro	73
Fase 10: concedere SELECT sui tavoli	73
Passaggio 11: interrogare il data lake utilizzando Amazon Athena	74
Passaggio 12: interroga i dati nel data lake utilizzando Amazon Redshift Spectrum	75
Passaggio 13: concedere o revocare le autorizzazioni di Lake Formation utilizzando Amazon Redshift Spectrum	79
Configurazione delle autorizzazioni per i formati di tabelle aperte in Lake Formation	79
Destinatari principali	80
Prerequisiti	81
Fase 1: Fornisci le tue risorse	82
Passaggio 2: imposta le autorizzazioni per una tabella Iceberg	84
Passaggio 3: configura le autorizzazioni per una tabella Hudi	90
Passaggio 4: configura le autorizzazioni per una tabella Delta Lake	93
Fase 5: Pulizia delle risorse AWS	95
Gestione di un data lake utilizzando il controllo degli accessi basato su tag	95
Destinatari principali	97
Prerequisiti	98
Fase 1: Fornisci le tue risorse	98
Fase 2: Registrare la posizione dei dati, creare un'ontologia LF-tag e concedere le autorizzazioni	99

Fase 3: Creare database Lake Formation	103
Passaggio 4: concedere le autorizzazioni per la tabella	113
Passaggio 5: esegui una query in Amazon Athena per verificare le autorizzazioni	115
Fase 6: Pulizia delle risorse AWS	116
Protezione dei data lake con il controllo degli accessi a livello di riga	116
Destinatari principali	117
Prerequisiti	118
Fase 1: Fornisci le tue risorse	119
Fase 2: Interrogazione senza filtri di dati	120
Passaggio 3: configura i filtri di dati e concedi le autorizzazioni	122
Passaggio 4: Interrogazione con filtri di dati	124
Fase 5: Pulire le AWS risorse	125
Condividi i tuoi dati in modo sicuro con Lake Formation	125
Destinatari principali	126
Configura le impostazioni di Lake Formation	128
Passaggio 1: Fornisci le tue risorse utilizzando modelli AWS CloudFormation	130
Fase 2: Prerequisiti per la condivisione tra account di Lake Formation	133
Fase 3: Implementare la condivisione tra account utilizzando il metodo di controllo degli accessi basato su tag	135
Fase 4: Implementazione del metodo di risorsa denominato	141
Passaggio 5: Pulisci AWS le risorse	145
Condivisione di risorse del Data Catalog con risorse esterne Account AWS utilizzando un controllo degli accessi granulare	146
Destinatari principali	147
Prerequisiti	148
Fase 1: Fornire un accesso dettagliato a un altro account	149
Passaggio 2: fornire un accesso dettagliato a un utente nello stesso account	150
Accedere ai permessi di Lake Formation	152
Panoramica delle autorizzazioni di Lake Formation	153
Metodi per il controllo granulare degli accessi	155
Controllo dell'accesso ai metadati	158
Controllo sottostante dell'accesso ai dati	162
Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM	168
AWS Lake Formation persone	168
AWS politiche gestite per Lake Formation	170
Autorizzazioni suggerite da Personas	178

Modifica delle impostazioni predefinite per il data lake	189
Autorizzazioni implicite di Lake Formation	192
Riferimento alle autorizzazioni di Lake Formation	194
Autorizzazioni Lake Formation per tipo di risorsa	194
Comandi di concessione e AWS CLI revoca di Lake Formation	197
Autorizzazioni Lake Formation	202
Integrazione di IAM Identity Center	217
Prerequisiti per l'integrazione di IAM Identity Center con Lake Formation	218
Connessione di Lake Formation con IAM Identity Center	222
Aggiornamento dell'integrazione con IAM Identity Center	225
Eliminazione di una connessione Lake Formation con IAM Identity Center	226
Concessione di autorizzazioni a utenti e gruppi	227
Inclusione del contesto utente di IAM Identity Center nei log CloudTrail	230
Aggiungere una posizione Amazon S3 al tuo data lake	232
Requisiti per i ruoli utilizzati per registrare le sedi	233
Registrazione di una sede Amazon S3	241
Registrazione di una posizione Amazon S3 crittografata	244
Registrazione di una sede Amazon S3 in un altro account AWS	249
Registrazione di una posizione Amazon S3 crittografata tra più account AWS	252
Annullamento della registrazione di una sede Amazon S3	256
Modalità di accesso ibrida	257
Casi d'uso comuni della modalità di accesso ibrido	259
Come funziona la modalità di accesso ibrida	260
Configurazione della modalità di accesso ibrida: scenari comuni	262
Rimozione di principi e risorse dalla modalità di accesso ibrida	282
Visualizzazione dei principali e delle risorse in modalità di accesso ibrida	283
Risorse aggiuntive	284
Creazione di oggetti in AWS Glue Data Catalog	284
Creazione di un catalogo	285
Creazione di un database	285
Creazione di tabelle	286
Visualizzazioni del Building Data Catalog	295
Importazione di dati tramite flussi di lavoro	343
Progetti e flussi di lavoro	344
Creazione di un flusso di lavoro	345
Esecuzione di un workflow	349

Bringing your data into the Data Catalog	351
Federazione dei cataloghi	353
Caratteristiche di Catalog Federation	353
Federati a Snowflake	354
Federare a Databricks	362
Bringing Amazon Redshift data into the Data Catalog	370
Vantaggi principali	372
Ruoli e responsabilità	372
Prerequisiti	373
Creazione di cataloghi federati Amazon Redshift	380
Visualizzazione degli oggetti del catalogo	389
Aggiornamento di un catalogo federato	391
Accesso a un catalogo federato condiviso	393
Eliminazione di un catalogo federato	398
Interrogazione di cataloghi federati	399
Risorse aggiuntive	399
Federazione in fonti di dati esterne	399
Flusso di lavoro	400
Prerequisiti	401
Creazione di un catalogo federato	405
Visualizzazione degli oggetti del catalogo	410
Eliminazione di un catalogo federato	413
Interrogazione di cataloghi federati	414
Risorse aggiuntive	414
Creazione di un catalogo di tabelle Amazon S3 nel Data Catalog	414
Come funziona l'integrazione tra Data Catalog e Lake Formation	415
Prerequisiti	416
Abilitazione dell'integrazione con Amazon S3 Tables	419
Creazione di database e tabelle	422
Registrazione di un bucket da tavolo Amazon S3 in un altro account AWS	425
Concessione di autorizzazioni	428
Creazione di un catalogo gestito da Amazon Redshift	431
Gestione delle autorizzazioni per i dati in un datashare Amazon Redshift	435
Prerequisiti	437
Configurazione delle autorizzazioni per le condivisioni di dati Amazon Redshift	437
Interrogazione di database federati	442

Gestione delle autorizzazioni sui set di dati che utilizzano metastore esterni	442
Flusso di lavoro	445
Prerequisiti	446
Connessione del Data Catalog a un metastore Hive esterno	448
Risorse aggiuntive	451
Gestione delle autorizzazioni di Lake Formation	452
Concessione delle autorizzazioni per la localizzazione dei dati	452
Concessione delle autorizzazioni per la localizzazione dei dati (stesso account)	453
Concessione delle autorizzazioni per la localizzazione dei dati (account esterno)	456
Concessione delle autorizzazioni su una posizione di dati condivisa con l'account	459
Concessione delle autorizzazioni relative ai dati	460
Autorizzazioni IAM necessarie per concedere le autorizzazioni Lake Formation	462
Utilizzo del metodo delle risorse denominate	465
Controllo degli accessi basato su tag	491
Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC	553
Controllo dell'accesso basato sugli attributi	560
Scenario di esempio di autorizzazioni	564
Filtraggio dei dati e sicurezza a livello di cella	566
Filtri di dati	568
Supporto PartiQL nelle espressioni di filtro di riga	572
Autorizzazioni necessarie per eseguire interrogazioni su tabelle con filtraggio a livello di cella	575
Gestione dei filtri di dati	575
Visualizzazione delle autorizzazioni per database e tabelle	591
Revoca delle autorizzazioni tramite la console	595
Condivisione dei dati tra account	596
Prerequisiti	599
Aggiornamento delle impostazioni della versione di condivisione dei dati tra account	603
Condivisione delle tabelle e dei database del Data Catalog tra i nostri principali IAM provenienti Account AWS da account esterni	609
Concessione delle autorizzazioni su un database o una tabella condivisa con il tuo account	615
Concessione delle autorizzazioni per i collegamenti alle risorse	617
Accesso ai dati sottostanti di una tabella condivisa	620
Registrazione su più account CloudTrail	622

Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation	627
Visualizzazione di tutte le sovvenzioni tra account utilizzando l'operazione API GetResourceShares	629
Accesso e visualizzazione di tabelle e database di Data Catalog condivisi	631
Accettazione di un invito alla condivisione di AWS RAM risorse	632
Visualizzazione di tabelle e database condivisi del Data Catalog	634
Creazione di collegamenti alle risorse	636
Come funzionano i collegamenti alle risorse	637
Creazione di un link di risorsa a una tabella condivisa	639
Creazione di un collegamento di risorsa a un database condiviso	642
Gestione dei link alle risorse in AWS Glue APIs	647
Accesso alle tabelle in tutte le regioni	651
Flussi di lavoro	653
Configurazione dell'accesso alle tabelle tra regioni	657
Sicurezza	660
Protezione dei dati	660
Crittografia dei dati inattivi	661
Sicurezza dell'infrastruttura	662
Prevenzione del confused deputy tra servizi	663
Registrazione degli eventi di sicurezza AWS Lake Formation	664
Integrazione con Lake Formation	665
Utilizzo dell'integrazione delle applicazioni Lake Formation	665
Come funziona l'integrazione delle applicazioni Lake Formation	666
Ruoli e responsabilità nell'integrazione delle applicazioni Lake Formation	668
Lake Formation flusso di lavoro per le operazioni API di integrazione delle applicazioni	669
Registrazione di un motore di query di terze parti	671
Abilitazione delle autorizzazioni per un motore di query di terze parti per chiamare le operazioni dell'API di integrazione delle applicazioni	672
Integrazione delle applicazioni per l'accesso completo alle tabelle	677
Collaborazione con altri AWS servizi	680
Amazon Athena	684
Support per i formati di tabelle transazionali	686
Risorse aggiuntive	689
Amazon Redshift Spectrum	689
Support per tipi di tabelle transazionali	690

Risorse aggiuntive	691
AWS Glue	692
Support per tipi di tabelle transazionali	693
Risorse aggiuntive	694
Amazon EMR	694
Support per i formati di tabelle transazionali	695
Risorse aggiuntive	696
Quick Suite	696
Risorse aggiuntive	697
AWS CloudTrail Lago	697
Registrazione delle chiamate API AWS Lake Formation utilizzando AWS CloudTrail	698
Informazioni su Lake Formation in CloudTrail	698
Comprendere gli eventi di Lake Formation	699
Le migliori pratiche, considerazioni e limitazioni di Lake Formation	702
Buone pratiche e considerazioni sulla condivisione dei dati tra account	702
Limitazioni dei ruoli legati ai servizi	705
Limitazioni di accesso ai dati tra regioni	706
Visualizzazioni, considerazioni e limitazioni di Data Catalog	706
Limitazioni del filtraggio dei dati	708
Note e restrizioni per il filtraggio a livello di colonna	708
Limitazioni del filtraggio a livello di cella	710
Considerazioni e limitazioni della modalità di accesso ibrido	711
Limitazioni per l'inserimento dei dati del data warehouse di Amazon Redshift nel AWS Glue	
Data Catalog	713
Limitazioni dell'integrazione del catalogo S3 Tables	714
I metadati di Hive archiviano: considerazioni e limitazioni sulla condivisione dei dati	715
Limitazioni della condivisione dei dati di Amazon Redshift	717
Limitazioni dell'integrazione di IAM Identity Center	718
Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation	719
Considerazioni, limitazioni e aree supportate sul controllo degli accessi basato sugli attributi ...	722
Risoluzione dei problemi relativi a Lake Formation	724
Risoluzione dei problemi generali	724
Errore: autorizzazioni Lake Formation insufficienti su <Amazon S3 location>	724
Errore: «Autorizzazioni per la chiave di crittografia insufficienti per l'API Glue»	725
La mia query Amazon Athena o quella di Amazon Redshift che utilizza i manifesti non riesce	725

Errore: «Autorizzazioni di Lake Formation insufficienti: è richiesta la creazione di un tag sul catalogo»	725
Errore durante l'eliminazione di amministratori di data lake non validi	725
Risoluzione dei problemi di accesso tra account	725
Ho concesso un'autorizzazione Lake Formation per più account ma il destinatario non può vedere la risorsa	726
I responsabili dell'account del destinatario possono vedere la risorsa Data Catalog ma non possono accedere ai dati sottostanti	727
Errore: «Associazione non riuscita perché il chiamante non era autorizzato» quando si accetta un AWS RAM invito alla condivisione delle risorse	727
Errore: «Non autorizzato a concedere le autorizzazioni per la risorsa»	728
Errore: «Accesso negato per recuperare le informazioni AWS sull'organizzazione»	728
Errore: «Organizzazione <organization-ID>non trovata»	728
Errore: «Autorizzazioni Lake Formation insufficienti: combinazione non valida»	728
ConcurrentModificationException sulle richieste di concessione/revoca ad account esterni ..	728
Errore durante l'utilizzo di Amazon EMR per accedere ai dati condivisi tramite più account ..	729
Risoluzione dei problemi relativi a blueprint e flussi di lavoro	730
<role-ARN>Il mio progetto non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:»	730
Il mio flusso di lavoro non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:<role-ARN>»	731
Un crawler del mio flusso di lavoro non è riuscito con il messaggio «La risorsa non esiste o il richiedente non è autorizzato ad accedere alle autorizzazioni richieste»	731
Un crawler del mio flusso di lavoro non è riuscito con il messaggio «Si è verificato un errore (AccessDeniedException) durante la chiamata dell' CreateTable operazione...»	731
Problemi noti per AWS Lake Formation	731
Limitazione al filtraggio dei metadati delle tabelle	732
Problema con la ridenominazione di una colonna esclusa	733
Problema con l'eliminazione delle colonne nelle tabelle CSV	733
Le partizioni delle tabelle devono essere aggiunte in un percorso comune	733
Problema con la creazione di un database durante la creazione del flusso di lavoro	733
Problema con l'eliminazione e la successiva creazione di un utente	734
Le operazioni dell'API Data Catalog non aggiornano il valore del parametro IsRegisteredWithLakeFormation	734
Le operazioni di Lake Formation non supportano AWS Glue Schema Registry	734
Messaggio di errore aggiornato	735

Lake Formation API	736
Autorizzazioni	737
— operazioni —	737
— tipi di dati —	737
Impostazioni Data Lake	738
— operazioni —	738
— tipi di dati —	738
Integrazione con IAM Identity Center	738
— operazioni —	738
— tipi di dati —	738
Modalità di accesso ibrida	739
— operazioni —	739
— tipi di dati —	737
Vendita di credenziali	739
— operazioni —	739
— tipi di dati —	740
Assegnazione di tag	740
— operazioni —	740
— tipi di dati —	741
Filtro dati APIs	741
— operazioni —	741
— tipi di dati —	741
Tipi di dati comuni	741
ErrorDetail	742
Modelli di stringa	742
Regioni supportate	743
Disponibilità generale	743
AWS GovCloud (US)	743
Ottimizzazione delle transazioni e dello storage	743
Cronologia dei documenti	746
AWS Glossario	761
.....	dcclxii

Che cos'è AWS Lake Formation?

Benvenuto nella Guida per AWS Lake Formation gli sviluppatori.

AWS Lake Formation ti aiuta a governare, proteggere e condividere a livello globale i dati per l'analisi e l'apprendimento automatico. Con Lake Formation, puoi gestire un controllo granulare degli accessi per i dati del tuo data lake su Amazon Simple Storage Service (Amazon S3) e i relativi metadati. AWS Glue Data Catalog

Lake Formation fornisce il proprio modello di autorizzazioni che amplia il modello di autorizzazioni IAM. Il modello di autorizzazioni di Lake Formation consente un accesso granulare ai dati archiviati nei data lake e a fonti di dati esterne come data warehouse Amazon Redshift, Amazon DynamoDB database e fonti di dati di terze parti attraverso un semplice meccanismo di concessione o revoca, proprio come un sistema di gestione di database relazionali (RDBMS). Le autorizzazioni di Lake Formation vengono applicate utilizzando controlli granulari a livello di colonna, riga e cella nei servizi di AWS analisi e apprendimento automatico, tra cui Amazon Athena, Amazon Amazon Quick Suite Redshift Spectrum, Amazon EMR e. AWS Glue

Con la modalità di accesso ibrido di Lake Formation per AWS Glue Data Catalog (Data Catalog), puoi proteggere e accedere ai dati catalogati utilizzando sia le autorizzazioni di Lake Formation che le politiche di autorizzazione IAM per Amazon S3 e azioni. AWS Glue Con la modalità di accesso ibrida, gli amministratori dei dati possono integrare le autorizzazioni di Lake Formation in modo selettivo e incrementale, concentrandosi su un caso d'uso del data lake alla volta.

Lake Formation consente inoltre di condividere i dati internamente ed esternamente tra più AWS organizzazioni o direttamente con i responsabili IAM in un altro account Account AWS, fornendo un accesso granulare ai metadati del Data Catalog e ai dati sottostanti.

Argomenti

- [Caratteristiche di Lake Formation](#)
- [AWS Lake Formation: Come funziona](#)
- [Componenti di Lake Formation](#)
- [Terminologia dei Lake Formation](#)
- [AWS integrazioni di servizi con Lake Formation](#)
- [Risorse aggiuntive per Lake Formation](#)
- [Guida introduttiva a Lake Formation](#)

Caratteristiche di Lake Formation

Lake Formation ti aiuta a scomporre i silos di dati e a combinare diversi tipi di dati strutturati e non strutturati in un repository centralizzato. Innanzitutto, identifica gli archivi di dati esistenti in Amazon S3 o nei database relazionali e NoSQL e sposta i dati nel tuo data lake. Quindi scansiona, cataloga e prepara i dati per l'analisi. Successivamente, offri ai tuoi utenti un accesso self-service sicuro ai dati tramite i servizi di analisi di loro scelta.

Puoi utilizzare la console Lake Formation per creare cataloghi federati a più livelli nel Data Catalog e unificare i dati tra i data lake Amazon S3 e i data warehouse Amazon Redshift. Puoi anche integrare i dati dei tuoi database operativi come Amazon DynamoDB e fonti di dati di terze parti come Google BigQuery, MySQL, tra le altre. Il Data Catalog fornisce un archivio centralizzato di metadati che semplifica la gestione e la scoperta dei dati su sistemi diversi.

Per ulteriori informazioni, consulta [Inserire i dati nel AWS Glue Data Catalog](#).

Argomenti

- [Inserimento e gestione dei dati](#)
- [Gestione della sicurezza](#)
- [Inserisci i tuoi dati nel Data Catalog](#)

Inserimento e gestione dei dati

Importazione di dati da database già presenti AWS

Dopo aver specificato dove si trovano i database esistenti e fornito le credenziali di accesso, Lake Formation legge i dati e i relativi metadati (schema) per comprendere il contenuto della fonte di dati. Quindi importa i dati nel nuovo data lake e registra i metadati in un catalogo centrale. Con Lake Formation, puoi importare dati da database MySQL, PostgreSQL, SQL Server, MariaDB e Oracle in esecuzione su Amazon RDS o ospitati su Amazon EC2. Sono supportati sia il caricamento di dati in blocco che quello incrementale.

Importa dati da altre fonti esterne

Puoi utilizzare Lake Formation per spostare i dati dai database locali connettendoti a Java Database Connectivity (JDBC). Identifica le fonti di destinazione e fornisci le credenziali di accesso nella console e Lake Formation legge e carica i tuoi dati nel data lake. Per importare dati da database diversi da quelli sopra elencati, puoi creare lavori ETL personalizzati con AWS Glue

Cataloga ed etichetta i tuoi dati

Puoi usare AWS Glue i crawler per leggere i tuoi dati in Amazon S3 ed estrarre schemi di database e tabelle e archiviare tali dati in un catalogo dati ricercabile. Quindi, usa Lake Formation [Controllo degli accessi basato su tag Lake Formation](#) (TBAC) per gestire le autorizzazioni su database, tabelle e colonne. Per ulteriori informazioni sull'aggiunta di tabelle al Data Catalog, consulta. [Creazione di oggetti in AWS Glue Data Catalog](#)

Gestione della sicurezza

Definisci e gestisci i controlli di accesso

Lake Formation offre un unico posto per gestire i controlli di accesso per i dati nel tuo data lake. È possibile definire politiche di sicurezza che limitano l'accesso ai dati a livello di database, tabella, colonna, riga e cella. Queste policy si applicano agli utenti e ai ruoli IAM e agli utenti e ai gruppi durante la federazione tramite un provider di identità esterno. Puoi utilizzare controlli granulari per accedere ai dati protetti da Lake Formation all'interno di Amazon Redshift Spectrum, Athena, ETL AWS Glue e Amazon EMR per Apache Spark. Ogni volta che crei identità IAM, assicurati di seguire le best practice IAM. Per ulteriori informazioni, consulta le [best practice di sicurezza](#) nella Guida per l'utente IAM.

Modalità di accesso ibrida

La modalità di accesso ibrido di Lake Formation offre la flessibilità necessaria per abilitare selettivamente le autorizzazioni di Lake Formation per database e tabelle nel tuo Data Catalog. Con la modalità di accesso ibrido, ora disponi di un percorso incrementale che ti consente di impostare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti. Per ulteriori informazioni, consulta [Modalità di accesso ibrida](#).

Implementa la registrazione degli audit

Lake Formation fornisce registri di controllo completi CloudTrail per monitorare l'accesso e dimostrare la conformità con le politiche definite a livello centrale. Puoi controllare la cronologia di accesso ai dati attraverso i servizi di analisi e machine learning che leggono i dati nel tuo data lake tramite Lake Formation. In questo modo puoi vedere quali utenti o ruoli hanno tentato di accedere a quali dati, con quali servizi e quando. È possibile accedere ai registri di controllo nello stesso modo in cui si accede a qualsiasi altro CloudTrail registro utilizzando la CloudTrail APIs console and. Per ulteriori informazioni sui CloudTrail log, vedere. [Registrazione delle chiamate API AWS Lake Formation utilizzando AWS CloudTrail](#)

Sicurezza a livello di riga e cella

Lake Formation fornisce filtri di dati che consentono di limitare l'accesso a una combinazione di colonne e righe. Utilizza la sicurezza a livello di riga e cella per proteggere i dati sensibili come le informazioni personali identificabili (PII). Per ulteriori informazioni sulla sicurezza a livello di riga, consulta [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#)

Controllo degli accessi basato su tag

Usa il [controllo degli accessi basato sugli attributi](#) di Lake Formation per gestire centinaia o addirittura migliaia di autorizzazioni per i dati creando etichette personalizzate chiamate LF-Tags. Ora puoi definire i tag LF e allegarli a database, tabelle o colonne. Quindi, condividi l'accesso controllato tra i servizi di analisi, machine learning (ML) ed estrazione, trasformazione e caricamento (ETL) per il consumo. I tag LF assicurano che la governance dei dati possa essere scalata facilmente sostituendo le definizioni delle politiche di migliaia di risorse con alcuni tag logici. Lake Formation fornisce una ricerca testuale su questi metadati, in modo che gli utenti possano trovare rapidamente i dati che devono analizzare.

Controllo degli accessi basato sugli attributi

Utilizza il [controllo degli accessi basato sugli attributi per concedere l'accesso agli oggetti del](#) Data Catalog. Il controllo degli accessi basato sugli attributi (ABAC) è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. AWS chiama questi attributi tag. È possibile utilizzare ABAC per concedere l'accesso ai principali all'interno dello stesso account o di un altro account sulle risorse del Data Catalog. Qualsiasi principale IAM con chiavi e valori del tag IAM o del tag di sessione corrispondenti ottiene l'accesso alla risorsa. È necessario disporre di autorizzazioni concedibili sulle risorse per effettuare queste sovvenzioni.

Accesso tra account

Le funzionalità di gestione delle autorizzazioni di Lake Formation semplificano la protezione e la gestione dei data lake distribuiti su più AWS account attraverso un approccio centralizzato, fornendo un controllo granulare degli accessi al Data Catalog e alle sedi Amazon S3. Per ulteriori informazioni, consulta [Condivisione dei dati tra account in Lake Formation](#).

Inserisci i tuoi dati nel Data Catalog

La funzionalità di federazione consente di creare cataloghi federati e impostare autorizzazioni per set di dati archiviati in diverse fonti di dati come Amazon Redshift senza migrare dati o metadati in

Amazon S3 o. AWS Glue Data Catalog Puoi utilizzare i seguenti metodi per importare dati e gestire le autorizzazioni su set di dati esterni in Lake Formation:

Per ulteriori informazioni, consulta [Portare i dati in](#). AWS Glue Data Catalog

- Trasferimento dei dati nei data warehouse di Amazon Redshift in AWS Glue Data Catalog: registra un namespace o un cluster [Amazon Redshift esistente](#) con Data Catalog e crea un catalogo federato a più livelli nel Data Catalog.

Puoi accedere ai tuoi dati utilizzando qualsiasi motore di query compatibile con le specifiche OpenAPI del catalogo REST di Apache Iceberg, come Amazon EMR Serverless e Amazon Athena.

Per ulteriori informazioni, consulta [Inserimento dei dati di Amazon Redshift nel AWS Glue Data Catalog](#).

- Federazione nel Data Catalog da fonti di dati esterne: collega il Data Catalog a fonti di dati esterne utilizzando AWS Glue connessioni e crea cataloghi federati per gestire centralmente le autorizzazioni di accesso sui set di dati utilizzando Lake Formation. Non è necessaria alcuna migrazione dei metadati nel Data Catalog.

Per ulteriori informazioni, consulta [La federazione in fonti di dati esterne in AWS Glue Data Catalog](#).

- Integrazione di Amazon S3 Table Bucket con Data Catalog: puoi pubblicare e catalogare le tabelle Amazon S3 come oggetti Data Catalog e registrare il catalogo come posizione dati di Lake Formation dalla console di Lake Formation o utilizzando. AWS Glue APIs

Per ulteriori informazioni, consulta [Creazione di un catalogo Amazon S3 Tables nel AWS Glue Data Catalog](#).

- Crea cataloghi per gestire le tabelle Amazon Redshift nel Data Catalog: potresti non avere un cluster di produttori Amazon Redshift o un datashare Amazon Redshift disponibile oggi, ma desideri creare e gestire tabelle Amazon Redshift utilizzando Data Catalog. Puoi iniziare creando un catalogo AWS Glue gestito utilizzando l'API `glue:CreateCatalogAPI` o la AWS Lake Formation console impostando il tipo di catalogo come Managed e Catalog source come Redshift.

Per ulteriori informazioni, consulta [Creazione di un catalogo gestito di Amazon Redshift nel AWS Glue Data Catalog](#).

- Integrazione di Lake Formation con la condivisione dei dati di Amazon Redshift: utilizza Lake Formation per gestire centralmente le autorizzazioni di accesso a livello di database, tabelle,

colonne e righe delle condivisioni di dati [Amazon Redshift](#) e limitare l'accesso degli utenti agli oggetti all'interno di un datashare.

- Connessione di Data Catalog a metastore esterni: connettiti AWS Glue Data Catalog a metastore esterni per gestire le autorizzazioni di accesso ai set di dati in Amazon S3 utilizzando Lake Formation. Non è necessaria alcuna migrazione dei metadati nel Data Catalog.

Per ulteriori informazioni, consulta [Gestione delle autorizzazioni sui set di dati che utilizzano metastore esterni](#).

- Integrazione di Lake Formation con AWS Data Exchange — Lake Formation supporta la concessione di licenze di accesso ai dati tramite AWS Data Exchange. Se sei interessato a concedere in licenza i tuoi dati di Lake Formation, consulta [Cosa c'è AWS Data Exchange](#) nella Guida per l'AWS Data Exchange utente.

AWS Lake Formation: Come funziona

AWS Lake Formation fornisce un modello di autorizzazioni del sistema di gestione dei database relazionali (RDBMS) per concedere o revocare l'accesso alle risorse del Data Catalog come database, tabelle e colonne con dati sottostanti in Amazon S3. Le autorizzazioni Lake Formation, facili da gestire, sostituiscono le complesse policy dei bucket di Amazon S3 e le corrispondenti policy IAM.

In Lake Formation, puoi implementare le autorizzazioni su due livelli:

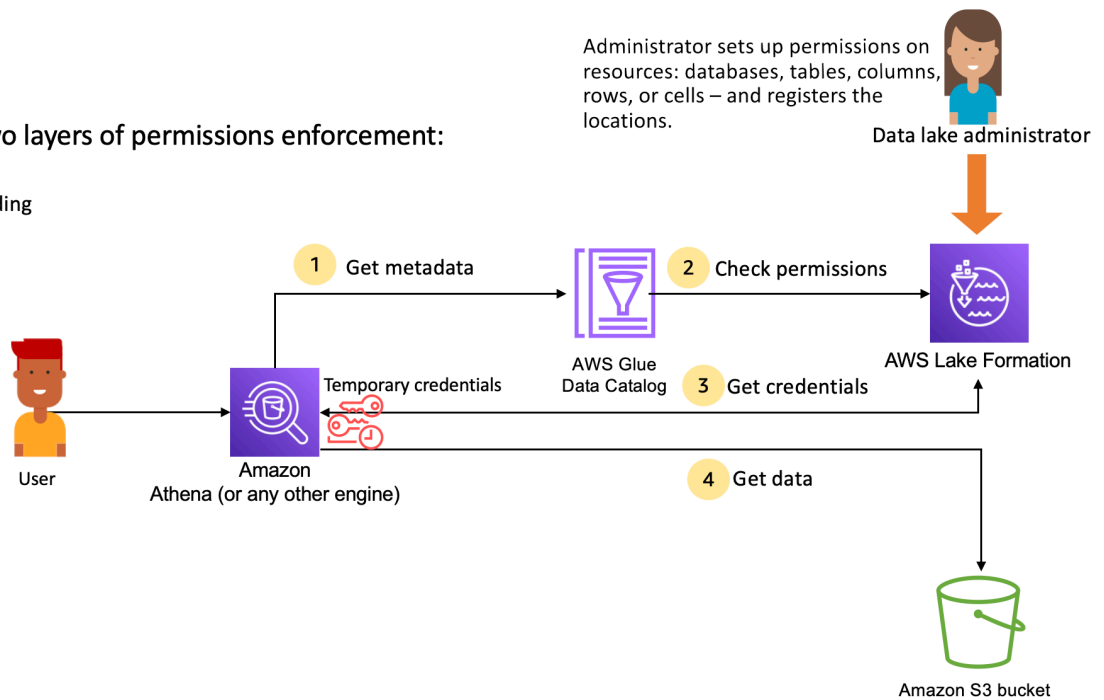
- Applicazione delle autorizzazioni a livello di metadati sulle risorse del Data Catalog come database e tabelle
- Gestione delle autorizzazioni di accesso allo storage sui dati sottostanti archiviati in Amazon S3 per conto di motori integrati

Flusso di lavoro per la gestione delle autorizzazioni di Lake Formation

Lake Formation si integra con i motori analitici per interrogare gli archivi di dati e gli oggetti di metadati di Amazon S3 registrati con Lake Formation. Il diagramma seguente illustra come funziona la gestione delle autorizzazioni in Lake Formation.

Lake Formation provides two layers of permissions enforcement:

- Metadata layer – Data Catalog
- Storage layer – Credential vending



Fasi di alto livello per la gestione dei permessi di Lake Formation

Prima che Lake Formation possa fornire controlli di accesso per i dati nel tuo data lake, un [amministratore del data lake](#) o un utente con autorizzazioni amministrative imposta le politiche utente delle singole tabelle Data Catalog per consentire o negare l'accesso alle tabelle Data Catalog utilizzando le autorizzazioni di Lake Formation.

Quindi, l'amministratore del data lake o un utente delegato dall'amministratore concede le autorizzazioni di Lake Formation agli utenti sui database e sulle tabelle di Data Catalog e registra la posizione Amazon S3 della tabella con Lake Formation.

1. Ottieni metadati: un principale (utente) invia una query o uno script ETL a un [motore di analisi integrato](#) come Amazon Athena, Amazon EMR o AWS Glue Amazon Redshift Spectrum. Il motore analitico integrato identifica la tabella richiesta e invia una richiesta di metadati al Data Catalog.
2. Controlla le autorizzazioni: il Data Catalog controlla le autorizzazioni dell'utente con Lake Formation e, se l'utente è autorizzato ad accedere alla tabella, restituisce al motore i metadati che l'utente può vedere.
3. Ottieni credenziali: il Data Catalog consente al motore di sapere se la tabella è gestita da Lake Formation o meno. Se i dati sottostanti sono registrati con Lake Formation, il motore analitico richiede a Lake Formation di fornire l'accesso ai dati concedendo un accesso temporaneo.
4. Ottieni dati: se l'utente è autorizzato ad accedere alla tabella, Lake Formation fornisce l'accesso temporaneo al motore analitico integrato. Utilizzando l'accesso temporaneo, il motore analitico

recupera i dati da Amazon S3 ed esegue i filtri necessari come il filtraggio di colonne, righe o celle. Quando il motore termina l'esecuzione del lavoro, restituisce i risultati all'utente. Questo processo è chiamato [vendita di credenziali](#).

Se la tabella non è gestita da Lake Formation, la seconda chiamata dal motore di analisi viene effettuata direttamente ad Amazon S3. La policy del bucket Amazon S3 interessata e la politica utente IAM vengono valutate per l'accesso ai dati.

Ogni volta che si utilizzano le policy IAM, assicurati di seguire le best practice IAM. Per ulteriori informazioni, consulta [Best Practice di sicurezza in IAM](#) nella Guida per l'utente di IAM.

Argomenti

- [Autorizzazioni per i metadati](#)
- [Gestione degli accessi allo storage](#)
- [Condivisione dei dati tra account in Lake Formation](#)

Autorizzazioni per i metadati

Lake Formation fornisce l'autorizzazione e il controllo degli accessi per il Data Catalog. Quando un ruolo IAM effettua una chiamata all'API Data Catalog da qualsiasi sistema, Data Catalog verifica le autorizzazioni relative ai dati dell'utente e restituisce solo i metadati a cui l'utente dispone delle autorizzazioni di accesso. Ad esempio, se un ruolo IAM ha accesso a una sola tabella all'interno di un database e un servizio o un utente che assume il ruolo esegue l'GetTablesoperazione, la risposta conterrà solo una tabella, indipendentemente dal numero di tabelle nel database.

Impostazioni predefinite: autorizzazioni **IAMAllowedPrincipal** di gruppo

AWS Lake Formation, per impostazione predefinita, imposta le autorizzazioni per tutti i database e le tabelle su un gruppo virtuale denominato. **IAMAllowedPrincipal** Questo gruppo è unico e visibile solo all'interno di Lake Formation. Il **IAMAllowedPrincipal** gruppo include tutti i responsabili IAM che hanno accesso alle risorse di Data Catalog tramite le politiche principali e le politiche AWS Glue delle risorse IAM. Se queste autorizzazioni esistono su un database o una tabella, a tutti i principali verrà concesso l'accesso al database o alla tabella.

Se desideri fornire autorizzazioni più granulari su un database o una tabella, rimuovi **IAMAllowedPrincipal** l'autorizzazione e Lake Formation applica tutte le altre politiche associate a quel database o tabella. Ad esempio, se esiste una politica che consente all'utente A

di accedere al database A con DESCRIBE le autorizzazioni ed IAMAllowedPrincipal esiste con tutte le autorizzazioni, l'utente A continuerà a eseguire tutte le altre azioni, fino alla revoca dell'autorizzazione. IAMAllowedPrincipal

Inoltre, per impostazione predefinita, il IAMAllowedPrincipal gruppo dispone delle autorizzazioni su tutti i nuovi database e tabelle al momento della creazione. Esistono due configurazioni che controllano questo comportamento. La prima è a livello di account e regione, il che consente questa operazione per i database appena creati, mentre la seconda è a livello di database. Per modificare l'impostazione predefinita, vedere. [Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida](#)

Concessione di autorizzazioni

Gli amministratori di Data Lake possono concedere le autorizzazioni di Data Catalog ai responsabili in modo che i responsabili possano creare e gestire database e tabelle e possano accedere ai dati sottostanti.

Autorizzazioni a livello di database e tabella

Quando concedi le autorizzazioni all'interno di Lake Formation, il concedente deve specificare il principale a cui concedere le autorizzazioni, le risorse a cui concedere le autorizzazioni e le azioni che il beneficiario deve avere accesso a eseguire. Per la maggior parte delle risorse all'interno di Lake Formation, l'elenco principale e le risorse per concedere le autorizzazioni sono simili, ma le azioni che un beneficiario può eseguire variano in base al tipo di risorsa. Ad esempio, SELECT le autorizzazioni sono disponibili per le tabelle per leggere le tabelle, ma le SELECT autorizzazioni non sono consentite per i database. L'CREATE_TABLE autorizzazione è consentita sui database, ma non sulle tabelle.

È possibile concedere AWS Lake Formation le autorizzazioni utilizzando due metodi:

- [Metodo di risorsa denominato](#): consente di scegliere i nomi di database e tabelle concedendo le autorizzazioni agli utenti.
- [Controllo degli accessi basato su tag LF \(LF-TBAC\)](#): gli utenti creano tag LF, li associano alle risorse del catalogo dati, concedono l'Describe autorizzazione sui tag LF, associano le autorizzazioni ai singoli utenti e scrivono politiche di autorizzazione LF utilizzando i tag LF a diversi utenti. Tali LF-Tag-based politiche si applicano a tutte le risorse del Data Catalog associate a tali valori LF-Tag.

 Note

I tag LF sono esclusivi di Lake Formation. Sono visibili solo in Lake Formation e non devono essere confusi con i tag AWS delle risorse.

LF-TBAC è una funzionalità che consente agli utenti di raggruppare le risorse in categorie di tag LF definite dall'utente e di applicare le autorizzazioni a tali gruppi di risorse. Pertanto, è il modo migliore per scalare le autorizzazioni su un numero enorme di risorse del Data Catalog.

Per ulteriori informazioni, consulta [Controllo degli accessi basato su tag Lake Formation](#).

Quando concedi le autorizzazioni a un responsabile, Lake Formation valuta le autorizzazioni come un'unione di tutte le politiche per quell'utente. Ad esempio, se si dispone di due criteri in una tabella per un principale in cui un criterio concede le autorizzazioni alle colonne col1, col2 e col3 tramite il metodo di risorsa denominato e l'altro criterio concede le autorizzazioni alla stessa tabella e principale a col5 e col6 tramite LF-tags, le autorizzazioni effettive saranno un'unione delle autorizzazioni che sarebbero col1, col2, col3, col5 e col6. Ciò include anche filtri e righe di dati.

Autorizzazioni per la localizzazione dei dati

Le autorizzazioni di localizzazione dei dati offrono agli utenti non amministrativi la possibilità di creare database e tabelle in posizioni Amazon S3 specifiche. Se un utente tenta di creare un database o una tabella in una posizione per cui non dispone delle autorizzazioni necessarie, l'operazione di creazione ha esito negativo. Questo serve a impedire agli utenti di creare tabelle in posizioni arbitrarie all'interno del data lake e consente di controllare dove tali utenti possono leggere e scrivere i dati. Esiste un'autorizzazione implicita durante la creazione di tabelle nella posizione Amazon S3 all'interno del database in cui viene creata. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

Crea autorizzazioni per tabelle e database

Per impostazione predefinita, gli utenti non amministrativi non dispongono delle autorizzazioni per creare database o tabelle all'interno di un database. La creazione del database è controllata a livello di account utilizzando le impostazioni di Lake Formation in modo che solo i responsabili autorizzati possano creare database. Per ulteriori informazioni, consulta [Creazione di un database](#). Per creare una tabella, un principale richiede l'CREATE_TABLE autorizzazione sul database in cui viene creata la tabella. Per ulteriori informazioni, consulta [Creazione di tabelle](#).

Autorizzazioni implicite ed esplicite

Lake Formation fornisce autorizzazioni implicite a seconda della persona e delle azioni che la persona compie. Ad esempio, gli amministratori del data lake ottengono automaticamente le DESCRIBE autorizzazioni per tutte le risorse all'interno del Data Catalog, le autorizzazioni per la localizzazione dei dati per tutte le posizioni, le autorizzazioni per creare database e tabelle in tutte le posizioni e le autorizzazioni per qualsiasi risorsa. Grant Revoke I creatori di database ottengono automaticamente tutte le autorizzazioni di database sui database che creano e i creatori di tabelle ottengono tutte le autorizzazioni sulle tabelle che creano. Per ulteriori informazioni, consulta [Autorizzazioni implicite di Lake Formation](#).

Autorizzazioni concedibili

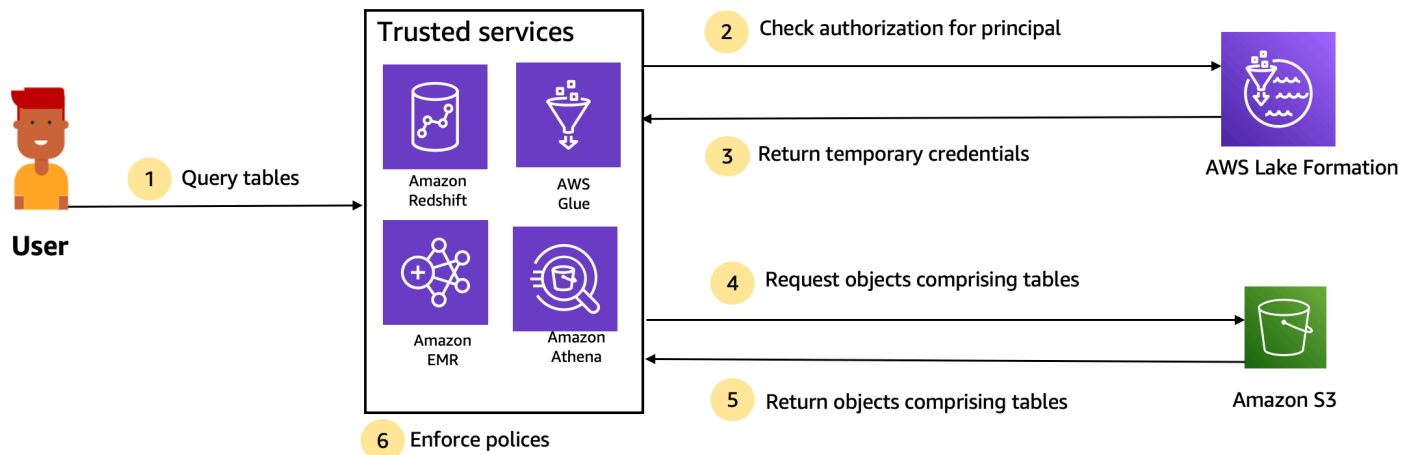
Gli amministratori di Data Lake hanno la possibilità di delegare la gestione delle autorizzazioni a utenti non amministrativi fornendo autorizzazioni concedibili. Quando a un committente vengono concesse autorizzazioni su una risorsa e un insieme di autorizzazioni, tale principale ottiene la capacità di concedere le autorizzazioni ad altri responsabili su quella risorsa.

Gestione degli accessi allo storage

Lake Formation utilizza la funzionalità di [vendita di credenziali](#) per fornire un accesso temporaneo ai dati di Amazon S3. La vendita di credenziali, o vendita di token, è uno schema comune che fornisce credenziali temporanee a utenti, servizi o altre entità allo scopo di concedere l'accesso a breve termine a una risorsa.

Lake Formation sfrutta questo modello per fornire un accesso a breve termine a servizi di AWS analisi come Athena per accedere ai dati per conto del committente chiamante. Quando concedono le autorizzazioni, gli utenti non devono aggiornare le policy dei bucket Amazon S3 o le policy IAM e non hanno bisogno dell'accesso diretto ad Amazon S3.

Il diagramma seguente mostra come Lake Formation fornisce l'accesso temporaneo alle località registrate:



Trusted services enforce AWS Lake Formation policies (distributed enforcement with fail close).

1. Un principale (utente) inserisce una query o una richiesta di dati per una tabella tramite un servizio integrato affidabile come Athena, Amazon EMR, Redshift Spectrum o. AWS Glue
2. Il servizio integrato verifica l'autorizzazione di Lake Formation per la tabella e le colonne richieste e determina l'autorizzazione. Se l'utente non è autorizzato, Lake Formation nega l'accesso ai dati e la query ha esito negativo.
3. Una volta completata l'autorizzazione e attivata l'autorizzazione all'archiviazione per la tabella e l'utente, il servizio integrato recupera le credenziali temporanee da Lake Formation per accedere ai dati.
4. Il servizio integrato utilizza le credenziali temporanee di Lake Formation per richiedere oggetti da Amazon S3.
5. Amazon S3 fornisce gli oggetti Amazon S3 al servizio integrato. Gli oggetti Amazon S3 contengono tutti i dati della tabella.
6. Il servizio integrato esegue la necessaria applicazione delle politiche di Lake Formation, come il filtraggio a livello di colonna, a livello di riga e/o a livello di cella. Il servizio integrato elabora le richieste e restituisce i risultati all'utente.

Abilita l'applicazione delle autorizzazioni a livello di storage per le tabelle del Data Catalog

Per impostazione predefinita, l'applicazione a livello di storage non è abilitata per le tabelle all'interno del Data Catalog. Per abilitare l'applicazione a livello di storage, devi registrare la posizione Amazon S3 dei tuoi dati di origine con Lake Formation e fornire un ruolo IAM. Le autorizzazioni a livello di storage verranno abilitate per tutte le tabelle con lo stesso percorso di posizione della tabella o lo stesso prefisso della posizione Amazon S3.

Quando un servizio integrato richiede l'accesso alla posizione dei dati per conto di un utente, il servizio Lake Formation assume questo ruolo e restituisce le credenziali al servizio richiesto con autorizzazioni limitate alla risorsa in modo che sia possibile effettuare l'accesso ai dati. Il ruolo IAM registrato deve disporre di tutti gli accessi necessari alla posizione Amazon S3, comprese AWS KMS le chiavi.

Per ulteriori informazioni, consulta [Registrazione di una sede Amazon S3](#).

Servizi supportati AWS

AWS servizi di analisi come Athena, Redshift Spectrum, Amazon AWS Glue EMR e Amazon SageMaker AI integrazione con AWS Lake Formation utilizzando le Amazon Quick Suite operazioni API di vendita delle credenziali di Lake Formation. Per un elenco completo dei AWS servizi che si integrano con Lake Formation e il livello di granularità e i formati di tabella che supportano, consulta. [Collaborazione con altri AWS servizi](#)

Condivisione dei dati tra account in Lake Formation

Con Lake Formation, puoi condividere le risorse del Data Catalog (database e tabelle) all'interno di un AWS account e tra account con una semplice configurazione utilizzando il metodo della risorsa denominata o i tag LF. Puoi condividere un intero database o selezionare tabelle da un database con qualsiasi principale IAM (ruoli e utenti IAM) in un account, con altri account a livello di AWS account o direttamente con i principali IAM in un altro account.

Puoi anche condividere le tabelle del Data Catalog con filtri di dati per limitare l'accesso ai dettagli a livello di riga e cella. Lake Formation utilizza AWS Resource Access Manager (AWS RAM) per facilitare la concessione di autorizzazioni tra account. Quando una risorsa viene condivisa tra due account, AWS RAM invia gli inviti all'account del destinatario. Quando un utente accetta un invito alla AWS RAM condivisione, AWS RAM fornisce le autorizzazioni necessarie a Lake Formation per avere a disposizione le risorse del Data Catalog e abilitare l'applicazione a livello di storage. Per ulteriori informazioni, consulta [Condivisione dei dati tra account in Lake Formation](#).

Quando l'amministratore del data lake dell'account del destinatario accetta la AWS RAM condivisione, le risorse condivise sono disponibili nell'account del destinatario. L'amministratore del data lake concede ulteriori autorizzazioni Lake Formation sulla risorsa condivisa ai principali IAM aggiuntivi nell'account del destinatario, se l'amministratore dispone delle GRANTABLE autorizzazioni sulla risorsa condivisa.

Tuttavia, i responsabili non possono interrogare le risorse condivise utilizzando Athena o Redshift Spectrum senza un collegamento alla risorsa. Un collegamento a una risorsa è un'entità nel Data Catalog ed è simile a un concetto Linux-Symlink.

L'amministratore del data lake dell'account del destinatario crea un link alla risorsa condivisa. L'amministratore concede `Describe` le autorizzazioni sul collegamento alla risorsa con le autorizzazioni richieste sulla risorsa condivisa originale ad altri utenti. Un utente nell'account destinatario può quindi utilizzare il collegamento alla risorsa per interrogare la risorsa condivisa utilizzando Athena e Redshift Spectrum. Per ulteriori informazioni sui collegamenti alle risorse, consulta [Creazione di collegamenti alle risorse](#)

Componenti di Lake Formation

AWS Lake Formation si basa sull'interazione di diversi componenti per creare e gestire il tuo data lake.

Console Lake Formation

Utilizzi la console Lake Formation per definire e gestire il tuo data lake e concedere e revocare le autorizzazioni di Lake Formation. Puoi utilizzare i blueprint sulla console per scoprire, pulire, trasformare e inserire dati. Puoi anche abilitare o disabilitare l'accesso alla console per i singoli utenti di Lake Formation.

API e interfaccia a riga di comando di Lake Formation

Lake Formation fornisce operazioni API tramite diversi linguaggi specifici SDKs e il AWS Command Line Interface (). AWS CLI L'API Lake Formation funziona in combinazione con AWS Glue API. L'API Lake Formation si concentra principalmente sulla gestione delle autorizzazioni di Lake Formation, mentre AWS Glue L'API fornisce un'API di catalogo dati e un'infrastruttura gestita per la definizione, la pianificazione e l'esecuzione di operazioni ETL sui dati.

Per informazioni su AWS Glue API, consulta la [Guida per AWS Glue gli sviluppatori](#). Per informazioni sull'utilizzo di AWS CLI, consulta il [AWS CLI Command Reference](#).

Altri AWS servizi

Lake Formation utilizza i seguenti servizi:

- [AWS Glue](#) per orchestrare job e crawler per trasformare i dati utilizzando AWS Glue trasforma.

- [IAM](#) concederà politiche di autorizzazione ai dirigenti di Lake Formation. Il modello di autorizzazione Lake Formation amplia il modello di autorizzazione IAM per proteggere il tuo data lake.

Terminologia dei Lake Formation

Di seguito sono riportati alcuni termini importanti che incontrerai in questa guida.

Lago di dati

Il data lake è costituito dai tuoi dati persistenti archiviati in Amazon S3 e gestiti da Lake Formation utilizzando un Data Catalog. Un data lake in genere archivia quanto segue:

- Dati strutturati e non strutturati
- Dati grezzi e dati trasformati

Affinché un percorso Amazon S3 si trovi all'interno di un data lake, deve essere registrato presso Lake Formation.

Accesso ai dati

Lake Formation fornisce un accesso sicuro e granulare ai dati attraverso un nuovo modello di autorizzazioni di concessione/revoca che aumenta le politiche (IAM). AWS Identity and Access Management

Analisti e data scientist possono utilizzare l'intero portafoglio di servizi di AWS analisi e apprendimento automatico, come Amazon Athena, per accedere ai dati. Le politiche di sicurezza configurate di Lake Formation aiutano a garantire che gli utenti possano accedere solo ai dati a cui sono autorizzati ad accedere.

Modalità di accesso ibrida

La modalità di accesso ibrida consente di proteggere e accedere ai dati catalogati utilizzando sia le autorizzazioni Lake Formation che le autorizzazioni IAM e Amazon S3. La modalità di accesso ibrido consente agli amministratori dei dati di integrare le autorizzazioni di Lake Formation in modo selettivo e incrementale, concentrandosi su un caso d'uso del data lake alla volta.

Piano

Un blueprint è un modello di gestione dei dati che consente di inserire facilmente i dati in un data lake. Lake Formation fornisce diversi modelli, ciascuno per un tipo di sorgente predefinito, come un database relazionale o dei log. AWS CloudTrail Da un blueprint, puoi creare un flusso di lavoro. I flussi di lavoro sono costituiti da AWS Glue crawler, job e trigger generati per orchestrare il caricamento e l'aggiornamento dei dati. I blueprint utilizzano l'origine dei dati, la destinazione dei dati e la pianificazione come input per configurare il flusso di lavoro.

Flusso di lavoro

Un flusso di lavoro è un contenitore per un insieme di elementi correlati AWS Glue lavori, crawler e trigger. Il flusso di lavoro viene creato in Lake Formation e viene eseguito in AWS Glue servizio. Lake Formation è in grado di tracciare lo stato di un flusso di lavoro come singola entità.

Quando si definisce un flusso di lavoro, si seleziona il progetto su cui si basa. È quindi possibile eseguire flussi di lavoro su richiesta o in base a una pianificazione.

I flussi di lavoro creati in Lake Formation sono visibili nel AWS Glue console come grafo aciclico diretto (DAG). Utilizzando il DAG, è possibile tenere traccia dell'avanzamento del flusso di lavoro ed eseguire la risoluzione dei problemi.

Catalogo dati

Il Data Catalog è il tuo archivio di metadati persistente. È un servizio gestito che ti consente di archiviare, annotare e condividere i metadati nel AWS cloud nello stesso modo in cui faresti in un metastore Apache Hive. Fornisce un repository uniforme in cui diversi sistemi possono archiviare e trovare metadati per tenere traccia dei dati in silos di dati e quindi utilizzare tali metadati per interrogare e trasformare i dati. Lake Formation utilizza il AWS Glue Data Catalog per archiviare metadati su data lake, fonti di dati, trasformazioni e destinazioni.

I metadati sulle fonti di dati e sulle destinazioni sono sotto forma di database e tabelle. Le tabelle memorizzano informazioni sullo schema, informazioni sulla posizione e altro ancora. I database sono raccolte di tabelle. Lake Formation fornisce una gerarchia di autorizzazioni per controllare l'accesso ai database e alle tabelle nel Data Catalog.

Ogni AWS account dispone di un catalogo dati per regione. AWS

Dati sottostanti

I dati sottostanti si riferiscono ai dati di origine o ai dati all'interno dei data lake a cui fanno riferimento le tabelle del Data Catalog.

Principale

Un principale è un utente o un ruolo AWS Identity and Access Management (IAM) o un utente di Active Directory.

Amministratore del data lake

Un amministratore del data lake è un responsabile che può concedere a qualsiasi principale (incluso se stesso) qualsiasi autorizzazione su qualsiasi risorsa o posizione dei dati del Data Catalog. Designare un amministratore del data lake come primo utente del Data Catalog. Questo utente può quindi concedere autorizzazioni più granulari sulle risorse ad altri responsabili.

Note

Gli utenti amministrativi IAM, ovvero gli utenti con la policy AdministratorAccess AWS gestita, non sono automaticamente amministratori di data lake. Ad esempio, non possono concedere le autorizzazioni di Lake Formation sugli oggetti del catalogo a meno che non abbiano ottenuto le autorizzazioni per farlo. Tuttavia, possono utilizzare la console o l'API di Lake Formation per designarsi amministratori di data lake.

Per informazioni sulle funzionalità di un amministratore di data lake, consulta [Autorizzazioni implicite di Lake Formation](#). Per informazioni sulla designazione di un utente come amministratore del data lake, consulta [Crea un amministratore del data lake](#).

AWS integrazioni di servizi con Lake Formation

Puoi utilizzare Lake Formation per gestire le autorizzazioni di accesso a livello di database, tabelle e colonne sui dati archiviati in Amazon S3. Dopo aver registrato i dati con Lake Formation, puoi utilizzare servizi di AWS analisi come Amazon Athena AWS Glue, Amazon Redshift Spectrum, Amazon EMR per interrogare i dati. I seguenti AWS servizi si integrano AWS Lake Formation e rispettano le autorizzazioni di Lake Formation.

AWS Servizio	Dettagli sull'integrazione
AWS Glue	Argomento di riferimento: AWS Lake Formation Utilizzo con AWS Glue AWS Glue e Lake Formation condividono lo stesso Data Catalog. Per le operazioni della console (come la visualizzazione di un elenco di tabelle) e tutte le operazioni API, AWS Glue gli utenti possono accedere solo ai database e alle tabelle su cui dispongono delle autorizzazioni Lake Formation.
Amazon Athena	Argomento di riferimento: Utilizzo AWS Lake Formation con Amazon Athena Usa Lake Formation per consentire o negare le autorizzazioni alla lettura dei dati in Amazon S3. Quando Amazon Athena gli utenti selezionano il AWS Glue catalogo nell'editor di query, possono interrogare solo i database, le tabelle e le colonne su cui dispongono delle autorizzazioni Lake Formation. Le query che utilizzano i manifesti non sono supportate. Attualmente, Lake Formation non supporta la gestione delle autorizzazioni su operazioni di scrittura come VACUUM UPDATE e OPTIMIZE sulle tabelle in Open Table Formats. MERGE Oltre ai principali che si autenticano con Athena tramite AWS Identity and Access Management (IAM), Lake Formation supporta gli utenti Athena che si connettono tramite il driver JDBC o ODBC e si autenticano tramite SAML. I provider SAML supportati includono Okta e Microsoft Active Directory Federation Service (ADFS).
Amazon Redshift Spectrum	Argomento di riferimento: Utilizzo AWS Lake Formation con Amazon Redshift Spectrum Quando gli utenti di Amazon Redshift creano uno schema esterno su un database in AWS Glue Data Catalog, possono interrogare solo le tabelle e le colonne dello schema su cui dispongono delle autorizzazioni Lake Formation.

AWS Servizio	Dettagli sull'integrazione
Amazon Quick Suite Enterprise Edition	Riferimento: Utilizzo AWS Lake Formation con Quick Suite Quando un utente di Amazon Quick Suite Enterprise Edition esegue una query su un set di dati in una posizione Amazon S3, deve disporre dell'autorizzazione Lake SELECT Formation sui dati.
Amazon EMR	Riferimento: Utilizzo AWS Lake Formation con Amazon EMR Puoi integrare le autorizzazioni di Lake Formation quando crei un cluster Amazon EMR con un ruolo di runtime. Un ruolo di runtime è un ruolo IAM che associ ai job o alle query di Amazon EMR, quindi Amazon EMR utilizza questo ruolo per accedere alle risorse. AWS

Lake Formation collabora anche con [AWS Key Management Service](#) (AWS KMS) per consentirti di configurare più facilmente questi servizi integrati per crittografare e decrittografare i dati nelle sedi Amazon Simple Storage Service (Amazon S3).

Risorse aggiuntive per Lake Formation

Per ulteriori informazioni in merito AWS Lake Formation, ti consigliamo di continuare ad approfondire i concetti introdotti in questa guida utilizzando le seguenti risorse:

Argomenti

- [Blog](#)
- [Conferenze tecniche e webinar](#)
- [Architettura moderna](#)
- [Risorse Data Mesh](#)
- [Guide alle migliori pratiche](#)

Blog

- [AWS Lake Formation Riepilogo dell'anno 2022](#)

- [Architettura di dati moderna multiregionale altamente resiliente](#)
- [Condivisione tra account utilizzando i tag LF per indirizzare i principali IAM](#)
- [Dashboard dell'inventario delle autorizzazioni di Lake Formation](#)
- [Mesh di dati basata sugli eventi](#)

Conferenze tecniche e webinar

- re:Invent 2020 — [Data lake: crea, proteggi e condividi facilmente con AWS Lake Formation](#)
- re:Invent 2022 — [Creazione e gestione di un datalake](#) su Amazon S3
- AWS [Summit SF 2022 — Comprensione e realizzazione di un'architettura di dati moderna](#)
- AWS Summit ATL 2022: [data lake moderni con AWS Lake Formation Amazon Redshift](#) e AWS Glue
- AWS Summit ANZ 2022 — [Data lake, lake house e data mesh: cosa, perché e come?](#)
- AWS Online Tech Talks: [semplificazione delle autorizzazioni e della governance nel tuo data lake](#)

Architettura moderna

- [Modelli architetturici moderni](#)

Risorse Data Mesh

- [Crea un'architettura di dati moderna e un pattern di data mesh su larga scala utilizzando il controllo degli accessi AWS Lake Formation basato su tag](#)
- [In che modo JPMorgan Chase ha creato un'architettura data mesh per generare valore significativo e migliorare la propria piattaforma dati aziendale](#)
- [Crea una rete di dati su AWS](#)

Guide alle migliori pratiche

- [AWS Lake Formation guide alle migliori pratiche](#)

Guida introduttiva a Lake Formation

Ti consigliamo di iniziare con le sezioni seguenti:

- [AWS Lake Formation: Come funziona](#)— Scopri la terminologia essenziale e il modo in cui i vari componenti interagiscono.
- [Guida introduttiva a Lake Formation](#)— Ottieni informazioni sui prerequisiti e completa importanti attività di configurazione.
- [AWS Lake Formation tutorial](#)— Segui step-by-step i tutorial per imparare a usare Lake Formation.
- [Sicurezza in AWS Lake Formation](#)— Scopri come puoi contribuire a proteggere l'accesso ai dati in Lake Formation.

Guida introduttiva a Lake Formation

Se non ti sei registrato AWS o hai bisogno di assistenza per iniziare, assicurati di completare le seguenti attività.

Argomenti

- [Completa le attività AWS di configurazione iniziali](#)
- [Configurare AWS Lake Formation](#)
- [Aggiornamento delle AWS Glue autorizzazioni relative ai dati al modello AWS Lake Formation](#)
- [AWS Lake Formation e endpoint VPC di interfaccia \(\)AWS PrivateLink](#)

Completa le attività AWS di configurazione iniziali

Per utilizzare AWS Lake Formation, devi innanzitutto completare i seguenti passaggi:

Argomenti

- [Iscriviti per un Account AWS](#)
- [Crea un utente con accesso amministrativo](#)
- [Concessione dell'accesso programmatico](#)

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 AWS viene 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 le [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 [Console di gestione AWS](#) come 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 il Centro identità IAM.

Per istruzioni, consulta [Abilitazione del AWS IAM Identity Center](#) nella Guida per l'utente di AWS IAM Identity Center .

2. Nel Centro identità IAM, 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 come utente del Centro identità IAM, utilizza l'URL di accesso che è stato inviato al tuo indirizzo e-mail quando hai creato l'utente del Centro identità IAM.

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.

Assegnazione dell'accesso ad altri utenti

1. Nel Centro identità IAM, crea un set di autorizzazioni conforme alla best practice per l'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 .

Concessione dell'accesso programmatico

Gli utenti hanno bisogno di un accesso programmatico se vogliono interagire con l' AWS Console di gestione AWS esterno di. Il modo per concedere l'accesso programmatico dipende dal tipo di utente che accede. AWS

Per fornire agli utenti l'accesso programmatico, scegli una delle seguenti opzioni.

Quale utente necessita dell'accesso programmatico?	Per	Come
IAM	(Consigliato) Utilizza le credenziali della console come credenziali temporanee per firmare le richieste programmatiche a,, o. AWS CLI AWS SDKs AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, consulta Login for AWS local development nella Guida per l'AWS Command Line Interface utente. • Per AWS SDKs, consulta Login for AWS local development nella AWS

Quale utente necessita dell'accesso programmatico?	Per	Come
		SDKs and Tools Reference Guide.
Identità della forza lavoro (Utenti gestiti nel centro identità IAM)	Utilizza credenziali temporane e per firmare le richieste programmatiche a AWS CLI, AWS SDKs, o. AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, vedere Configurazione dell'uso AWS IAM Identity Center nella AWS CLI Guida per l'utente. AWS Command Line Interface • Per AWS SDKs gli strumenti e AWS APIs, consulta l'autenticazione di IAM Identity Center nella Guida di riferimento AWS SDKs and Tools.
IAM	Utilizza credenziali temporane e per firmare le richieste programmatiche a AWS CLI, AWS SDKs, o. AWS APIs	Seguendo le istruzioni riportate in Utilizzo delle credenziali temporanee con le AWS risorse nella Guida per l'utente IAM .

Quale utente necessita dell'accesso programmatico?	Per	Come
IAM	(Non consigliato) Utilizza credenziali a lungo termine per firmare richieste programmatiche a AWS CLI,, AWS SDKs o. AWS APIs	Segui le istruzioni per l'interfaccia che desideri utilizzare. • Per la AWS CLI, consulta Autenticazione tramite credenziali utente IAM nella Guida per l'utente.AWS Command Line Interface • Per gli strumenti AWS SDKs e gli strumenti, consulta Autenticazione tramite credenziali a lungo termine nella Guida di riferimento agli strumenti e agli AWS SDKs strumenti. • Per AWS APIs, consulta la sezione Gestione delle chiavi di accesso per gli utenti IAM nella Guida per l'utente IAM.

Configurare AWS Lake Formation

Le sezioni seguenti forniscono informazioni sulla configurazione di Lake Formation per la prima volta. Non tutti gli argomenti di questa sezione sono necessari per iniziare a usare Lake Formation. Puoi utilizzare le istruzioni per configurare il modello di autorizzazioni Lake Formation per gestire AWS Glue Data Catalog gli oggetti e le posizioni dei dati esistenti in Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3).

1. [Crea un amministratore del data lake](#)
2. [Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida](#)
3. [the section called “Configura una posizione Amazon S3 per il tuo data lake”](#)

4. [the section called “Assegna le autorizzazioni agli utenti di Lake Formation”](#)
5. [the section called “Integrazione di IAM Identity Center”](#)
6. [the section called “\(Facoltativo\) Impostazioni di filtraggio dei dati esterni”](#)
7. [the section called “\(Facoltativo\) Concedi l'accesso alla chiave di crittografia Data Catalog”](#)
8. [\(Facoltativo\) Crea un ruolo IAM per i flussi di lavoro](#)

Questa sezione mostra come configurare le risorse di Lake Formation in due modi diversi:

- Utilizzando un AWS CloudFormation modello
- Utilizzo della console Lake Formation

Per configurare Lake Formation tramite AWS console, vai a [Crea un amministratore del data lake](#).

Configura le risorse di Lake Formation utilizzando il CloudFormation modello

Note

Lo CloudFormation stack esegue i passaggi da 1 a 6 di cui sopra, ad eccezione dei passaggi 2 e 5. Esegui [Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida](#) e [the section called “Integrazione di IAM Identity Center”](#) manualmente dalla console Lake Formation.

1. Accedi alla AWS CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> come amministratore IAM nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Avanti nella schermata Crea stack.
4. Inserisci un nome per lo stack.
5. Per DatalakeAdminName e DatalakeAdminPassword, inserisci il nome utente e la password per l'utente amministratore di Data Lake.
6. Per DatalakeUser1Name e DatalakeUser1Password, inserisci il nome utente e la password per l'utente analista di data lake.

7. Per `DataLakeBucketName`, inserisci il nuovo nome del bucket che verrà creato.
8. Scegli Next (Successivo).
9. Nella pagina successiva, scegli `I acknowledge that CloudFormation might create IAM resources with custom names` e scegli Avanti.
10. Controlla i dettagli nella pagina finale e seleziona Riconosco che AWS CloudFormation potrebbe creare risorse IAM.
11. Scegli Create (Crea).

La creazione dello stack può richiedere fino a due minuti.

Eseguire la pulizia delle risorse

Se desideri ripulire le risorse dello CloudFormation stack:

1. Annulla la registrazione del bucket Amazon S3 creato dallo stack e registrato come posizione di data lake.
2. Elimina lo stack. CloudFormation Questo eliminerà tutte le risorse create dallo stack.

Crea un amministratore del data lake

Gli amministratori di Data Lake sono inizialmente gli unici utenti o ruoli AWS Identity and Access Management (IAM) che possono concedere le autorizzazioni di Lake Formation sulle posizioni dei dati e sulle risorse del Data Catalog a qualsiasi principale (incluso se stesso). Per ulteriori informazioni sulle funzionalità di amministratore del data lake, consulta [Autorizzazioni implicite di Lake Formation](#). Per impostazione predefinita, Lake Formation consente di creare fino a 30 amministratori di data lake.

Puoi creare un amministratore di data lake utilizzando la console Lake Formation o il `PutDataLakeSettings` funzionamento dell'API Lake Formation.

Le seguenti autorizzazioni sono necessarie per creare un amministratore del data lake. L'Administratorutente dispone implicitamente di queste autorizzazioni.

- `lakeformation:PutDataLakeSettings`
- `lakeformation:GetDataLakeSettings`

Se concedi la `AWSLakeFormationDataAdmin` policy a un utente, quell'utente non sarà in grado di creare altri utenti amministratori di Lake Formation.

Per creare un amministratore del data lake (console)

1. Se l'utente che deve essere amministratore del data lake non esiste ancora, utilizza la console IAM per crearlo. Altrimenti, scegli un utente esistente che sarà l'amministratore del data lake.

 Note

Ti consigliamo di non selezionare un utente amministrativo IAM (utente con la policy `AdministratorAccess` AWS gestita) come amministratore del data lake.

Allega le seguenti politiche AWS gestite all'utente:

Policy	Obbligatorio?	Note
<code>AWSLakeFormationDataAdmin</code>	Obbligatorio	Autorizzazioni di base per gli amministratori del data lake. Questa policy AWS gestita contiene un divieto esplicito per il funzionamento dell'API Lake Formation, <code>PutDataLakeSetting</code> che impedisce agli utenti di creare nuovi amministratori di data lake.

Policy	Obbligatorio?	Note
AWSGlueConsoleFullAccess , CloudWatchLogsReadOnlyAccess	Facoltativo	Allega queste politiche se l'amministratore del data lake intende risolvere i problemi dei flussi di lavoro creati dai blueprint di Lake Formation. Queste politiche consentono all'amministratore del data lake di visualizzare le informazioni sulla risoluzione dei problemi nella AWS Glue console e nella Amazon CloudWatch Logs console. Per informazioni sui flussi di lavoro, consulta the section called "Importazione di dati tramite flussi di lavoro" .
AWSLakeFormationCrossAccountManager	Facoltativo	Allega questa policy per consentire e all'amministratore del data lake di concedere e revocare le autorizzazioni tra più account sulle risorse di Data Catalog. Per ulteriori informazioni, consulta Condivisione dei dati tra account in Lake Formation .
AmazonAthenaFullAccess	Facoltativo	Allega questa policy se l'amministratore del data lake eseguirà le query in Amazon Athena

- Allega la seguente policy in linea, che concede all'amministratore del data lake l'autorizzazione a creare il ruolo collegato al servizio Lake Formation. Un nome suggerito per la policy è. LakeFormationSLR

Il ruolo collegato ai servizi consente all'amministratore del data lake di registrare più facilmente le sedi Amazon S3 con Lake Formation. Per ulteriori informazioni sul ruolo collegato ai servizi di Lake Formation, vedere. [the section called "Uso di ruoli collegati ai servizi"](#)

⚠ Important

In tutti i criteri seguenti, sostituiscili *<account-id>* con un numero di AWS account valido.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "iam:CreateServiceLinkedRole",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:AWSServiceName": "lakeformation.amazonaws.com"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:PutRolePolicy"
      ],
      "Resource": "arn:aws:iam::<account-id>:role/aws-service-role/
lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess"
    }
  ]
}
```

3. (Facoltativo) Allega la seguente politica PassRole in linea all'utente. Questa policy consente all'amministratore del data lake di creare ed eseguire flussi di lavoro. L'iam:PassRole autorizzazione consente al flusso di lavoro di assumere il ruolo di LakeFormationWorkflowRole creare crawler e job e di assegnare il ruolo ai crawler e ai job creati. Un nome suggerito per la policy è. UserPassRole

⚠ Important

Sostituiscilo *<account-id>* con un numero di AWS conto valido.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PassRolePermissions",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::111122223333:role/LakeFormationWorkflowRole"
      ]
    }
  ]
}
```

4. (Facoltativo) Allega questa politica in linea aggiuntiva se il tuo account concederà o riceverà autorizzazioni Lake Formation per più account. Questa policy consente all'amministratore del data lake di visualizzare e accettare AWS Resource Access Manager (AWS RAM) gli inviti alla condivisione di risorse. Inoltre, per gli amministratori del data lake dell'account di AWS Organizations gestione, la policy include l'autorizzazione ad abilitare le concessioni tra account alle organizzazioni. Per ulteriori informazioni, consulta [Condivisione dei dati tra account in Lake Formation](#).

Un nome suggerito per la policy è. RAMAccess

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ram:AcceptResourceShareInvitation",
        "ram:RejectResourceShareInvitation",
        "ec2:DescribeAvailabilityZones",

```

```
        "ram:EnableSharingWithAwsOrganization"  
      ],  
      "Resource": "*"   
    }  
  ]  
}
```

5. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come AdministratorAccess utente con policy AWS gestite dagli utenti.
6. Se viene visualizzata la finestra Welcome to Lake Formation, scegli l'utente IAM che hai creato o selezionato nello Step 1, quindi scegli Inizia.
7. Se non vedi la finestra Welcome to Lake Formation, esegui i seguenti passaggi per configurare un Lake Formation Administrator.
 - a. Nel pannello di navigazione, in Amministrazione, scegli Ruoli e attività amministrative. Nella sezione Amministratori di Data Lake della pagina della console, scegli Aggiungi.
 - b. Nella finestra di dialogo Aggiungi amministratori, in Tipo di accesso, scegli Amministratore del data lake.
 - c. Per gli utenti e i ruoli IAM, scegli l'utente IAM che hai creato o selezionato nel passaggio 1, quindi scegli Salva.

Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida

Lake Formation inizia con le impostazioni «Usa solo il controllo di accesso IAM» abilitate per la compatibilità con AWS Glue Data Catalog il comportamento esistente. Queste impostazioni ti consentono di gestire l'accesso ai tuoi dati nel data lake e ai relativi metadati tramite policy IAM e policy bucket di Amazon S3.

Per facilitare la transizione delle autorizzazioni del data lake da un modello IAM e Amazon S3 alle autorizzazioni Lake Formation, ti consigliamo di utilizzare la modalità di accesso ibrida per Data Catalog. Con la modalità di accesso ibrida, hai a disposizione un percorso incrementale in cui puoi abilitare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere altri utenti o carichi di lavoro esistenti.

Per ulteriori informazioni, consulta [Modalità di accesso ibrida](#).

Disabilita le impostazioni predefinite per spostare tutti gli utenti esistenti di una tabella su Lake Formation in un unico passaggio.

Important

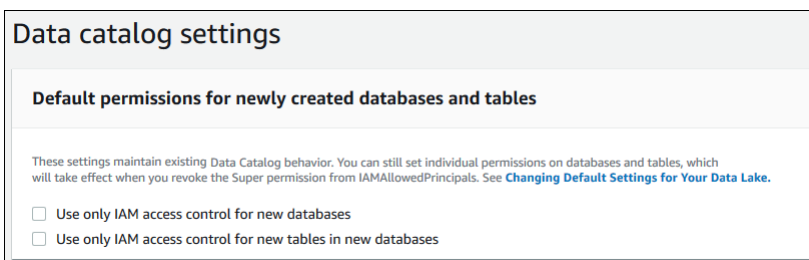
Se disponi di AWS Glue Data Catalog database e tabelle esistenti, non seguire le istruzioni in questa sezione. Seguire invece le istruzioni in [the section called “Aggiornamento delle autorizzazioni per AWS Glue i dati al modello Lake Formation”](#).

Warning

Se disponi di un'automazione che crea database e tabelle nel Data Catalog, i passaggi seguenti potrebbero causare il fallimento dei processi di automazione e di estrazione, trasformazione e caricamento (ETL) a valle. Procedi solo dopo aver modificato i processi esistenti o concesso autorizzazioni esplicite di Lake Formation ai responsabili richiesti. Per informazioni sulle autorizzazioni di Lake Formation, vedere [the section called “Riferimento alle autorizzazioni di Lake Formation”](#).

Per modificare le impostazioni predefinite del Data Catalog

1. Continua nella console di Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Assicurati di aver effettuato l'accesso come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come utente con la policy AdministratorAccess AWS gestita.
2. Modifica le impostazioni del Data Catalog:
 - a. Nel riquadro di navigazione, in Amministrazione, scegli Impostazioni Data Catalog.
 - b. Deseleziona entrambe le caselle di controllo e scegli Salva.



3. Revoca l'IAMAllowedPrincipals autorizzazione per i creatori di database.

- a. Nel riquadro di navigazione, in Amministrazione, scegli Ruoli e attività amministrative.
- b. Nella pagina della console Ruoli e attività amministrative, nella sezione Creatori di database, seleziona il `IAMAllowedPrincipals` gruppo e scegli Revoca.

Viene visualizzata la finestra di dialogo Revoca autorizzazioni, che indica che **IAMAllowedPrincipals** dispone dell'autorizzazione Crea database.

- c. Scegli Revoke.

Assegna le autorizzazioni agli utenti di Lake Formation

Crea un utente per avere accesso al data lake in. AWS Lake Formation Questo utente dispone dei privilegi minimi per interrogare il data lake.

Per ulteriori informazioni sulla creazione di utenti o gruppi, consulta le [identità IAM nella IAM User Guide](#).

Per assegnare a un utente non amministratore le autorizzazioni per accedere ai dati di Lake Formation

1. Apri la console IAM all'indirizzo <https://console.aws.amazon.com/iam> e accedi come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come utente con la policy `AdministratorAccess` AWS gestita.
2. Scegli Utenti o Gruppi di utenti.
3. Nell'elenco, scegli il nome dell'utente o del gruppo in cui integrare una policy.

Seleziona Autorizzazioni.

4. Scegli Aggiungi autorizzazioni e scegli Allega direttamente le politiche. Inserisci Athena nel campo di testo Filtra le politiche. Nell'elenco dei risultati, seleziona la casella `perAmazonAthenaFullAccess`.
5. Scegli il pulsante Crea politica. Nella pagina Crea policy, scegli la scheda JSON. Copia e incolla il seguente codice nell'editor delle politiche.

JSON

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

```
"Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "lakeformation:GetDataAccess",
      "glue:GetTable",
      "glue:GetTables",
      "glue:SearchTables",
      "glue:GetDatabase",
      "glue:GetDatabases",
      "glue:GetPartitions",
      "lakeformation:GetResourceLFTags",
      "lakeformation:ListLFTags",
      "lakeformation:GetLFTag",
      "lakeformation:SearchTablesByLFTags",
      "lakeformation:SearchDatabasesByLFTags"
    ],
    "Resource": "*"
  }
]
```

6. Scegli il pulsante Avanti in basso fino a visualizzare la pagina di revisione della politica. Inserisci un nome per la politica, ad esempio `DataLakeUserBasic`. Scegli Crea politica, quindi chiudi la scheda Politiche o la finestra del browser.

Configura una posizione Amazon S3 per il tuo data lake

Per utilizzare Lake Formation per gestire e proteggere i dati nel tuo data lake, devi prima registrare una sede Amazon S3. Quando registri una posizione, vengono registrati quel percorso Amazon S3 e tutte le cartelle in quel percorso, il che consente a Lake Formation di applicare le autorizzazioni a livello di storage. Quando l'utente richiede dati da un motore integrato come Amazon Athena, Lake Formation fornisce l'accesso ai dati anziché utilizzare le autorizzazioni degli utenti.

Quando registri una posizione, specifichi un ruolo IAM che concede read/write le autorizzazioni su quella posizione. Lake Formation assume questo ruolo quando fornisce credenziali temporanee a AWS servizi integrati che richiedono l'accesso ai dati nella posizione registrata di Amazon S3. Puoi specificare il ruolo collegato ai servizi (SLR) di Lake Formation o creare il tuo ruolo.

Utilizzate un ruolo personalizzato nelle seguenti situazioni:

- Prevedi di pubblicare i parametri in Amazon CloudWatch Logs. Il ruolo definito dall'utente deve includere una politica per l'aggiunta di log in CloudWatch Logs e la pubblicazione di metriche oltre alle autorizzazioni SLR. Per un esempio di politica in linea che concede le autorizzazioni necessarie, vedi. CloudWatch [Requisiti per i ruoli utilizzati per registrare le sedi](#)
- La posizione Amazon S3 si trova in un account diverso. Per informazioni dettagliate, vedi [the section called "Registrazione di una sede Amazon S3 in un altro account AWS"](#).
- La posizione Amazon S3 contiene dati crittografati con un. Chiave gestita da AWS Per informazioni dettagliate, consulta [Registrazione di una posizione Amazon S3 crittografata](#) e [Registrazione di una posizione Amazon S3 crittografata tra più account AWS](#).
- Prevedi di accedere alla posizione Amazon S3 utilizzando Amazon EMR. Per ulteriori informazioni sui requisiti dei ruoli, consulta [IAM roles for Lake Formation](#) nella Amazon EMR Management Guide.

Il ruolo scelto deve disporre delle autorizzazioni necessarie, come descritto in. [Requisiti per i ruoli utilizzati per registrare le sedi](#) Per istruzioni su come registrare una sede Amazon S3, consulta. [Aggiungere una posizione Amazon S3 al tuo data lake](#)

(Facoltativo) Impostazioni di filtraggio dei dati esterni

Se intendi analizzare ed elaborare i dati nel tuo data lake utilizzando motori di query di terze parti, devi attivare l'opzione per consentire ai motori esterni di accedere ai dati gestiti da Lake Formation. Se non effettui l'attivazione, i motori esterni non saranno in grado di accedere ai dati nelle sedi Amazon S3 registrate con Lake Formation.

Lake Formation supporta le autorizzazioni a livello di colonna per limitare l'accesso a colonne specifiche di una tabella. Servizi di analisi integrati come Amazon Athena Amazon Redshift Spectrum e Amazon EMR recuperano i metadati delle tabelle non filtrati da. AWS Glue Data Catalog L'effettivo filtraggio delle colonne nelle risposte alle query è responsabilità del servizio integrato. È responsabilità degli amministratori di terze parti gestire correttamente le autorizzazioni per evitare l'accesso non autorizzato ai dati.

Per consentire ai motori di terze parti di accedere e filtrare i dati (console)

1. Continua nella console di Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Assicurati di aver effettuato l'accesso come principale con l'autorizzazione IAM sull'operazione dell'PutDataLakeSettingsAPI Lake Formation. L'utente amministratore IAM in cui hai creato [Iscriviti per un Account AWS](#) dispone di questa autorizzazione.

2. Nel riquadro di navigazione, in Amministrazione, scegli Impostazioni di integrazione delle applicazioni.
3. Nella pagina delle impostazioni di integrazione dell'applicazione, procedi come segue:
 - a. Seleziona la casella Consenti ai motori esterni di filtrare i dati nelle località Amazon S3 registrate con Lake Formation.
 - b. Inserisci i valori dei tag di sessione definiti per i motori di terze parti.
 - c. Per AWS account IDs, inserisci l'account IDs da cui i motori di terze parti possono accedere alle località registrate con Lake Formation. Premi Invio dopo ogni ID account.
 - d. Scegli Save (Salva).

Per consentire ai motori esterni di accedere ai dati senza la convalida dei tag di sessione, vedi [Integrazione delle applicazioni per l'accesso completo alla tabella](#)

(Facoltativo) Concedi l'accesso alla chiave di crittografia Data Catalog

Se AWS Glue Data Catalog è crittografato, concedi le autorizzazioni AWS Identity and Access Management (IAM) sulla AWS KMS chiave a tutti i responsabili che devono concedere le autorizzazioni di Lake Formation sui database e sulle tabelle del Data Catalog.

Per ulteriori informazioni, consulta la Guida per gli sviluppatori di AWS Key Management Service .

(Facoltativo) Crea un ruolo IAM per i flussi di lavoro

Con AWS Lake Formation, puoi importare i tuoi dati utilizzando flussi di lavoro eseguiti dai AWS Glue crawler. Un flusso di lavoro definisce l'origine dei dati e la pianificazione per l'importazione dei dati nel data lake. Puoi definire facilmente i flussi di lavoro utilizzando i blueprint o i modelli forniti da Lake Formation.

Quando crei un flusso di lavoro, devi assegnargli un ruolo AWS Identity and Access Management (IAM) che conceda a Lake Formation le autorizzazioni necessarie per importare i dati.

La procedura seguente presuppone la familiarità con IAM.

Per creare un ruolo IAM per i flussi di lavoro

1. Apri la console IAM all'indirizzo <https://console.aws.amazon.com/iam> e accedi come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come utente con la policy AdministratorAccess AWS gestita.

2. Nel pannello di navigazione, scegli Ruoli, quindi Crea ruolo.
3. Nella pagina Crea ruolo, scegli AWS service, quindi scegli Glue. Scegli Next (Successivo).
4. Nella pagina Aggiungi autorizzazioni, cerca la politica AWSGlueServiceRolegestita e seleziona la casella di controllo accanto al nome della politica nell'elenco. Quindi completa la procedura guidata di creazione del ruolo, assegnando un nome al ruolo. LFWorkflowRole Per finire, scegli Crea ruolo.
5. Torna alla pagina RuoliLFWorkflowRole, cerca e scegli il nome del ruolo.
6. Nella pagina di riepilogo del ruolo, nella scheda Autorizzazioni, scegli Crea policy in linea. Nella schermata Crea policy, vai alla scheda JSON e aggiungi la seguente policy in linea. Un nome suggerito per la policy è. LakeFormationWorkflow

Important

Nella seguente politica, sostituiscila *<account-id>* con un Account AWS numero valido.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "lakeformation:GrantPermissions"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": ["iam:PassRole"],
      "Resource": [
        "arn:aws:iam::111122223333:role/LakeFormationWorkflowRole"
      ]
    }
  ]
}
```



```
}
```

Di seguito sono riportate brevi descrizioni delle autorizzazioni contenute in questa politica:

- `lakeformation:GetDataAccess` consente ai lavori creati dal flusso di lavoro di scrivere nella posizione di destinazione.
 - `lakeformation:GrantPermissions` consente al flusso di lavoro di concedere l'`SELECT` autorizzazione sulle tabelle di destinazione.
 - `iam:PassRole` consente al servizio di assumere il ruolo di `LakeFormationWorkflowRole` creare crawler e job (istanze di flussi di lavoro) e di assegnare il ruolo ai crawler e ai job creati.
7. Verifica che al ruolo siano associate due policy. `LakeFormationWorkflowRole`
 8. Se stai importando dati che si trovano all'esterno della posizione del data lake, aggiungi una policy in linea che conceda le autorizzazioni per leggere i dati di origine.

Aggiornamento delle AWS Glue autorizzazioni relative ai dati al modello AWS Lake Formation

AWS Lake Formation le autorizzazioni consentono un controllo granulare degli accessi per i dati nel data lake. Puoi utilizzare il modello di autorizzazioni Lake Formation per gestire AWS Glue Data Catalog gli oggetti e le posizioni dei dati esistenti in Amazon Simple Storage Service (Amazon S3).

Il modello di autorizzazioni Lake Formation utilizza autorizzazioni a grana grossa AWS Identity and Access Management (IAM) per l'accesso ai servizi API. Lake Formation utilizza la [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#) funzionalità per limitare l'accesso alla tabella a livello di colonna, riga e cella per gli utenti e le relative applicazioni. In confronto, il AWS Glue modello garantisce l'accesso ai dati tramite politiche IAM [basate sull'identità e sulle risorse](#).

Per effettuare il passaggio, segui i passaggi di questa guida.

Per ulteriori informazioni, consulta [Panoramica delle autorizzazioni di Lake Formation](#).

Informazioni sulle autorizzazioni predefinite

Per mantenere la compatibilità con le versioni precedenti AWS Glue, per impostazione predefinita, AWS Lake Formation concede l'Superautorizzazione al `IAMAllowedPrincipals` gruppo su tutte le risorse del AWS Glue Data Catalog esistenti e concede l'Superautorizzazione sulle nuove risorse

del Data Catalog se le impostazioni di controllo di accesso Use only IAM sono abilitate. Ciò fa sì che l'accesso alle risorse del Data Catalog e alle sedi Amazon S3 sia controllato esclusivamente da policy AWS Identity and Access Management (IAM). Il IAMAllowedPrincipals gruppo include tutti gli utenti e i ruoli IAM a cui è consentito l'accesso agli oggetti del Data Catalog dalle politiche IAM. L'Superautorizzazione consente a un principale di eseguire tutte le operazioni di Lake Formation supportate sul database o sulla tabella su cui è concessa.

Puoi iniziare a utilizzare Lake Formation per gestire l'accesso ai tuoi dati registrando le posizioni delle risorse Data Catalog esistenti in Lake Formation o utilizzando la modalità di accesso ibrida. Quando registri una posizione Amazon S3 in modalità di accesso ibrido, puoi abilitare le autorizzazioni Lake Formation optando per i principali database e tabelle in quella posizione.

Per facilitare la transizione delle autorizzazioni del data lake da un modello IAM e Amazon S3 alle autorizzazioni Lake Formation, ti consigliamo di utilizzare la modalità di accesso ibrida per Data Catalog. Con la modalità di accesso ibrida, hai a disposizione un percorso incrementale in cui puoi abilitare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere altri utenti o carichi di lavoro esistenti.

Per ulteriori informazioni, consulta [Modalità di accesso ibrida](#).

Disabilita le impostazioni predefinite del Data Catalog per spostare tutti gli utenti esistenti di una tabella su Lake Formation in un unico passaggio.

Per iniziare a utilizzare le autorizzazioni di Lake Formation con i database e le tabelle AWS Glue Data Catalog esistenti, devi fare quanto segue:

1. Determina le autorizzazioni IAM esistenti degli utenti per ogni database e tabella.
2. Replica queste autorizzazioni in Lake Formation.
3. Per ogni posizione Amazon S3 che contiene dati:
 - a. Revoca l'Superautorizzazione del IAMAllowedPrincipals gruppo su ogni risorsa del catalogo dati che fa riferimento a quella posizione.
 - b. Registra la posizione con Lake Formation.
4. Pulisci le politiche IAM di controllo degli accessi a grana fine esistenti.

Important

Per aggiungere nuovi utenti durante il processo di transizione del Data Catalog, devi configurare le AWS Glue autorizzazioni granulari in IAM come in precedenza. È inoltre

necessario replicare tali autorizzazioni in Lake Formation come descritto in questa sezione. Se i nuovi utenti dispongono delle politiche IAM dettagliate descritte in questa guida, possono elencare tutti i database o le tabelle a cui è stata concessa l'autorizzazione. Super IAMAllowedPrincipals Possono anche visualizzare i metadati di tali risorse.

Segui i passaggi in questa sezione per eseguire l'aggiornamento al modello di autorizzazioni Lake Formation.

Argomenti

- [Passaggio 1: Elenca le autorizzazioni esistenti di utenti e ruoli](#)
- [Fase 2: Impostare le autorizzazioni equivalenti di Lake Formation](#)
- [Passaggio 3: concedere agli utenti le autorizzazioni IAM per utilizzare Lake Formation](#)
- [Passaggio 4: Passa i tuoi archivi dati al modello di autorizzazioni Lake Formation](#)
- [Fase 5: Proteggi le nuove risorse del Data Catalog](#)
- [Fase 6: Offri agli utenti una nuova policy IAM per i futuri accessi ai data lake](#)
- [Fase 7: Pulisci le policy IAM esistenti](#)

Passaggio 1: Elenca le autorizzazioni esistenti di utenti e ruoli

Per iniziare a utilizzare AWS Lake Formation le autorizzazioni con i AWS Glue database e le tabelle esistenti, devi prima determinare le autorizzazioni esistenti degli utenti.

Important

Prima di iniziare, assicurati di aver completato le attività in. [Nozioni di base](#)

Argomenti

- [Utilizzo dell'operazione API](#)
- [Usando il Console di gestione AWS](#)
- [Usando AWS CloudTrail](#)

Utilizzo dell'operazione API

Utilizza l'operazione [ListPoliciesGrantingServiceAccess](#) API AWS Identity and Access Management (IAM) per determinare le policy IAM associate a ciascun principale (utente o ruolo). In base alle policy restituite nei risultati, puoi determinare le autorizzazioni IAM concesse al principale. È necessario richiamare l'API per ogni principale separatamente.

Example

L' AWS CLI esempio seguente restituisce le politiche allegate all'utente `glue_user1`.

```
aws iam list-policies-granting-service-access --arn arn:aws:iam::111122223333:user/glue_user1 --service-namespaces glue
```

Il comando restituisce risultati simili ai seguenti.

```
{
  "PoliciesGrantingServiceAccess": [
    {
      "ServiceNamespace": "glue",
      "Policies": [
        {
          "PolicyType": "INLINE",
          "PolicyName": "GlueUserBasic",
          "EntityName": "glue_user1",
          "EntityType": "USER"
        },
        {
          "PolicyType": "MANAGED",
          "PolicyArn": "arn:aws:iam::aws:policy/AmazonAthenaFullAccess",
          "PolicyName": "AmazonAthenaFullAccess"
        }
      ]
    }
  ],
  "IsTruncated": false
}
```

Usando il Console di gestione AWS

Puoi anche visualizzare queste informazioni sulla console AWS Identity and Access Management (IAM), nella scheda Access Advisor nella pagina di riepilogo dell'utente o del ruolo:

1. Aprire la console IAM all'indirizzo <https://console.aws.amazon.com/iam/>.
2. Nel riquadro di navigazione, selezionare Users (Utenti) o Roles (Ruoli).
3. Scegli un nome nell'elenco per aprire la relativa pagina di riepilogo e scegli la scheda Access Advisor.
4. Ispeziona ciascuna politica per determinare la combinazione di database, tabelle e azioni per cui ogni utente dispone delle autorizzazioni.

Ricorda di controllare i ruoli oltre agli utenti durante questo processo, perché i tuoi processi di elaborazione dei dati potrebbero assumere ruoli per accedere ai dati.

Usando AWS CloudTrail

Un altro modo per determinare le autorizzazioni esistenti consiste nel AWS CloudTrail cercare le chiamate AWS Glue API in cui il `additionalEventData` campo dei log contiene una `insufficientLakeFormationPermissions` voce. Questa voce elenca il database e la tabella su cui l'utente necessita delle autorizzazioni di Lake Formation per eseguire la stessa azione.

Si tratta di registri di accesso ai dati, quindi non è garantito che producano un elenco completo degli utenti e delle relative autorizzazioni. Ti consigliamo di scegliere un ampio intervallo di tempo per registrare la maggior parte dei modelli di accesso ai dati degli utenti, ad esempio diverse settimane o mesi.

Per ulteriori informazioni, consulta [Visualizzazione degli eventi con cronologia degli CloudTrail eventi](#) nella Guida per l'AWS CloudTrail utente.

Successivamente, puoi impostare le autorizzazioni di Lake Formation in modo che corrispondano alle AWS Glue autorizzazioni. Per informazioni, consulta [Fase 2: Impostare le autorizzazioni equivalenti di Lake Formation](#).

Fase 2: Impostare le autorizzazioni equivalenti di Lake Formation

Utilizzando le informazioni raccolte in [Passaggio 1: Elenca le autorizzazioni esistenti di utenti e ruoli](#), concedi le AWS Lake Formation autorizzazioni corrispondenti alle autorizzazioni. AWS Glue Utilizzate uno dei seguenti metodi per eseguire le concessioni:

- Usa la console Lake Formation o il AWS CLI.

Per informazioni, consulta [the section called “Concessione delle autorizzazioni relative ai dati”](#).

- Usa le operazioni `GrantPermissions` o `BatchGrantPermissions` API.

Per informazioni, consulta [Autorizzazioni APIs](#).

Per ulteriori informazioni, consulta [Panoramica delle autorizzazioni di Lake Formation](#).

Dopo aver impostato le autorizzazioni di Lake Formation, procedi a [Passaggio 3: concedere agli utenti le autorizzazioni IAM per utilizzare Lake Formation](#).

Passaggio 3: concedere agli utenti le autorizzazioni IAM per utilizzare Lake Formation

Per utilizzare il modello di AWS Lake Formation autorizzazioni, i mandanti devono disporre delle autorizzazioni AWS Identity and Access Management (IAM) sulla Lake Formation. APIs

Crea la seguente policy in IAM e collegala a ogni utente che ha bisogno di accedere al tuo data lake. Assegnare un nome alla policy `LakeFormationDataAccess`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationDataAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": "*"
    }
  ]
}
```

Successivamente, esegui l'upgrade alle autorizzazioni di Lake Formation una posizione dati alla volta. Per informazioni, consulta [Passaggio 4: Passa i tuoi archivi dati al modello di autorizzazioni Lake Formation](#).

Passaggio 4: Passa i tuoi archivi dati al modello di autorizzazioni Lake Formation

Passa alle autorizzazioni di Lake Formation una posizione dati alla volta. A tale scopo, ripeti l'intera sezione fino a registrare tutti i percorsi di Amazon Simple Storage Service (Amazon S3) a cui fa riferimento il tuo Data Catalog.

Argomenti

- [Verifica le autorizzazioni di Lake Formation](#)
- [Proteggi le risorse del Data Catalog esistenti](#)
- [Attiva le autorizzazioni di Lake Formation per la tua sede Amazon S3](#)

Verifica le autorizzazioni di Lake Formation

Prima di registrare una sede, esegui una procedura di verifica per assicurarti che i mandanti corretti dispongano delle autorizzazioni Lake Formation richieste e che non vengano concesse autorizzazioni Lake Formation ai mandanti che non dovrebbero averle. Utilizzando l'operazione `GetEffectivePermissionsForPath` API Lake Formation, identifica le risorse del Data Catalog che fanno riferimento alla posizione Amazon S3, insieme ai responsabili che dispongono delle autorizzazioni su tali risorse.

L' AWS CLI esempio seguente restituisce i database e le tabelle del catalogo dati che fanno riferimento al bucket Amazon S3. `products`

```
aws lakeformation get-effective-permissions-for-path --resource-arn
arn:aws:s3:::products --profile datalake_admin
```

Nota l'opzione. `profile` Ti consigliamo di eseguire il comando come amministratore del data lake.

Di seguito è riportato un estratto dei risultati restituiti.

```
{
  "PermissionsWithGrantOption": [
    "SELECT"
  ],
  "Resource": {
    "TableWithColumns": {
      "Name": "inventory_product",
      "ColumnWildcard": {},

```

```
        "DatabaseName": "inventory"
      },
      "Permissions": [
        "SELECT"
      ],
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1",
        "DataLakePrincipalType": "IAM_USER"
      }
    }, ...
```

Important

Se il tuo AWS Glue Data Catalog è crittografato, `GetEffectivePermissionsForPath` restituisce solo database e tabelle che sono stati creati o modificati dopo la disponibilità generale di Lake Formation.

Proteggi le risorse del Data Catalog esistenti

Successivamente, revoca l'Superautorizzazione da `IAMAllowedPrincipals` ogni tabella e database che hai identificato per la posizione.

Warning

Se disponi di un'automazione che crea database e tabelle nel Data Catalog, i passaggi seguenti potrebbero causare il fallimento dei processi di automazione e di estrazione, trasformazione e caricamento (ETL) a valle. Procedi solo dopo aver modificato i processi esistenti o concesso autorizzazioni esplicite di Lake Formation ai responsabili richiesti. Per informazioni sulle autorizzazioni di Lake Formation, vedere [the section called “Riferimento alle autorizzazioni di Lake Formation”](#).

Per revocare un account **Super** da un tavolo **IAMAllowedPrincipals**

1. Apri la AWS Lake Formation console all'indirizzo. <https://console.aws.amazon.com/lakeformation/> Accedi come amministratore del data lake.

2. Nel pannello di navigazione, seleziona Tabelle.
3. Nella pagina Tabelle, seleziona il pulsante di opzione accanto alla tabella desiderata.
4. Nel menu Azioni, scegli Revoca.
5. Nella finestra di dialogo Revoca le autorizzazioni, nell'elenco Utenti e ruoli IAM, scorri verso il basso fino all'intestazione Gruppo e scegli Principal. IAMAllowed
6. Nella sezione Autorizzazioni della tabella, assicurati che Super sia selezionato, quindi scegli Revoke.

Per revocare da **Super** un database **IAMAllowedPrincipals**

1. Apri la AWS Lake Formation console all'indirizzo. <https://console.aws.amazon.com/lakeformation/> Accedi come amministratore del data lake.
2. Nel pannello di navigazione, seleziona Database.
3. Nella pagina Database, seleziona il pulsante di opzione accanto al database desiderato.
4. Nel menu Operazioni, scegliere Modifica.
5. Nella pagina Modifica database, deseleziona Usa solo il controllo di accesso IAM per le nuove tabelle in questo database, quindi scegli Salva.
6. Tornando alla pagina Database, assicurati che il database sia ancora selezionato, quindi nel menu Azioni, scegli Revoca.
7. Nella finestra di dialogo Revoca le autorizzazioni, nell'elenco Utenti e ruoli IAM, scorri verso il basso fino all'intestazione Gruppo e scegli Principali. IAMAllowed
8. In Autorizzazioni del database, assicurati che Super sia selezionato, quindi scegli Revoke.

Attiva le autorizzazioni di Lake Formation per la tua sede Amazon S3

Quindi, registra la sede Amazon S3 con Lake Formation. A tale scopo, puoi utilizzare la procedura descritta in [Aggiungere una posizione Amazon S3 al tuo data lake](#). In alternativa, utilizzate l'operazione RegisterResource API come descritto in [Vendita di credenziali APIs](#).

Note

Se la sede di un genitore è registrata, non è necessario registrare le sedi dei figli.

Dopo aver completato questi passaggi e aver verificato che i tuoi utenti possano accedere ai propri dati, hai eseguito con successo l'upgrade alle autorizzazioni di Lake Formation. Continua con il passaggio successivo, [Fase 5: Proteggi le nuove risorse del Data Catalog](#)

Fase 5: Proteggi le nuove risorse del Data Catalog

Successivamente, proteggi tutte le nuove risorse del Data Catalog modificando le impostazioni predefinite del Data Catalog. Disattiva le opzioni per utilizzare solo il controllo di accesso AWS Identity and Access Management (IAM) per nuovi database e tabelle.

Warning

Se disponi di un'automazione che crea database e tabelle nel Data Catalog, i passaggi seguenti potrebbero causare il fallimento dei processi di automazione e di estrazione, trasformazione e caricamento (ETL) a valle. Procedi solo dopo aver modificato i processi esistenti o concesso autorizzazioni esplicite di Lake Formation ai responsabili richiesti. Per informazioni sulle autorizzazioni di Lake Formation, vedere [the section called “Riferimento alle autorizzazioni di Lake Formation”](#).

Per modificare le impostazioni predefinite del Data Catalog

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come utente amministrativo IAM (l'utente Administrator o un altro utente con la policy AdministratorAccess AWS gestita).
2. Nel pannello di navigazione scegli Impostazioni.
3. Nella pagina delle impostazioni del catalogo dati, deseleziona entrambe le caselle di controllo, quindi scegli Salva.

Il passaggio successivo consiste nel concedere agli utenti l'accesso a database o tabelle aggiuntivi in futuro. Per informazioni, consulta [Fase 6: Offri agli utenti una nuova policy IAM per i futuri accessi ai data lake](#).

Fase 6: Offri agli utenti una nuova policy IAM per i futuri accessi ai data lake

Per concedere ai tuoi utenti l'accesso a database o tabelle di Data Catalog aggiuntivi in futuro, devi fornire loro la policy in linea a grana grossa AWS Identity and Access Management (IAM) che segue. Assegnare un nome alla policy GlueFullReadAccess.

 Important

Se alleggi questa policy a un utente prima di revocarla Super IAMAllowedPrincipals su ogni database e tabella del tuo Data Catalog, quell'utente può visualizzare tutti i metadati di qualsiasi risorsa a cui è concesso. Super IAMAllowedPrincipals

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "GlueFullReadAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions"
      ],
      "Resource": "*"
    }
  ]
}
```

 Note

Le politiche in linea indicate in questo passaggio e nei passaggi precedenti contengono autorizzazioni IAM minime. Per le policy suggerite per gli amministratori di data lake, gli analisti di dati e altri personaggi, consulta [the section called “Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM”](#)

Quindi, procedi a [Fase 7: Pulisci le policy IAM esistenti](#)

Fase 7: Pulisci le policy IAM esistenti

Dopo aver configurato le AWS Lake Formation autorizzazioni e aver creato e collegato le politiche di controllo degli accessi AWS Identity and Access Management (IAM) granulari, completa il seguente passaggio finale:

- Rimuovi da utenti, gruppi e ruoli le vecchie policy IAM di [controllo degli accessi granulari](#) che hai replicato in Lake Formation.

In questo modo, ti assicuri che tali responsabili non abbiano più accesso diretto ai dati in Amazon Simple Storage Service (Amazon S3). Puoi quindi gestire l'accesso al data lake per questi principali interamente tramite Lake Formation.

AWS Lake Formation e endpoint VPC di interfaccia ()AWS PrivateLink

Amazon VPC è un AWS servizio che puoi utilizzare per avviare AWS risorse in una rete virtuale definita dall'utente. Con un VPC, detieni il controllo delle impostazioni della rete, come l'intervallo di indirizzi IP, le sottoreti, le tabelle di routing e i gateway di rete.

Se utilizzi Amazon Virtual Private Cloud (Amazon VPC) per ospitare AWS le tue risorse, puoi stabilire una connessione privata tra il tuo VPC e Lake Formation. Utilizzi questa connessione in modo che Lake Formation possa comunicare con le risorse del tuo VPC senza passare attraverso la rete Internet pubblica.

Puoi stabilire una connessione privata tra il tuo VPC e creare un AWS Lake Formation endpoint VPC di interfaccia. Gli endpoint di interfaccia sono alimentati da [AWS PrivateLink](#), una tecnologia che consente di accedere privatamente a Lake Formation APIs senza un gateway Internet, un dispositivo NAT, una connessione VPN o una connessione. Direct Connect Le istanze del tuo VPC non necessitano di indirizzi IP pubblici per comunicare con Lake Formation. APIs Il traffico tra il tuo VPC e Lake Formation non esce dalla rete Amazon.

Ogni endpoint dell'interfaccia è rappresentato da una o più [interfacce di rete elastiche](#) nelle sottoreti.

Per ulteriori informazioni, consultare [Endpoint VPC di interfaccia \(AWS PrivateLink\)](#) nella Guida per l'utente di Amazon VPC.

Considerazioni sugli endpoint VPC di Lake Formation

Prima di configurare un endpoint VPC di interfaccia per Lake Formation, assicurati di esaminare le [proprietà e le limitazioni degli endpoint dell'interfaccia nella](#) Amazon VPC User Guide.

Lake Formation supporta l'esecuzione di chiamate a tutte le sue azioni API dal tuo VPC. Puoi utilizzare Lake Formation con endpoint VPC in tutti gli endpoint Regioni AWS che supportano sia gli endpoint Lake Formation che Amazon VPC.

Creazione di un endpoint VPC di interfaccia per Lake Formation

Puoi creare un endpoint VPC per il servizio Lake Formation utilizzando la console Amazon VPC o il (. AWS Command Line Interface AWS CLI Per ulteriori informazioni, consultare [Creazione di un endpoint di interfaccia](#) nella Guida per l'utente di Amazon VPC.

Crea un endpoint VPC per Lake Formation utilizzando il seguente nome di servizio:

- com.amazonaws. *region*. formazione di laghi

Se abiliti il DNS privato per l'endpoint, puoi effettuare richieste API a Lake Formation utilizzando il nome DNS predefinito per la regione, ad esempio. lakeformation.us-east-1.amazonaws.com

Per ulteriori informazioni, consultare [Accesso a un servizio tramite un endpoint di interfaccia](#) in Guida per l'utente di Amazon VPC.

Creazione di una policy sugli endpoint VPC per Lake Formation

Lake Formation supporta le policy degli endpoint VPC. Una policy per gli endpoint è una policy basata sulle risorse che si collega a un endpoint VPC per controllare quali AWS responsabili possono utilizzare l'endpoint per accedere a un servizio. AWS

Puoi allegare una policy per gli endpoint al tuo endpoint VPC che controlla l'accesso a Lake Formation. La policy specifica le informazioni riportate di seguito:

- Il principale che può eseguire operazioni.
- Le azioni che possono essere eseguite.
- Le risorse sui cui si possono eseguire operazioni.

Per ulteriori informazioni, consulta [Controllo degli accessi ai servizi con endpoint VPC](#) in Guida per l'utente di Amazon VPC.

Esempio: policy sugli endpoint VPC per le azioni di Lake Formation

Il seguente esempio di politica degli endpoint VPC per Lake Formation consente la vendita di credenziali utilizzando le autorizzazioni di Lake Formation. Puoi utilizzare questa policy per eseguire query utilizzando le autorizzazioni di Lake Formation da un cluster Amazon Redshift o da un cluster situato in Amazon EMR una sottorete privata.

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "lakeformation:GetDataAccess",
      "Resource": "*",
      "Principal": "*"
    }
  ]
}
```

 Note

Se non alleggi una policy quando crei un endpoint, viene allegata una policy predefinita che consente l'accesso completo al servizio.

Per ulteriori informazioni, consulta questi argomenti nella documentazione di Amazon VPC:

- [What Is Amazon VPC?](#)
- [Crea un endpoint di interfaccia](#)
- [Usa le policy degli endpoint VPC](#)

AWS Lake Formation tutorial

I seguenti tutorial sono organizzati in tre tracce e forniscono step-by-step istruzioni su come creare un data lake, importare dati, condividere e proteggere i data lake utilizzando: AWS Lake Formation

1. Crea un data lake e inserisci dati: impara a creare un data lake e usa i blueprint per spostare, archiviare, catalogare, pulire e organizzare i dati. Imparerai anche a configurare tabelle gestite. Una tabella governata è un nuovo tipo di tabella Amazon S3 che supporta transazioni atomiche, consistenti, isolate e durevoli (ACID).

Prima di iniziare, assicurati di aver completato i passaggi indicati. [Guida introduttiva a Lake Formation](#)

- [Creare un data lake da una fonte AWS CloudTrail](#)

Crea e carica il tuo primo data lake utilizzando i tuoi CloudTrail log come fonte di dati.

- [Creazione di un data lake da una sorgente JDBC in Lake Formation](#)

Crea un data lake utilizzando uno dei tuoi data store accessibili con JDBC, ad esempio un database relazionale, come origine dati.

2. Protezione dei data lake: impara a utilizzare controlli di accesso basati su tag e a livello di riga per proteggere e gestire efficacemente l'accesso ai tuoi data lake.

- [Impostazione delle autorizzazioni per i formati di archiviazione a tabella aperta in Lake Formation](#)

Questo tutorial dimostra come impostare le autorizzazioni per i formati di tabelle transazionali open source (tabelle Apache Iceberg, Apache Hudi e Linux Foundation Delta Lake) in Lake Formation.

- [Gestione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation](#)

Impara a gestire l'accesso ai dati all'interno di un data lake utilizzando il controllo degli accessi basato su tag in Lake Formation.

- [Protezione dei data lake con il controllo degli accessi a livello di riga](#)

Scopri come configurare le autorizzazioni a livello di riga che ti consentono di limitare l'accesso a righe specifiche in base alle politiche di conformità e governance dei dati in Lake Formation.

3. Condivisione dei dati: impara a condividere in modo sicuro i tuoi dati Account AWS utilizzando il controllo degli accessi basato su tag (TBAC) e gestisci le autorizzazioni granulari sui set di dati condivisi tra Account AWS

- [Condivisione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation e risorse denominate](#)

In questo tutorial, imparerai come condividere in modo sicuro i tuoi dati Account AWS utilizzando Lake Formation.

- [Condivisione di un data lake utilizzando il controllo granulare degli accessi di Lake Formation](#)

In questo tutorial, imparerai come condividere in modo rapido e semplice set di dati utilizzando Lake Formation quando gestisci più set Account AWS con AWS Organizations.

Argomenti

- [Creare un data lake da una fonte AWS CloudTrail](#)
- [Creazione di un data lake da una sorgente JDBC in Lake Formation](#)
- [Impostazione delle autorizzazioni per i formati di archiviazione a tabella aperta in Lake Formation](#)
- [Gestione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation](#)
- [Protezione dei data lake con il controllo degli accessi a livello di riga](#)
- [Condivisione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation e risorse denominate](#)
- [Condivisione di un data lake utilizzando il controllo granulare degli accessi di Lake Formation](#)

Creare un data lake da una fonte AWS CloudTrail

Questo tutorial ti guida attraverso le azioni da intraprendere sulla console Lake Formation per creare e caricare il tuo primo data lake da una AWS CloudTrail fonte.

Passaggi di alto livello per la creazione di un data lake

1. Registra un percorso Amazon Simple Storage Service (Amazon S3) come data lake.
2. Concedi a Lake Formation le autorizzazioni per scrivere nel Data Catalog e nelle posizioni Amazon S3 nel data lake.
3. Crea un database per organizzare le tabelle di metadati nel Data Catalog.

4. Usa un blueprint per creare un flusso di lavoro. Esegui il flusso di lavoro per importare dati da una fonte di dati.
5. Configura le tue autorizzazioni Lake Formation per consentire ad altri di gestire i dati nel Data Catalog e nel data lake.
6. Configura Amazon Athena per interrogare i dati che hai importato nel tuo data lake Amazon S3.
7. Per alcuni tipi di data store, configura Amazon Redshift Spectrum per interrogare i dati che hai importato nel tuo data lake Amazon S3.

Argomenti

- [Destinatari principali](#)
- [Prerequisiti](#)
- [Fase 1: Creare un utente analista di dati](#)
- [Passaggio 2: Aggiungere le autorizzazioni per leggere i AWS CloudTrail registri al ruolo del flusso di lavoro](#)
- [Fase 3: creare un bucket Amazon S3 per il data lake](#)
- [Fase 4: Registrare un percorso Amazon S3](#)
- [Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati](#)
- [Fase 6: Creare un database nel Data Catalog](#)
- [Passaggio 7: concedere le autorizzazioni per i dati](#)
- [Fase 8: Utilizzare un blueprint per creare un flusso di lavoro](#)
- [Passaggio 9: Esegui il flusso di lavoro](#)
- [Fase 10: concedere SELECT sui tavoli](#)
- [Passaggio 11: interrogare il data lake utilizzando Amazon Athena](#)

Destinatari principali

La tabella seguente elenca i ruoli utilizzati in questo tutorial per creare un data lake.

Destinatari principali

Ruolo	Description
Amministratore IAM	Ha la politica AWS gestita: <code>AdministratorAccess</code> . Può creare ruoli IAM e bucket Amazon S3.
Amministratore del data lake	Utente che può accedere al catalogo dati, creare database e concedere autorizzazioni Lake Formation ad altri utenti. Dispone di meno autorizzazioni IAM rispetto all'amministratore IAM, ma sufficienti per amministrare il data lake.
Analista dei dati	Utente che può eseguire query sul data lake. Dispone solo delle autorizzazioni sufficienti per eseguire le query.
Ruolo del workflow	Ruolo con le politiche IAM richieste per eseguire un flusso di lavoro. Per ulteriori informazioni, consulta (Facoltativo) Crea un ruolo IAM per i flussi di lavoro .

Prerequisiti

Prima di iniziare:

- Assicuratevi di aver completato le attività in [Configurare AWS Lake Formation](#).
- Conosci la posizione dei tuoi CloudTrail registri.
- Athena richiede all'analista di dati di creare un bucket Amazon S3 per archiviare i risultati delle query prima di utilizzare Athena.

Si presume la familiarità con AWS Identity and Access Management (IAM). Per informazioni su IAM, consulta la [IAM User Guide](#).

Fase 1: Creare un utente analista di dati

Questo utente dispone del set minimo di autorizzazioni per interrogare il data lake.

1. Aprire la console IAM all'indirizzo <https://console.aws.amazon.com/iam>. Accedi come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come utente con la policy AdministratorAccess AWS gestita.
2. Crea un utente denominato `dataLake_user` con le seguenti impostazioni:
 - Abilita Console di gestione AWS l'accesso.
 - Imposta una password e non richiedi la reimpostazione della password.
 - Allega la politica AmazonAthenaFullAccess AWS gestita.
 - Allega la seguente politica in linea. Assegnare un nome alla policy `DataLakeUserBasic`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
      ],
      "Resource": "*"
    }
  ]
}
```

Passaggio 2: Aggiungere le autorizzazioni per leggere i AWS CloudTrail registri al ruolo del flusso di lavoro

1. Allega la seguente politica in linea al ruolo. LakeFormationWorkflowRole La policy concede il permesso di leggere i tuoi AWS CloudTrail log. Assegnare un nome alla policy DatalakeGetCloudTrail.

Per creare il ruolo LakeFormationWorkflowRole, consulta [\(Facoltativo\) Crea un ruolo IAM per i flussi di lavoro](#).

Important

Sostituisci *<your-s3-cloudtrail-bucket>* con la posizione dei tuoi CloudTrail dati in Amazon S3.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "s3:GetObject",
      "Resource": ["arn:aws:s3:::<your-s3-cloudtrail-bucket>/*"]
    }
  ]
}
```

2. Verifica che al ruolo siano associate tre policy.

Fase 3: creare un bucket Amazon S3 per il data lake

Crea il bucket Amazon S3 che sarà la posizione principale del tuo data lake.

1. Apri la console Amazon S3 all'indirizzo <https://console.aws.amazon.com/s3/> e accedi come utente amministratore con cui hai creato. [Crea un utente con accesso amministrativo](#)

2. Scegli Crea bucket e segui la procedura guidata per creare un bucket denominato **<yourName>-datalake-cloudtrail**, dove **<yourName>** sono il tuo nome e cognome. Ad esempio: **jdoe-datalake-cloudtrail**.

Per istruzioni dettagliate sulla creazione di un bucket Amazon S3, consulta [Creazione](#) di un bucket.

Fase 4: Registrare un percorso Amazon S3

Registra un percorso Amazon S3 come posizione principale del tuo data lake.

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake.
2. Nel riquadro di navigazione, in Registra e inserisci, scegli Data lake locations.
3. Scegli Registra posizione e poi Sfoglia.
4. Seleziona il **<yourName>-datalake-cloudtrail** bucket che hai creato in precedenza, accetta il ruolo IAM predefinito **AWSServiceRoleForLakeFormationDataAccess**, quindi scegli Registra posizione.

Per ulteriori informazioni sulla registrazione delle sedi, consulta [Aggiungere una posizione Amazon S3 al tuo data lake](#)

Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati

I responsabili devono disporre delle autorizzazioni di localizzazione dei dati su una posizione di data lake per creare tabelle o database di Data Catalog che puntano a tale posizione. È necessario concedere le autorizzazioni per la localizzazione dei dati al ruolo IAM per i flussi di lavoro in modo che il flusso di lavoro possa scrivere nella destinazione di inserimento dei dati.

1. Nel riquadro di navigazione, in Autorizzazioni, scegli Posizioni dei dati.
2. Scegli Concedi e, nella finestra di dialogo Concedi autorizzazioni, effettua le seguenti selezioni:
 - a. Per utenti e ruoli IAM, scegli. **LakeFormationWorkflowRole**
 - b. Per le posizioni di archiviazione, scegli il tuo **<yourName>-datalake-cloudtrail** bucket.
3. Scegli Concessione.

Per ulteriori informazioni sulle autorizzazioni per la localizzazione dei dati, consulta [Underlying data access control](#)

Fase 6: Creare un database nel Data Catalog

Le tabelle di metadati nel Lake Formation Data Catalog sono archiviate all'interno di un database.

1. Nel riquadro di navigazione, in Catalogo dati, scegli Database.
2. Scegli Crea database e in Dettagli del database, inserisci il `nomelakeformation_cloudtrail`.
3. Lascia vuoti gli altri campi e scegli Crea database.

Passaggio 7: concedere le autorizzazioni per i dati

È necessario concedere le autorizzazioni per creare tabelle di metadati nel Catalogo dati. Poiché il flusso di lavoro verrà eseguito con il ruolo `LakeFormationWorkflowRole`, è necessario concedere queste autorizzazioni al ruolo.

1. Nella console Lake Formation, nel riquadro di navigazione, in Catalogo dati, scegli Databases.
2. Scegli il `lakeformation_cloudtrail` database, quindi, dall'elenco a discesa Azioni, scegli Concedi sotto la voce Autorizzazioni.
3. Nella finestra di dialogo Concedi le autorizzazioni per i dati, effettua le seguenti selezioni:
 - a. In Principali, per utenti e ruoli IAM, scegli `LakeFormationWorkflowRole`
 - b. In LF-Tags o risorse del catalogo, scegli Risorse Catalogo dati denominato.
 - c. Per i database, dovresti vedere che il `lakeformation_cloudtrail` database è già stato aggiunto.
 - d. In Autorizzazioni del database, seleziona Crea tabella, Alter e Drop e deseleziona Super se è selezionato.
4. Scegli Concessione.

Per ulteriori informazioni sulla concessione delle autorizzazioni di Lake Formation, vedere [Gestione delle autorizzazioni di Lake Formation](#)

Fase 8: Utilizzare un blueprint per creare un flusso di lavoro

Per leggere CloudTrail i log, comprenderne la struttura, creare le tabelle appropriate nel Data Catalog, dobbiamo impostare un flusso di lavoro composto da AWS Glue crawler, job, trigger e workflow. I progetti di Lake Formation semplificano questo processo.

Il flusso di lavoro genera i job, i crawler e i trigger che rilevano e inseriscono i dati nel tuo data lake. Crei un flusso di lavoro basato su uno dei blueprint predefiniti di Lake Formation.

1. Nella console di Lake Formation, nel pannello di navigazione, scegli Blueprint in Ingestion, quindi scegli Usa blueprint.
2. Nella pagina Usa un blueprint, in Tipo di progetto, scegli AWS CloudTrail
3. In Importa fonte, scegli una CloudTrail fonte e una data di inizio.
4. In Import target, specifica questi parametri:

Database di destinazione	lakeformation_cloudtrail
Posizione di archiviazione di destinazione	s3://<yourName> -datalake-cloudtrail
Formato dei dati	Parquet

5. Per la frequenza di importazione, scegli Esegui su richiesta.
6. In Opzioni di importazione, specificate questi parametri:

Nome del flusso di lavoro	lakeformationcloudtrailtest
Ruolo IAM	LakeFormationWorkflowRole
Prefisso della tabella	cloudtrailtest



Note

Deve essere in lettere minuscole.

7. Scegli Crea e attendi che la console segnali che il flusso di lavoro è stato creato correttamente.

 Tip

Hai ricevuto il seguente messaggio di errore?

```
User: arn:aws:iam::<account-id>:user/<datalake_administrator_user> is not authorized to
perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/
LakeFormationWorkflowRole...
```

In tal caso, verifica di aver sostituito `<account-id>` nella politica in linea per l'utente amministratore del data lake un numero di AWS account valido.

Passaggio 9: Esegui il flusso di lavoro

Poiché hai specificato che il flusso di lavoro è run-on-demand, devi avviarlo manualmente.

- Nella pagina Blueprint, seleziona il flusso di lavoro `lakeformationcloudtrailtest` e nel menu Azioni scegli Avvia.

Durante l'esecuzione del flusso di lavoro, è possibile visualizzarne l'avanzamento nella colonna Stato dell'ultima esecuzione. Scegli il pulsante di aggiornamento di tanto in tanto.

Lo stato va da RUNNING, a Discovering, a Importing, a COMPLETED.

Al termine del flusso di lavoro:

- Il Data Catalog avrà nuove tabelle di metadati.
- CloudTrail I log verranno inseriti nel data lake.

Se il flusso di lavoro fallisce, procedi come segue:

- a. Seleziona il flusso di lavoro e nel menu Azioni scegli Visualizza grafico.

Il flusso di lavoro viene aperto nella AWS Glue console.

- b. Verifica che il flusso di lavoro sia selezionato e scegli la scheda History (Cronologia).
- c. In Cronologia, seleziona l'esecuzione più recente e scegli Visualizza i dettagli dell'esecuzione.

- d. Seleziona un processo o un crawler non riuscito nel grafico dinamico (di runtime) ed esamina il messaggio di errore. I nodi con errori sono rossi o gialli.

Fase 10: concedere SELECT sui tavoli

È necessario concedere l'AUTHORIZZAZIONE per le nuove tabelle del Catalogo dati in modo che l'analista dei dati possa interrogare i dati a cui fanno riferimento le tabelle.

Note

Un flusso di lavoro concede automaticamente l'AUTHORIZZAZIONE per le tabelle che crea all'utente che lo ha eseguito. Poiché l'amministratore del data lake ha eseguito questo flusso di lavoro, è necessario concederle SELECT all'analista dei dati.

1. Nella console Lake Formation, nel riquadro di navigazione, in Catalogo dati, scegli Databases.
2. Scegli il `lakeformation_cloudtrail` database, quindi, dall'elenco a discesa Azioni, scegli Concedi sotto la voce Autorizzazioni.
3. Nella finestra di dialogo Concedi le autorizzazioni per i dati, effettua le seguenti selezioni:
 - a. In Principali, per utenti e ruoli IAM, scegli. `datalake_user`
 - b. In LF-Tags o Catalog resources, scegli Named data catalog resources.
 - c. Per Databases, il `lakeformation_cloudtrail` database dovrebbe essere già selezionato.
 - d. Per Tabelle, scegli `tecloudtrailtest-cloudtrail`.
 - e. In Autorizzazioni per tabelle e colonne, scegli Seleziona.
4. Scegli Concessione.

Il passaggio successivo viene eseguito come analista dei dati.

Passaggio 11: interrogare il data lake utilizzando Amazon Athena

Usa la Amazon Athena console per interrogare i CloudTrail dati nel tuo data lake.

1. Apri la console Athena all'indirizzo <https://console.aws.amazon.com/athena/> e accedi come analista di dati, utente. `datalake_user`

2. Se necessario, scegli **Inizia** per passare all'editor di query Athena.
3. Per Origine dati scegliere **AwsDataCatalog**.
4. Per Database, scegliere **lakeformation_cloudtrail**.

L'elenco delle tabelle viene compilato.

5. Nel menu a discesa (3 punti disposti orizzontalmente) accanto alla tabella **cloudtrailtest-cloudtrail**, scegliete **Anteprima tabella**, quindi scegliete **Esegui**.

La query viene eseguita e visualizza 10 righe di dati.

Se non hai mai usato Athena prima, devi prima configurare una posizione Amazon S3 nella console Athena per archiviare i risultati delle query. `datalake_user` È necessario disporre delle autorizzazioni necessarie per accedere al bucket Amazon S3 scelto.

Note

Ora che hai completato il tutorial, concedi le autorizzazioni per i dati e le autorizzazioni per la localizzazione dei dati ai responsabili della tua organizzazione.

Creazione di un data lake da una sorgente JDBC in Lake Formation

Questo tutorial illustra i passaggi da eseguire sulla AWS Lake Formation console per creare e caricare il primo data lake da una sorgente JDBC utilizzando Lake Formation.

Argomenti

- [Destinatari principali](#)
- [Prerequisiti del tutorial JDBC](#)
- [Fase 1: Creare un utente analista di dati](#)
- [Fase 2: Creare una connessione in AWS Glue](#)
- [Fase 3: creare un bucket Amazon S3 per il data lake](#)
- [Fase 4: Registrare un percorso Amazon S3](#)
- [Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati](#)
- [Fase 6: Creare un database nel Data Catalog](#)
- [Passaggio 7: concedere le autorizzazioni per i dati](#)

- [Fase 8: Utilizzare un blueprint per creare un flusso di lavoro](#)
- [Passaggio 9: Esegui il flusso di lavoro](#)
- [Fase 10: concedere SELECT sui tavoli](#)
- [Passaggio 11: interrogare il data lake utilizzando Amazon Athena](#)
- [Passaggio 12: interroga i dati nel data lake utilizzando Amazon Redshift Spectrum](#)
- [Passaggio 13: concedere o revocare le autorizzazioni di Lake Formation utilizzando Amazon Redshift Spectrum](#)

Destinatari principali

La tabella seguente elenca i ruoli utilizzati in questo tutorial [AWS Lake Formation JDBC](#).

Ruolo	Descrizione
Amministratore IAM	Un utente che può creare utenti e ruoli AWS Identity and Access Management (IAM) e bucket Amazon Simple Storage Service (Amazon S3). Ha la politica AdministratorAccess AWS gestita.
Amministratore del data lake	Un utente che può accedere al Data Catalog, creare database e concedere autorizzazioni Lake Formation ad altri utenti. Dispone di meno autorizzazioni IAM rispetto all'amministratore IAM, ma sufficienti per amministrare il data lake.
Analista dei dati	Un utente che può eseguire query sul data lake. Dispone solo delle autorizzazioni sufficienti per eseguire le query.
Ruolo del workflow	Un ruolo con le politiche IAM richieste per eseguire un flusso di lavoro.

Per informazioni sui prerequisiti per il completamento del tutorial, consulta [Prerequisiti del tutorial JDBC](#).

Prerequisiti del tutorial JDBC

Prima di iniziare il [tutorial su AWS Lake Formation JDBC](#), assicurati di aver fatto quanto segue:

- Completare le operazioni descritte in [Guida introduttiva a Lake Formation](#).
- Scegli un archivio dati accessibile con JDBC che desideri utilizzare per il tutorial.
- Raccogli le informazioni necessarie per creare una AWS Glue connessione di tipo JDBC. Questo oggetto Data Catalog include l'URL del data store, le credenziali di accesso e, se il data store è stato creato in un Amazon Virtual Private Cloud (Amazon VPC), informazioni di configurazione aggiuntive specifiche per VPC. Per ulteriori informazioni, consulta [Definizione delle connessioni nel catalogo AWS Glue dati nella Guida per gli sviluppatori](#). AWS Glue

Il tutorial presuppone che tu abbia familiarità con AWS Identity and Access Management (IAM). Per informazioni su IAM, consulta la [IAM User Guide](#).

Per iniziare, procedi [alla sezione chiamata "Fase 1: Creare un utente analista di dati"](#).

Fase 1: Creare un utente analista di dati

In questo passaggio, crei un utente AWS Identity and Access Management (IAM) che funga da analista di dati per il tuo data lake in AWS Lake Formation.

Questo utente dispone del set minimo di autorizzazioni per interrogare il data lake.

1. Aprire la console IAM all'indirizzo <https://console.aws.amazon.com/iam>. Accedi come utente amministratore che hai creato [Crea un utente con accesso amministrativo](#) o come utente con la policy AdministratorAccess AWS gestita.
2. Crea un utente denominato `dataLake_user` con le seguenti impostazioni:
 - Abilita Console di gestione AWS l'accesso.
 - Imposta una password e non richiedi la reimpostazione della password.
 - Allega la politica AmazonAthenaFullAccess AWS gestita.
 - Allega la seguente politica in linea. Assegnare un nome alla policy `DataLakeUserBasic`.

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

```
{
  "Effect": "Allow",
  "Action": [
    "lakeformation:GetDataAccess",
    "glue:GetTable",
    "glue:GetTables",
    "glue:SearchTables",
    "glue:GetDatabase",
    "glue:GetDatabases",
    "glue:GetPartitions",
    "lakeformation:GetResourceLFTags",
    "lakeformation:ListLFTags",
    "lakeformation:GetLFTag",
    "lakeformation:SearchTablesByLFTags",
    "lakeformation:SearchDatabasesByLFTags"
  ],
  "Resource": "*"
}
```

Fase 2: Creare una connessione in AWS Glue

Note

Salta questo passaggio se disponi già di una AWS Glue connessione alla tua fonte di dati JDBC.

AWS Lake Formation accede alle sorgenti dati JDBC tramite una connessione. AWS Glue Una connessione è un oggetto Data Catalog che contiene tutte le informazioni necessarie per connettersi all'origine dati. È possibile creare una connessione utilizzando la AWS Glue console.

Per creare una connessione

1. Apri AWS Glue la console all'indirizzo <https://console.aws.amazon.com/glue/> e accedi come utente amministratore con cui hai creato il file [Crea un utente con accesso amministrativo](#).
2. Nel riquadro di navigazione, in Data catalog (Catalogo dati), seleziona Connections (Connessioni).
3. Nella pagina Connettori, scegli Crea connessione.

4. Nella pagina Scegli l'origine dati, scegli JDBC come tipo di connessione. Quindi scegli Successivo.
5. Continua con la procedura guidata di connessione e salva la connessione.

Per informazioni sulla creazione di una connessione, consulta le [proprietà della connessione AWS Glue JDBC](#) nella Guida per gli AWS Glue sviluppatori.

Fase 3: creare un bucket Amazon S3 per il data lake

In questo passaggio, crei il bucket Amazon Simple Storage Service (Amazon S3) che deve essere la posizione principale del tuo data lake.

1. Apri la console Amazon S3 all'indirizzo <https://console.aws.amazon.com/s3/> e accedi come utente amministratore con cui hai creato. [Crea un utente con accesso amministrativo](#)
2. Scegli Crea bucket e segui la procedura guidata per creare un bucket denominato `<yourName>-datalake-tutorial`, dove `<yourName>` sono il tuo nome e cognome. Ad esempio: `jdoe-datalake-tutorial`.

Per istruzioni dettagliate sulla creazione di un bucket Amazon S3, vedi [Come si crea un bucket S3?](#) nella Guida per l'utente di Amazon Simple Storage Service.

Fase 4: Registrare un percorso Amazon S3

In questa fase, registri un percorso Amazon Simple Storage Service (Amazon S3) come posizione principale del tuo data lake.

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake.
2. Nel riquadro di navigazione, in Amministrazione, scegli Posizioni dei data lake.
3. Scegli Registra posizione, quindi scegli Sfoglia.
4. Seleziona il `<yourName>-datalake-tutorial` bucket che hai creato in precedenza, accetta il ruolo IAM predefinito `AWSServiceRoleForLakeFormationDataAccess`, quindi scegli Registra posizione.

Per ulteriori informazioni sulla registrazione delle sedi, consulta. [Aggiungere una posizione Amazon S3 al tuo data lake](#)

Passaggio 5: concedere le autorizzazioni per la localizzazione dei dati

I responsabili devono disporre delle autorizzazioni per la localizzazione dei dati su una posizione di data lake per creare tabelle o database di Data Catalog che puntano a tale posizione. È necessario concedere le autorizzazioni per la localizzazione dei dati al ruolo IAM per i flussi di lavoro in modo che il flusso di lavoro possa scrivere nella destinazione di inserimento dei dati.

1. Sulla console Lake Formation, nel riquadro di navigazione, in Autorizzazioni, scegli Posizioni dati.
2. Scegliete Concedi e, nella finestra di dialogo Concedi autorizzazioni, effettuate le seguenti operazioni:
 - a. Per utenti e ruoli IAM, scegli `LakeFormationWorkflowRole`.
 - b. Per le posizioni di archiviazione, scegli il tuo `<yourName>-datalake-tutorial` bucket.
3. Scegli Concessione.

Per ulteriori informazioni sulle autorizzazioni per la localizzazione dei dati, consulta [Underlying data access control](#)

Fase 6: Creare un database nel Data Catalog

Le tabelle di metadati nel Lake Formation Data Catalog sono archiviate all'interno di un database.

1. Nella console Lake Formation, nel riquadro di navigazione, in Catalogo dati, scegli Databases.
2. Scegli Crea database e in Dettagli del database, inserisci il nome `lakeformation_tutorial`.
3. Lascia vuoti gli altri campi e scegli Crea database.

Passaggio 7: concedere le autorizzazioni per i dati

È necessario concedere le autorizzazioni per creare tabelle di metadati nel Data Catalog. Poiché il flusso di lavoro viene eseguito con il ruolo `LakeFormationWorkflowRole`, è necessario concedere queste autorizzazioni al ruolo.

1. Nella console Lake Formation, nel riquadro di navigazione, in Autorizzazioni, scegli Autorizzazioni Data lake.
2. Scegli Concedi e nella finestra di dialogo Concedi le autorizzazioni per i dati, procedi come segue:

- a. In Principal, per utenti e ruoli IAM, scegli. LakeFormationWorkflowRole
 - b. In LF-Tags o Catalog resources, scegli Named data catalog resources.
 - c. Per Database, scegli il database che hai creato in precedenza,.
lakeformation_tutorial
 - d. In Autorizzazioni del database, seleziona Crea tabella, Alter e Drop e deseleziona Super se è selezionato.
3. Scegli Concessione.

Per ulteriori informazioni sulla concessione delle autorizzazioni di Lake Formation, vedere.

[Panoramica delle autorizzazioni di Lake Formation](#)

Fase 8: Utilizzare un blueprint per creare un flusso di lavoro

Il AWS Lake Formation flusso di lavoro genera i AWS Glue job, i crawler e i trigger che rilevano e inseriscono dati nel tuo data lake. Crei un flusso di lavoro basato su uno dei blueprint predefiniti di Lake Formation.

1. Nella console Lake Formation, nel pannello di navigazione, scegli Blueprints, quindi scegli Usa blueprint.
2. Nella pagina Usa un blueprint, in Tipo di blueprint, scegli Database snapshot.
3. In Origine di importazione, per Connessione al database, scegli la connessione che hai appena creato o scegli una connessione esistente per la tua origine dati. datalake-tutorial
4. Per Percorso dei dati di origine, inserisci nel modulo `<database>/<schema>/<table>` il percorso da cui importare i dati.

È possibile sostituire lo schema o la tabella con il carattere jolly percentuale (%). Per i database che supportano gli schemi, inserisci `<database>/<schema>/%` per abbinare tutte le tabelle contenute all'interno. `<schema>` `<database>` Oracle Database e MySQL non supportano lo schema nel percorso; inserisci invece `/%`. `<database>` Per Oracle Database, `<database>` è l'identificatore di sistema (SID).

Ad esempio, se un database Oracle ha orcl come SID, immettilo in modo che orcl/% corrisponda a tutte le tabelle a cui ha accesso l'utente specificato nella connessione JDBC.

⚠ Important

Questo campo fa distinzione tra minuscole e maiuscole.

5. In Import target, specifica questi parametri:

Database di destinazione	lakeformation_tutorial
Posizione di archiviazione di destinazione	s3://<yourName> -datalake-tutorial
Formato dei dati	(Scegli Parquet o CSV)

6. Per la frequenza di importazione, scegli Esegui su richiesta.

7. In Opzioni di importazione, specificate questi parametri:

Nome del flusso di lavoro	lakeformationjdbctest
Ruolo IAM	LakeFormationWorkflowRole
Prefisso della tabella	jdbctest

📘 Note

Deve essere in minuscolo.

8. Scegli Crea e attendi che la console segnali che il flusso di lavoro è stato creato correttamente.

📘 Tip

Hai ricevuto il seguente messaggio di errore?

```
User: arn:aws:iam::<account-id>:user/<datalake_administrator_user> is not authorized to
perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/
LakeFormationWorkflowRole...
```

In tal caso, verifica di aver sostituito `<account-id>` nella politica in linea per l'utente amministratore del data lake un numero di AWS account valido.

Passaggio 9: Esegui il flusso di lavoro

Poiché hai specificato che il flusso di lavoro è run-on-demand, devi avviarlo manualmente in AWS Lake Formation.

1. Nella console Lake Formation, nella pagina Blueprints, seleziona il flusso di lavoro `lakeformationjdbctest`.
2. Scegli Azioni, quindi scegli Avvia.
3. Durante l'esecuzione del flusso di lavoro, visualizzane l'avanzamento nella colonna Stato dell'ultima esecuzione. Scegli il pulsante di aggiornamento di tanto in tanto.

Lo stato va da RUNNING, a Discovering, a Importing, a COMPLETED.

Quando il flusso di lavoro è completo:

- Il Data Catalog ha nuove tabelle di metadati.
- I tuoi dati vengono inseriti nel data lake.

Se il flusso di lavoro fallisce, procedi come segue:

- a. Seleziona un flusso di lavoro. Scegliete Azioni, quindi scegliete Visualizza grafico.

Il flusso di lavoro si apre nella AWS Glue console.

- b. Seleziona il flusso di lavoro e scegli la scheda Cronologia.
- c. Seleziona l'esecuzione più recente e scegli Visualizza i dettagli della corsa.
- d. Seleziona un processo o un crawler non riuscito nel grafico dinamico (runtime) ed esamina il messaggio di errore. I nodi con errori sono rossi o gialli.

Fase 10: concedere SELECT sui tavoli

È necessario concedere l'AUTHORIZZAZIONE per le nuove tabelle del Data Catalog in AWS Lake Formation modo che l'analista dei dati possa interrogare i dati a cui puntano le tabelle.

 Note

Un flusso di lavoro concede automaticamente l'SELECT autorizzazione per le tabelle che crea all'utente che lo ha eseguito. Poiché l'amministratore del data lake ha eseguito questo flusso di lavoro, è necessario concederlo SELECT all'analista dei dati.

1. Nella console Lake Formation, nel riquadro di navigazione, in Autorizzazioni, scegli Autorizzazioni Data lake.
2. Scegli Concedi e nella finestra di dialogo Concedi le autorizzazioni per i dati, procedi come segue:
 - a. In Principal, per utenti e ruoli IAM, scegli. `dataLake_user`
 - b. In LF-Tags o Catalog resources, scegli Named data catalog resources.
 - c. Per Database, scegliete. `lakeformation_tutorial`

Viene compilato l'elenco Tabelle.

 - d. Per Tabelle, scegli una o più tabelle dalla tua fonte di dati.
 - e. In Autorizzazioni per tabelle e colonne, scegli Seleziona.
3. Scegli Concessione.

Il passaggio successivo viene eseguito come analista dei dati.

Passaggio 11: interrogare il data lake utilizzando Amazon Athena

Usa la Amazon Athena console per interrogare i dati nel tuo data lake.

1. Apri la console Athena all'indirizzo e accedi come analista di dati <https://console.aws.amazon.com/athena/>, utente. `dataLake_user`
2. Se necessario, scegli Inizia per passare all'editor di query Athena.
3. Per Origine dati scegliere AwsDataCatalog.
4. Per Database, scegliere `lakeformation_tutorial`.

L'elenco delle tabelle viene compilato.

5. Nel menu pop-up accanto a una delle tabelle, scegli Anteprima tabella.

La query viene eseguita e visualizza 10 righe di dati.

Passaggio 12: interroga i dati nel data lake utilizzando Amazon Redshift Spectrum

Puoi configurare Amazon Redshift Spectrum per interrogare i dati che hai importato nel tuo data lake Amazon Simple Storage Service (Amazon S3). Innanzitutto, crea un ruolo AWS Identity and Access Management (IAM) da utilizzare per avviare il cluster Amazon Redshift e per interrogare i dati di Amazon S3. Quindi, concedi a questo ruolo le `Select` autorizzazioni sulle tabelle che desideri interrogare. Quindi, concedi all'utente le autorizzazioni per utilizzare l'editor di query di Amazon Redshift. Infine, crea un cluster Amazon Redshift ed esegui le query.

Crei il cluster come amministratore e interroghi il cluster come analista di dati.

Per ulteriori informazioni su Amazon Redshift Spectrum, [consulta Using Amazon Redshift Spectrum to External Data nella Amazon Redshift Database Developer Guide](#).

Per configurare le autorizzazioni per eseguire query Amazon Redshift

1. Aprire la console IAM all'indirizzo <https://console.aws.amazon.com/iam/>. Accedi come utente amministratore in cui hai creato [Crea un utente con accesso amministrativo](#) (nome utenteAdministrator) o come utente con la AdministratorAccess AWS policy gestita.
2. Nel pannello di navigazione, selezionare Policy.

Se è la prima volta che selezioni Policy, verrà visualizzata la pagina Benvenuto nelle policy gestite. Seleziona Inizia.

3. Scegli Crea policy.
4. Scegliere la scheda JSON.
5. Incolla il seguente documento di policy JSON.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
```

```

        "glue:GetTable",
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetPartitions",
        "lakeformation:GetResourceLFTags",
        "lakeformation:ListLFTags",
        "lakeformation:GetLFTag",
        "lakeformation:SearchTablesByLFTags",
        "lakeformation:SearchDatabasesByLFTags"
    ],
    "Resource": "*"
}
]
}
```

6. Una volta terminato, selezionare Review (Rivedi) per rivedere la policy. In Validatore di policy vengono segnalati eventuali errori di sintassi.
7. Nella pagina Rivedi la politica, inserisci **RedshiftLakeFormationPolicy** il nome della politica che stai creando. (Opzionale) Immettere una Description (descrizione). Consulta il Summary (Riepilogo) della policy per visualizzare le autorizzazioni concesse dalla policy. Seleziona Create policy (Crea policy) per salvare il proprio lavoro.
8. Nel pannello di navigazione della console IAM, scegliere Ruoli e quindi Crea ruolo.
9. In Seleziona tipo di entità attendibile, scegli Servizio AWS .
10. Scegliere il servizio Amazon Redshift per assumere questo ruolo.
11. Selezionare il caso d'uso Redshift Customizable (Redshift personalizzabile) per il servizio. Quindi scegliere Next: Permissions (Successivo: Autorizzazioni).
12. Cerca la politica di autorizzazione che hai creato e seleziona la casella di controllo accanto al nome della politica nell'elenco. RedshiftLakeFormationPolicy
13. Scegliere Next: Tags (Successivo: Tag).
14. Scegliere Next:Review (Successivo: Rivedi).
15. In Role name (Nome ruolo), inserire il nome **RedshiftLakeFormationRole**.
16. (Facoltativo) In Descrizione ruolo, immettere una descrizione per il nuovo ruolo.
17. Verificare il ruolo e quindi scegliere Create role (Crea ruolo).

Per concedere **Select** le autorizzazioni sulla tabella da interrogare nel database Lake Formation

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake.
2. Nel riquadro di navigazione, in Autorizzazioni, scegli Autorizzazioni Data lake, quindi scegli Concedi.
3. Inserisci le informazioni che seguono:
 - Per gli utenti e i ruoli IAM, scegli il ruolo IAM che hai creato, `RedshiftLakeFormationRole`. Quando si esegue l'editori di query di Amazon Redshift, utilizzare il ruolo IAM per le autorizzazioni ai dati.
 - Per Database, scegliere `lakeformation_tutorial`.

L'elenco delle tabelle viene compilato.

 - Per Tabella, scegli una tabella all'interno della fonte di dati da interrogare.
 - Scegliete l'autorizzazione Seleziona tabella.
4. Scegli Concessione.

Per configurare Amazon Redshift Spectrum ed eseguire query

1. Apri la console Amazon Redshift all'indirizzo. <https://console.aws.amazon.com/redshift> Accedi come utenteAdministrator.
2. Scegliere Crea cluster.
3. Nella pagina Crea cluster, inserisci `redshift-lakeformation-demo` l'identificatore del cluster.
4. Per il tipo di nodo, seleziona `dc2.large`.
5. Scorri verso il basso e, in Configurazioni del database, inserisci o accetta questi parametri:
 - Nome utente dell'amministratore: `awsuser`
 - Password utente amministratore: (*Choose a password*)
6. Espandi le autorizzazioni del cluster e, per i ruoli IAM disponibili, scegli `RedshiftLakeFormationRole`. Scegli Add IAM role (Aggiungi ruolo IAM).
7. Se devi utilizzare una porta diversa dal valore predefinito 5439, accanto a Configurazioni aggiuntive, disattiva l'opzione Usa valori predefiniti. Espandi la sezione relativa alle configurazioni del database e inserisci un nuovo numero di porta del database.

8. Scegliere Crea cluster.

Viene caricata la pagina Cluster.

9. Attendi che lo stato del cluster diventi Disponibile. Scegli periodicamente l'icona di aggiornamento.
10. Concedi all'analista dei dati l'autorizzazione a eseguire query sul cluster. Per fare ciò, completare questa procedura:
 - a. Apri la console IAM all'indirizzo <https://console.aws.amazon.com/iam/> e accedi come utente. Administrator
 - b. Nel riquadro di navigazione, scegli Utenti e allega all'utente le seguenti politiche gestitedatalake_user.
 - AmazonRedshiftQueryEditor
 - AmazonRedshiftReadOnlyAccess
11. Esci dalla console Amazon Redshift e accedi nuovamente come utente. datalake_user
12. Nella barra degli strumenti verticale a sinistra, scegli l'icona EDITOR per aprire l'editor di query e connetterti al cluster. Se viene visualizzata la finestra di dialogo Connetti al database, scegliete il nome redshift-lakeformation-demo del cluster e immettete il nome **dev** del database **awsuser**, il nome utente e la password che avete creato. Quindi scegli Connect to database (Collegati al database).

Note

Se non ti vengono richiesti i parametri di connessione e nell'editor di query è già selezionato un altro cluster, scegli Cambia connessione per aprire la finestra di dialogo Connetti al database.

13. Nella casella di testo New Query 1, inserisci ed esegui la seguente istruzione per mappare il database lakeformation_tutorial in Lake Formation al nome dello schema Amazon Redshift: redshift_jdbc

Important

Sostituiscilo *<account-id>* con un numero di AWS account valido e *<region>* con un nome di AWS regione valido (ad esempio, us-east-1).

```
create external schema if not exists redshift_jdbc from DATA CATALOG
  database 'lakeformation_tutorial' iam_role 'arn:aws:iam::<account-id>:role/
  RedshiftLakeFormationRole' region '<region>';
```

14. Nell'elenco degli schemi in Seleziona schema, scegliete redshift_jdbc.

L'elenco delle tabelle viene compilato. L'editor di query mostra solo le tabelle su cui ti sono state concesse le autorizzazioni per il data lake di Lake Formation.

15. Nel menu a comparsa accanto al nome di una tabella, scegli Anteprima dati.

Amazon Redshift restituisce le prime 10 righe.

Ora puoi eseguire query sulle tabelle e sulle colonne per le quali disponi delle autorizzazioni.

Passaggio 13: concedere o revocare le autorizzazioni di Lake Formation utilizzando Amazon Redshift Spectrum

Amazon Redshift supporta la possibilità di concedere e revocare le autorizzazioni Lake Formation su database e tabelle utilizzando istruzioni SQL modificate. Queste dichiarazioni sono simili alle dichiarazioni esistenti di Amazon Redshift. Per ulteriori informazioni, consulta [GRANT](#) and [REVOKE](#) nella Amazon Redshift Database Developer Guide.

Impostazione delle autorizzazioni per i formati di archiviazione a tabella aperta in Lake Formation

AWS Lake Formation [supporta la gestione delle autorizzazioni di accesso per Open Table Formats \(OTFs\) come Apache Iceberg, ApacheHudi e Linux foundation Delta Lake](#). In questo tutorial, imparerai come creare Iceberg, Hudi e Delta Lake utilizzando tabelle [manifest](#) symlink AWS Glue, configurare autorizzazioni AWS Glue Data Catalog granulari utilizzando Lake Formation e interrogare i dati utilizzando Amazon Athena.

Note

AWS i servizi di analisi non supportano tutti i formati di tabelle transazionali. Per ulteriori informazioni, consulta [Collaborazione con altri AWS servizi](#). Questo tutorial illustra

manualmente la creazione di un nuovo database e di una tabella nel Data Catalog utilizzando solo i AWS Glue job.

Questo tutorial include un AWS CloudFormation modello per una configurazione rapida. Puoi vederlo e personalizzarlo in base alle tue esigenze.

Argomenti

- [Destinatari principali](#)
- [Prerequisiti](#)
- [Fase 1: Fornisci le tue risorse](#)
- [Passaggio 2: imposta le autorizzazioni per una tabella Iceberg](#)
- [Passaggio 3: configura le autorizzazioni per una tabella Hudi](#)
- [Passaggio 4: configura le autorizzazioni per una tabella Delta Lake](#)
- [Fase 5: Pulizia delle risorse AWS](#)

Destinatari principali

Questo tutorial è destinato agli amministratori IAM, agli amministratori di data lake e agli analisti aziendali. La tabella seguente elenca i ruoli utilizzati in questo tutorial per creare una tabella governata utilizzando Lake Formation.

Ruolo	Descrizione
Amministratore IAM	Un utente che può creare utenti e ruoli IAM e bucket Amazon S3. Ha la politica AdministratorAccess AWS gestita.
Amministratore del data lake	Un utente che può accedere al Data Catalog, creare database e concedere autorizzazioni Lake Formation ad altri utenti. Dispone di meno autorizzazioni IAM rispetto all'amministratore IAM, ma sufficienti per amministrare il data lake.

Ruolo	Descrizione
Business analyst	Un utente che può eseguire query sul data lake. Dispone delle autorizzazioni per eseguire query.

Prerequisiti

Prima di iniziare questo tutorial, devi disporre di un Account AWS account per poter accedere come utente con le autorizzazioni corrette. Per ulteriori informazioni, consulta [Iscriviti per un Account AWS](#) e [Crea un utente con accesso amministrativo](#).

Il tutorial presuppone che tu abbia familiarità con i ruoli e le politiche di IAM. Per informazioni su IAM, consulta la [IAM User Guide](#).

È necessario configurare le seguenti AWS risorse per completare questo tutorial:

- Utente amministratore di Data Lake
- Impostazioni del data lake Lake Formation
- Motore Amazon Athena versione 3

Per creare un amministratore di data lake

1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> come utente amministratore. Creerai risorse nella regione Stati Uniti orientali (Virginia settentrionale) per questo tutorial.
2. Nella console di Lake Formation, nel riquadro di navigazione, in Autorizzazioni, scegli Ruoli e attività amministrativi.
3. Seleziona Scegli amministratori in Amministratori di Data lake.
4. Nella finestra pop-up, Gestisci gli amministratori del data lake, in Utenti e ruoli IAM, scegli Utente amministratore IAM.
5. Seleziona Salva.

Per abilitare le impostazioni del data lake

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Nel pannello di navigazione, in Catalogo dati, scegli Impostazioni. Deseleziona quanto segue:
 - Utilizza solo il controllo degli accessi IAM per i nuovi database.
 - Utilizza solo il controllo di accesso IAM per nuove tabelle in nuovi database.
2. In Impostazioni della versione per più account, scegli Versione 3 come versione per più account.
3. Seleziona Salva.

Per aggiornare il motore Amazon Athena alla versione 3

1. Apri la console Athena all'indirizzo. <https://console.aws.amazon.com/athena/>
2. Seleziona il gruppo di lavoro e seleziona il gruppo di lavoro principale.
3. Assicurati che il gruppo di lavoro abbia una versione minima di 3. In caso contrario, modifica il gruppo di lavoro, scegli Manual for Upgrade Query Engine e seleziona la versione 3.
4. Scegli Save changes (Salva modifiche).

Fase 1: Fornisci le tue risorse

Questa sezione mostra come configurare le AWS risorse utilizzando un CloudFormation modello.

Per creare le tue risorse utilizzando CloudFormation un modello

1. Accedi alla AWS CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> come amministratore IAM nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Avanti nella schermata Crea stack.
4. Inserisci un nome per lo stack.
5. Scegli Next (Successivo).
6. Nella pagina successiva, scegli Avanti.
7. Controlla i dettagli nella pagina finale e seleziona Riconosco che AWS CloudFormation potrebbe creare risorse IAM.
8. Scegli Create (Crea) .

La creazione dello stack può richiedere fino a due minuti.

L'avvio dello stack di formazione cloud crea le seguenti risorse:

- If-otf-datalake-123456789012 — Bucket Amazon S3 per archiviare dati

 Note

L'ID dell'account aggiunto al nome del bucket Amazon S3 viene sostituito con l'ID dell'account.

- If-otf-tutorial-123456789012 — Bucket Amazon S3 per archiviare i risultati delle query e gli script di lavoro AWS Glue
- AWS Glue Ificebergdb — Database Iceberg
- Ifhudidb — Banca dati AWS Glue Hudi
- Ifdeltadb — Banca AWS Glue dati Delta
- native-iceberg-create — AWS Glue lavoro che crea una tabella Iceberg nel Data Catalog
- native-hudi-create — AWS Glue lavoro che crea una tabella Hudi nel Data Catalog
- native-delta-create — AWS Glue lavoro che crea una tabella Delta nel Data Catalog
- LF-OTF-GlueServiceRole — Ruolo IAM a cui si passa AWS Glue per eseguire i job. A questo ruolo sono allegate le politiche necessarie per accedere a risorse come Data Catalog, Amazon S3 bucket ecc.
- LF-OTF-RegisterRole — Ruolo IAM per registrare la posizione Amazon S3 con Lake Formation. Questo ruolo è LF-Data-Lake-Storage-Policy associato al ruolo.
- If-consumer-analystuser — Utente IAM per interrogare i dati utilizzando Athena
- If-consumer-analystuser-credentials — Password per l'utente dell'analista di dati memorizzata in Gestione dei segreti AWS

Una volta completata la creazione dello stack, vai alla scheda Output e annota i valori per:

- AthenaQueryResultLocation — Posizione Amazon S3 per l'output delle query Athena
- BusinessAnalystUserCredentials — Password per l'utente dell'analista di dati

Per recuperare il valore della password:

1. Scegli il `lf-consumer-analystuser-credentials` valore accedendo alla console Secrets Manager.
2. Nella sezione Secret value (Valore segreto) scegli Retrieve secret value (Recupera valore segreto).
3. Annota il valore segreto della password.

Passaggio 2: imposta le autorizzazioni per una tabella Iceberg

In questa sezione, imparerai come creare una tabella Iceberg in AWS Glue Data Catalog, impostare le autorizzazioni per i dati e interrogare i dati utilizzando Amazon Athena. AWS Lake Formation

Per creare una tabella Iceberg

In questo passaggio, eseguirai un AWS Glue job che crea una tabella transazionale Iceberg nel Data Catalog.

1. Apri il AWS Glue console <https://console.aws.amazon.com/glue/> nella regione Stati Uniti orientali (Virginia settentrionale) come utente amministratore del data lake.
2. Scegli i lavori dal riquadro di navigazione a sinistra.
3. Seleziona `native-iceberg-create`.

Create job [Info](#) Create

☒ **Visual with a source and target**
 Start with a source, ApplyMapping transform, and target.

☐ **Visual with a blank canvas**
 Author using an interactive visual interface.

☐ **Spark script editor**
 Write or upload your own Spark code.

☐ **Python Shell script editor**
 Write or upload your own Python shell script.

☐ **Jupyter Notebook**
 Write your own code in a Jupyter Notebook for interactive development.

☐ **Ray script editor** New
 Write your own code to run on Ray.

Source Amazon S3
 JSON, CSV, or Parquet files stored in S3.

Target Amazon S3
 S3 bucket by specifying a bucket path as the data target.

Your jobs (24) [Info](#)

Find jobs

	Job name	Type	Last modified	
☐	native-delta-create	Glue ETL	2/24/2023, 9:22:31 AM	
☒	native-iceberg-create	Glue ETL	2/24/2023, 9:22:31 AM	3.0
☐	native-hudi-create	Glue ETL	2/24/2023, 9:22:30 AM	3.0

Actions: Edit job, Clone job, Schedule job, Delete job(s), Reset job bookmark

- In Azioni, scegli Modifica lavoro.
- In Dettagli Job, espandi Proprietà avanzate e seleziona la casella accanto a Usa AWS Glue Data Catalog come metastore Hive per aggiungere i metadati della tabella in. AWS Glue Data Catalog. Questo specifica AWS Glue Data Catalog come metastore per le risorse del Data Catalog utilizzate nel job e consente di applicare successivamente le autorizzazioni di Lake Formation alle risorse del catalogo.
- Seleziona Salva.
- Seleziona Esegui. È possibile visualizzare lo stato del job mentre è in esecuzione.

Per ulteriori informazioni sui AWS Glue job, consulta [Lavorare con i job sulla AWS Glue console](#) nella AWS Glue Developer Guide.

Questo lavoro crea una tabella Iceberg denominata product nel l1cebergddb database. Verifica la tabella dei prodotti nella console Lake Formation.

Per registrare la posizione dei dati con Lake Formation

Successivamente, registra il percorso Amazon S3 come posizione del tuo data lake.

1. Apri la console Lake Formation <https://console.aws.amazon.com/lakeformation/> come utente amministratore del data lake.
2. Nel pannello di navigazione, sotto Registra e inserisci, scegli Posizione dei dati.
3. In alto a destra della console, scegli Registra posizione.
4. Nella pagina Registra posizione, inserisci quanto segue:
 - Percorso Amazon S3: scegli Sfoglia e seleziona. lf-otf-datalake-123456789012 Fai clic sulla freccia destra (>) accanto alla posizione principale di Amazon S3 per accedere alla s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-iceberg posizione.
 - Ruolo IAM: scegli LF-OTF-RegisterRole come ruolo IAM.
 - Scegli la posizione di registrazione.

Register location

Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

Amazon S3 path

Choose an Amazon S3 path for your data lake.

s3://lf-otf-datalake- /transactionaldata/native-iceberg

Browse

Review location permissions - strongly recommended

Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

Review location permissions

IAM role

To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

LF-OTF-GlueServiceRole ▼

☐ Enable Catalog Federation

Lake Formation will only assume a role to access a registered location when accessing a table under a federated database

Cancel

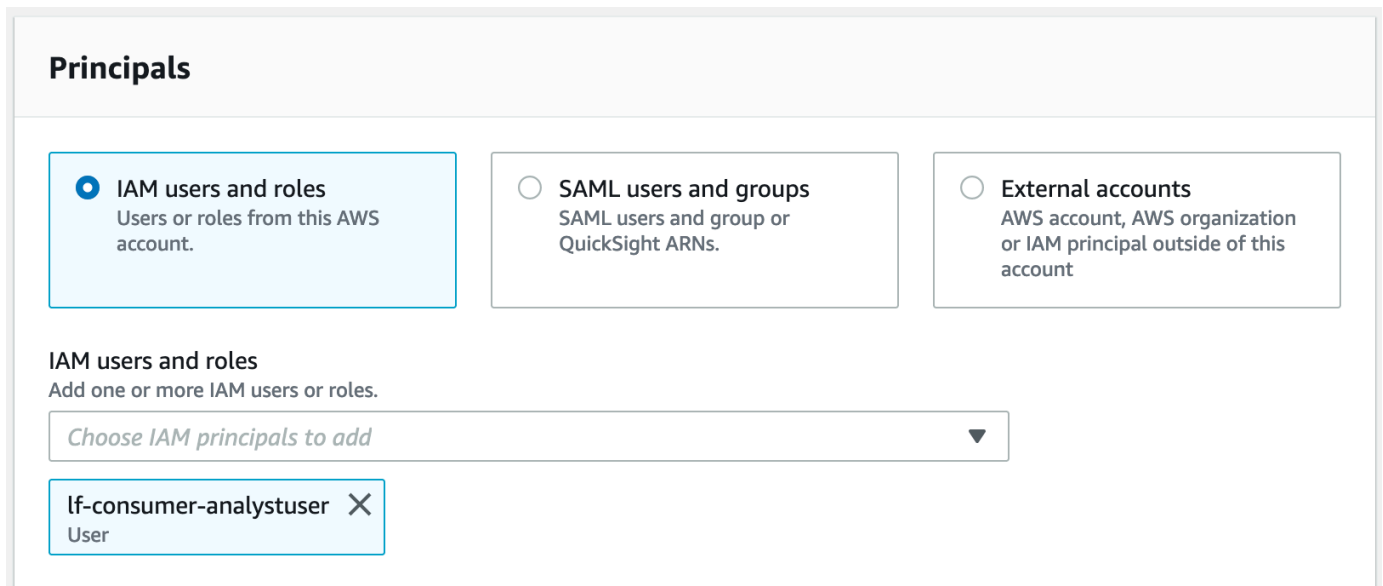
Register location

Per ulteriori informazioni sulla registrazione di una posizione dati con Lake Formation, vedere [Aggiungere una posizione Amazon S3 al tuo data lake](#).

Per concedere i permessi di Lake Formation sulla tabella Iceberg

In questo passaggio, concederemo le autorizzazioni di data lake all'utente business analyst.

1. In Autorizzazioni Data lake, scegli Concedi.
2. Nella schermata Concedi le autorizzazioni per i dati, scegli, utenti e ruoli IAM.
3. Scegli `lf-consumer-analystuser` dal menu a discesa.



Principals

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles
Add one or more IAM users or roles.

Choose IAM principals to add ▼

lf-consumer-analystuser X
User

4. Scegli la risorsa Named data catalog.
5. Per Database scegli `lficebergdb`.
6. Per Tabelle, scegli `product`.

LF-Tags or catalog resources

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named data catalog resources
Manager permissions for specific databases or tables, in addition to fine-grained data access.

Databases

Select one or more databases.

Choose databases ▼

Load more

lficebergdb X

Tables - optional

Select one or more tables.

Choose tables ▼

Load more

product X

Data filters - optional

Select one or more data filters.

Choose data filters ▼

Load more

Create new

[Manage data filters](#)

7. Successivamente, puoi concedere l'accesso basato su colonne specificando le colonne.
 - a. In Autorizzazioni per la tabella, scegli Seleziona.
 - b. In Autorizzazioni dati, scegli Accesso basato su colonne, scegli Includi colonne.
 - c. Scegli product_name e colonneprice. category
 - d. Scegli Concessione.

Table permissions

Table permissions
Choose specific access permissions to grant.

☒ Select

☐ Insert

☐ Delete

☐ Describe

☐ Alter

☐ Drop

☐ Super
This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions
Choose the permission that may be granted to others.

☐ Select

☐ Insert

☐ Delete

☐ Describe

☐ Alter

☐ Drop

☐ Super
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Data permissions

☐ All data access
Grant access to all data without any restrictions.

☒ Column-based access
Grant data access to specific columns only.

Choose permission filter
Choose whether to include or exclude columns.

☒ Include columns
Grant permissions to access specific columns.

☐ Exclude columns
Grant permissions to access all but specific columns.

Select columns

Choose one or more columns ▼

product_name
string

price
bigint

category
string

CancelGrant

Per interrogare la tabella Iceberg usando Athena

Ora puoi iniziare a interrogare la tabella Iceberg che hai creato usando Athena. Se è la prima volta che esegui query in Athena, devi configurare una posizione dei risultati della query. Per ulteriori informazioni, vedere [Specificazione della posizione dei risultati di una query](#).

1. Esci come utente amministratore del data lake e accedi come `lf-consumer-analystuser` nella regione Stati Uniti orientali (Virginia settentrionale) utilizzando la password indicata in precedenza nell' CloudFormation output.
2. Apri la console Athena all'indirizzo <https://console.aws.amazon.com/athena/>.
3. Scegli Impostazioni e seleziona Gestisci.
4. Nella casella Posizione dei risultati della query, inserisci il percorso del bucket che hai creato negli CloudFormation output. Copia il valore di **AthenaQueryResultLocation** (`s3://lf-ottutorial-123456789012/athena-results/`) e scegli Salva.
5. Esegui la seguente query per visualizzare in anteprima 10 record memorizzati nella tabella Iceberg:

```
select * from lficebergdb.product limit 10;
```

Per ulteriori informazioni sull'interrogazione delle tabelle Iceberg utilizzando Athena, consulta [Interrogare le tabelle Iceberg nella Amazon Athena User Guide](#).

Passaggio 3: configura le autorizzazioni per una tabella Hudi

In questa sezione, imparerai come creare una tabella Hudi in AWS Glue Data Catalog, impostare le autorizzazioni per i dati e interrogare i dati utilizzando Amazon Athena. AWS Lake Formation

Per creare una tabella Hudi

In questo passaggio, eseguirai un AWS Glue job che crea una tabella transazionale Hudi nel Data Catalog.

1. Accedi al AWS Glue console <https://console.aws.amazon.com/glue/> nella regione Stati Uniti orientali (Virginia settentrionale)
come utente amministratore del data lake.
2. Scegli i lavori dal riquadro di navigazione a sinistra.
3. Seleziona `native-hudi-create`.
4. In Azioni, scegli Modifica lavoro.
5. In Dettagli Job, espandi Proprietà avanzate e seleziona la casella accanto a Usa AWS Glue Data Catalog come metastore Hive per aggiungere i metadati della tabella in. AWS Glue Data Catalog Questo specifica AWS Glue Data Catalog come metastore per le risorse del Data Catalog

utilizzate nel job e consente di applicare successivamente le autorizzazioni di Lake Formation alle risorse del catalogo.

6. Seleziona Salva.
7. Seleziona Esegui. È possibile visualizzare lo stato del job mentre è in esecuzione.

Per ulteriori informazioni sui AWS Glue job, consulta [Lavorare con i job sulla AWS Glue console](#) nella AWS Glue Developer Guide.

Questo job crea una tabella Hudi (cow) nel database:lfhudidb. Verifica la product tabella nella console Lake Formation.

Per registrare la posizione dei dati con Lake Formation

Successivamente, registra un percorso Amazon S3 come posizione principale del tuo data lake.

1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> come utente amministratore del data lake.
2. Nel pannello di navigazione, sotto Registra e inserisci, scegli Data location.
3. In alto a destra della console, scegli Registra posizione.
4. Nella pagina Registra posizione, inserisci quanto segue:
 - Percorso Amazon S3: scegli Sfoglia e seleziona. lf-otf-datalake-123456789012 Fai clic sulla freccia destra (>) accanto alla posizione principale di Amazon S3 per accedere alla s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-hudi posizione.
 - Ruolo IAM: scegli LF-OTF-RegisterRole come ruolo IAM.
 - Scegli la posizione di registrazione.

Per concedere le autorizzazioni del data lake sulla tabella Hudi

In questo passaggio, concederemo le autorizzazioni del data lake all'utente business analyst.

1. In Autorizzazioni Data lake, scegli Concedi.
2. Nella schermata Concedi le autorizzazioni per i dati, scegli, utenti e ruoli IAM.
3. lf-consumer-analystuserdal menu a discesa.
4. Scegli la risorsa Named data catalog.

5. Per Database scegli `lfhudidb`.
6. Per Tabelle, scegli `product`.
7. Successivamente, puoi concedere l'accesso basato su colonne specificando le colonne.
 - a. In Autorizzazioni per la tabella, scegli Seleziona.
 - b. In Autorizzazioni dati, scegli Accesso basato su colonne, scegli Includi colonne.
 - c. Scegli `product_name` e colonne `price`, `category`
 - d. Scegli Concessione.

Per interrogare la tabella Hudi usando Athena

Ora iniziate a interrogare la tabella Hudi che avete creato usando Athena. Se è la prima volta che esegui query in Athena, devi configurare una posizione dei risultati della query. Per ulteriori informazioni, vedere [Specificazione della posizione dei risultati di una query](#).

1. Esci come utente amministratore del data lake e accedi come `lf-consumer-analystuser` nella regione Stati Uniti orientali (Virginia settentrionale) utilizzando la password indicata in precedenza nell' CloudFormation output.
2. Apri la console Athena all'indirizzo <https://console.aws.amazon.com/athena/>.
3. Scegli Impostazioni e seleziona Gestisci.
4. Nella casella Posizione dei risultati della query, inserisci il percorso del bucket che hai creato negli CloudFormation output. Copia il valore di **AthenaQueryResultLocation** (`s3://lf-off-tutorial-123456789012/athena-results/`) e salva.
5. Esegui la seguente query per visualizzare in anteprima 10 record memorizzati nella tabella Hudi:

```
select * from lfhudidb.product limit 10;
```

Per ulteriori informazioni sull'interrogazione delle tabelle Hudi, consulta la sezione [Interrogazione delle tabelle Hudi](#) nella Guida per l'utente di Amazon Athena.

Passaggio 4: configura le autorizzazioni per una tabella Delta Lake

In questa sezione, imparerai come creare una tabella Delta Lake con il file manifest symlink nel file AWS Glue Data Catalog, impostare le autorizzazioni per i dati AWS Lake Formation e interrogare i dati utilizzando Amazon Athena.

Per creare una tabella Delta Lake

In questo passaggio, eseguirai un AWS Glue processo che crea una tabella transazionale Delta Lake nel Data Catalog.

1. Accedi al AWS Glue console <https://console.aws.amazon.com/glue/> nella regione Stati Uniti orientali (Virginia settentrionale)

come utente amministratore del data lake.
2. Scegli i lavori dal riquadro di navigazione a sinistra.
3. Seleziona `native-delta-create`.
4. In Azioni, scegli Modifica lavoro.
5. In Dettagli Job, espandi Proprietà avanzate e seleziona la casella accanto a Usa AWS Glue Data Catalog come metastore Hive per aggiungere i metadati della tabella in. AWS Glue Data Catalog Questo specifica AWS Glue Data Catalog come metastore per le risorse del Data Catalog utilizzate nel job e consente di applicare successivamente le autorizzazioni di Lake Formation alle risorse del catalogo.
6. Seleziona Salva.
7. Scegli Esegui in Azioni.

Questo lavoro crea una tabella Delta Lake denominata `product` nel `lfdeltadb` database. Verifica la `product` tabella nella console Lake Formation.

Per registrare la posizione dei dati con Lake Formation

Successivamente, registra il percorso Amazon S3 come posizione principale del tuo data lake.

1. Apri la console Lake Formation presso <https://console.aws.amazon.com/lakeformation/> l'utente amministratore del data lake.
2. Nel pannello di navigazione, sotto Registra e inserisci, scegli Posizione dei dati.
3. In alto a destra della console, scegli Registra posizione.

4. Nella pagina Registra posizione, inserisci quanto segue:

- Percorso Amazon S3: scegli Sfoglia e seleziona. `lf-otf-datalake-123456789012` Fai clic sulla freccia destra (>) accanto alla posizione principale di Amazon S3 per accedere alla `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-delta` posizione.
- Ruolo IAM: scegli `LF-OTF-RegisterRole` come ruolo IAM.
- Scegli la posizione di registrazione.

Per concedere le autorizzazioni del data lake sulla tabella Delta Lake

In questo passaggio, concederemo le autorizzazioni del data lake all'utente `business analyst`.

1. In Autorizzazioni Data lake, scegli Concedi.
2. Nella schermata Concedi le autorizzazioni per i dati, scegli, utenti e ruoli IAM.
3. `lf-consumer-analystuser` dal menu a discesa.
4. Scegli la risorsa `Named data catalog`.
5. Per Database scegli `lfdeltadb`.
6. Per Tabelle, scegli `product`.
7. Successivamente, puoi concedere l'accesso basato su colonne specificando le colonne.
 - a. In Autorizzazioni per la tabella, scegli Seleziona.
 - b. In Autorizzazioni dati, scegli Accesso basato su colonne, scegli Includi colonne.
 - c. Scegli `product_name` e colonne `price`. `category`
 - d. Scegli Concessione.

Per interrogare la tabella Delta Lake usando Athena

Ora iniziate a interrogare la tabella Delta Lake che avete creato usando Athena. Se è la prima volta che esegui query in Athena, devi configurare una posizione dei risultati della query. Per ulteriori informazioni, vedere [Specificazione della posizione dei risultati di una query](#).

1. Esci come utente amministratore del data lake e accedi come `BusinessAnalystUser` nella regione Stati Uniti orientali (Virginia settentrionale) utilizzando la password indicata in precedenza nell' `CloudFormation` output.

2. Apri la console Athena all'indirizzo <https://console.aws.amazon.com/athena/>.
3. Scegli Impostazioni e seleziona Gestisci.
4. Nella casella Posizione dei risultati della query, inserisci il percorso del bucket che hai creato negli CloudFormation output. Copia il valore di **AthenaQueryResultLocation** (s3://lf-ottutorial-123456789012/athena-results/) e salva.
5. Esegui la seguente query per visualizzare in anteprima 10 record memorizzati nella tabella Delta Lake:

```
select * from lfdeltadb.product limit 10;
```

Per ulteriori informazioni sull'interrogazione delle tabelle Delta Lake, consulta la sezione [Interrogazione delle tabelle Delta Lake](#) nella Guida per l'utente di Amazon Athena.

Fase 5: Pulizia delle risorse AWS

Per eliminare le risorse

Per evitare addebiti indesiderati Account AWS, elimina le AWS risorse che hai utilizzato per questo tutorial.

1. Accedi alla CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> come amministratore IAM.
2. [Elimina lo stack di formazione del cloud](#). Le tabelle che hai creato vengono eliminate automaticamente con lo stack.

Gestione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation

Migliaia di clienti stanno costruendo data lake su scala petabyte. AWS Molti di questi clienti lo utilizzano AWS Lake Formation per creare e condividere facilmente i propri data lake all'interno dell'organizzazione. Con l'aumento del numero di tabelle e utenti, i data steward e gli amministratori sono alla ricerca di modi per gestire le autorizzazioni sui data lake in modo semplice e su larga scala. Il controllo degli accessi basato su Lake Formation Tag (LF-TBAC) risolve questo problema consentendo ai data steward di creare tag LF (in base alla classificazione e all'ontologia dei dati) che possono quindi essere collegati alle risorse.

LF-TBAC è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. In Lake Formation, questi attributi sono chiamati LF-tag. Puoi allegare i tag LF alle risorse del Data Catalog e ai principali di Lake Formation. Gli amministratori del Data Lake possono assegnare e revocare le autorizzazioni sulle risorse di Lake Formation utilizzando i tag LF. Per ulteriori informazioni su see, [Controllo degli accessi basato su tag Lake Formation](#)

Questo tutorial dimostra come creare una politica di controllo degli accessi basata su tag Lake Formation utilizzando un set di dati AWS pubblico. Inoltre, mostra come interrogare tabelle, database e colonne a cui sono associate politiche di accesso basate su tag Lake Formation.

È possibile utilizzare LF-TBAC per i seguenti casi d'uso:

- Hai un gran numero di tabelle e principali a cui l'amministratore del data lake deve concedere l'accesso
- Vuoi classificare i tuoi dati in base a un'ontologia e concedere le autorizzazioni in base alla classificazione
- L'amministratore del data lake desidera assegnare le autorizzazioni in modo dinamico, in modo vagamente accoppiato

Di seguito sono riportati i passaggi di alto livello per la configurazione delle autorizzazioni utilizzando LF-TBAC:

1. Il data steward definisce l'ontologia dei tag con due tag LF: e. Confidential Sensitive Data with Confidential=True ha controlli di accesso più rigorosi. I dati Sensitive=True richiedono un'analisi specifica da parte dell'analista.
2. Il data steward assegna diversi livelli di autorizzazione al data engineer per creare tabelle con tag LF diversi.
3. L'ingegnere dei dati crea due database: e. tag_database col_tag_database Tutte le tabelle tag_database sono configurate con Confidential=True. Tutte le tabelle di col_tag_database sono configurate con Confidential=False. Alcune colonne della tabella col_tag_database sono contrassegnate con tag Sensitive=True per esigenze di analisi specifiche.
4. L'ingegnere dei dati concede il permesso di lettura all'analista per le tabelle con condizioni di espressione specifiche Confidential=True e Confidential=False, Sensitive=True
5. Con questa configurazione, l'analista di dati può concentrarsi sull'esecuzione dell'analisi con i dati giusti.

Argomenti

- [Destinatari principali](#)
- [Prerequisiti](#)
- [Fase 1: Fornisci le tue risorse](#)
- [Fase 2: Registrare la posizione dei dati, creare un'ontologia LF-tag e concedere le autorizzazioni](#)
- [Fase 3: Creare database Lake Formation](#)
- [Passaggio 4: concedere le autorizzazioni per la tabella](#)
- [Passaggio 5: esegui una query in Amazon Athena per verificare le autorizzazioni](#)
- [Fase 6: Pulizia delle risorse AWS](#)

Destinatari principali

Questo tutorial è destinato a amministratori di dati, ingegneri di dati e analisti di dati. Per quanto riguarda la gestione AWS Glue Data Catalog e l'amministrazione delle autorizzazioni in Lake Formation, gli amministratori dei dati all'interno degli account di produzione hanno la proprietà funzionale in base alle funzioni che supportano e possono concedere l'accesso a vari consumatori, organizzazioni esterne e account.

La tabella seguente elenca i ruoli utilizzati in questo tutorial:

Ruolo	Descrizione
Data steward (amministratore)	L'lf-data-steward utente dispone dei seguenti accessi: • Accesso in lettura a tutte le risorse del Data Catalog • Può creare tag LF e associarsi al ruolo di ingegnere dei dati per concedere l'autorizzazione ad altri responsabili
Ingegnere dei dati	lf-data-engineer l'utente ha i seguenti accessi:

Ruolo	Descrizione
	• Accesso completo in lettura, scrittura e aggiornamento a tutte le risorse del Data Catalog • Autorizzazioni per la localizzazione dei dati nel data lake • Può associare tag LF e associarsi al Data Catalog • Può allegare tag LF alle risorse, il che fornisce l'accesso ai principali in base a qualsiasi politica creata dai data steward
Analista dei dati	L'lf-data-analyst utente ha il seguente accesso: • Accesso granulare alle risorse condivise dalle politiche di accesso basate su tag di Lake Formation

Prerequisiti

Prima di iniziare questo tutorial, è necessario disporre di un file Account AWS da utilizzare per accedere come utente amministrativo con le autorizzazioni corrette. Per ulteriori informazioni, consulta [Completa le attività AWS di configurazione iniziali](#).

Il tutorial presuppone che tu abbia familiarità con IAM. Per informazioni su IAM, consulta la [IAM User Guide](#).

Fase 1: Fornisci le tue risorse

Questo tutorial include un AWS CloudFormation modello per una configurazione rapida. Puoi rivederlo e personalizzarlo in base alle tue esigenze. Il modello crea tre ruoli diversi (elencati in [Destinatari principali](#)) per eseguire questo esercizio e copia il nyc-taxi-data set di dati nel bucket Amazon S3 locale.

- Un bucket Amazon S3

- Le impostazioni appropriate di Lake Formation
- Le EC2 risorse Amazon appropriate
- Tre ruoli IAM con credenziali

Crea le tue risorse

1. Accedi alla AWS CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Next (Successivo).
4. Nella sezione Configurazione utente, inserisci la password per tre ruoli: `DataStewardUserPassword`, `DataEngineerUserPassword` e `DataAnalystUserPassword`.
5. Controlla i dettagli nella pagina finale e seleziona Riconosco che AWS CloudFormation potrebbe creare risorse IAM.
6. Scegli Create (Crea) .

La creazione dello stack può richiedere fino a cinque minuti.

Note

Dopo aver completato il tutorial, potresti voler eliminare lo stack per evitare di continuare CloudFormation a incorrere in addebiti. Verifica che le risorse siano state eliminate correttamente nello stato dell'evento per lo stack.

Fase 2: Registrare la posizione dei dati, creare un'ontologia LF-tag e concedere le autorizzazioni

In questo passaggio, l'utente Data Steward definisce l'ontologia dei tag con due LF-Tag: `Confidential` e `Sensitive` fornisce a specifici principi IAM la possibilità di allegare tag LF appena creati alle risorse.

Registra una posizione dei dati e definisci l'ontologia LF-Tag

1. Esegui il primo passaggio come utente amministratore dei dati (lf-data-steward) per verificare i dati in Amazon S3 e nel Data Catalog in Lake Formation.
 - a. Accedi alla console di Lake Formation lf-data-steward con la <https://console.aws.amazon.com/lakeformation/> stessa password utilizzata durante la distribuzione dello CloudFormation stack.
 - b. Nel riquadro di navigazione, in Autorizzazioni, scegli Ruoli e attività amministrative.
 - c. Scegli Aggiungi nella sezione Amministratori di Data Lake.
 - d. Nella pagina Aggiungi amministratore, per gli utenti e i ruoli IAM, scegli l'utente lf-data-steward.
 - e. Scegli Salva per aggiungerlo lf-data-steward come amministratore di Lake Formation.
2. Successivamente, aggiorna le impostazioni del Data Catalog per utilizzare l'autorizzazione Lake Formation per controllare le risorse del catalogo anziché il controllo degli accessi basato su IAM.
 - a. Nel pannello di navigazione, in Amministrazione, scegli Impostazioni Data Catalog.
 - b. Deseleziona Usa solo il controllo di accesso IAM per i nuovi database.
 - c. Deseleziona Usa solo il controllo di accesso IAM per nuove tabelle in nuovi database.
 - d. Fai clic su Save (Salva).
3. Quindi, registra la posizione dei dati per il data lake.
 - a. Nel riquadro di navigazione, in Amministrazione, scegli Posizioni dei data lake.
 - b. Scegli Registra posizione.
 - c. Nella pagina Registra posizione, per il percorso Amazon S3, inserisci. s3://lf-tagbased-demo-*Account-ID*
 - d. Per il ruolo IAM, lascia il valore predefinito `AWSServiceRoleForLakeFormationDataAccess` così com'è.
 - e. Scegli Lake Formation come modalità di autorizzazione.
 - f. Scegli Registra posizione.
4. Quindi, crea l'ontologia definendo un tag LF.
 - a. In Autorizzazioni nel riquadro di navigazione, scegli LF-tags e permessi. .
 - b. Scegli Aggiungi tag LF.
 - c. In Chiave, inserire `Confidential`.

- d. Per Valori, aggiungi True e False
- e. Scegli Aggiungi tag LF.
- f. Ripeti i passaggi per creare il tag LF Sensitive con il valore True

Avete creato tutti i tag LF necessari per questo esercizio.

Concedi le autorizzazioni agli utenti IAM

1. Successivamente, offri a specifici principi IAM la possibilità di allegare tag LF appena creati alle risorse.
 - a. In Autorizzazioni nel pannello di navigazione, scegli LF-tags e permessi.
 - b. Nella sezione Autorizzazioni LF-Tag, scegli Concedi autorizzazioni.
 - c. Per Tipo di autorizzazione, scegli le autorizzazioni per la coppia chiave-valore LF-Tag.
 - d. Seleziona utenti e ruoli IAM.
 - e. Per gli utenti e i ruoli IAM, cerca e scegli il lf-data-engineer ruolo.
 - f. Nella sezione LF-Tags, aggiungi la chiave Confidential con valori True e False e key Sensitive with value True
 - g. In Autorizzazioni, seleziona Descrivi e associa per le autorizzazioni e le autorizzazioni concedibili.
 - h. Scegli Concessione.
2. Successivamente, concedi le autorizzazioni lf-data-engineer per creare database nel nostro catalogo dati e nel bucket Amazon S3 sottostante creato da AWS CloudFormation
 - a. In Amministrazione nel riquadro di navigazione, scegli Ruoli e attività amministrative.
 - b. Nella sezione Creatori di database, scegli Concedi.
 - c. Per gli utenti e i ruoli IAM, scegli il lf-data-engineer ruolo.
 - d. Per le autorizzazioni del catalogo, seleziona Crea database.
 - e. Scegli Concessione.
3. Successivamente, concedi all'utente le autorizzazioni sul (s3://lf-tagbased-demo-*Account-ID*) bucket Amazon S3. lf-data-engineer
 - a. Nel riquadro di navigazione, in Autorizzazioni, scegli Posizioni dei dati.
 - b. Scegli Concessione.

- c. Seleziona Il mio account.
 - d. Per gli utenti e i ruoli IAM, scegli il `lf-data-engineer` ruolo.
 - e. Per le posizioni di archiviazione, inserisci il bucket Amazon S3 creato dal modello. CloudFormation (`s3://lf-tagbased-demo-Account-ID`)
 - f. Scegli Concessione.
4. Successivamente, concedi autorizzazioni **lf-data-engineer** concedibili sulle risorse associate all'espressione LF-Tag. `Confidential=True`
 - a. Nel riquadro di navigazione, sotto Autorizzazioni, scegli Autorizzazioni Data lake.
 - b. Scegli Concessione.
 - c. Seleziona utenti e ruoli IAM.
 - d. Scegli il ruolo `lf-data-engineer`.
 - e. Nella sezione LF-tags o risorse del catalogo, seleziona Risorse abbinate ai tag LF.
 - f. Scegliete Aggiungi coppia chiave-valore LF-Tag.
 - g. Aggiungi la chiave con i valori `Confidential. True`
 - h. Nella sezione Autorizzazioni del database, seleziona Descrivi per le autorizzazioni del database e le autorizzazioni concedibili.
 - i. Nella sezione Autorizzazioni per la tabella, seleziona Descrivi, Seleziona e Alter sia per le autorizzazioni Table che per le autorizzazioni Grantable.
 - j. Scegli Concessione.
5. Successivamente, concedi autorizzazioni `lf-data-engineer` concedibili sulle risorse associate all'espressione LF-Tag. `Confidential=False`
 - a. Nel riquadro di navigazione, sotto Autorizzazioni, scegli Autorizzazioni Data lake.
 - b. Scegli Concessione.
 - c. Seleziona utenti e ruoli IAM.
 - d. Scegli il ruolo `lf-data-engineer`.
 - e. Seleziona Risorse abbinate ai tag LF.
 - f. Scegli Aggiungi tag LF.
 - g. Aggiungi la chiave `Confidential` con il valore. `False`
 - h. Nella sezione Autorizzazioni del database, seleziona Descrivi per le autorizzazioni del database e le autorizzazioni concedibili.

- i. Nella sezione Autorizzazioni per tabelle e colonne, non selezionare nulla.
 - j. Scegli Concessione.
6. Successivamente, concediamo autorizzazioni `lf-data-engineer` concedibili sulle risorse associate alle coppie chiave-valore del tag LF e. `Confidential=False Sensitive=True`
- a. Nel pannello di navigazione, in Autorizzazioni, scegli Autorizzazioni dati.
 - b. Scegli Concessione.
 - c. Seleziona utenti e ruoli IAM.
 - d. Scegli il ruolo `lf-data-engineer`.
 - e. Nella sezione LF-tags o nella sezione delle risorse del catalogo, seleziona Risorse abbinate ai tag LF.
 - f. Scegli Aggiungi tag LF.
 - g. Aggiungi la chiave `Confidential` con il valore. `False`
 - h. Scegli Aggiungi coppia chiave-valore del tag LF.
 - i. Aggiungi la chiave con il valore `Sensitive. True`
 - j. Nella sezione Autorizzazioni del database, seleziona Descrivi per le autorizzazioni del database e le autorizzazioni concedibili.
 - k. Nella sezione Autorizzazioni per la tabella, seleziona Descrivi, Seleziona e Alter sia per le autorizzazioni Table che per le autorizzazioni Grantable.
 - l. Scegli Concessione.

Fase 3: Creare database Lake Formation

In questo passaggio, create due database e allegate i tag LF ai database e a colonne specifiche a scopo di test.

Crea i tuoi database e le tue tabelle per l'accesso a livello di database

1. Innanzitutto, crea il database `tag_database`, la tabella e allega i tag `source_data` LF appropriati.
 - a. Nella console Lake Formation (<https://console.aws.amazon.com/lakeformation/>), in Data Catalog, scegli Databases.
 - b. Scegliere Crea database.

- c. Per Nome, immetti tag_database.
 - d. Per Posizione, inserisci la posizione Amazon S3 creata dal CloudFormation modello.
(s3://1f-tagbased-demo-*Account-ID*/tag_database/)
 - e. Deseleziona Usa solo il controllo di accesso IAM per le nuove tabelle in questo database.
 - f. Scegliere Crea database.
2. Quindi, crea una nuova tabella all'interno tag_database di.
- a. Nella pagina Database, seleziona il databasetag_database.
 - b. Scegliete Visualizza tabelle e fate clic su Crea tabella.
 - c. Per Nome, immetti source_data.
 - d. Per Database (Database), scegli il database tag_database.
 - e. Per Formato tabella, scegli AWS Glue Tabella standard.
 - f. Se i dati si trovano in, seleziona Percorso specificato nel mio account.
 - g. In Includi percorso, inserisci il percorso tag_database creato dal CloudFormation modello(s3://1f-tagbased-demo*Account-ID*/tag_database/).
 - h. Per Formato dati, seleziona CSV.
 - i. In Upload schema, inserisci la seguente matrice JSON di struttura di colonne per creare uno schema:

```
[
    {
        "Name": "vendorid",
        "Type": "string"
    },
    {
        "Name": "lpep_pickup_datetime",
        "Type": "string"
    },
    {
        "Name": "lpep_dropoff_datetime",
        "Type": "string"
    },
    {
        "Name": "store_and_fwd_flag",
        "Type": "string"
    },
]
```

```
    {
      "Name": "ratecodeid",
      "Type": "string"
    },
    {
      "Name": "pulocationid",
      "Type": "string"
    },
    {
      "Name": "dolocationid",
      "Type": "string"
    },
    {
      "Name": "passenger_count",
      "Type": "string"
    },
    {
      "Name": "trip_distance",
      "Type": "string"
    },
    {
      "Name": "fare_amount",
      "Type": "string"
    },
    {
      "Name": "extra",
      "Type": "string"
    },
    {
      "Name": "mta_tax",
      "Type": "string"
    },
    {
      "Name": "tip_amount",
      "Type": "string"
    }
  ]
}
```

```
    },  
    {  
      "Name": "tolls_amount",  
      "Type": "string"  
    },  
    {  
      "Name": "ehail_fee",  
      "Type": "string"  
    },  
    {  
      "Name": "improvement_surcharge",  
      "Type": "string"  
    },  
    {  
      "Name": "total_amount",  
      "Type": "string"  
    },  
    {  
      "Name": "payment_type",  
      "Type": "string"  
    }  
  ]
```

- j. Scegli Carica. Dopo aver caricato lo schema, lo schema della tabella dovrebbe apparire come nella schermata seguente:

#	Column Name	▼	Data type
1	vendorid		string
2	lpep_pickup_datetime		string
3	lpep_dropoff_datetime		string
4	store_and_fwd_flag		string
5	ratecodeid		string
6	pulocationid		string
7	dolocationid		string
8	passenger_count		string
9	trip_distance		string
10	fare_amount		string
11	extra		string
12	mta_tax		string
13	tip_amount		string
14	tolls_amount		string
15	ehail_fee		string
16	improvement_surcharge		string
17	total_amount		string
18	payment_type		string

- k. Scegli Invia.
3. Successivamente, allega i tag LF a livello di database.
 - a. Nella pagina Database, trova e seleziona. tag_database
 - b. Nel menu Azioni, scegliete Modifica tag LF.
 - c. Scegliete Assegna nuovo LF-tag.
 - d. Per le chiavi assegnate, scegliete il tag Confidential LF che avete creato in precedenza.
 - e. Per Valori, scegliete. True
 - f. Seleziona Salva.

Questo completa l'assegnazione del tag LF al database tag_database.

Crea il tuo database e la tua tabella per l'accesso a livello di colonna

Ripetete i seguenti passaggi per creare il database col_tag_database e la tabella e allegare source_data_col_lvl1 i tag LF a livello di colonna.

1. Nella pagina Database, scegli Crea database.
2. Per Nome, immetti col_tag_database.
3. Per Posizione, inserisci la posizione Amazon S3 creata dal CloudFormation modello. (s3://lf-tagbased-demo-*Account-ID*/col_tag_database/)
4. Deseleziona Usa solo il controllo di accesso IAM per le nuove tabelle in questo database.
5. Scegliere Crea database.
6. Nella pagina Database, seleziona il tuo nuovo database(col_tag_database).
7. Scegli Visualizza tabelle e fai clic su Crea tabella.
8. Per Nome, immetti source_data_col_lvl1.
9. Per Database, scegli il tuo nuovo database(col_tag_database).
10. Per Formato tabella, scegli AWS Glue Tabella standard.
11. Se i dati si trovano in, seleziona Percorso specificato nel mio account.
12. Inserisci il percorso Amazon S3 per. col_tag_database (s3://lf-tagbased-demo-*Account-ID*/col_tag_database/)
13. Per Formato dei dati, seleziona CSV.
14. In Upload schema, inserisci il seguente schema JSON:

```
[
    {
        "Name": "vendorid",
        "Type": "string"
    },
    {
        "Name": "lpep_pickup_datetime",
        "Type": "string"
    },
    {
        "Name": "lpep_dropoff_datetime",
        "Type": "string"
    },
    {
        "Name": "store_and_fwd_flag",
        "Type": "string"
    },
    {
        "Name": "ratecodeid",
        "Type": "string"
    },
    {
        "Name": "pulocationid",
        "Type": "string"
    },
    {
        "Name": "dolocationid",
        "Type": "string"
    },
    ],
```

```
        {
            "Name": "passenger_count",
            "Type": "string"
        },
        {
            "Name": "trip_distance",
            "Type": "string"
        },
        {
            "Name": "fare_amount",
            "Type": "string"
        },
        {
            "Name": "extra",
            "Type": "string"
        },
        {
            "Name": "mta_tax",
            "Type": "string"
        },
        {
            "Name": "tip_amount",
            "Type": "string"
        },
        {
            "Name": "tolls_amount",
            "Type": "string"
        },
        {
            "Name": "ehail_fee",
```

```
        "Type": "string"

    },
    {
        "Name": "improvement_surcharge",
        "Type": "string"

    },
    {
        "Name": "total_amount",
        "Type": "string"

    },
    {
        "Name": "payment_type",
        "Type": "string"

    }

]
```

15. Scegli Upload. Dopo aver caricato lo schema, lo schema della tabella dovrebbe apparire come nella schermata seguente.

#	Column Name	▼	Data type
1	vendorid		string
2	lpep_pickup_datetime		string
3	lpep_dropoff_datetime		string
4	store_and_fwd_flag		string
5	ratecodeid		string
6	pulocationid		string
7	dolocationid		string
8	passenger_count		string
9	trip_distance		string
10	fare_amount		string
11	extra		string
12	mta_tax		string
13	tip_amount		string
14	tolls_amount		string
15	ehail_fee		string
16	improvement_surcharge		string
17	total_amount		string
18	payment_type		string

16. Scegli Invia per completare la creazione della tabella.
17. Ora, associa il Sensitive=True tag LF alle colonne vendorid e. fare_amount
 - a. Nella pagina Tabelle, seleziona la tabella che hai creato. (source_data_col_lvl1)
 - b. Nel menu Azioni, scegli Schema.
 - c. Selezionate la colonna vendorid e scegliete Modifica tag LF.
 - d. Per Chiavi assegnate, scegliete Sensibile.
 - e. Per Valori, scegli True.
 - f. Seleziona Salva.
18. Quindi, associa il Confidential=False tag LF a. col_tag_database Questo è necessario per poter lf-data-analyst descrivere il database col_tag_database quando si effettua l'accesso da. Amazon Athena
 - a. Nella pagina Database, trova e selezionacol_tag_database.
 - b. Nel menu Azioni, scegliete Modifica tag LF.
 - c. Scegliete Assegna nuovo LF-tag.
 - d. Per Tasti assegnati, scegli il tag Confidential LF che hai creato in precedenza.
 - e. Per Valori, scegliete. False
 - f. Seleziona Salva.

Passaggio 4: concedere le autorizzazioni per la tabella

Concedi le autorizzazioni agli analisti di dati per l'utilizzo dei database tag_database e della tabella col_tag_database utilizzando i tag LF e. Confidential Sensitive

1. Segui questi passaggi per concedere all'lf-data-analystutente le autorizzazioni sugli oggetti associati al tag LF Confidential=True (Database:TAG_database) per avere il database e l'autorizzazione sulle tabelle. Describe Select
 - a. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/aslf-data-engineer>.
 - b. In Autorizzazioni, seleziona Autorizzazioni Data lake.
 - c. Scegli Concessione.
 - d. In Principali, seleziona Utenti e ruoli IAM.

- e. Per gli utenti e i ruoli IAM, scegli `lf-data-analyst`.
 - f. In LF-tags o catalog resources, seleziona Risorse abbinate a LF-Tags.
 - g. Scegli Aggiungi tag LF.
 - h. Per Key, scegli. `Confidential`
 - i. Per Valori, scegli `True`.
 - j. Per le autorizzazioni del database, selezionare `Describe`.
 - k. Per le autorizzazioni relative alle tabelle, scegli Seleziona e descrivi.
 - l. Scegli Concessione.
2. Quindi, ripeti i passaggi per concedere le autorizzazioni agli analisti di dati per l'espressione LF-tag `for. Confidential=False` Questo tag LF viene utilizzato per descrivere la tabella `col_tag_database` e la tabella `source_data_col_lvl` quando si effettua l'accesso da Amazon `lf-data-analyst` Athena.
 - a. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/aslf-data-engineer>.
 - b. Nella pagina Database, seleziona il database `col_tag_database`.
 - c. Scegli Azione e Concedi.
 - d. In Principi, seleziona Utenti e ruoli IAM.
 - e. Per gli utenti e i ruoli IAM, scegli `lf-data-analyst`.
 - f. Seleziona Risorse abbinate ai tag LF.
 - g. Scegli Aggiungi tag LF.
 - h. Per Key, scegli. `Confidential`
 - i. Per Valori, scegli `False`.
 - j. Per le autorizzazioni del database, seleziona `Describe`.
 - k. Per le autorizzazioni relative alle tabelle, non selezionate nulla.
 - l. Scegli Concessione.
3. Quindi, ripeti i passaggi per concedere le autorizzazioni agli analisti di dati per l'espressione LF-tag `per e. Confidential=False Sensitive=True` Questo tag LF viene utilizzato per descrivere la tabella `col_tag_database` e `source_data_col_lvl` (a livello di colonna) quando si effettua l'accesso da Amazon Athena. `lf-data-analyst`
 - a. Accedi alla console di Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/aslf-data-engineer>.

- b. Nella pagina Database, seleziona il database `col_tag_database`.
- c. Scegli Azione e Concedi.
- d. In Principi, seleziona Utenti e ruoli IAM.
- e. Per gli utenti e i ruoli IAM, scegli `lf-data-analyst`.
- f. Seleziona Risorse abbinate ai tag LF.
- g. Scegli Aggiungi tag LF.
- h. Per Key, scegli `Confidential`.
- i. Per Valori, scegli `False`.
- j. Scegli Aggiungi tag LF.
- k. Per Key, scegli `Sensitive`.
- l. Per Valori, scegli `True`.
- m. Per le autorizzazioni del database, seleziona `Describe`.
- n. Per le autorizzazioni relative alle tabelle, seleziona `Select` e `Describe`.
- o. Scegli Concessione.

Passaggio 5: esegui una query in Amazon Athena per verificare le autorizzazioni

Per questo passaggio, usa Amazon Athena per eseguire `SELECT` query sulle due tabelle. (`source_data` and `source_data_col_lvl1`) Usa il percorso Amazon S3 come posizione dei risultati della query. (`s3://lf-tagbased-demo-Account-ID/athena-results/`)

1. Accedi alla console Athena all'indirizzo <https://console.aws.amazon.com/athena/as.1f-data-analyst>
2. Nell'editor di query Athena, scegli `tag_database` nel pannello a sinistra.
3. Scegliete l'icona delle opzioni di menu aggiuntive (tre punti verticali) accanto `source_data` e scegliete la tabella di anteprima.
4. Scegli Esegui query.

L'esecuzione della query dovrebbe richiedere alcuni minuti. La query visualizza tutte le colonne dell'output perché il tag LF è associato a livello di database e la `source_data` tabella lo eredita automaticamente LF-tag dal database. `tag_database`

5. Esegui un'altra query utilizzando `col_tag_database` `source_data_col_lvl1`

La seconda query restituisce le due colonne contrassegnate come Non-Confidential e Sensitive.

6. Puoi anche controllare il comportamento della politica di accesso basata sui tag di Lake Formation nelle colonne per le quali non hai concessioni politiche. Quando viene selezionata una colonna senza tag dalla tabella, `source_data_col_lvl1` Athena restituisce un errore. Ad esempio, puoi eseguire la seguente query per scegliere colonne senza tag: `geolocationid`

```
SELECT geolocationid FROM "col_tag_database"."source_data_col_lvl1" limit 10;
```

Fase 6: Pulizia delle risorse AWS

Per evitare addebiti indesiderati Account AWS, puoi eliminare le AWS risorse che hai utilizzato per questo tutorial.

1. Accedi alla console di Lake Formation come `lf-data-engineer` ed elimina i database `tag_database` e `col_tag_database`.
2. Successivamente, accedi come `lf-data-steward` e ripulisci tutte le autorizzazioni LF-Tag, le autorizzazioni per i dati e le autorizzazioni per la localizzazione dei dati concesse in precedenza e che sono state concesse e. `lf-data-engineer` `lf-data-analyst`.
3. Accedi alla console Amazon S3 come proprietario dell'account utilizzando le credenziali IAM che hai usato per distribuire lo stack. CloudFormation
4. Elimina i seguenti bucket:
 - `lf-tagbased-demo-accesslogs-acct-id`
 - `lf-tagbased-demo-acct-id`
5. Accedi alla CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> ed elimina lo stack che hai creato. Attendi che lo stato dello stack cambi a. `DELETE_COMPLETE`

Protezione dei data lake con il controllo degli accessi a livello di riga

AWS Lake Formation le autorizzazioni a livello di riga consentono di fornire l'accesso a righe specifiche di una tabella in base alla conformità dei dati e alle politiche di governance. Se disponi di tabelle di grandi dimensioni che contengono miliardi di record, hai bisogno di un modo per consentire a diversi utenti e team di accedere solo ai dati che sono autorizzati a vedere. Il controllo degli accessi

a livello di riga è un modo semplice e performante per proteggere i dati, offrendo al contempo agli utenti l'accesso ai dati di cui hanno bisogno per svolgere il proprio lavoro. Lake Formation fornisce audit centralizzati e report di conformità identificando quali responsabili hanno avuto accesso a quali dati, quando e attraverso quali servizi.

In questo tutorial, imparerai come funzionano i controlli di accesso a livello di riga in Lake Formation e come configurarli.

Questo tutorial include un AWS CloudFormation modello per configurare rapidamente le risorse richieste. Puoi rivederlo e personalizzarlo in base alle tue esigenze.

Argomenti

- [Destinatari principali](#)
- [Prerequisiti](#)
- [Fase 1: Fornisci le tue risorse](#)
- [Fase 2: Interrogazione senza filtri di dati](#)
- [Passaggio 3: configura i filtri di dati e concedi le autorizzazioni](#)
- [Passaggio 4: Interrogazione con filtri di dati](#)
- [Fase 5: Pulire le AWS risorse](#)

Destinatari principali

Questo tutorial è destinato a amministratori di dati, ingegneri dei dati e analisti di dati. La tabella seguente elenca i ruoli e le responsabilità di un proprietario e di un consumatore di dati.

Ruolo	Descrizione
Amministratore IAM	Un utente che può creare utenti e ruoli e bucket Amazon Simple Storage Service (Amazon S3). Dispone della politica AdministratorAccess AWS gestita.
Amministratore del data lake	Un utente responsabile della configurazione del data lake, della creazione di filtri di dati e della concessione delle autorizzazioni agli analisti di dati.

Ruolo	Descrizione
Analista dei dati	Un utente che può eseguire query sul data lake. Gli analisti di dati che risiedono in diversi paesi (per il nostro caso d'uso, Stati Uniti e Giappone) possono analizzare solo le recensioni dei prodotti per i clienti residenti nel proprio paese e, per motivi di conformità, non dovrebbero essere in grado di vedere i dati dei clienti che si trovano in altri paesi.

Prerequisiti

Prima di iniziare questo tutorial, è necessario disporre di un file Account AWS da utilizzare per accedere come utente amministrativo con le autorizzazioni corrette. Per ulteriori informazioni, consulta [Completa le attività AWS di configurazione iniziali](#).

Il tutorial presuppone che tu abbia familiarità con IAM. Per informazioni su IAM, consulta la [IAM User Guide](#).

Modifica le impostazioni di Lake Formation

Important

Prima di avviare il CloudFormation modello, disabilita l'opzione Usa solo il controllo di accesso IAM per nuovi database/tabelle in Lake Formation seguendo i passaggi seguenti:

1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> nella regione Stati Uniti orientali (Virginia settentrionale) o nella regione Stati Uniti occidentali (Oregon).
2. In Data Catalog, scegli Impostazioni.
3. Deseleziona Usa solo il controllo di accesso IAM per i nuovi database e Usa solo il controllo di accesso IAM per le nuove tabelle nei nuovi database.
4. Seleziona Salva.

Fase 1: Fornisci le tue risorse

Questo tutorial include un CloudFormation modello per una configurazione rapida. Puoi rivederlo e personalizzarlo in base alle tue esigenze. Il CloudFormation modello genera le seguenti risorse:

- Utenti e politiche per:
 - DataLakeAdmin
 - DataAnalystUSA
 - DataAnalystJP
- Impostazioni e autorizzazioni del data lake Lake Formation
- Una funzione Lambda (per risorse CloudFormation personalizzate supportate da Lambda) utilizzata per copiare file di dati di esempio dal bucket Amazon S3 pubblico al bucket Amazon S3
- Un bucket Amazon S3 che funge da data lake
- Un AWS Glue Data Catalog database, una tabella e una partizione

Crea le tue risorse

Segui questi passaggi per creare le tue risorse utilizzando il CloudFormation modello.

1. Accedi alla CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Avanti nella schermata Crea stack.
4. Inserisci un nome per lo stack.
5. Per DatalakeAdminUsernamee DatalakeAdminUserPassword, inserisci il nome utente e la password IAM per l'utente amministratore di data lake.
6. Per DataAnalystUsUsernamee DataAnalystUsUserPassword, inserisci il nome utente e la password per il nome utente e la password che desideri per l'utente analista di dati responsabile del mercato statunitense.
7. Per DataAnalystJpUsernamee DataAnalystJpUserPassword, inserisci il nome utente e la password per il nome utente e la password che desideri assegnare all'analista di dati responsabile del mercato giapponese.
8. Per DataLakeBucketName, inserisci il nome del tuo bucket di dati.
9. Per DatabaseName, e TableNamelascia come impostazione predefinita.

10. Seleziona Next (Successivo).
11. Nella pagina successiva, scegli Avanti.
12. Controlla i dettagli nella pagina finale e seleziona Riconosco che CloudFormation potrebbe creare risorse IAM.
13. Scegli Create (Crea) .

Il completamento della creazione dello stack può richiedere un minuto.

Fase 2: Interrogazione senza filtri di dati

Dopo aver configurato l'ambiente, puoi interrogare la tabella delle recensioni dei prodotti. Per prima cosa interroga la tabella senza controlli di accesso a livello di riga per assicurarvi di poter vedere i dati. Se esegui query in Amazon Athena per la prima volta, devi configurare la posizione dei risultati della query.

Interroga la tabella senza controllo degli accessi a livello di riga

1. Accedi a Athena console at <https://console.aws.amazon.com/athena/> come DataLakeAdmin utente ed esegui la seguente query:

```
SELECT *  
FROM lakeformation_tutorial_row_security.amazon_reviews  
LIMIT 10
```

La schermata seguente mostra il risultato della query. Questa tabella ha una sola partizione `product_category=Video`, quindi ogni record è un commento di recensione per un prodotto video.

New query 1

```
1 SELECT *
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 LIMIT 10
```

Run query **Save as** **Create** (Run time: 12.62 seconds, Data scanned: 64.57 MB) **Format query** **Clear**

Use Ctrl + Enter to run query, Ctrl + Space to autocomplete

Athena engine version 2 [Release versions](#)

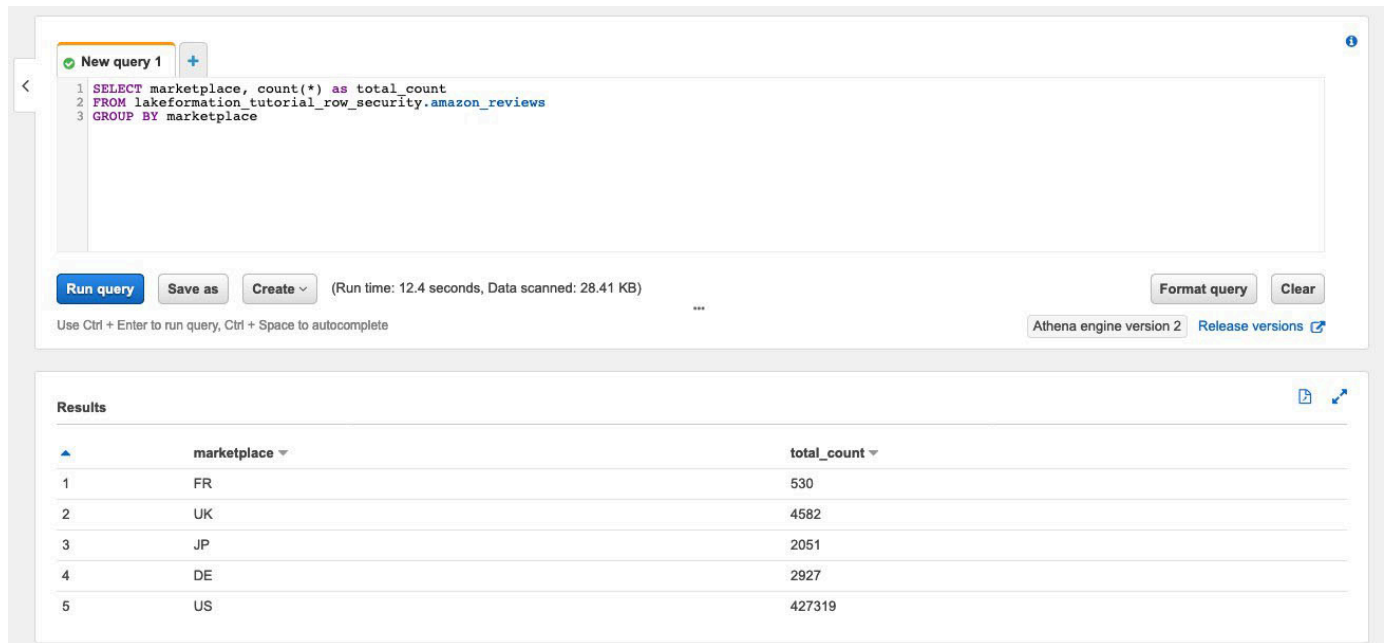
Results

	marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine
1	US	22066705	R3HZYXMJ5HEXIG	6304878621	928670802	The Thin Blue Line 3 [VHS]	5	0	0	N
2	US	20838467	RJC8PH4K3DVQB	630335663X	577032943	Covert Bailey: Fit Or Fat for the 90's [VHS]	1	0	0	N
3	US	15338666	R1OH4581ARVWNX	6300269434	266152594	Young Man With a Horn [VHS]	1	0	2	N
4	US	7080939	R3TWQ5OT8KW0E8	B000EKCQMQ	345913478	Madeline in London (Told By Christopher Plummer)	5	0	0	N
5	US	30548191	R3BK9ULGX82VG0	078311317X	38445970	2 Days in the Valley (Widescreen Edition) [VHS]	5	0	0	N
6	US	16052189	R1LV7NN89A38YT	6302862833	924318070	Zoltz [VHS]	4	0	0	N
7	US	43430756	R2IJAELO3PXEYM	B00027VBBI	51076382	Party Crasher	1	1	1	N
8	US	43539164	R3TNQJ9JANR9Q5	6303205542	69262780	Frugal Gourmet: Spanish Kitchen [VHS]	5	0	0	N
9	US	21187650	R2AVXCQOLI53IC	6302606713	934453987	Live [VHS]	5	0	0	N
10	US	7080939	RC71NIBDHR9KA	B00007ELHT	498552125	Golden Rules of Growing Up [VHS]	5	0	0	N

2. Quindi, esegui una query di aggregazione per recuperare il numero totale di record per marketplace

```
SELECT marketplace, count(*) as total_count
FROM lakeformation_tutorial_row_security.amazon_reviews
GROUP BY marketplace
```

La schermata seguente mostra il risultato della query. La marketplace colonna ha cinque valori diversi. Nei passaggi successivi, imposterai i filtri basati su righe utilizzando la marketplace colonna.



The screenshot shows the AWS Lake Formation console interface. At the top, there is a text area for a SQL query labeled 'New query 1'. The query is:

```
1 SELECT marketplace, count(*) as total_count
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 GROUP BY marketplace
```

Below the query area, there are buttons for 'Run query', 'Save as', and 'Create'. To the right of these buttons, it says '(Run time: 12.4 seconds, Data scanned: 28.41 KB)'. Further right are buttons for 'Format query' and 'Clear'. Below these buttons, a small text says 'Use Ctrl + Enter to run query, Ctrl + Space to autocomplete'. On the far right, it says 'Athena engine version 2' and 'Release versions' with a link icon.

Below the query area, there is a section titled 'Results'. It contains a table with two columns: 'marketplace' and 'total_count'. The table has 5 rows of data:

	marketplace	total_count
1	FR	530
2	UK	4582
3	JP	2051
4	DE	2927
5	US	427319

Passaggio 3: configura i filtri di dati e concedi le autorizzazioni

Questo tutorial utilizza due analisti di dati: uno responsabile del mercato statunitense e l'altro del mercato giapponese. Ogni analista utilizza Athena per analizzare le recensioni dei clienti solo per il proprio marketplace specifico. Crea due diversi filtri di dati, uno per l'analista responsabile del mercato statunitense e un altro per quello responsabile del mercato giapponese. Quindi, concedi agli analisti le rispettive autorizzazioni.

Crea filtri per i dati e concedi le autorizzazioni

1. Crea un filtro per limitare l'accesso ai US marketplace dati.
 - a. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> nella regione Stati Uniti orientali (Virginia settentrionale) come DataLakeAdmin utente.
 - b. Scegli Filtri dati.
 - c. Scegli Crea nuovo filtro.
 - d. Per il nome del filtro dati, inserisciamazon_reviews_US.
 - e. Per Database Target, scegli il databaselakeformation_tutorial_row_security.
 - f. Per la tabella Target, scegli la tabellaamazon_reviews.
 - g. Per l'accesso a livello di colonna, lascia come impostazione predefinita.
 - h. Per Espressione del filtro Row, immettere. marketplace= 'US '

- i. Scegli Create Filter (Crea filtro).
2. Create un filtro per limitare l'accesso ai marketplace dati giapponesi.
 - a. Nella pagina Filtri dati, scegli Crea nuovo filtro.
 - b. Per il nome del filtro dati, inserisciamazon_reviews_JP.
 - c. Per Database Target, scegli il databaselakeformation_tutorial_row_security.
 - d. Per la tabella Target, scegli table amazon_reviews.
 - e. Per l'accesso a livello di colonna, lascia come impostazione predefinita.
 - f. Per Espressione del filtro Row, immettere. marketplace= 'JP'
 - g. Scegli Create Filter (Crea filtro).
3. Successivamente, concedi le autorizzazioni agli analisti di dati utilizzando questi filtri di dati. Segui questi passaggi per concedere le autorizzazioni all'analista di dati statunitense (): DataAnalystUS
 - a. In Autorizzazioni, scegli Autorizzazioni Data lake.
 - b. In Autorizzazione dati, scegli Concedi.
 - c. Per Principal, scegli utenti e ruoli IAM e seleziona il ruoloDataAnalystUS.
 - d. Per i tag LF o le risorse del catalogo, scegli Named data catalog resources.
 - e. Per Database, scegliere lakeformation_tutorial_row_security.
 - f. Per Tabelle (facoltativo), scegliete. amazon_reviews
 - g. Per i filtri di dati, facoltativo, seleziona. amazon_reviews_US
 - h. Per le autorizzazioni del filtro dati, seleziona Seleziona.
 - i. Scegli Concessione.
4. Segui questi passaggi per concedere le autorizzazioni all'analista di dati giapponese (): DataAnalystJP
 - a. In Autorizzazioni, scegli Autorizzazioni Data Lake.
 - b. In Autorizzazione dati, scegli Concedi.
 - c. Per Principal, scegli utenti e ruoli IAM e seleziona il ruoloDataAnalystJP.
 - d. Per i tag LF o le risorse del catalogo, scegli Named data catalog resources.
 - e. Per Database, scegliere lakeformation_tutorial_row_security.
 - f. Per Tabelle (facoltativo), scegliete. amazon_reviews
 - g. Per i filtri di dati, facoltativo, seleziona. amazon_reviews_JP

- h. Per le autorizzazioni del filtro dati, seleziona Seleziona.
- i. Scegli Concessione.

Passaggio 4: Interrogazione con filtri di dati

Con i filtri dei dati allegati alla tabella delle recensioni dei prodotti, esegui alcune query e scopri come vengono applicate le autorizzazioni da Lake Formation.

1. Accedi alla console Athena <https://console.aws.amazon.com/athena/> come utente. DataAnalystUS
2. Esegui la seguente query per recuperare alcuni record, che vengono filtrati in base alle autorizzazioni a livello di riga che abbiamo definito:

```
SELECT *
FROM lakeformation_tutorial_row_security.amazon_reviews
LIMIT 10
```

La schermata seguente mostra il risultato della query.

The screenshot shows the AWS Athena console interface. At the top, there are tabs for 'New query 1' and 'New query 2'. The query editor contains the following SQL query:

```
1 SELECT *
2 FROM lakeformation_tutorial_row_security.amazon_reviews
3 LIMIT 10
```

Below the query editor, there are buttons for 'Run query', 'Save as', and 'Create'. A status bar indicates '(Run time: 11.9 seconds, Data scanned: 0 KB)'. To the right, there are buttons for 'Format query' and 'Clear'. Below the query editor, there is a section for 'Results' which displays a table with 10 rows of data. The table has the following columns: marketplace, customer_id, review_id, product_id, product_parent, product_title, star_rating, helpful_votes, total_votes, vine, verified_purchase, and review_text. The data is as follows:

marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine	verified_purchase	review_text
US	43836277	R2NUBTTUO60VYU	B00068S41I	653409458	The Notebook [VHS]	4	0	0	N	Y	KI
US	20261976	R2QTOLZUQERU5B	6303060013	176265879	American Cyborg: Steel Warrior [VHS]	5	0	1	N	Y	it
US	15947067	R1PHKR75RKZNSU	6303927319	850909689	Biography - Darryl Zanuck [VHS]	5	0	0	N	N	G
US	19288153	R1BL2WVE5X34UN	6304032153	479446069	Timon & Pumbaa: Quit Buggin Me [VHS]	5	0	0	N	N	Fi
US	19712967	R2DKOCIBS5FSP7	0784017743	35164822	Denise Austin - Hit the Spot: Arms & Bust [VHS]	5	0	0	N	Y	G
US	51047097	R2XF5HQATT4IVR	0793960142	233936597	I Love Lucy - Lucy's Italian Movie/Ballet [VHS]	5	0	0	N	N	Fi
US	43836277	R2NUBTTUO60VYU	B00068S41I	653409458	The Notebook [VHS]	4	0	0	N	Y	KI
US	51047097	R1C0H0G6NATZXO	6304872585	233936597	I Love Lucy: Lucy Meets Superman/Freez [VHS]	5	0	1	N	N	Fi
US	42808630	R2HXW7UD4IGZLN	6303060013	176265879	American Cyborg: Steel Warrior [VHS]	5	0	1	N	Y	M
US	11682952	R18IURLUPYI4DP	6302993717	42308924	Songs of Christmas [VHS]	1	0	0	N	Y	R

3. Allo stesso modo, esegui una query per contare il numero totale di record per marketplace.

```
SELECT marketplace , count ( * ) as total_count
```

```
FROM lakeformation_tutorial_row_security .amazon_reviews  
GROUP BY marketplace
```

Il risultato della query mostra solo marketplace US i risultati. Questo perché all'utente è consentito visualizzare solo le righe in cui il valore della marketplace colonna è uguale a US.

4. Passa all'`DataAnalystJP` utente ed esegui la stessa query.

```
SELECT *  
FROM lakeformation_tutorial_row_security.amazon_reviews  
LIMIT 10
```

Il risultato della query mostra che solo i record appartengono a JP marketplace.

5. Esegui la query per contare il numero totale di record per marketplace.

```
SELECT marketplace, count(*) as total_count  
FROM lakeformation_tutorial_row_security.amazon_reviews  
GROUP BY marketplace
```

Il risultato della query mostra solo la riga appartenente a JP marketplace.

Fase 5: Pulire le AWS risorse

Pulizia delle risorse

Per evitare addebiti indesiderati Account AWS, puoi eliminare le AWS risorse che hai utilizzato per questo tutorial.

- [Elimina lo stack di formazione del cloud.](#)

Condivisione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation e risorse denominate

Questo tutorial dimostra come configurare AWS Lake Formation la condivisione sicura dei dati archiviati all'interno di un data lake con più aziende, organizzazioni o unità aziendali, senza dover copiare l'intero database. Esistono due opzioni per condividere database e tabelle con altri utenti Account AWS utilizzando il controllo degli accessi tra account di Lake Formation:

- Controllo degli accessi basato su tag Lake Formation (consigliato)

Il controllo degli accessi basato su tag Lake Formation è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. In Lake Formation, questi attributi sono chiamati LF-tag. Per ulteriori informazioni, consulta [Gestione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation](#).

- Risorse denominate Lake Formation

Il metodo delle risorse denominato Lake Formation è una strategia di autorizzazione che definisce le autorizzazioni per le risorse. Le risorse includono database, tabelle e colonne. Gli amministratori del Data Lake possono assegnare e revocare le autorizzazioni sulle risorse di Lake Formation. Per ulteriori informazioni, consulta [Condivisione dei dati tra account in Lake Formation](#).

Consigliamo di utilizzare risorse denominate se l'amministratore del data lake preferisce concedere le autorizzazioni in modo esplicito alle singole risorse. Quando utilizzi il metodo della risorsa denominata per concedere le autorizzazioni Lake Formation su una risorsa Data Catalog a un account esterno, Lake Formation utilizza AWS Resource Access Manager (AWS RAM) per condividere la risorsa.

Argomenti

- [Destinatari principali](#)
- [Configura le impostazioni di Lake Formation Data Catalog nell'account produttore](#)
- [Passaggio 1: Fornisci le tue risorse utilizzando modelli AWS CloudFormation](#)
- [Fase 2: Prerequisiti per la condivisione tra account di Lake Formation](#)
- [Fase 3: Implementare la condivisione tra account utilizzando il metodo di controllo degli accessi basato su tag](#)
- [Fase 4: Implementazione del metodo di risorsa denominato](#)
- [Passaggio 5: Pulisci AWS le risorse](#)

Destinatari principali

Questo tutorial è destinato a amministratori di dati, ingegneri di dati e analisti di dati. Quando si tratta di condividere le tabelle di Data Catalog AWS Glue e amministrare le autorizzazioni in Lake Formation, gli amministratori dei dati all'interno degli account di produzione hanno la proprietà

funzionale in base alle funzioni supportate e possono concedere l'accesso a vari consumatori, organizzazioni esterne e account. La tabella seguente elenca i ruoli utilizzati in questo tutorial:

Ruolo	Description
DataLakeAdminProducer	L'utente IAM amministratore del data lake dispone dei seguenti accessi: • Accesso completo in lettura, scrittura e aggiornamento a tutte le risorse del Data Catalog • Possibilità di concedere autorizzazioni alle risorse • Può creare collegamenti a risorse per la tabella condivisa • Può allegare tag LF alle risorse, il che fornisce l'accesso ai principali in base a qualsiasi politica creata dagli amministratori dei dati
DataLakeAdminConsumer	L'utente IAM di data lake admin ha il seguente accesso: • Accesso completo in lettura, scrittura e aggiornamento a tutte le risorse del Data Catalog • Possibilità di concedere autorizzazioni alle risorse • Può creare collegamenti a risorse per la tabella condivisa • Può allegare tag LF alle risorse, il che fornisce l'accesso ai principali in base a qualsiasi politica creata dagli amministratori dei dati
DataAnalyst	L' DataAnalyst utente ha il seguente accesso:

Ruolo	Description
	• Accesso granulare alle risorse condivise dalle politiche di accesso basate su tag di Lake Formation o utilizzando il metodo delle risorse denominate

Configura le impostazioni di Lake Formation Data Catalog nell'account produttore

Prima di iniziare questo tutorial, è necessario disporre di un Account AWS file da utilizzare per accedere come utente amministrativo con le autorizzazioni corrette. Per ulteriori informazioni, consulta [Completa le attività AWS di configurazione iniziali](#).

Il tutorial presuppone che tu abbia familiarità con IAM. Per informazioni su IAM, consulta la [IAM User Guide](#).

Configura le impostazioni di Lake Formation Data Catalog nell'account produttore

Note

In questo tutorial, l'account che ha la tabella di origine è chiamato account produttore e l'account che deve accedere alla tabella di origine è chiamato account consumatore.

Lake Formation fornisce il proprio modello di gestione delle autorizzazioni. Per mantenere la retrocompatibilità con il modello di autorizzazione IAM, per impostazione predefinita l'Superautorizzazione viene concessa al gruppo IAMAllowedPrincipals su tutte AWS Glue Data Catalog le risorse esistenti. Inoltre, le impostazioni di controllo degli accessi Use only IAM sono abilitate per le nuove risorse del Data Catalog. Questo tutorial utilizza un controllo degli accessi granulare utilizzando le autorizzazioni di Lake Formation e utilizza le politiche IAM per un controllo degli accessi a grana grossolana. Per informazioni dettagliate, vedi [Metodi per il controllo granulare degli accessi](#). Pertanto, prima di utilizzare un AWS CloudFormation modello per una configurazione rapida, è necessario modificare le impostazioni del Lake Formation Data Catalog nell'account produttore.

 Important

Questa impostazione influisce su tutti i database e le tabelle appena creati, quindi consigliamo vivamente di completare questo tutorial in un account non di produzione o in un nuovo account. Inoltre, se utilizzi un account condiviso (ad esempio l'account di sviluppo della tua azienda), assicurati che ciò non influisca sulle risorse altrui. Se preferisci mantenere le impostazioni di sicurezza predefinite, devi completare un passaggio aggiuntivo durante la condivisione delle risorse con altri account, in cui revochi l'autorizzazione Super predefinita dal IAMAllowedPrincipals database o dalla tabella. Discuteremo i dettagli più avanti in questo tutorial.

Per configurare le impostazioni di Lake Formation Data Catalog nell'account produttore, completa i seguenti passaggi:

1. Accedi all'account Console di gestione AWS utilizzando l'account produttore come utente amministratore o come utente con l'autorizzazione dell'PutDataLakeSettingsAPI Lake Formation.
2. Sulla console Lake Formation, nel pannello di navigazione, in Data Catalog, scegli Impostazioni.
3. Deseleziona Usa solo il controllo di accesso IAM per i nuovi database e Usa solo il controllo di accesso IAM per le nuove tabelle nei nuovi database

Scegli Save (Salva).

AWS Lake Formation > Data catalog settings

Data catalog settings

Default permissions for newly created databases and tables

These settings maintain existing AWS Glue Data Catalog behavior. You can still set individual permissions on databases and tables, which will take effect when you revoke the Super permission from IAMAllowedPrincipals. See [Changing Default Settings for Your Data Lake](#).

- ☐ Use only IAM access control for new databases
- ☐ Use only IAM access control for new tables in new databases

Default permissions for AWS CloudTrail

These settings specify the information being shown in AWS CloudTrail.

Resource owners

Enter resource owners you wish to share your CloudTrail access details with.

Enter one or more AWS account IDs. Press Enter after each ID.

Cancel

Save

Inoltre, puoi rimuovere `CREATE_DATABASE` le autorizzazioni per la `IAMAllowedPrincipals` sezione Ruoli e attività amministrative, Creatori di database. Solo allora, puoi decidere chi può creare un nuovo database tramite le autorizzazioni di Lake Formation.

Passaggio 1: Fornisci le tue risorse utilizzando modelli AWS CloudFormation

Il CloudFormation modello per l'account produttore genera le seguenti risorse:

- Un bucket Amazon S3 che funge da data lake.
- Una funzione Lambda (per risorse personalizzate supportate da Lambda). CloudFormation Utilizziamo la funzione per copiare file di dati di esempio dal bucket Amazon S3 pubblico al tuo bucket Amazon S3.
- Utenti e politiche IAM: `DataLakeAdminProducer`

- Le impostazioni e le autorizzazioni appropriate di Lake Formation, tra cui:
 - Definizione dell'amministratore del data lake di Lake Formation nell'account produttore
 - Registrazione di un bucket Amazon S3 come ubicazione del data lake Lake Formation (account produttore)
- Un AWS Glue Data Catalog database, una tabella e una partizione. Poiché esistono due opzioni per la condivisione delle risorse Account AWS, questo modello crea due set separati di database e tabelle.

Il CloudFormation modello per l'account consumatore genera le seguenti risorse:

- Utenti e politiche IAM:
 - DataLakeAdminConsumer
 - DataAnalyst
- Un AWS Glue Data Catalog database. Questo database serve per creare collegamenti di risorse a risorse condivise.

Crea le tue risorse nell'account del produttore

1. Accedi alla AWS CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Next (Successivo).
4. Per il nome dello stack, inserisci un nome per lo stack, ad esempio. stack-producer
5. Nella sezione Configurazione utente, inserisci il nome utente e la password per ProducerDataLakeAdminUserName e. ProducerDataLakeAdminUserPassword
6. Per DataLakeBucketName, inserisci il nome del tuo bucket data lake. Questo nome deve essere univoco a livello globale.
7. Per DatabaseName e TableName, lascia i valori predefiniti.
8. Scegli Next (Successivo).
9. Nella pagina successiva, scegli Avanti.
10. Controlla i dettagli nella pagina finale e seleziona Riconosco che AWS CloudFormation potrebbe creare risorse IAM.
11. Scegli Create (Crea).

La creazione dello stack può richiedere fino a un minuto.

Crea le tue risorse nell'account consumatore

1. Accedi alla AWS CloudFormation console all'indirizzo <https://console.aws.amazon.com/cloudformation> nella regione Stati Uniti orientali (Virginia settentrionale).
2. [Scegli Launch Stack.](#)
3. Scegli Next (Successivo).
4. Per il nome dello stack, inserisci un nome per lo stack, ad esempio. `stack-consumer`
5. Nella sezione Configurazione utente, inserisci il nome utente e la password per `ConsumerDataLakeAdminUserName` e `ConsumerDataLakeAdminUserPassword`
6. Per `DataAnalystUserName` e `DataAnalystUserPassword`, inserisci il nome utente e la password che desideri per l'utente IAM dell'analista di dati.
7. Per `DataLakeBucketName`, inserisci il nome del tuo bucket data lake. Questo nome deve essere univoco a livello globale.
8. Per `DatabaseName`, lascia i valori predefiniti.
9. Per `AthenaQueryResultS3BucketName`, inserisci il nome del bucket Amazon S3 in cui sono archiviati i risultati delle query di Amazon Athena. Se non ne hai uno, [crea un bucket Amazon S3](#).
10. Scegli Next (Successivo).
11. Nella pagina successiva, scegli Avanti.
12. Controlla i dettagli nella pagina finale e seleziona Riconosco che AWS CloudFormation potrebbe creare risorse IAM.
13. Scegli Create (Crea).

La creazione dello stack può richiedere fino a un minuto.

Note

Dopo aver completato il tutorial, elimina lo stack CloudFormation per evitare di incorrere in addebiti. Verifica che le risorse siano state eliminate correttamente nello stato dell'evento per lo stack.

Fase 2: Prerequisiti per la condivisione tra account di Lake Formation

Prima di condividere le risorse con Lake Formation, esistono dei prerequisiti sia per il metodo di controllo degli accessi basato su tag che per il metodo delle risorse denominate.

Prerequisiti completi per il controllo degli accessi basato su tag per la condivisione dei dati tra account

- Per ulteriori informazioni sui requisiti di condivisione dei dati tra account, consulta la [Prerequisiti](#) sezione del capitolo Condivisione dei dati tra account.

Per condividere le risorse di Data Catalog con la versione 3 o successiva delle impostazioni della versione dell'account Cross, il concedente deve disporre delle autorizzazioni IAM definite nella policy AWS gestita del tuo account. `AWSLakeFormationCrossAccountManager`

Se utilizzi la versione 1 o la versione 2 delle impostazioni della versione dell'account Cross, prima di poter utilizzare il metodo di controllo dell'accesso basato su tag per concedere l'accesso alle risorse tra account diversi, devi aggiungere il seguente oggetto di JSON autorizzazione alla politica delle risorse di Data Catalog nell'account produttore. Ciò concede all'account consumatore l'autorizzazione ad accedere al Data Catalog quando `glue:EvaluatedByLakeFormationTags` è vero. Inoltre, questa condizione diventa vera per le risorse a cui hai concesso l'autorizzazione utilizzando i tag di autorizzazione di Lake Formation sull'account del consumatore. Questa politica è necessaria per tutte le persone Account AWS a cui concedi le autorizzazioni.

La seguente politica deve essere inclusa in un Statement elemento. Discuteremo la politica IAM completa nella prossima sezione.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:*"
  ],
  "Principal": {
    "AWS": [
      "consumer-account-id"
    ]
  },
  "Resource": [
    "arn:aws:glue:region:account-id:table/*",
  ]
}
```

```

        "arn:aws:glue:region:account-id:database/*",
        "arn:aws:glue:region:account-id:catalog"
    ],
    "Condition": {
        "Bool": {
            "glue:EvaluatedByLakeFormationTags": true
        }
    }
}

```

Completa i prerequisiti per la condivisione tra account del metodo delle risorse denominate

1. Se nel tuo account non è presente alcuna politica sulle risorse di Data Catalog, il cross-account di Lake Formation ti consente di procedere come al solito. Tuttavia, se esiste una politica in materia di risorse di Data Catalog, devi aggiungere la seguente dichiarazione per consentire che le concessioni tra account abbiano esito positivo se effettuate con il metodo della risorsa denominata. Se prevedi di utilizzare solo il metodo della risorsa denominata o solo il metodo di controllo dell'accesso basato su tag, puoi saltare questo passaggio. In questo tutorial, valutiamo entrambi i metodi e dobbiamo aggiungere la seguente politica.

La seguente politica deve trovarsi all'interno di un Statement elemento. Discuteremo la politica IAM completa nella prossima sezione.

```

{
    "Effect": "Allow",
    "Action": [
        "glue:ShareResource"
    ],
    "Principal": {
        "Service": "ram.amazonaws.com"
    },
    "Resource": [
        "arn:aws:glue:region:account-id:table/*/*",
        "arn:aws:glue:region:account-id:database/*",
        "arn:aws:glue:region:account-id:catalog"
    ]
}

```

2. Quindi, aggiungi la politica AWS Glue Data Catalog delle risorse utilizzando AWS Command Line Interface (AWS CLI).

Se concedi autorizzazioni per più account utilizzando sia il metodo di controllo degli accessi basato su tag che il metodo Named Resource, devi impostare l'EnableHybridargomento su «true» quando aggiungi le politiche precedenti. Perché questa opzione non è attualmente supportata sulla console ed è necessario utilizzare l'API e. `glue:PutResourcePolicy` AWS CLI

Innanzitutto, crea un documento di policy (come `policy.json`) e aggiungi le due politiche precedenti. Sostituiscilo `consumer-account-id` con quello `account ID` del Account AWS destinatario della concessione, `region` con la regione del catalogo dati contenente i database e le tabelle per i quali stai concedendo le autorizzazioni e con l'ID del produttore. `account-id` Account AWS

Immettete il seguente comando AWS CLI . Sostituisci `glue-resource-policy` con i valori corretti (ad esempio file: `//policy.json`).

```
aws glue put-resource-policy --policy-in-json glue-resource-policy --enable-hybrid TRUE
```

Per ulteriori informazioni, consulta [put-resource-policy](#).

Fase 3: Implementare la condivisione tra account utilizzando il metodo di controllo degli accessi basato su tag

In questa sezione, ti illustreremo i seguenti passaggi di alto livello:

1. Definisci un tag LF.
2. Assegna il tag LF alla risorsa di destinazione.
3. Concedi le autorizzazioni LF-Tag all'account consumatore.
4. Concedi le autorizzazioni relative ai dati all'account consumatore.
5. Facoltativamente, revoca le autorizzazioni per il database, le IAMAllowedPrincipals tabelle e le colonne.
6. Crea un link di risorsa alla tabella condivisa.
7. Crea un tag LF e assegnalo al database di destinazione.

8. Concedi le autorizzazioni relative ai dati LF-Tag all'account consumatore.

Definisci un tag LF

Note

Se hai effettuato l'accesso al tuo account produttore, esci prima di completare i seguenti passaggi.

1. Accedi all'account produttore come amministratore del data lake all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Usa il numero di account del produttore, il nome utente IAM (l'impostazione predefinita è `DataLakeAdminProducer`) e la password che hai specificato durante la creazione CloudFormation dello stack.
2. Sulla console di Lake Formation (<https://console.aws.amazon.com/lakeformation/>), nel pannello di navigazione, in Autorizzazioni, scegli LF-tags e permessi.
3. Scegliete Aggiungi LF-tag.

Assegna il tag LF alla risorsa di destinazione

Assegna il tag LF alla risorsa di destinazione e concedi le autorizzazioni relative ai dati a un altro account

In qualità di amministratore del data lake, puoi allegare tag alle risorse. Se prevedi di utilizzare un ruolo separato, potresti dover concedere autorizzazioni di descrizione e allegare al ruolo separato.

1. Nel riquadro di navigazione, in Data Catalog, seleziona Database.
2. Seleziona il database di destinazione (`lakeformation_tutorial_cross_account_database_tbac`) e nel menu Azioni, scegli Modifica tag LF.

In questo tutorial, puoi assegnare un tag LF a un database, ma puoi anche assegnare tag LF a tabelle e colonne.

3. Scegliete Assegna nuovo LF-tag.
4. Aggiungi la chiave e il valore `Confidentiality: public`
5. Scegli Save (Salva).

Concedi l'autorizzazione LF-Tag all'account consumatore

Sempre nell'account produttore, concedi le autorizzazioni all'account consumer per accedere all'LF-Tag.

1. Nel pannello di navigazione, sotto Autorizzazioni, scegli LF-tags e permessi.
2. Scegliete la scheda LF-Tags e scegliete la chiave e i valori del tag LF che viene condiviso con l'account consumatore (chiave e valore). **Confidentiality** public
3. Scegli Concedi autorizzazioni.
4. Per Tipo di autorizzazione, scegli le autorizzazioni della coppia chiave-valore LF-Tag.
5. Per Principali, scegli Account esterni.
6. Inserisci l'Account AWS ID di destinazione.

Account AWS all'interno della stessa organizzazione appaiono automaticamente. Altrimenti, devi inserire manualmente l' Account AWS ID.

7. In Autorizzazioni, seleziona Descrivi.

Si tratta delle autorizzazioni concesse all'account consumatore. Le autorizzazioni concedibili sono autorizzazioni che l'account consumatore può concedere ad altri destinatari.

8. Scegli Concessione.

A questo punto, l'amministratore del consumer data lake dovrebbe essere in grado di trovare il policy tag condiviso tramite la console dell'account consumer Lake Formation, sotto Permissions, LF-tags and permissions.

Concedi l'autorizzazione ai dati all'account consumatore

Ora forniremo l'accesso ai dati all'account consumatore specificando un'espressione LF-Tag e concedendo all'account consumatore l'accesso a qualsiasi tabella o database che corrisponda all'espressione.

1. Nel pannello di navigazione, in Autorizzazioni, Autorizzazioni Data Lake, scegli Concedi.
2. Per Principali, scegli Account esterni e inserisci l'ID di destinazione. Account AWS
3. Per i tag LF o le risorse del catalogo, scegli la chiave e i valori del tag LF da condividere con l'account consumatore (chiave e valore). **Confidentiality** public
4. Per Autorizzazioni, in Risorse abbinate ai tag LF (consigliato) scegli Aggiungi tag LF.

5. Seleziona la chiave e il valore del tag che viene condiviso con l'account consumatore (chiave e valore). `Confidentiality public`
6. Per le autorizzazioni del database, seleziona `Descrivi in Autorizzazioni del database` per concedere le autorizzazioni di accesso a livello di database.
7. L'amministratore del consumer data lake dovrebbe essere in grado di trovare il policy tag condiviso tramite l'account consumer sulla console di Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>, in Autorizzazioni, ruoli e attività amministrativi, LF-Tags.
8. Seleziona `Descrivi in Autorizzazioni concesse` in modo che l'account consumer possa concedere autorizzazioni a livello di database ai propri utenti.
9. Per le autorizzazioni relative alle tabelle e alle colonne, seleziona `Seleziona e descrivi in Autorizzazioni per le tabelle`.
10. Seleziona `Seleziona e descrivi in Autorizzazioni concedibili`.
11. Scegli `Concessione`.

Revoca l'autorizzazione per **IAMAllowedPrincipals** il database, le tabelle e le colonne (facoltativo).

All'inizio di questo tutorial, hai modificato le impostazioni del Lake Formation Data Catalog. Se hai saltato quella parte, questo passaggio è obbligatorio. Se hai modificato le impostazioni del Lake Formation Data Catalog, puoi saltare questo passaggio.

In questo passaggio, dobbiamo revocare l'autorizzazione Super predefinita dal database o **IAMAllowedPrincipals** dalla tabella. Per informazioni dettagliate, vedi [Passaggio 4: Passa i tuoi archivi dati al modello di autorizzazioni Lake Formation](#).

Prima di revocare l'autorizzazione **IAMAllowedPrincipals**, assicurati di aver concesso ai principali IAM esistenti l'autorizzazione necessaria tramite Lake Formation. Ciò include tre passaggi:

1. Aggiungi l'autorizzazione IAM all'utente o al ruolo IAM di destinazione con `GetDataAccess` Lake Formation (con policy IAM).
2. Concedi all'utente o al ruolo IAM di destinazione le autorizzazioni relative ai dati di Lake Formation (modifica, selezione e così via).
3. Quindi, revoca le autorizzazioni per **IAMAllowedPrincipals**. Altrimenti, dopo aver revocato le autorizzazioni per **IAMAllowedPrincipals**, i principali IAM esistenti potrebbero non essere più in grado di accedere al database di destinazione o al Data Catalog.

La revoca dell'autorizzazione Super for IAMAllowedPrincipals è necessaria quando si desidera applicare il modello di autorizzazione Lake Formation (anziché il modello di policy IAM) per gestire l'accesso degli utenti all'interno di un singolo account o tra più account utilizzando il modello di autorizzazione Lake Formation. Non è necessario revocare l'autorizzazione IAMAllowedPrincipals per altre tabelle in cui si desidera mantenere il modello di policy IAM tradizionale.

A questo punto, l'amministratore di Data Lake dell'account consumer dovrebbe essere in grado di trovare il database e la tabella condivisi tramite l'account consumer sulla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>, sotto Data Catalog, database. In caso contrario, verifica se quanto segue è configurato correttamente:

1. Il tag di policy e i valori corretti vengono assegnati ai database e alle tabelle di destinazione.
2. L'autorizzazione corretta per il tag e l'autorizzazione per i dati vengono assegnate all'account consumatore.
3. Revoca la super autorizzazione predefinita dal IAMAllowedPrincipals database o dalla tabella.

Crea un link di risorsa alla tabella condivisa

Quando una risorsa viene condivisa tra account e le risorse condivise non vengono inserite nel catalogo dati degli account consumatori. Per renderli disponibili e interrogare i dati sottostanti di una tabella condivisa utilizzando servizi come Athena, dobbiamo creare un collegamento di risorsa alla tabella condivisa. Un link a una risorsa è un oggetto del Data Catalog che è un collegamento a un database o a una tabella locale o condivisa. Per informazioni dettagliate, vedi [Creazione di collegamenti alle risorse](#). Creando un collegamento a una risorsa, puoi:

- Assegna un nome diverso a un database o a una tabella in linea con le politiche di denominazione delle risorse del Data Catalog.
- Utilizza servizi come Athena e Redshift Spectrum per interrogare database o tabelle condivisi.

Per creare un collegamento a una risorsa, completa i seguenti passaggi:

1. Se hai effettuato l'accesso al tuo account consumatore, esci.

2. Accedi come amministratore di Data Lake dell'account consumer. Utilizza l'ID dell'account consumer, il nome utente IAM (predefinito `DatalakeAdminConsumer`) e la password che hai specificato durante la creazione CloudFormation dello stack.
3. Sulla console Lake Formation (<https://console.aws.amazon.com/lakeformation/>), nel pannello di navigazione, in Data Catalog, Databases, scegli il database `condivisolakeformation_tutorial_cross_account_database_tbac`.

Se non vedi il database, rivedi i passaggi precedenti per vedere se tutto è configurato correttamente.

4. Scegli Visualizza tabelle.
5. Scegli la tabella condivisa `amazon_reviews_table_tbac`.
6. Nel menu Azioni, scegli Crea link a una risorsa.
7. Per il nome del link alla risorsa, inserisci un nome (per questo `tutorial_amazon_reviews_table_tbac_resource_link`).
8. In Database, seleziona il database in cui viene creato il collegamento alla risorsa (per questo post, lo stack CloudFormation n ha creato il `database_lakeformation_tutorial_cross_account_database_consumer`).
9. Scegli Create (Crea).

Il link alla risorsa viene visualizzato in Catalogo dati, Tabelle.

Create un tag LF e assegnatelo al database di destinazione

I tag Lake Formation si trovano nello stesso Data Catalog delle risorse. Ciò significa che i tag creati nell'account produttore non possono essere utilizzati quando si concede l'accesso ai link alle risorse nell'account consumatore. È necessario creare un set separato di tag LF nell'account consumatore per utilizzare il controllo degli accessi basato sui tag LF quando si condividono i link alle risorse nell'account consumatore.

1. Definisci il tag LF nell'account consumatore. Per questo tutorial, utilizziamo chiavi `Division` e valori `sales` `emarketing`. `analyst`
2. Assegna la chiave `Division` e il valore del tag LF `analyst` al `database_lakeformation_tutorial_cross_account_database_consumer`, dove viene creato il collegamento alla risorsa.

Concedi l'autorizzazione ai dati LF-Tag al consumatore

Come ultimo passaggio, concedi l'autorizzazione ai dati di LF-Tag al consumatore.

1. Nel pannello di navigazione, in Autorizzazioni, Autorizzazioni Data lake, scegli Concedi.
2. Per Principal, scegli utenti e ruoli IAM e scegli l'utente. `DataAnalyst`
3. Per i tag LF o le risorse del catalogo, scegli Risorse abbinate ai tag LF (consigliato).
4. Scegliete Key Division and Value Analyst.
5. Per le autorizzazioni del database, seleziona Descrivi in Autorizzazioni del database.
6. Per le autorizzazioni per tabelle e colonne, seleziona Seleziona e descrivi in Autorizzazioni per le tabelle.
7. Scegli Concessione.
8. Ripeti questi passaggi per l'utente `DataAnalyst`, dove si trova la chiave LF-Tag e il valore `èConfidentiality.public`

A questo punto, l'utente analista di dati nell'account consumatore dovrebbe essere in grado di trovare il database e il collegamento alla risorsa e di interrogare la tabella condivisa tramite la console Athena all'indirizzo. <https://console.aws.amazon.com/athena/> In caso contrario, verifica se quanto segue è configurato correttamente:

- Il link alla risorsa viene creato per la tabella condivisa
- Hai concesso all'utente l'accesso al tag LF condiviso dall'account produttore
- Hai concesso all'utente l'accesso al tag LF associato al link alla risorsa e al database in cui è stato creato il link alla risorsa
- Controlla di aver assegnato il tag LF corretto al link alla risorsa e al database in cui è stato creato il link alla risorsa

Fase 4: Implementazione del metodo di risorsa denominato

Per utilizzare il metodo della risorsa denominata, ti spieghiamo i seguenti passaggi di alto livello:

1. Facoltativamente, revoca l'autorizzazione per il database, `IAMAllowedPrincipals` le tabelle e le colonne.
2. Concedi l'autorizzazione ai dati all'account consumatore.
3. Accetta una condivisione di risorse da AWS Resource Access Manager.

4. Crea un link alla risorsa per la tabella condivisa.
5. Concedi al consumatore l'autorizzazione ai dati per la tabella condivisa.
6. Concedi al consumatore l'autorizzazione ai dati per il collegamento alla risorsa.

Revoca l'autorizzazione per **IAMAllowedPrincipals** il database, le tabelle e le colonne (facoltativo)

- All'inizio di questo tutorial, abbiamo modificato le impostazioni del Lake Formation Data Catalog. Se hai saltato quella parte, questo passaggio è obbligatorio. Per istruzioni, consulta il passaggio facoltativo nella sezione precedente.

Concedi l'autorizzazione ai dati all'account consumatore

1.

 **Note**

Se hai effettuato l'accesso all'account produttore come altro utente, esci prima.

Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> utilizzando l'account producer data lake administrator utilizzando l' Account AWS ID, il nome utente IAM (l'impostazione predefinita è `DataLakeAdminProducer`) e la password specificati durante la creazione CloudFormation dello stack.

2. Nella pagina Autorizzazioni, in Autorizzazioni Data lake scegli Concedi.
3. In Principali, scegli Account esterni e inserisci una o più Account AWS IDs organizzazioni. AWS IDs Per ulteriori informazioni, vedere: [AWS Organizations](#).

Le organizzazioni a cui appartiene l'account produttore e che Account AWS fanno parte della stessa organizzazione vengono visualizzate automaticamente. Altrimenti, inserisci manualmente l'ID dell'account o l'ID dell'organizzazione.

4. Per i tag LF o le risorse del catalogo, scegliete. `Named data catalog resources`
5. In Database, scegliete il database.
`lakeformation_tutorial_cross_account_database_named_resource`
6. Scegli Aggiungi tag LF.
7. In Tabelle, scegli Tutte le tabelle.

8. Per le autorizzazioni per le colonne della tabella, scegli Seleziona e Descrivi in Autorizzazioni per la tabella.
9. Seleziona Seleziona e descrivi, in Autorizzazioni concedibili.
10. Facoltativamente, per le autorizzazioni relative ai dati, scegli Accesso semplice basato su colonne se è richiesta la gestione delle autorizzazioni a livello di colonna.
11. Scegli Concessione.

Se non hai revocato l'autorizzazione per **IAMAllowedPrincipals**, viene visualizzato l'errore Concedi autorizzazioni non riuscite. A questo punto, dovresti vedere la tabella di destinazione condivisa AWS RAM con l'account consumatore in Autorizzazioni, Autorizzazioni dati.

Accetta una condivisione di risorse da AWS RAM

 Note

Questo passaggio è necessario solo per la condivisione Account AWS basata, non per la condivisione basata sull'organizzazione.

1. Accedi alla AWS console <https://console.aws.amazon.com/connect/> utilizzando l'amministratore del data lake dell'account consumer utilizzando il nome utente IAM (l'impostazione predefinita è DatalakeAdminConsumer) e la password specificati durante la creazione CloudFormation dello stack.
2. Sulla AWS RAM console, nel pannello di navigazione, sotto Shared with me, Resource shares, scegli la risorsa condivisa Lake Formation. Lo stato dovrebbe essere In sospeso.
3. Scegli Azione e Concedi.
4. Conferma i dettagli della risorsa e scegli Accetta condivisione delle risorse.

A questo punto, l'amministratore del data lake dell'account consumer dovrebbe essere in grado di trovare la risorsa condivisa sulla console Lake Formation (<https://console.aws.amazon.com/lakeformation/>) in Data Catalog, Databases.

Crea un link alla risorsa per la tabella condivisa

- Segui le istruzioni riportate nel [Fase 3: Implementare la condivisione tra account utilizzando il metodo di controllo degli accessi basato su tag](#) passaggio 6 per

creare un link a una risorsa per una tabella condivisa. Assegna un nome al link alla risorsa `amazon_reviews_table_named_resource_resource_link`. Crea il link alla risorsa nel database `lakeformation_tutorial_cross_account_database_consumer`.

Concedi al consumatore l'autorizzazione ai dati per la tabella condivisa

Per concedere al consumatore l'autorizzazione ai dati per la tabella condivisa, completa i seguenti passaggi:

1. Nella console di Lake Formation console (<https://console.aws.amazon.com/lakeformation/>), in Autorizzazioni, Autorizzazioni Data lake, scegli Concedi.
2. Per Principal, scegli utenti e ruoli IAM e scegli l'utente. `DataAnalyst`
3. Per i tag LF o le risorse del catalogo, scegli Risorse del catalogo dati denominato.
4. In Database, scegliete il database.
`lakeformation_tutorial_cross_account_database_named_resource` Se non vedi il database nell'elenco a discesa, scegli Carica altro.
5. In Tabelle, scegli la tabella. `amazon_reviews_table_named_resource`
6. Per le autorizzazioni per tabelle e colonne, seleziona Seleziona e descrivi in Autorizzazioni per le tabelle.
7. Scegli Concessione.

Concedi al consumatore l'autorizzazione ai dati per il collegamento alla risorsa

Oltre a concedere all'utente del data lake l'autorizzazione ad accedere alla tabella condivisa, è necessario concedere all'utente del data lake l'autorizzazione ad accedere al collegamento alla risorsa.

1. Nella console Lake Formation (<https://console.aws.amazon.com/lakeformation/>), in Autorizzazioni, Autorizzazioni Data lake, scegli Concedi.
2. Per Principal, scegli utenti e ruoli IAM e scegli l'utente. `DataAnalyst`
3. Per i tag LF o le risorse del catalogo, scegli Risorse del catalogo dati denominato.
4. In Database, scegliete il database.
`lakeformation_tutorial_cross_account_database_consumer` Se non vedi il database nell'elenco a discesa, scegli Carica altro.
5. In Tabelle, scegli la tabella. `amazon_reviews_table_named_resource_resource_link`

6. Per le autorizzazioni relative ai collegamenti alle risorse, seleziona Descrivi in Autorizzazioni per i collegamenti alle risorse.
7. Scegli Concessione.

A questo punto, l'utente analista di dati nell'account consumatore dovrebbe essere in grado di trovare il database e il collegamento alla risorsa e di interrogare la tabella condivisa tramite la console Athena.

In caso contrario, verifica se quanto segue è configurato correttamente:

- Il link alla risorsa viene creato per la tabella condivisa
- Hai concesso all'utente l'accesso alla tabella condivisa dall'account produttore
- Hai concesso all'utente l'accesso al collegamento alla risorsa e al database per i quali è stato creato il collegamento alla risorsa

Passaggio 5: Pulisci AWS le risorse

Per evitare addebiti indesiderati Account AWS, puoi eliminare le AWS risorse che hai utilizzato per questo tutorial.

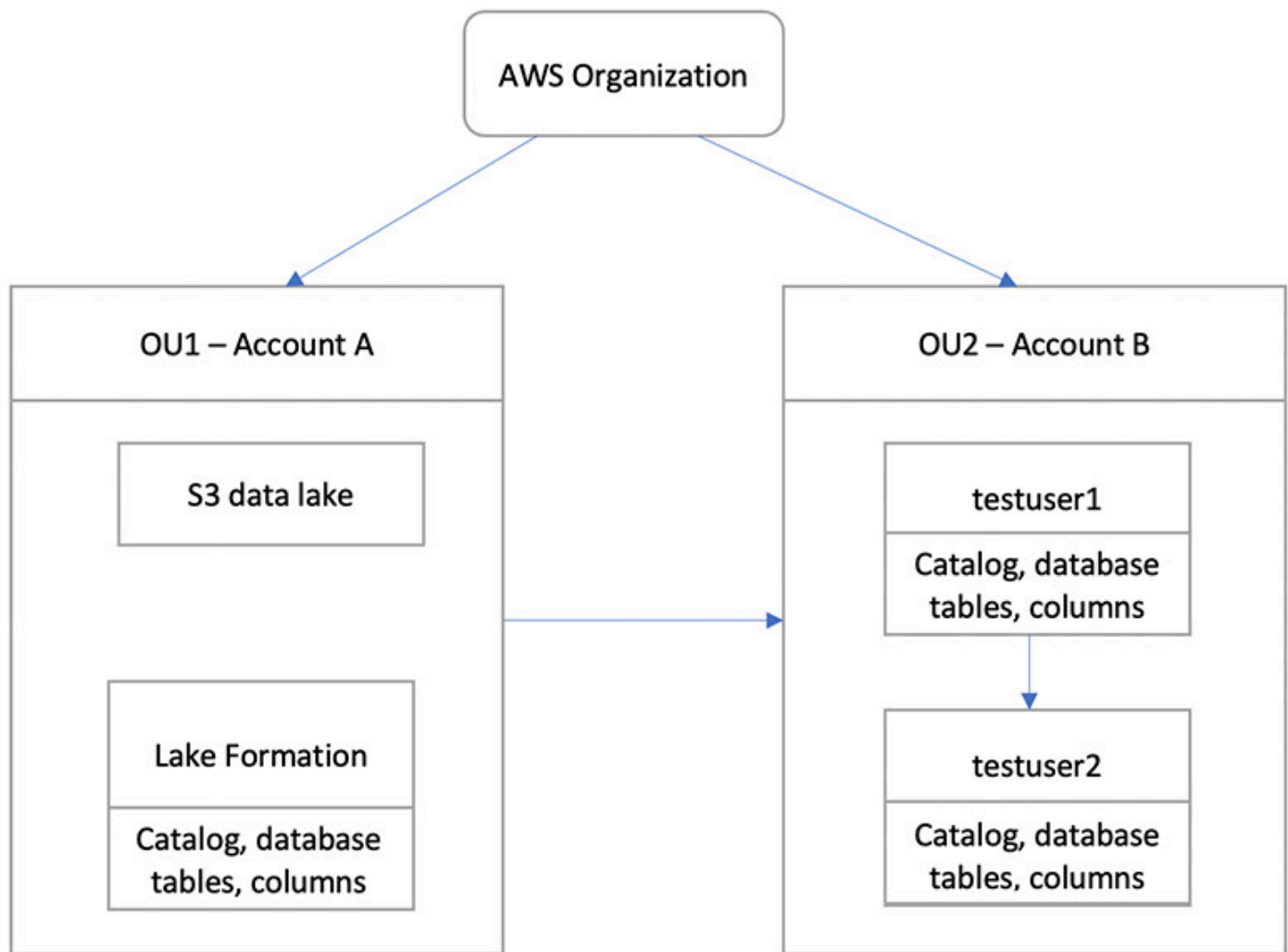
1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> utilizzando l'account produttore ed elimina o modifica quanto segue:
 - AWS Resource Access Manager condivisione di risorse
 - Etichette di Lake Formation
 - CloudFormation pila
 - Impostazioni Lake Formation
 - AWS Glue Data Catalog
2. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> utilizzando l'account utente ed elimina o modifica quanto segue:
 - Etichette di Lake Formation
 - CloudFormation pila

Condivisione di un data lake utilizzando il controllo granulare degli accessi di Lake Formation

Questo tutorial fornisce step-by-step istruzioni su come condividere in modo rapido e semplice set di dati utilizzando Lake Formation quando si gestiscono più set Account AWS con AWS Organizations. Definisci autorizzazioni granulari per controllare l'accesso ai dati sensibili.

Le seguenti procedure mostrano anche come un amministratore di data lake dell'Account A può fornire un accesso granulare all'Account B e come un utente dell'Account B, che funge da amministratore dei dati, può concedere un accesso granulare alla tabella condivisa ad altri utenti del proprio account. I data steward di ciascun account possono delegare in modo indipendente l'accesso ai propri utenti, dando autonomia a ciascun team o linea di business (LOB).

Il caso d'uso presuppone che tu stia utilizzando AWS Organizations per gestire il tuo Account AWS. L'utente dell'Account A in un'unità organizzativa (OU1) concede l'accesso agli utenti dell'Account B in OU2. Puoi usare lo stesso approccio quando non usi Organizations, ad esempio quando hai solo pochi account. Il diagramma seguente illustra il controllo granulare degli accessi dei set di dati in un data lake. Il data lake è disponibile nell'Account A. L'amministratore del data lake dell'Account A fornisce un accesso granulare all'Account B. Il diagramma mostra anche che un utente dell'Account B fornisce l'accesso a livello di colonna alla tabella del data lake Account A a un altro utente dell'Account B.



Argomenti

- [Destinatari principali](#)
- [Prerequisiti](#)
- [Fase 1: Fornire un accesso dettagliato a un altro account](#)
- [Passaggio 2: fornire un accesso dettagliato a un utente nello stesso account](#)

Destinatari principali

Questo tutorial è destinato a amministratori di dati, ingegneri di dati e analisti di dati. La tabella seguente elenca i ruoli utilizzati in questo tutorial:

Ruolo	Descrizione
Amministratore IAM	Utente che dispone della politica AWS gestita: <code>AdministratorAccess</code> .
Amministratore del data lake	Utente che dispone della policy AWS gestita: <code>AWSLakeFormationDataAdmin</code> associato al ruolo.
Analista dei dati	Utente che dispone della politica AWS gestita: <code>AmazonAthenaFullAccess</code> allegato.

Prerequisiti

Prima di iniziare questo tutorial, è necessario disporre di un Account AWS file da utilizzare per accedere come utente amministrativo con le autorizzazioni corrette. Per ulteriori informazioni, consulta [Completa le attività AWS di configurazione iniziali](#).

Il tutorial presuppone che tu abbia familiarità con IAM. Per informazioni su IAM, consulta la [IAM User Guide](#).

Per questo tutorial sono necessarie le seguenti risorse:

- Due unità organizzative:
 - OU1 — Contiene l'account A
 - OU2 — Contiene l'account B
- Una posizione del data lake Amazon S3 (bucket) nell'account A.
- Un utente amministratore del data lake nell'Account A. È possibile creare un amministratore del data lake utilizzando la console Lake Formation (<https://console.aws.amazon.com/lakeformation/>) o il `PutDataLakeSettings` funzionamento dell'API Lake Formation.
- Lake Formation configurata nell'Account A e la posizione del data lake Amazon S3 registrata con Lake Formation nell'Account A.
- Due utenti nell'Account B con le seguenti politiche gestite da IAM:
 - `testuser1`: ha le politiche AWS `AWSLakeFormationDataAdmin` gestite allegate.
 - `testuser2` — Ha la policy gestita allegata AWS . `AmazonAthenaFullAccess`
- Un database `testdb` nel database Lake Formation per l'account B.

Fase 1: Fornire un accesso dettagliato a un altro account

Scopri in che modo un amministratore di data lake dell'Account A fornisce un accesso granulare all'Account B.

Concedi l'accesso dettagliato a un altro account

1. Accedi Console di gestione AWS all'Account <https://console.aws.amazon.com/connect/> A come amministratore del data lake.
2. Apri la console Lake Formation (<https://console.aws.amazon.com/lakeformation/>) e scegli Inizia.
3. nel pannello di navigazione, scegli Database.
4. Scegliere Create database (Crea database).
5. Nella sezione Dettagli del database, seleziona Database.
6. Per Nome, inserisci un nome (per questo tutorial, lo usiamo `sampledb01`).
7. Assicurati che l'opzione Usa solo il controllo di accesso IAM per le nuove tabelle in questo database non sia selezionata. Se questa opzione non è selezionata, possiamo controllare l'accesso da Lake Formation.
8. Scegliere Crea database.
9. Nella pagina Database, scegli il tuo database `sampledb01`.
10. Nel menu Azioni, scegli Concedi.
11. Nella sezione Concedi autorizzazioni, seleziona Account esterno.
12. Per Account AWS ID o ID AWS dell'organizzazione, inserisci l'ID dell'account per l'account B. OU2
13. Per Tabella, scegli la tabella a cui desideri che l'Account B abbia accesso (per questo post, usiamo la tabella `acc_a_area`). Facoltativamente, puoi concedere l'accesso alle colonne all'interno della tabella, cosa che facciamo in questo post.
14. Per Includi colonne, scegli le colonne a cui desideri che l'Account B abbia accesso (per questo post, concediamo le autorizzazioni per tipo, nome e identificatori).
15. Per Colonne, scegli Includi colonne.
16. Per le autorizzazioni relative alle tabelle, seleziona Seleziona.
17. Per le autorizzazioni concesse, seleziona Seleziona. Le autorizzazioni concesse sono necessarie per consentire agli utenti amministratori dell'Account B di concedere le autorizzazioni ad altri utenti dell'Account B.
18. Scegli Concessione.

19. Nel pannello di navigazione, seleziona Tabelle.
20. È possibile visualizzare una connessione attiva nella sezione Account AWS e AWS organizzazioni con accesso.

Crea un link alla risorsa

I servizi integrati come Amazon Athena non possono accedere direttamente a database o tabelle tra account. Pertanto, è necessario creare un collegamento alle risorse in modo che Athena possa accedere ai collegamenti alle risorse del proprio account ai database e alle tabelle di altri account. Crea un link di risorsa alla tabella (`acc_a_area`) in modo che gli utenti dell'Account B possano interrogarne i dati con Athena.

1. Accedi alla AWS console all'indirizzo <https://console.aws.amazon.com/connect>/Account B `cometestuser1`.
2. Nella console Lake Formation (<https://console.aws.amazon.com/lakeformation/>), nel pannello di navigazione, scegli Tabelle. Dovresti vedere le tabelle a cui l'Account A ha fornito l'accesso.
3. Seleziona la tabella `acc_a_area`.
4. Nel menu Azioni, scegli Crea collegamento alla risorsa.
5. Per il nome del link alla risorsa, inserisci un nome (per questo tutorial, `acc_a_area_r1`).
6. Per Database, scegli il tuo database (`testdb`).
7. Scegli Create (Crea) .
8. Nel pannello di navigazione, seleziona Tabelle.
9. Seleziona la tabella `acc_b_area_r1`.
10. Nel menu Azioni, scegli Visualizza dati.

Verrai reindirizzato alla console Athena, dove dovresti vedere il database e la tabella.

Ora puoi eseguire una query sulla tabella per vedere il valore della colonna per cui è stato fornito l'accesso a `testuser1` dall'account B.

Passaggio 2: fornire un accesso dettagliato a un utente nello stesso account

Questa sezione mostra come un utente dell'Account B (`testuser1`), che funge da amministratore dei dati, fornisce un accesso dettagliato a un altro utente nello stesso account (`testuser2`) al nome della colonna nella tabella condivisa. `aac_b_area_r1`

Concedi un accesso granulare a un utente nello stesso account

1. Accedi alla AWS console all'indirizzo Account <https://console.aws.amazon.com/connect/> B come `testuser1`.
2. Nella console Lake Formation, nel riquadro di navigazione, scegli Tabelle.

Puoi concedere le autorizzazioni su una tabella tramite il relativo link alla risorsa. A tale scopo, nella pagina Tabelle, seleziona il link alla risorsa `acc_b_area_r1`, nel menu Azioni, scegli Concedi sulla destinazione.

3. Nella sezione Concedi autorizzazioni, seleziona Il mio account.
4. Per gli utenti e i ruoli IAM, scegli l'utente `testuser2`.
5. Per Colonna, scegli il nome della colonna.
6. Per le autorizzazioni relative alla tabella, seleziona Seleziona.
7. Scegli Concessione.

Quando crei un link a una risorsa, solo tu puoi visualizzarlo e accedervi. Per consentire agli altri utenti del tuo account di accedere al link alla risorsa, devi concedere le autorizzazioni sul link alla risorsa stesso. È necessario concedere le autorizzazioni `DESCRIBE` o `DROP`. Nella pagina Tabelle, seleziona nuovamente la tua tabella e nel menu Azioni, scegli Concedi.

8. Nella sezione Concedi autorizzazioni, seleziona Il mio account.
9. Per gli utenti e i ruoli IAM, seleziona l'utente `testuser2`.
10. Per le autorizzazioni del collegamento alle risorse, seleziona Descrivi.
11. Scegli Concessione.
12. Accedi alla AWS console nell'account B come `testuser2`.

Sulla console Athena (<https://console.aws.amazon.com/athena/>), dovresti vedere il database e la tabella `acc_b_area_r1`. Ora puoi eseguire una query sulla tabella per vedere il valore della colonna a cui `testuser2` hai accesso.

Accedere ai permessi di Lake Formation

AWS Lake Formation utilizza AWS Glue Data Catalog (Data Catalog) per archiviare i metadati per i data lake Amazon S3 e le fonti di dati esterne come Amazon Redshift sotto forma di cataloghi, database e tabelle. I metadati nel Data Catalog sono organizzati in una gerarchia di dati a tre livelli che comprende cataloghi, database e tabelle. Organizza i dati provenienti da varie fonti in contenitori logici chiamati cataloghi. I database sono raccolte di tabelle. Il Data Catalog contiene anche collegamenti a risorse, che sono collegamenti a database e tabelle condivisi in account esterni e vengono utilizzati per l'accesso tra account diversi ai dati nel data lake. Ogni AWS account dispone di un catalogo dati per AWS regione.

Lake Formation fornisce un modello di autorizzazioni del sistema di gestione dei database relazionali (RDBMS) per concedere o revocare l'accesso a cataloghi, database, tabelle e colonne nel Data Catalog con i dati sottostanti in Amazon S3.

Prima di conoscere i dettagli del modello di autorizzazioni di Lake Formation, è utile esaminare le seguenti informazioni di base:

- I data lake gestiti da Lake Formation risiedono in luoghi designati in Amazon Simple Storage Service (Amazon S3). Il Data Catalog contiene anche oggetti di catalogo. Ogni catalogo rappresenta dati provenienti da fonti come data warehouse Amazon Redshift, Amazon DynamoDB database e fonti di dati di terze parti come Snowflake, MySQL e oltre 30 fonti di dati esterne, integrate tramite connettori federati.
- Lake Formation gestisce un Data Catalog che contiene metadati sui dati di origine da importare nei tuoi data lake, come i dati nei log e nei database relazionali, e sui dati nei tuoi data lake in Amazon S3. Il catalogo dati contiene anche metadati sui dati provenienti da fonti di dati esterne diverse da Amazon S3. I metadati sono organizzati in cataloghi, database e tabelle. Le tabelle di metadati contengono schema, posizione, partizionamento e altre informazioni sui dati che rappresentano. I database di metadati sono raccolte di tabelle.
- Il Lake Formation Data Catalog è lo stesso Data Catalog utilizzato da AWS Glue. Puoi utilizzare AWS Glue i crawler per creare tabelle del Data Catalog e puoi utilizzare i processi di AWS Glue estrazione, trasformazione e caricamento (ETL) per popolare i dati sottostanti nei tuoi data lake.
- I cataloghi, i database e le tabelle del Data Catalog sono denominati risorse del Data Catalog. Le tabelle nel catalogo dati vengono chiamate tabelle di metadati per distinguerle dalle tabelle nelle fonti di dati o dai dati tabulari in Amazon S3. I dati a cui puntano le tabelle di metadati in Amazon S3 o nelle fonti di dati vengono definiti dati sottostanti.

- Un principale è un utente o un ruolo, un utente o gruppo Amazon Quick Suite, un utente o un gruppo che si autentica con Lake Formation tramite un provider SAML o, per il controllo degli accessi tra account, ID AWS account, ID organizzazione o ID unità organizzativa.
- AWS Gluei crawler creano tabelle di metadati, ma puoi anche creare manualmente tabelle di metadati con la console Lake Formation, l'API o (). AWS Command Line Interface AWS CLI Quando si crea una tabella di metadati, è necessario specificare una posizione. Quando si crea un database, la posizione è facoltativa. Le posizioni delle tabelle possono essere posizioni Amazon S3 o posizioni di origini dati come un database Amazon Relational Database Service (Amazon RDS). Le posizioni dei database sono sempre sedi Amazon S3.
- I servizi che si integrano con Lake Formation, come Amazon Athena e Amazon Redshift, possono accedere al Data Catalog per ottenere metadati e verificare l'autorizzazione per l'esecuzione di query. Per un elenco completo dei servizi integrati, consulta. [AWS integrazioni di servizi con Lake Formation](#)

Argomenti

- [Panoramica delle autorizzazioni di Lake Formation](#)
- [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#)
- [Modifica delle impostazioni predefinite per il data lake](#)
- [Autorizzazioni implicite di Lake Formation](#)
- [Riferimento alle autorizzazioni di Lake Formation](#)
- [Integrazione di IAM Identity Center](#)
- [Aggiungere una posizione Amazon S3 al tuo data lake](#)
- [Modalità di accesso ibrida](#)
- [Creazione di oggetti in AWS Glue Data Catalog](#)
- [Importazione di dati utilizzando i flussi di lavoro in Lake Formation](#)

Panoramica delle autorizzazioni di Lake Formation

Esistono due tipi principali di autorizzazioni in AWS Lake Formation:

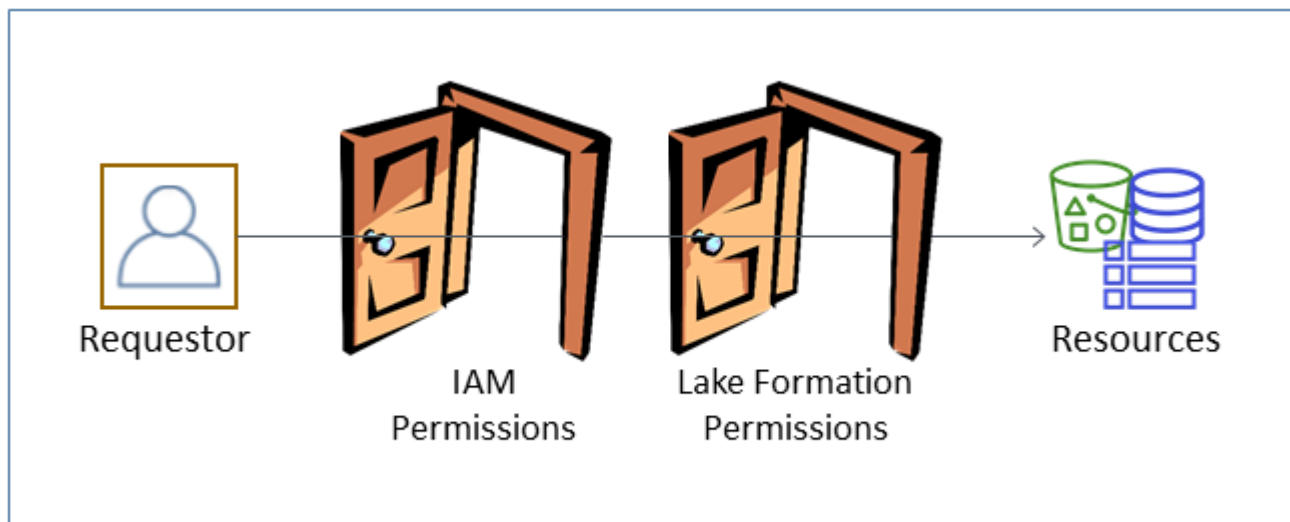
- Accesso ai metadati: autorizzazioni per le risorse del catalogo dati (autorizzazioni del catalogo dati).

Queste autorizzazioni consentono ai responsabili di creare, leggere, aggiornare ed eliminare database e tabelle di metadati nel Catalogo dati.

- Accesso ai dati di base: autorizzazioni su sedi in Amazon Simple Storage Service (Amazon S3) (autorizzazioni di accesso ai dati e autorizzazioni di localizzazione dei dati).
- Le autorizzazioni Data Lake consentono ai responsabili di leggere e scrivere dati nelle posizioni Amazon S3 sottostanti, dati a cui fanno riferimento le risorse del Data Catalog.
- Le autorizzazioni di localizzazione dei dati consentono ai responsabili di creare e modificare database e tabelle di metadati che puntano a posizioni Amazon S3 specifiche.

Per entrambe le aree, Lake Formation utilizza una combinazione di autorizzazioni Lake Formation e autorizzazioni AWS Identity and Access Management (IAM). Il modello di autorizzazioni IAM è costituito da politiche IAM. Il modello di autorizzazioni di Lake Formation è implementato come GRANT/REVOKE comandi in stile DBMS, ad esempio. `Grant SELECT on tableName to userName`

Quando un principale effettua una richiesta di accesso alle risorse del Data Catalog o ai dati sottostanti, affinché la richiesta abbia esito positivo, deve superare i controlli di autorizzazione sia di IAM che di Lake Formation.



Le autorizzazioni di Lake Formation controllano l'accesso alle risorse di Data Catalog, alle sedi Amazon S3 e ai dati sottostanti in tali sedi. Le autorizzazioni IAM controllano l'accesso a Lake Formation AWS Glue APIs e alle risorse. Quindi, anche se potresti avere l'autorizzazione Lake Formation per creare una tabella di metadati nel Data Catalog (`CREATE_TABLE`),

l'operazione fallisce se non disponi dell'autorizzazione IAM sull'`glue:CreateTableAPI`. (Perché un'`glue:autorizzazione`? Perché Lake Formation utilizza il AWS Glue Data Catalog.)

Note

Le autorizzazioni di Lake Formation si applicano solo nella regione in cui sono state concesse.

AWS Lake Formation richiede che ogni principale (utente o ruolo) sia autorizzato a eseguire azioni sulle risorse gestite da Lake Formation. A un principale vengono concesse le autorizzazioni necessarie dall'amministratore del data lake o da un altro principale con le autorizzazioni per concedere le autorizzazioni di Lake Formation.

Quando concedi l'autorizzazione di Lake Formation a un preside, puoi facoltativamente concedere la possibilità di passare tale autorizzazione a un altro preside.

Puoi utilizzare l'API Lake Formation, il AWS Command Line Interface (AWS CLI) o le pagine Data permissions e Data locations della console Lake Formation per concedere e revocare le autorizzazioni Lake Formation.

Metodi per il controllo granulare degli accessi

Con un data lake, l'obiettivo è avere un controllo granulare degli accessi ai dati. In Lake Formation, ciò significa un controllo granulare degli accessi alle risorse del Data Catalog e alle sedi Amazon S3. Puoi ottenere un controllo granulare degli accessi con uno dei seguenti metodi.

Metodo	Autorizzazioni Lake Formation	Autorizzazioni IAM	Commenti
Metodo 1	Aperta	A grana fine	Questo è il metodo predefinito per la compatibilità con le versioni precedenti con. AWS Glue • Aperto significa che l'autorizzazione speciale Super viene concessa al gruppo <code>IAMAllowedPrincipals</code>, dove <code>IAMAllowedPrincipals</code>

Metodo	Autorizzazioni Lake Formation	Autorizzazioni IAM	Commenti
			Is viene creata automaticamente e include tutti gli utenti e i ruoli IAM a cui è consentito l'accesso alle risorse del tuo Data Catalog dalle tue politiche IAM, e l'Superautorizzazione consente a un principale di eseguire ogni operazione Lake Formation supportata sul database o sulla tabella su cui è concessa. Ciò fa sì che l'accesso alle risorse del Data Catalog e alle sedi Amazon S3 sia controllato esclusivamente dalle policy IAM. Per ulteriori informazioni, consultare e Modifica delle impostazioni predefinite per il data lake e Aggiornamento delle AWS Glue autorizzazioni relative ai dati al modello AWS Lake Formation. • Con criteri granulari si intende che le policy IAM controllano tutti gli accessi alle risorse del Data Catalog e ai singoli bucket Amazon S3. Sulla console Lake Formation, questo metodo ha il nome Usa solo il controllo degli accessi IAM.

Metodo	Autorizzazioni Lake Formation	Autorizzazioni IAM	Commenti
Metodo 2	A grana fine	A grana grossa	Questo è il metodo consigliato. • L'accesso granulare significa concedere autorizzazioni limitate di Lake Formation a singoli responsabili sulle risorse Data Catalog, sulle sedi Amazon S3 e sui dati sottostanti in tali sedi. • Graduale significa autorizzazioni più ampie per le singole operazioni e per l'accesso alle sedi Amazon S3. Ad esempio, una policy IAM a grana grossolana potrebbe includere <code>"glue:*"</code> o <code>"glue:Create*"</code> piuttosto <code>"glue:CreateTables"</code> che lasciare, le autorizzazioni di Lake Formation per controllare se un principal e può creare o meno oggetti di catalogo. Significa anche dare ai direttori l'accesso a ciò di APIs cui hanno bisogno per svolgere il loro lavoro, bloccando però altre risorse. APIs Ad esempio, potresti creare una policy IAM che consenta a un responsabile di creare risorse del Data Catalog e creare ed eseguire flussi di lavoro, ma non consenta la creazione di AWS Glue connessioni o funzioni definite dall'utente. Vedi gli esempi più avanti in questa sezione.

Important

Ricorda quanto segue:

- Per impostazione predefinita, Lake Formation ha le impostazioni di controllo degli accessi Use only IAM abilitate per la compatibilità con il comportamento esistente del AWS Glue Data Catalog. Ti consigliamo di disabilitare queste impostazioni dopo la transizione all'utilizzo delle autorizzazioni di Lake Formation. Per ulteriori informazioni, consulta [Modifica delle impostazioni predefinite per il data lake](#).
- Gli amministratori di Data Lake e i creatori di database dispongono di autorizzazioni implicite di Lake Formation che devi comprendere. Per ulteriori informazioni, consulta [Autorizzazioni implicite di Lake Formation](#).

Controllo dell'accesso ai metadati

Per il controllo degli accessi per le risorse del Data Catalog, la discussione seguente presuppone un controllo degli accessi a grana fine con le autorizzazioni di Lake Formation e un controllo degli accessi a grana grossa con le policy IAM.

Esistono due metodi distinti per concedere le autorizzazioni di Lake Formation sulle risorse del Data Catalog:

- Controllo dell'accesso alle risorse con nome: con questo metodo, si concedono le autorizzazioni su database o tabelle specifici specificando i nomi dei database o delle tabelle. Le sovvenzioni hanno questa forma:

Concedi le autorizzazioni ai responsabili sulle risorse [con opzione di concessione].

Con l'opzione di concessione, puoi consentire al beneficiario di concedere le autorizzazioni ad altri mandanti.

- Controllo dell'accesso basato su tag: con questo metodo, si assegnano uno o più tag LF ai database, alle tabelle e alle colonne del Data Catalog e si concedono le autorizzazioni su uno o più tag LF ai principali. Ogni tag LF è una coppia chiave-valore, ad esempio. department=sales Un principale con tag LF che corrispondono ai tag LF su una risorsa del catalogo dati può accedere a tale risorsa. Questo metodo è consigliato per i data lake con un gran numero di database e tabelle. È spiegato in dettaglio in [Controllo degli accessi basato su tag Lake Formation](#).

Le autorizzazioni che un principale ha su una risorsa sono l'unione delle autorizzazioni concesse da entrambi i metodi.

La tabella seguente riassume le autorizzazioni di Lake Formation disponibili sulle risorse di Data Catalog. Le intestazioni delle colonne indicano la risorsa a cui è concessa l'autorizzazione.

Catalogo	Database	Tabella
CREATE_DATABASE	CREATE_TABLE	ALTER
	ALTER	DROP
	DROP	DESCRIBE
	DESCRIBE	SELECT*
		INSERT*
		DELETE*

Ad esempio, l'`CREATE_TABLE` autorizzazione viene concessa a un database. Ciò significa che il principale può creare tabelle in quel database.

Le autorizzazioni con un asterisco (*) sono concesse alle risorse del Data Catalog, ma si applicano ai dati sottostanti. Ad esempio, l'`DROP` autorizzazione per una tabella di metadati consente di eliminare la tabella dal Data Catalog. Tuttavia, l'`DELETE` autorizzazione concessa sulla stessa tabella consente di eliminare i dati sottostanti della tabella in Amazon S3, utilizzando, ad esempio, un'istruzione `SQLDELETE`. Con queste autorizzazioni, puoi anche visualizzare la tabella sulla console Lake Formation e recuperare informazioni sulla tabella con l'`AWS GlueAPI`. Pertanto `SELECT`, `INSERT`, e `DELETE` sono sia le autorizzazioni di Data Catalog che le autorizzazioni di accesso ai dati.

Quando si concede la `SELECT` licenza su una tabella, è possibile aggiungere un filtro che includa o escluda una o più colonne. Ciò consente un controllo granulare degli accessi alle colonne della tabella di metadati, limitando le colonne che gli utenti dei servizi integrati possono vedere durante l'esecuzione delle query. Questa funzionalità non è disponibile utilizzando solo le policy IAM.

Esiste anche un'autorizzazione speciale denominata `Super`. L'`Super` autorizzazione consente a un principale di eseguire tutte le operazioni di Lake Formation supportate sul database o sulla tabella su cui è concessa. Questa autorizzazione può coesistere con le altre autorizzazioni di Lake Formation. Ad esempio, puoi concedere `Super` e `INSERT` su una tabella di metadati. `SELECT` Il principale può eseguire tutte le azioni supportate sulla tabella e, in caso di revoca, le autorizzazioni `SELECT` e `INSERT` rimangono `Super` invariate.

Per i dettagli su ciascuna autorizzazione, vedere. [Riferimento alle autorizzazioni di Lake Formation](#)

Important

Per poter visualizzare una tabella Data Catalog creata da un altro utente, è necessario disporre di almeno un'autorizzazione Lake Formation sulla tabella. Se ti viene concessa almeno un'autorizzazione sulla tabella, puoi anche vedere il database che contiene la tabella.

Puoi concedere o revocare le autorizzazioni di Data Catalog utilizzando la console Lake Formation, l'API o (). AWS Command Line Interface AWS CLI Di seguito è riportato un esempio di AWS CLI comando che concede all'utente il `datalake_user1` permesso di creare tabelle nel database. `retail`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"} }'
```

Di seguito è riportato un esempio di policy IAM di controllo degli accessi a grana grossa che integra il controllo degli accessi a grana fine con le autorizzazioni di Lake Formation. Consente tutte le operazioni su qualsiasi database o tabella di metadati.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:*Database*",
        "glue:*Table*",
        "glue:*Partition*"
      ],
      "Resource": "*"
    }
  ]
}
```

L'esempio successivo è anch'esso a grana grossa ma leggermente più restrittivo. Consente operazioni di sola lettura su tutti i database e le tabelle di metadati nel Catalogo dati nell'account e nella regione designati.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetTable",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ],
      "Resource": "arn:aws:glue:us-east-1:111122223333:*"
    }
  ]
}
```

Confronta queste politiche con la seguente politica, che implementa il controllo granulare degli accessi basato su IAM. Concede le autorizzazioni solo su un sottoinsieme di tabelle nel database di metadati di gestione delle relazioni con i clienti (CRM) nell'account e nella regione designati.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTables",
        "glue:SearchTables",
        "glue:GetTable",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ]
    }
  ]
}
```

```
    ],
    "Resource": [
        "arn:aws:glue:us-east-1:111122223333:catalog",
        "arn:aws:glue:us-east-1:111122223333:database/CRM",
        "arn:aws:glue:us-east-1:111122223333:table/CRM/P*"
    ]
}
]
```

Per altri esempi di politiche di controllo degli accessi a grana grossolana, vedere. [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#)

Controllo sottostante dell'accesso ai dati

Quando un AWS servizio integrato richiede l'accesso ai dati in una posizione Amazon S3 con accesso controllato da Lake AWS Lake Formation, Lake Formation fornisce credenziali temporanee per accedere ai dati.

Per consentire a Lake Formation di controllare l'accesso ai dati sottostanti in una sede Amazon S3, devi registrare tale posizione con Lake Formation.

Dopo aver registrato una sede Amazon S3, puoi iniziare a concedere le seguenti autorizzazioni Lake Formation:

- Autorizzazioni di accesso ai dati (SELECT, INSERT, e DELETE) sulle tabelle di Data Catalog che puntano a quella posizione.
- Autorizzazioni per l'ubicazione dei dati in quella posizione.

Le autorizzazioni di localizzazione dei dati di Lake Formation controllano la capacità di creare risorse Data Catalog che puntano a particolari posizioni Amazon S3. Le autorizzazioni di localizzazione dei dati forniscono un ulteriore livello di sicurezza alle posizioni all'interno del data lake. Quando concedi l'autorizzazione CREATE_TABLE o l'autorizzazione a un'entità principale, concedi anche le autorizzazioni per la localizzazione dei dati per limitare le posizioni per le quali l'ente principale può creare o modificare tabelle di metadati.

Le sedi Amazon S3 sono bucket o prefissi all'interno di un bucket, ma non singoli oggetti Amazon S3.

Puoi concedere le autorizzazioni per la localizzazione dei dati a un principale utilizzando la console Lake Formation, l'API o il AWS CLI. La forma generale di una sovvenzione è la seguente:

```
grant DATA_LOCATION_ACCESS to principal on S3 location [with grant option]
```

Se lo includi con `with grant option`, il beneficiario può concedere le autorizzazioni ad altri committenti.

Ricorda che le autorizzazioni di Lake Formation funzionano sempre in combinazione con le autorizzazioni AWS Identity and Access Management (IAM) per un controllo granulare degli accessi. Per read/write le autorizzazioni sui dati Amazon S3 sottostanti, le autorizzazioni IAM vengono concesse come segue:

Quando registri una posizione, specifichi un ruolo IAM che concede autorizzazioni di lettura/scrittura su quella posizione. Lake Formation assume questo ruolo quando fornisce credenziali temporanee ai servizi integrati. AWS Un ruolo tipico potrebbe avere la seguente politica allegata, in cui la posizione registrata è il bucket. `awsexamplebucket`

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket"
      ]
    }
  ]
}
```

```

    ]
  }
}

```

Lake Formation offre un ruolo collegato al servizio che puoi utilizzare durante la registrazione per creare automaticamente politiche come questa. Per ulteriori informazioni, consulta [Utilizzo di ruoli collegati ai servizi per Lake Formation](#).

Pertanto, la registrazione di una sede Amazon S3 concede le autorizzazioni s3: IAM richieste su quella posizione, dove le autorizzazioni sono specificate dal ruolo utilizzato per registrare la posizione.

Important

Evita di registrare un bucket Amazon S3 con Requester pay abilitato. Per i bucket registrati con Lake Formation, il ruolo utilizzato per registrare il bucket viene sempre visualizzato come richiedente. Se al bucket si accede da un altro AWS account, al proprietario del bucket viene addebitato l'accesso ai dati se il ruolo appartiene allo stesso account del proprietario del bucket.

Per read/write accedere ai dati sottostanti, oltre alle autorizzazioni di Lake Formation, i mandanti necessitano anche dell'autorizzazione `lakeformation:GetDataAccess` IAM. Con questa autorizzazione, Lake Formation concede la richiesta di credenziali temporanee per accedere ai dati.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "lakeformation:GetDataAccess",
      "Resource": "*"
    }
  ]
}

```

Nella politica precedente, è necessario impostare il parametro Resource su '*' (all). La specificazione di altre risorse per questa autorizzazione non è supportata. Questa configurazione garantisce che Lake Formation possa gestire l'accesso ai dati nell'intero ambiente di data lake in modo efficiente.

Note

Amazon Athena richiede che l'utente disponga dell'`lakeformation:GetDataAccess` autorizzazione. Altri servizi integrati richiedono che il ruolo di esecuzione sottostante disponga dell'`lakeformation:GetDataAccess` autorizzazione.

Questa autorizzazione è inclusa nelle politiche suggerite in [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

Riassumendo, per consentire ai dirigenti di Lake Formation di leggere e scrivere i dati sottostanti con accesso controllato dai permessi di Lake Formation:

- Registra le sedi Amazon S3 che contengono i dati con Lake Formation.
- I responsabili che creano tabelle Data Catalog che puntano alle posizioni dei dati sottostanti devono disporre delle autorizzazioni per la localizzazione dei dati.
- I responsabili che leggono e scrivono i dati sottostanti devono disporre delle autorizzazioni di accesso ai dati di Lake Formation sulle tabelle del Data Catalog che puntano alle posizioni dei dati sottostanti.
- I responsabili che leggono e scrivono i dati sottostanti devono disporre dell'autorizzazione `lakeformation:GetDataAccess` IAM quando la posizione dei dati sottostanti è registrata con Lake Formation.

Note

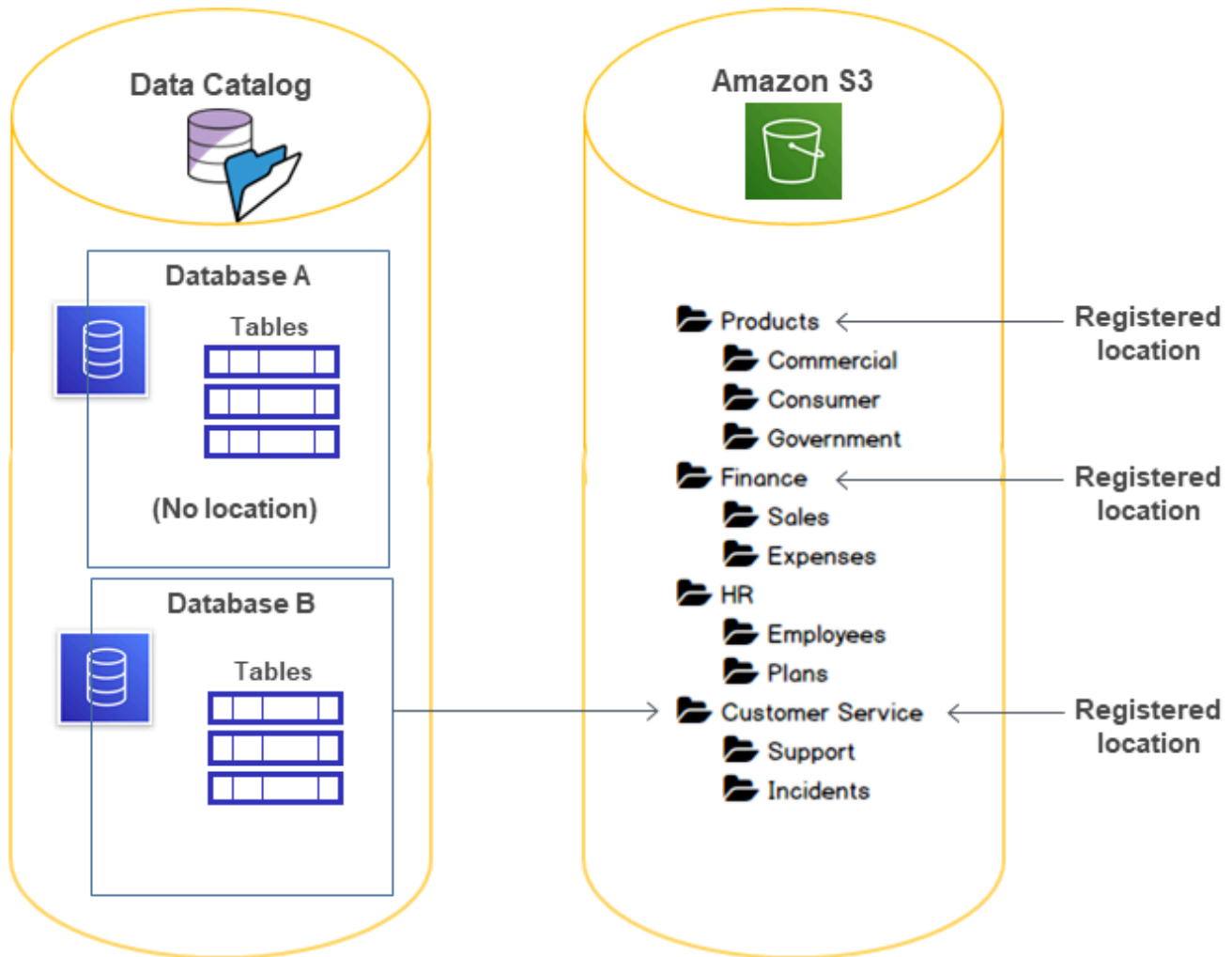
Il modello di autorizzazioni Lake Formation non impedisce l'accesso alle sedi Amazon S3 tramite l'API o la console di Amazon S3 se hai accesso ad esse tramite policy IAM o Amazon S3. Puoi collegare le policy IAM ai principali per bloccare questo accesso.

Ulteriori informazioni sulle autorizzazioni di localizzazione dei dati

Le autorizzazioni di localizzazione dei dati regolano il risultato delle operazioni di creazione e aggiornamento su database e tabelle Data Catalog. Le regole sono le seguenti:

- Un principale deve disporre di autorizzazioni esplicite o implicite per la localizzazione dei dati su una posizione Amazon S3 per creare o aggiornare un database o una tabella che specifichi tale posizione.
- L'autorizzazione esplicita `DATA_LOCATION_ACCESS` viene concessa utilizzando la console, l'API o. AWS CLI
- Le autorizzazioni implicite vengono concesse quando un database ha una proprietà `location` che punta a una posizione registrata, l'entità principale dispone dell'`CREATE_TABLE` autorizzazione sul database e l'entità principale tenta di creare una tabella in quella posizione o in una posizione secondaria.
- Se a un responsabile vengono concesse le autorizzazioni per la localizzazione dei dati su una posizione, il responsabile dispone delle autorizzazioni per la localizzazione dei dati su tutte le sedi secondarie.
- Un responsabile non necessita delle autorizzazioni di localizzazione dei dati per eseguire read/write operazioni sui dati sottostanti. È sufficiente disporre delle `SELECT` o delle autorizzazioni di accesso ai `INSERT` dati. Le autorizzazioni per la localizzazione dei dati si applicano solo alla creazione di risorse del Catalogo Dati che puntano alla posizione.

Considerate lo scenario illustrato nel diagramma seguente.



In questo diagramma:

- I bucket Products Amazon S3 Customer Service sono registrati presso Lake Formation. Finance
- Database Anon ha una proprietà di posizione e Database B dispone di una proprietà di posizione che punta al Customer Service bucket.
- L'utente `datalake_user` ha `CREATE_TABLE` su entrambi i database.
- All'utente sono `datalake_user` state concesse le autorizzazioni per la localizzazione dei dati solo nel Products bucket.

Di seguito sono riportati i risultati ottenuti quando l'utente `datalake_user` tenta di creare una tabella di catalogo in un particolare database in una posizione particolare.

Posizione in cui **datalake_user** tenta di creare una tabella

Database e posizione	Ha successo o fallisce	Motivo
Database A in Finance/Sales	Fallisce	Nessuna autorizzazione per la localizzazione dei dati
Database A in Products	Ha successo	Dispone dell'autorizzazione alla localizzazione dei dati
Database A in HR/Plans	Ha successo	La posizione non è registrata
Database B in Customer Service/Incidents	Ha successo	Il database ha la proprietà di localizzazione in Customer Service

Per ulteriori informazioni, consulta gli argomenti seguenti:

- [Aggiungere una posizione Amazon S3 al tuo data lake](#)
- [Riferimento alle autorizzazioni di Lake Formation](#)
- [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#)

Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM

Questa sezione elenca alcuni personaggi suggeriti di Lake Formation e le relative autorizzazioni suggerite AWS Identity and Access Management (IAM). Per informazioni sulle autorizzazioni di Lake Formation, vedere [the section called “Riferimento alle autorizzazioni di Lake Formation”](#).

AWS Lake Formation persone

La tabella seguente elenca i AWS Lake Formation personaggi suggeriti.

Personaggi di Lake Formation

Utente	Description
Amministratore IAM (superutente)	(Obbligatorio) Utente che può creare utenti e ruoli IAM. Dispone della politica AdministratorAccess AWS gestita. Dispone di tutte le autorizzazioni su tutte le risorse di Lake Formation. Può aggiungere amministratori di data lake. Non può concedere le autorizzazioni di Lake Formation se non è stato designato anche un amministratore del data lake.
Amministratore del data lake	(Obbligatorio) Utente in grado di registrare sedi Amazon S3, accedere al Data Catalog, creare database, creare ed eseguire flussi di lavoro, concedere autorizzazioni Lake Formation ad altri utenti e visualizzare i log. AWS CloudTrail Dispone di meno autorizzazioni IAM rispetto all'amministratore IAM, ma sufficienti per amministrare il data lake. Impossibile aggiungere altri amministratori di data lake.
Amministratore di sola lettura	(Facoltativo) Utente che può visualizzare i principali, le risorse del Data Catalog, le autorizzazioni e AWS CloudTrail i registri, senza le autorizzazioni per effettuare aggiornamenti.
Ingegnere dei dati	(Facoltativo) Utente che può creare database, creare ed eseguire crawler e flussi di lavoro e concedere autorizzazioni Lake Formation sulle tabelle del Data Catalog create dai crawler e dai flussi di lavoro. Ti consigliamo di nominare tutti i data engineer creatori di database. Per ulteriori informazioni, consulta Creazione di un database .
Analista dei dati	(Facoltativo) Utente che può eseguire query sul data lake utilizzando, ad esempio, Amazon Athena Dispone solo delle autorizzazioni sufficienti per eseguire le query.
Ruolo del workflow	(Obbligatorio) Ruolo che esegue un flusso di lavoro per conto di un utente. Questo ruolo viene specificato quando si crea un flusso di lavoro da un blueprint.

 Note

In Lake Formation, gli amministratori di data lake aggiunti dopo la creazione del database possono concedere autorizzazioni ma non dispongono automaticamente di autorizzazioni di accesso ai dati come SELECT o DESCRIBE. Gli amministratori che creano database ricevono le SUPER autorizzazioni per tali database. Questo comportamento è intenzionale: sebbene tutti gli amministratori possano concedersi le autorizzazioni necessarie, tali autorizzazioni non vengono applicate automaticamente alle risorse preesistenti. Pertanto, gli amministratori devono concedersi esplicitamente l'accesso ai database esistenti prima dell'assegnazione dei privilegi di amministratore.

AWS politiche gestite per Lake Formation

Puoi concedere le autorizzazioni AWS Identity and Access Management (IAM) necessarie per lavorare AWS Lake Formation utilizzando policy AWS gestite e policy in linea. Le seguenti politiche AWS gestite sono disponibili per Lake Formation.

AWS politica gestita: AWSLakeFormationDataAdmin

[AWSLakeFormationDataAdmin](#) la politica garantisce l'accesso amministrativo AWS Lake Formation ai servizi correlati, come AWS Glue la gestione dei data lake.

È possibile associare la policy AWSLakeFormationDataAdmin a utenti, gruppi e ruoli.

Dettagli delle autorizzazioni

- **CloudTrail**— Consente ai responsabili di visualizzare i AWS CloudTrail registri. Ciò è necessario per esaminare eventuali errori nella configurazione del data lake.
- **Glue**— Consente ai responsabili di visualizzare, creare e aggiornare tabelle e database di metadati in Data Catalog. Ciò include le operazioni API che iniziano con `Get`, `List`, `Create`, `Update`, `Delete`, e `Search`. Ciò è necessario per gestire i metadati delle tabelle del data lake.
- **IAM**— Consente ai responsabili di recuperare informazioni sugli utenti, i ruoli e le policy IAM associati ai ruoli. Ciò è necessario affinché l'amministratore dei dati riveda ed elenchi gli utenti e i ruoli IAM a cui concedere le autorizzazioni di Lake Formation.
- **Lake Formation**— Concede agli amministratori dei data lake le autorizzazioni necessarie per Lake Formation per gestire i data lake.

- S3— Consente ai responsabili di recuperare informazioni sui bucket Amazon S3 e sulle loro posizioni per configurare la posizione dei dati per i data lake.

```
"Statement": [  
  {  
    "Sid": "AWSLakeFormationDataAdminAllow",  
    "Effect": "Allow",  
    "Action": [  
      "lakeformation:*",  
      "cloudtrail:DescribeTrails",  
      "cloudtrail:LookupEvents",  
      "glue:CreateCatalog",  
      "glue:UpdateCatalog",  
      "glue>DeleteCatalog",  
      "glue:GetCatalog",  
      "glue:GetCatalogs",  
      "glue:GetDatabase",  
      "glue:GetDatabases",  
      "glue>CreateDatabase",  
      "glue:UpdateDatabase",  
      "glue>DeleteDatabase",  
      "glue:GetConnections",  
      "glue:SearchTables",  
      "glue:GetTable",  
      "glue:CreateTable",  
      "glue:UpdateTable",  
      "glue>DeleteTable",  
      "glue:GetTableVersions",  
      "glue:GetPartitions",  
      "glue:GetTables",  
      "glue:ListWorkflows",  
      "glue:BatchGetWorkflows",  
      "glue>DeleteWorkflow",  
      "glue:GetWorkflowRuns",  
      "glue:StartWorkflowRun",  
      "glue:GetWorkflow",  
      "s3:ListBucket",  
      "s3:GetBucketLocation",  
      "s3:ListAllMyBuckets",  
      "s3:GetBucketAcl",  
      "iam:ListUsers",  
      "iam:ListRoles",  
    ]  
  }  
]
```

```

        "iam:GetRole",
        "iam:GetRolePolicy"
    ],
    "Resource": "*"
  },
  {
    "Sid": "AWSLakeFormationDataAdminDeny",
    "Effect": "Deny",
    "Action": [
      "lakeformation:PutDataLakeSettings"
    ],
    "Resource": "*"
  }
]
}

```

Note

La `AWSLakeFormationDataAdmin` policy non concede tutte le autorizzazioni necessarie agli amministratori dei data lake. Sono necessarie autorizzazioni aggiuntive per creare ed eseguire flussi di lavoro e registrare posizioni con il ruolo collegato al servizio. `AWSServiceRoleForLakeFormationDataAccess` Per ulteriori informazioni, consultare [Crea un amministratore del data lake](#) e [Utilizzo di ruoli collegati ai servizi per Lake Formation](#).

AWS politica gestita: `AWSLakeFormationCrossAccountManager`

[AWSLakeFormationCrossAccountManager](#) la politica fornisce l'accesso multiaccount alle AWS Glue risorse tramite Lake Formation e concede l'accesso in lettura ad altri servizi richiesti come AWS Organizations e AWS RAM.

È possibile associare la policy `AWSLakeFormationCrossAccountManager` a utenti, gruppi e ruoli.

Dettagli delle autorizzazioni

Questa policy include le seguenti autorizzazioni:

- **Glue**— Consente ai responsabili di impostare o eliminare la politica delle risorse del Data Catalog per il controllo degli accessi.

- **Organizations**— Consente ai responsabili di recuperare le informazioni sull'account e sull'unità organizzativa (OU) di un'organizzazione.
- **ram:CreateResourceShare**— Consente ai dirigenti di creare una condivisione di risorse.
- **ram:UpdateResourceShare**— Consente ai principali di modificare alcune proprietà della condivisione di risorse specificata.
- **ram>DeleteResourceShare**— Consente ai principali di eliminare la condivisione di risorse specificata.
- **ram:AssociateResourceShare**— Consente ai responsabili di aggiungere l'elenco di principali e l'elenco di risorse specificati a una condivisione di risorse.
- **ram:DisassociateResourceShare**— Consente ai principali di rimuovere i principali o le risorse specificati dalla partecipazione alla condivisione di risorse specificata.
- **ram:GetResourceShares**— Consente ai responsabili di recuperare i dettagli sulle condivisioni di risorse che possiedi o che sono condivise con te.
- **ram:RequestedResourceType**— Consente ai responsabili di recuperare il tipo di risorsa (database, tabella o catalogo).
- **AssociateResourceSharePermission**— Consente ai responsabili di aggiungere o sostituire l' AWS RAM autorizzazione per un tipo di risorsa incluso in una condivisione di risorse. È possibile associare esattamente un'autorizzazione a ciascun tipo di risorsa nella condivisione di risorse.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "AllowCreateResourceShare",
    "Effect": "Allow",
    "Action": [
      "ram:CreateResourceShare"
    ],
    "Resource": "*",
    "Condition": {
      "StringLikeIfExists": {
        "ram:RequestedResourceType": [
          "glue:Table",
          "glue:Database",
          "glue:Catalog"
        ]
      }
    }
  ]
}
```

```

    }
  },
  {
    "Sid": "AllowManageResourceShare",
    "Effect": "Allow",
    "Action": [
      "ram:UpdateResourceShare",
      "ram:DeleteResourceShare",
      "ram:AssociateResourceShare",
      "ram:DisassociateResourceShare",
      "ram:GetResourceShares"
    ],
    "Resource": "*",
    "Condition": {
      "StringLike": {
        "ram:ResourceShareName": [
          "LakeFormation*"
        ]
      }
    }
  },
  {
    "Sid": "AllowManageResourceSharePermissions",
    "Effect": "Allow",
    "Action": [
      "ram:AssociateResourceSharePermission"
    ],
    "Resource": "*",
    "Condition": {
      "ArnLike": {
        "ram:PermissionArn": [
          "arn:aws:ram::aws:permission/AWSRAMLFEnabled*"
        ]
      }
    }
  },
  {
    "Sid": "AllowXAcctManagerPermissions",
    "Effect": "Allow",
    "Action": [
      "glue:PutResourcePolicy",
      "glue:DeleteResourcePolicy",
      "organizations:DescribeOrganization",

```

```

        "organizations:DescribeAccount",
        "ram:Get*",
        "ram:List*"
    ],
    "Resource": "*"
},
{
    "Sid": "AllowOrganizationsPermissions",
    "Effect": "Allow",
    "Action": [
        "organizations:ListRoots",
        "organizations:ListAccountsForParent",
        "organizations:ListOrganizationalUnitsForParent"
    ],
    "Resource": "*"
}
]
}

```

AWS politica gestita: AWSGlue ConsoleFullAccess

[AWSGlueConsoleFullAccess](#) la politica garantisce l'accesso completo alle AWS Glue risorse quando un'identità a cui è associata la politica utilizza il Console di gestione AWS. Se si segue la convenzione di denominazione per le risorse specificate nella policy, gli utenti hanno la piena funzionalità della console. Questo criterio è in genere associato agli utenti della AWS Glue console.

Inoltre, AWS Glue Lake Formation assume il ruolo di servizio `AWSGlueServiceRole` per consentire l'accesso ai servizi correlati, tra cui Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3) e Amazon CloudWatch.

AWS managed policy: LakeFormationDataAccessServiceRolePolicy

Questa policy è associata a un ruolo collegato al servizio denominato `ServiceRoleForLakeFormationDataAccess` che consente al servizio di eseguire azioni sulle risorse su richiesta dell'utente. Non puoi collegare questa policy alle tue identità IAM.

Questa policy consente ai AWS servizi integrati di Lake Formation come Amazon Athena Amazon Redshift di utilizzare il ruolo collegato ai servizi per scoprire le risorse di Amazon S3.

Per ulteriori informazioni, consultare [Utilizzo di ruoli collegati ai servizi per Lake Formation](#).

Dettagli di autorizzazione

Questa policy include la seguente autorizzazione.

- `s3:ListAllMyBuckets`— Restituisce un elenco di tutti i bucket di proprietà del mittente autenticato della richiesta.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationDataAccessServiceRolePolicy",
      "Effect": "Allow",
      "Action": [
        "s3:ListAllMyBuckets"
      ],
      "Resource": [
        "arn:aws:s3:::*"
      ]
    }
  ]
}
```

Lake Formation: aggiornamenti alle politiche AWS gestite

Visualizza i dettagli sugli aggiornamenti alle politiche AWS gestite per Lake Formation da quando questo servizio ha iniziato a tracciare queste modifiche.

Modifica	Descrizione	Data
AWSLakeFormationCrossAccountManager Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationCrossAccountManager politica sostituendo l'operatore delle StringLike condizioni con l'ArnLikeoperatore che consente a	gennaio 2025

Modifica	Descrizione	Data
	IAM di eseguire il controllo del formato ARN.	
AWSLakeFormationDataAdmin Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationDataAdmin politica aggiungendo il seguente AWS Glue Data Catalog CRUD APIs come parte della funzionalità multicatalogo. • colla: CreateCatalog • colla: UpdateCatalog • colla: DeleteCatalog • colla: GetCatalog • colla: GetCatalogs Questa modifica gestita delle politiche serve a garantire che l'amministratore di Lake Formation disponga per impostazione predefinita dell'autorizzazione IAM per queste nuove operazioni.	dicembre 2024
AWSLakeFormationCrossAccountManager Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationCrossAccountManager politica aggiungendo elementi Sid alla dichiarazione politica.	marzo 2024
AWSLakeFormationDataAdmin Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationDataAdmin politica aggiungendo un elemento Sid alla dichiarazione politica e rimuovendo un'azione ridondante.	marzo 2024

Modifica	Descrizione	Data
LakeFormationDataAccessServiceRolePolicy Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la LakeFormationDataAccessServiceRolePolicy politica aggiungendo un elemento Sid alla dichiarazione politica.	febbraio 2024
AWSLakeFormationCrossAccountManager Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationCrossAccountManager politica aggiungendo una nuova autorizzazione per abilitare la condivisione dei dati tra account in modalità di accesso ibrido.	ottobre 2023
AWSLakeFormationCrossAccountManager Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationCrossAccountManager politica per creare una sola condivisione di risorse per account destinatario quando la risorsa viene condivisa per la prima volta. Tutte le risorse condivise successivamente con lo stesso account vengono allegate alla stessa condivisione di risorse.	6 maggio 2022
Lake Formation ha iniziato a tracciare le modifiche.	Lake Formation ha iniziato a tracciare le modifiche per le sue politiche AWS gestite.	6 maggio 2022

Autorizzazioni suggerite da Personas

Di seguito sono riportate le autorizzazioni suggerite per ogni persona. L'amministratore IAM non è incluso perché quell'utente dispone di tutte le autorizzazioni su tutte le risorse.

Argomenti

- [Autorizzazioni di amministratore di Data Lake](#)
- [Autorizzazioni di amministratore di sola lettura](#)

- [Autorizzazioni del tecnico dei dati](#)
- [Autorizzazioni per analisti di dati](#)
- [Autorizzazioni relative ai ruoli del workflow](#)

Autorizzazioni di amministratore di Data Lake

Important

Nei seguenti criteri, sostituiscili `<account-id>` con un numero di AWS account valido e sostituiscili `<workflow_role>` con il nome di un ruolo che dispone delle autorizzazioni per eseguire un flusso di lavoro, come definito in [Autorizzazioni relative ai ruoli del workflow](#)

Tipo di policy	Policy
AWS politiche gestite	• AWSLakeFormationDataAdmin • LakeFormationDataAccessServiceRolePolicy (politica relativa ai ruoli legati ai servizi) • AWSGlueConsoleFullAccess (facoltativo). • CloudWatchLogsReadOnlyAccess (facoltativo) • AWSLakeFormationCrossAccountManager (facoltativo) • AmazonAthenaFullAccess (facoltativo).

Per informazioni sulle politiche AWS gestite opzionali, vedere [the section called “Crea un amministratore del data lake”](#)

Politica in linea (per la creazione del ruolo collegato ai servizi di Lake Formation)	<pre>{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": "iam:CreateServiceLinkedRole", "Resource": "*", "Condition": {</pre>

Tipo di policy	Policy
	<pre> "StringEquals": { "iam:AWSServiceName": "lakeformation.amazonaws.com" } }, { "Effect": "Allow", "Action": ["iam:PutRolePolicy"], "Resource": "arn:aws:iam:: <account-id> :role/aws-service-role/lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess" }] }</pre>
(Facoltativo) Politica in linea (politica passrole per il ruolo del flusso di lavoro). Questa operazione è necessaria solo se l'amministratore del data lake crea ed esegue flussi di lavoro.	<pre> { "Version": "2012-10-17", "Statement": [{ "Sid": "PassRolePermissions", "Effect": "Allow", "Action": ["iam:PassRole"], "Resource": ["arn:aws:iam:: 111122223333 :role/<workflow_role> "] }] }</pre>

Tipo di policy	Policy
(Facoltativo) Politica in linea (se il tuo account concede o riceve autorizzazioni Lake Formation per più account). Questa politica serve ad accettare o rifiutare gli inviti alla condivisione di AWS RAM risorse e a consentire la concessione di autorizzazioni per più account alle organizzazioni. <code>ram:EnableSharingWithAwsOrganization</code> è richiesto solo per gli amministratori del data lake presenti nell'account di gestione. AWS Organizations	<pre>{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["ram:AcceptResourceShareInvitation", "ram:RejectResourceShareInvitation", "ec2:DescribeAvailabilityZones", "ram:EnableSharingWithAwsOrganization"], "Resource": "*" }] }</pre>

Autorizzazioni di amministratore di sola lettura

Tipo di policy	Policy
Politica in linea (di base)	<pre>{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:GetEffectivePermissionsForPath", "lakeformation:ListPermissions", "lakeformation:ListDataCellsFilter", "lakeformation:GetDataCellsFilter", "lakeformation:SearchDatabasesByLFTags", </pre>

Tipo di policy	Policy
	<pre> "lakeformation:SearchTablesByLFTags" , "lakeformation:GetLFTag", "lakeformation:ListLFTags", "lakeformation:GetResourceLFTags", "lakeformation:ListLakeFormationOpti ns", "cloudtrail:DescribeTrails", "cloudtrail:LookupEvents", "glue:GetDatabase", "glue:GetDatabases", "glue:GetConnections", "glue:SearchTables", "glue:GetTable", "glue:GetTableVersions", "glue:GetPartitions", "glue:GetTables", "glue:GetWorkflow", "glue:ListWorkflows", "glue:BatchGetWorkflows", "glue:GetWorkflowRuns", "glue:GetWorkflow", "s3:ListBucket", "s3:GetBucketLocation", "s3:ListAllMyBuckets", "s3:GetBucketAcl", "iam:ListUsers", "iam:ListRoles", "iam:GetRole", "iam:GetRolePolicy"], "Resource": "*" }, { "Effect": "Deny", "Action": ["lakeformation:PutDataLakeSettings"], "Resource": "*" }] </pre>

Tipo di policy	Policy
	<code>}</code>

Autorizzazioni del tecnico dei dati

Important

Nelle seguenti politiche, sostituiscilo `<account-id>` con un numero di AWS account valido e sostituiscilo `<workflow_role>` con il nome del ruolo del flusso di lavoro.

Tipo di policy	Policy
AWS politica gestita	AWSGlueConsoleFullAccess
politica in linea (di base)	<pre>{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", "lakeformation:GrantPermissions", "lakeformation:RevokePermissions", "lakeformation:BatchGrantPermissions", "lakeformation:BatchRevokePermissions", "lakeformation:ListPermissions", "lakeformation:AddLFTagsToResource", "lakeformation:RemoveLFTagsFromResource", "lakeformation:GetResourceLFTags", "lakeformation:ListLFTags", "lakeformation:GetLFTag", "lakeformation:SearchTablesByLFTags"] }] }</pre>

Tipo di policy	Policy
	<pre> "lakeformation:SearchDatabasesByLFTa gs", "lakeformation:GetWorkUnits", "lakeformation:GetWorkUnitResults", "lakeformation:StartQueryPlanning", "lakeformation:GetQueryState", "lakeformation:GetQueryStatistics"], "Resource": "*" } </pre>
Politica in linea (per le operazioni su tabelle gestite, incluse le operazioni all'interno delle transazioni)	<pre> { "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:StartTransaction", "lakeformation:CommitTransaction", "lakeformation:CancelTransaction", "lakeformation:ExtendTransaction", "lakeformation:DescribeTransaction", "lakeformation:ListTransactions", "lakeformation:GetTableObjects", "lakeformation:UpdateTableObjects", "lakeformation>DeleteObjectsOnCancel "], "Resource": "*" }] } </pre>

Tipo di policy	Policy
Policy in linea (per il controllo dell'accesso ai metadati utilizzando il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC))	<pre data-bbox="574 247 1507 1125"> { "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:AddLFTagsToResource", "lakeformation:RemoveLFTagsFromResource", "lakeformation:GetResourceLFTags", "lakeformation:ListLFTags", "lakeformation:GetLFTag", "lakeformation:SearchTablesByLFTags", "lakeformation:SearchDatabasesByLFTags"], "Resource": "*" }] }</pre>
Politica in linea (politica passrole per il ruolo del flusso di lavoro)	<pre data-bbox="574 1188 1507 1856"> { "Version": "2012-10-17", "Statement": [{ "Sid": "PassRolePermissions", "Effect": "Allow", "Action": ["iam:PassRole"], "Resource": ["arn:aws:iam:: 111122223333 :role/<workflow_role> "] }] }</pre>

Autorizzazioni per analisti di dati

Tipo di policy	Policy
AWS politica gestita	AmazonAthenaFullAccess
politica in linea (di base)	<pre> { "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", "glue:GetTable", "glue:GetTables", "glue:SearchTables", "glue:GetDatabase", "glue:GetDatabases", "glue:GetPartitions", "lakeformation:GetResourceLFTags", "lakeformation:ListLFTags", "lakeformation:GetLFTag", "lakeformation:SearchTablesByLFTags", "lakeformation:SearchDatabasesByLFTags"], "Resource": "*" }] }</pre>

(Facoltativo) Politica in linea (per le operazioni su tabelle gestite, incluse le operazioni all'interno delle transazioni)

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:StartTransaction",
        "lakeformation:CommitTransaction",
        "lakeformation:CancelTransaction",
        "lakeformation:ExtendTransaction",
```

Tipo di policy	Policy
	<pre> "lakeformation:DescribeTransaction", "lakeformation:ListTransactions", "lakeformation:GetTableObjects", "lakeformation:UpdateTableObjects", "lakeformation>DeleteObjectsOnCancel "], "Resource": "*" }] } </pre>

Autorizzazioni relative ai ruoli del workflow

Questo ruolo dispone delle autorizzazioni necessarie per eseguire un flusso di lavoro. Si specifica un ruolo con queste autorizzazioni quando si crea un flusso di lavoro.

Important

Nelle seguenti politiche, sostituiscilo *<region>* con un identificatore di AWS regione valido (ad esempio *us-east-1*), *<account-id>* con un numero di AWS account valido, *<workflow_role>* con il nome del ruolo del flusso di lavoro e *<your-s3-cloudtrail-bucket>* con il percorso Amazon S3 verso AWS CloudTrail i log.

Tipo di policy	Policy
AWS politica gestita	AWSGlueServiceRole
policy in linea (accesso ai dati)	<pre> { "Version": "2012-10-17", "Statement": [{ "Sid": "Lakeformation", "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", </pre>

Tipo di policy	Policy
	<pre> "lakeformation:GrantPermissions"], "Resource": "*" }] }</pre>

Politica in linea (politica passrole per il ruolo del flusso di lavoro)

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PassRolePermissions",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam:: 111122223333 :role/<workflow_role> "
      ]
    }
  ]
}
```

Policy in linea (per l'acquisizione di dati all'esterno del data lake, ad esempio log) AWS CloudTrail

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": ["s3:GetObject", "s3:ListB
ucket"],
      "Resource": ["arn:aws:s3::: <your-s3-
cloudtrail-bucket> /*"]
    }
  ]
}
```

Modifica delle impostazioni predefinite per il data lake

Per mantenere la retrocompatibilità con AWS Glue, AWS Lake Formation dispone delle seguenti impostazioni di sicurezza iniziali:

- L'Superautorizzazione viene concessa al gruppo `IAMAllowedPrincipals` su tutte le risorse del AWS Glue Data Catalog esistenti.
- Le impostazioni «Usa solo il controllo di accesso IAM» sono abilitate per le nuove risorse del Data Catalog.

Queste impostazioni fanno sì che l'accesso alle risorse di Data Catalog e alle sedi Amazon S3 sia controllato esclusivamente da policy AWS Identity and Access Management (IAM). Le autorizzazioni individuali di Lake Formation non sono valide.

Il `IAMAllowedPrincipals` gruppo include tutti gli utenti e i ruoli IAM a cui è consentito l'accesso alle risorse del Data Catalog in base alle politiche IAM. L'Superautorizzazione consente a un principale di eseguire tutte le operazioni di Lake Formation supportate sul database o sulla tabella su cui è concessa.

Per modificare le impostazioni di sicurezza in modo che l'accesso alle risorse del Data Catalog (database e tabelle) sia gestito dalle autorizzazioni di Lake Formation, procedi come segue:

1. Modifica le impostazioni di sicurezza predefinite per le nuove risorse. Per istruzioni, consulta [Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida](#).
2. Modifica le impostazioni per le risorse esistenti del Data Catalog. Per istruzioni, consulta [Aggiornamento delle AWS Glue autorizzazioni relative ai dati al modello AWS Lake Formation](#).

Modifica delle impostazioni di sicurezza predefinite utilizzando l'operazione dell'**PutDataLakeSettings** API Lake Formation

Puoi anche modificare le impostazioni di sicurezza predefinite utilizzando l'operazione [PutDataLakeSettings](#) API Lake Formation. Questa azione utilizza come argomenti un ID di catalogo e una [DataLakeSettings](#) struttura opzionali.

Per applicare i metadati e il controllo dell'accesso ai dati sottostanti da parte di Lake Formation su nuovi database e tabelle, codifica la `DataLakeSettings` struttura come segue.

 Note

Sostituiscilo *<AccountID>* con un ID AWS account valido e *<Username>* con un nome utente IAM valido. Puoi specificare più di un utente come amministratore del data lake.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ],
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": []
  }
}
```

È inoltre possibile codificare la struttura come segue. Omettere il `CreateTableDefaultPermissions` parametro `CreateDatabaseDefaultPermissions` o equivale a passare un elenco vuoto.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ]
  }
}
```

Questa azione revoca efficacemente tutte le autorizzazioni di Lake Formation dal `IAMAllowedPrincipals` gruppo su nuovi database e tabelle. Quando crei un database, puoi sovrascrivere questa impostazione.

Per applicare i metadati e il controllo dell'accesso ai dati sottostanti solo da parte di IAM su nuovi database e tabelle, codifica la `DataLakeSettings` struttura come segue.

```
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier":
"arn:aws:iam::<AccountId>:user/<Username>"
      }
    ],
    "CreateDatabaseDefaultPermissions": [
      {
        "Principal": {
          "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
        },
        "Permissions": [
          "ALL"
        ]
      }
    ],
    "CreateTableDefaultPermissions": [
      {
        "Principal": {
          "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
        },
        "Permissions": [
          "ALL"
        ]
      }
    ]
  }
}
```

Ciò concede a Super Lake Formation l'autorizzazione al IAMAllowedPrincipals gruppo su nuovi database e tabelle. Quando crei un database, puoi sovrascrivere questa impostazione.

Note

Nella DataLakeSettings struttura precedente, l'unico valore consentito per DataLakePrincipalIdentifier è IAM_ALLOWED_PRINCIPALS l'unico valore consentito per Permissions è ALL

Autorizzazioni implicite di Lake Formation

AWS Lake Formation concede le seguenti autorizzazioni implicite agli amministratori di data lake, ai creatori di database e ai creatori di tabelle.

amministratori di data lake

- Hanno `Describe` accesso a tutte le risorse del Data Catalog ad eccezione delle risorse condivise da un altro account direttamente con un altro principale. Questo accesso non può essere revocato a un amministratore.
- Disponi delle autorizzazioni per la localizzazione dei dati ovunque nel data lake.
- Può concedere o revocare l'accesso a qualsiasi risorsa del Data Catalog a qualsiasi principale (incluso se stesso). Questo accesso non può essere revocato a un amministratore.
- Può creare database nel Data Catalog.
- Può concedere il permesso di creare un database a un altro utente.

Note

Gli amministratori di Data Lake possono registrare le sedi Amazon S3 solo se dispongono delle autorizzazioni IAM necessarie per farlo. Le politiche di amministrazione dei data lake suggerite in questa guida concedono tali autorizzazioni. Inoltre, gli amministratori dei data lake non dispongono delle autorizzazioni implicite per eliminare database o alter/drop tabelle creati da altri. Tuttavia, possono concedersi le autorizzazioni per farlo.

Per ulteriori informazioni sugli amministratori del data lake, consulta [Crea un amministratore del data lake](#)

Creatori di cataloghi

- Dispongono di tutte le autorizzazioni di catalogo sui cataloghi che creano, delle autorizzazioni sui database e sulle tabelle che creano nel catalogo e possono concedere ad altri responsabili dello stesso AWS account il permesso di creare database e tabelle nel catalogo. Un creatore di catalogo che dispone anche della politica `AWSLakeFormationCrossAccountManager` AWS gestita può concedere le autorizzazioni sul catalogo ad altri account o organizzazioni. AWS

Gli amministratori di Data Lake possono utilizzare la console o l'API di Lake Formation per designare i creatori del catalogo.

 Note

I creatori di cataloghi non dispongono implicitamente delle autorizzazioni sui database e sulle tabelle che altri creano nel catalogo.

Per ulteriori informazioni sulla creazione di cataloghi, consulta [Inserire i dati nel AWS Glue Data Catalog](#)

Creatori di database

- Dispongono di tutte le autorizzazioni di database sui database che creano, delle autorizzazioni sulle tabelle che creano nel database e possono concedere ad altri responsabili dello stesso AWS account il permesso di creare tabelle nel database. Un creatore di database che dispone anche della politica `AWSLakeFormationCrossAccountManager` AWS gestita può concedere le autorizzazioni sul database ad altri AWS account o organizzazioni.

Gli amministratori di Data Lake possono utilizzare la console o l'API di Lake Formation per designare i creatori di database.

 Note

I creatori di database non dispongono implicitamente delle autorizzazioni sulle tabelle create da altri nel database.

Per ulteriori informazioni, consulta [Creazione di un database](#).

I creatori di tabelle

- Disponi di tutte le autorizzazioni sulle tabelle che creano.
- Possono concedere le autorizzazioni su tutte le tabelle che creano ai responsabili dello stesso account. AWS
- Può concedere le autorizzazioni su tutte le tabelle che crea ad altri AWS account o organizzazioni se dispongono della `AWSLakeFormationCrossAccountManager` AWS politica gestita.
- Può visualizzare i database che contengono le tabelle che crea.

Riferimento alle autorizzazioni di Lake Formation

Per eseguire AWS Lake Formation le operazioni, i mandanti necessitano sia delle autorizzazioni di Lake Formation che delle autorizzazioni AWS Identity and Access Management (IAM). In genere concedi le autorizzazioni IAM utilizzando politiche di controllo degli accessi granulari, come descritto in [the section called “Panoramica delle autorizzazioni di Lake Formation”](#). Puoi concedere le autorizzazioni di Lake Formation utilizzando la console, l'API o AWS Command Line Interface (AWS CLI).

Per informazioni su come concedere o revocare le autorizzazioni di Lake Formation, vedere e. [the section called “Concessione delle autorizzazioni relative ai dati”](#) [the section called “Concessione delle autorizzazioni per la localizzazione dei dati”](#)

Note

Gli esempi in questa sezione mostrano come concedere le autorizzazioni ai responsabili dello stesso account. AWS Per esempi di sovvenzioni su più account, vedere. [the section called “Condivisione dei dati tra account”](#)

Autorizzazioni Lake Formation per tipo di risorsa

Di seguito sono riportate le autorizzazioni valide di Lake Formation disponibili per ogni tipo di risorsa:

Risorsa	Autorizzazione	
Catalog	ALL(Super), Super utente	
	ALTER	
	CREATE_DATABASE	
	DESCRIBE	
	DROP	
Database	ALL (Super)	
	ALTER	

Risorsa	Autorizzazione	
	CREATE_TABLE	
	DESCRIBE	
	DROP	
Table	ALL (Super)	
	ALTER	
	DELETE	
	DESCRIBE	
	DROP	
	INSERT	
	SELECT	
View	ALL (Super)	
	SELECT	
	DESCRIBE	
	DROP	
Data Catalog	CREATE_DATABASE	
Amazon S3 location	DATA_LOCATION_ACCESS	
LF-Tags	DROP	
	ALTER	
LF-Tag values	ASSOCIATE	
	DESCRIBE	

Risorsa	Autorizzazione	
	GrantWithLFTagExpression	
LF-Tag policy - Database	ALL (Super)	
	ALTER	
	CREATE_TABLE	
	DESCRIBE	
	DROP	
LF-Tag policy - Table	ALL (Super)	
	ALTER	
	DESCRIBE	
	DELETE	
	DROP	
	INSERT	
	SELECT	
Resource link - Database or Table	DESCRIBE	
	DROP	
Table with data filters	DESCRIBE	
	DROP	
	SELECT	
Table with column filter	SELECT	

Argomenti

- [Comandi di concessione e AWS CLI revoca di Lake Formation](#)
- [Autorizzazioni Lake Formation](#)

Comandi di concessione e AWS CLI revoca di Lake Formation

Ogni descrizione dell'autorizzazione in questa sezione include esempi di concessione dell'autorizzazione tramite un comando. AWS CLI Di seguito sono riportate le sinossi `grant-permissions` e `revoke-permissions` AWS CLI i comandi del Lake Formation.

```
grant-permissions
[--catalog-id <value>]
--principal <value>
--resource <value>
--permissions <value>
[--permissions-with-grant-option <value>]
[--cli-input-json <value>]
[--generate-cli-skeleton <value>]
```

```
revoke-permissions
[--catalog-id <value>]
--principal <value>
--resource <value>
--permissions <value>
[--permissions-with-grant-option <value>]
[--cli-input-json <value>]
[--generate-cli-skeleton <value>]
```

[Per una descrizione dettagliata di questi comandi, vedete `grant-permissions` e `revoke-permissions` nel \[Command Reference.AWS CLI\]\(#\)](#) Questa sezione fornisce informazioni aggiuntive sull'opzione. `--principal`

Il valore dell'`--principal` opzione è uno dei seguenti:

- Amazon Resource Name (ARN) per un utente o un AWS Identity and Access Management ruolo (IAM)
- ARN per un utente o un gruppo che esegue l'autenticazione tramite un provider SAML, come Microsoft Active Directory Federation Service (ADFS)

- ARN per un utente o un gruppo Amazon Quick Suite
- Per le autorizzazioni su più account, un ID AWS account, un ID dell'organizzazione o un ID di unità organizzativa
- Per utente o gruppo IAM Identity Center, ARN per utenti o gruppi IAM Identity Center.

Di seguito sono riportati la sintassi e gli esempi per tutti i `--principal` tipi.

Principal è un utente IAM

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
```

Esempio:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/  
datalake_user1
```

Principal è un ruolo IAM

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
```

Esempio:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/workflowrole
```

Principal è un utente che si autentica tramite un provider SAML

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-  
provider/<SAMLproviderName>:user/<user-name>
```

Esempi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/  
idp1:user/datalake_user1
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/  
AthenaLakeFormation0kta:user/athena-user@example.com
```

Principal è un gruppo che si autentica tramite un provider SAML

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-  
provider/<SAMLproviderName>:group/<group-name>
```

Esempi:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/  
idp1:group/data-scientists
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/  
AthenaLakeFormation0kta:group/my-group
```

Principal è un utente di Amazon Quick Suite Enterprise Edition

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-  
id>:user/<namespace>/<user-name>
```

 Note

Per *<namespace>*, è necessario specificare default.

Esempio:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-  
east-1:111122223333:user/default/bi_user1
```

Principal è un gruppo Amazon Quick Suite Enterprise Edition

Sintassi:


```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-id>:group/<namespace>/<group-name>
```

Note

Per *<namespace>*, è necessario specificare default.

Esempio:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-east-1:111122223333:group/default/data_scientists
```

Principal è un AWS account

Sintassi:

```
--principal DataLakePrincipalIdentifier=<account-id>
```

Esempio:

```
--principal DataLakePrincipalIdentifier=111122223333
```

Il preside è un'organizzazione

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-id>:organization/<organization-id>
```

Esempio:

```
--principal  
DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/o-  
abcdefghijkl
```

Il preside è un'unità organizzativa

Sintassi:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-id>:ou/<organization-id>/<organizational-unit-id>
```

Esempio:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:ou/o-abcdefghijklmnopqrstuvwxyz/ou-ab00-cdefghij
```

Il preside è un utente o un gruppo di identità di IAM Identity Center

Esempio: utente

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserID>
```

Esempio: Gruppo:

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::group/<GroupID>
```

Principal è un gruppo IAM - **IAMAllowedPrincipals**

Lake Formation imposta le Super autorizzazioni su tutti i database e le tabelle del Data Catalog a un gruppo chiamato per IAMAllowedPrincipals impostazione predefinita. Se questa autorizzazione di gruppo esiste su un database o una tabella, tutti i principali del tuo account avranno accesso alla risorsa tramite le politiche principali IAM per AWS Glue. Fornisce la compatibilità con le versioni precedenti quando inizi a utilizzare le autorizzazioni di Lake Formation per proteggere le risorse del Data Catalog per cui in precedenza erano protette dalle politiche IAM. AWS Glue

Quando utilizzi Lake Formation per gestire le autorizzazioni per le tue risorse Data Catalog, devi prima revocare l'IAMAllowedPrincipals autorizzazione sulle risorse o attivare i principali e le risorse per la modalità di accesso ibrido affinché le autorizzazioni di Lake Formation funzionino.

Esempio:

```
--principal DataLakePrincipalIdentifier=IAM_Allowed_Principals
```

Principal è un gruppo IAM - **ALLIAMPprincipals**

Quando concedi le autorizzazioni per il ALLIAMPprincipals gruppo su una risorsa Data Catalog, ogni principale dell'account ottiene l'accesso alla risorsa Data Catalog utilizzando le autorizzazioni Lake Formation e le autorizzazioni IAM.

Esempio:

```
--principal DataLakePrincipalIdentifier=123456789012:IAMPrincipals
```

Autorizzazioni Lake Formation

Questa sezione contiene i permessi disponibili di Lake Formation che puoi concedere ai presidi.

ALTER

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
ALTER	DATABASE	glue:UpdateDatabase
ALTER	TABLE	glue:UpdateTable
ALTER	LF-Tag	lakeformation:UpdateLFTag

Un principale con questa autorizzazione può modificare i metadati di un database o di una tabella nel Catalogo dati. Per le tabelle, è possibile modificare lo schema delle colonne e aggiungere parametri di colonna. Non è possibile modificare le colonne nei dati sottostanti a cui fa riferimento una tabella di metadati.

Se la proprietà da modificare è una posizione registrata di Amazon Simple Storage Service (Amazon S3), il principale deve disporre delle autorizzazioni per la localizzazione dei dati nella nuova posizione.

Example

L'esempio seguente concede l'ALTER autorizzazione all'utente `datalake_user1` sul database `retail` nell'account 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "ALTER" --resource '{ "Database": {"Name":"retail"}}'
```

Example

L'esempio seguente concede ALTER a un utente l'`datalake_user1` accesso alla tabella del database `inventory retail`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

CREATE_DATABASE

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
CREATE_DATABASE	Catalogo dati	glue:CreateDatabase

Un responsabile con questa autorizzazione può creare un database di metadati o un collegamento a una risorsa nel Catalogo dati. Il principale può anche creare tabelle nel database.

Example

L'esempio seguente concede CREATE_DATABASE a un utente `datalake_user1` nell' AWS account 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {} }'
```

Quando un principale crea un database nel Data Catalog, non viene concessa alcuna autorizzazione per i dati sottostanti. Vengono concesse le seguenti autorizzazioni aggiuntive per i metadati (oltre alla possibilità di concedere tali autorizzazioni ad altri):

- CREATE_TABLE nel database

- Database ALTER
- Database DROP

Quando crea un database, il principale può facoltativamente specificare una posizione Amazon S3. A seconda che il principale disponga delle autorizzazioni per la localizzazione dei dati, l'`CREATE_DATABASE` autorizzazione potrebbe non essere sufficiente per creare database in tutti i casi. È importante tenere a mente i tre casi seguenti.

Crea un caso d'uso del database	Autorizzazioni necessarie
La proprietà <code>location</code> non è specificata.	<code>CREATE_DATABASE</code> è sufficiente.
La proprietà <code>location</code> è specificata e la posizione non è gestita da Lake Formation (non è registrata).	<code>CREATE_DATABASE</code> è sufficiente.
La proprietà <code>location</code> è specificata e la posizione è gestita da Lake Formation (è registrata).	<code>CREATE_DATABASE</code> è obbligatorio più le autorizzazioni di localizzazione dei dati nella posizione specificata.

CREATE_TABLE

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
<code>CREATE_TABLE</code>	<code>DATABASE</code>	<code>glue:CreateTable</code>

Un principale con questa autorizzazione può creare una tabella di metadati o un collegamento a una risorsa nel Catalogo dati all'interno del database specificato.

Example

L'esempio seguente concede all'utente `datalake_user1` autorizzazione a creare tabelle nel `retail` database nell' AWS account 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
```

```
--permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"} }'
```

Quando un principale crea una tabella nel Data Catalog, tutti i permessi di Lake Formation sulla tabella vengono concessi al principale, con la possibilità di concedere tali autorizzazioni ad altri.

Sovvenzioni per più account

Se un account del proprietario del database concede CREATE_TABLE a un account destinatario e un utente dell'account destinatario crea correttamente una tabella nel database dell'account proprietario, si applicano le seguenti regole:

- Gli amministratori utente e data lake nell'account del destinatario dispongono di tutte le autorizzazioni di Lake Formation sulla tabella. Possono concedere le autorizzazioni sulla tabella ad altri responsabili del loro account. Non possono concedere autorizzazioni ai responsabili dell'account proprietario o di qualsiasi altro account.
- Gli amministratori di Data Lake dell'account proprietario possono concedere le autorizzazioni relative alla tabella ad altri responsabili del proprio account.

Autorizzazioni per la localizzazione dei dati

Quando tenti di creare una tabella che punti a una posizione Amazon S3, a seconda che tu disponga delle autorizzazioni per la localizzazione dei dati, l'CREATE_TABLE autorizzazione potrebbe non essere sufficiente per creare una tabella. È importante tenere a mente i seguenti tre casi.

Crea un caso d'uso della tabella	Autorizzazioni necessarie
La località specificata non è gestita da Lake Formation (non è registrata).	CREATE_TABLE è sufficiente.
La posizione specificata è gestita da Lake Formation (è registrata) e il database che lo contiene non ha alcuna proprietà di location o ha una proprietà di location che non sia un prefisso Amazon S3 della posizione della tabella.	CREATE_TABLE è richiesta più le autorizzazioni di localizzazione dei dati nella posizione specificata.
La posizione specificata è gestita da Lake Formation (è registrata) e il database che lo	CREATE_TABLE è sufficiente.

Crea un caso d'uso della tabella	Autorizzazioni necessarie
contiene ha una proprietà location che punta a una posizione registrata ed è un prefisso Amazon S3 della posizione della tabella.	

DATA_LOCATION_ACCESS

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
DATA_LOCATION_ACCESS	Posizione di Amazon S3.	(Autorizzazioni Amazon S3 sulla posizione, che devono essere specificate dal ruolo utilizzato per registrare la posizione.)

Questa è l'unica autorizzazione per la localizzazione dei dati. Un principale con questa autorizzazione può creare un database o una tabella di metadati che punti alla posizione Amazon S3 specificata. La sede deve essere registrata. Un preside che dispone delle autorizzazioni per la localizzazione dei dati su una sede dispone anche delle autorizzazioni di localizzazione per le sedi dei figli.

Example

L'esempio seguente concede le autorizzazioni di localizzazione dei dati `s3://products/retail` all'utente `datalake_user1` nell'account 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3:::products/retail"} }'
```

`DATA_LOCATION_ACCESS` non è necessario per interrogare o aggiornare i dati sottostanti. Questa autorizzazione si applica solo alla creazione di risorse del catalogo dati.

Per ulteriori informazioni sulle autorizzazioni per la localizzazione dei dati, vedere [Underlying data access control](#).

DELETE

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
DELETE	TABLE	(Non sono necessarie autorizzazioni IAM aggiuntive se la posizione è registrata.)

Un principale con questa autorizzazione può inserire, aggiornare e leggere i dati sottostanti nella posizione Amazon S3 specificata dalla tabella. Il responsabile può anche visualizzare la tabella sulla console di Lake Formation e recuperare informazioni sulla tabella con l'AWS GlueAPI.

Example

L'esempio seguente concede l'DELETE autorizzazione all'utente `datalake_user1` sulla tabella del database `inventory retail` nell' AWS account 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DELETE" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

Questa autorizzazione si applica solo ai dati in Amazon S3 e non ai dati in altri archivi di dati come Amazon Relational Database Service (Amazon RDS).

DESCRIBE

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
DESCRIBE	Link alle risorse della tabella	<code>glue:GetTable</code>
	Collegamento alle risorse del database	<code>glue:GetDatabase</code>
DESCRIBE	DATABASE	<code>glue:GetDatabase</code>
DESCRIBE	TABLE	<code>glue:GetTable</code>

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
DESCRIBE	LF-Tag	<code>glue:GetTable</code> <code>glue:GetDatabase</code> <code>lakeformation:GetResourceLFTags</code> <code>lakeformation:ListLFTags</code> <code>lakeformation:GetLFTag</code> <code>lakeformation:SearchTablesByLFTags</code> <code>lakeformation:SearchDatabasesByLFTags</code>

Un principale con questa autorizzazione può visualizzare il database, la tabella o il link alla risorsa specificato. Non vengono concesse implicitamente altre autorizzazioni di Data Catalog e non vengono concesse implicitamente autorizzazioni di accesso ai dati. I database e le tabelle vengono visualizzati negli editor di query dei servizi integrati, ma non è possibile eseguire alcuna query su di essi a meno che non vengano concesse altre autorizzazioni di Lake Formation (ad esempio `SELECT`).

Ad esempio, un utente che ha un database può visualizzare il database e tutti i metadati del database (descrizione, posizione e così `DESCRIBE` via). Tuttavia, l'utente non può scoprire quali tabelle contiene il database e non può eliminare, modificare o creare tabelle nel database. Analogamente, un utente che `DESCRIBE` utilizza una tabella può visualizzare la tabella e i relativi metadati (descrizione, schema, posizione e così via), ma non può eliminare, modificare o eseguire query sulla tabella.

Di seguito sono riportate alcune regole aggiuntive per: `DESCRIBE`

- Se un utente dispone di altre autorizzazioni Lake Formation su un database, una tabella o un collegamento a una risorsa, `DESCRIBE` viene concessa implicitamente.

- Se un utente ha SELECT solo un sottoinsieme di colonne per una tabella (parzialeSELECT), l'utente è limitato a visualizzare solo quelle colonne.
- Non puoi concedere DESCRIBE a un utente che dispone di una selezione parziale su una tabella. Al contrario, non è possibile specificare elenchi di inclusione o esclusione di colonne per le tabelle a cui DESCRIBE è consentito.

Example

L'esempio seguente concede all'utente `datalake_user1` l'DESCRIBE autorizzazione per il collegamento alle risorse della tabella nel database `inventory-link retail` nell' AWS account `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory-link"}}'
```

DROP

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
DROP	DATABASE	glue:DeleteDatabase
DROP	TABLE	glue:DeleteTable
DROP	LF-Tag	lakeformation:DeleteLFTag
DROP	Collegamento alla risorsa del database Link alle risorse della tabella	glue:DeleteDatabase glue:DeleteTable

Un principale con questa autorizzazione può eliminare un collegamento a un database, una tabella o una risorsa nel Catalogo dati. Non puoi concedere DROP su un database a un account o un'organizzazione esterni.

 Warning

L'eliminazione di un database comporta l'eliminazione di tutte le tabelle del database.

Example

L'esempio seguente concede l'DROP autorizzazione all'utente del database `retail` nell' AWS account `datalake_user1 1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Database": {"Name":"retail"}}'
```

Example

L'esempio seguente concede DROP all'utente l'accesso alla tabella del database `datalake_user1.inventory retail`

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

Example

L'esempio che segue concede DROP all'utente `datalake_user1` della tabella un link di risorse `inventory-link` nel database. `retail`

```
aws lakeformation grant-permissions --principal
  DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail", "Name":"inventory-
link"}}'
```

INSERT

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
INSERT	TABLE	(Non sono necessarie autorizzazioni IAM aggiuntive se la posizione è registrata.)

Un principale con questa autorizzazione può inserire, aggiornare e leggere i dati sottostanti nella posizione Amazon S3 specificata dalla tabella. Il responsabile può anche visualizzare la tabella nella console Lake Formation e recuperare informazioni sulla tabella con l'AWS GlueAPI.

Example

L'esempio seguente concede l'INSERT autorizzazione all'utente `datalake_user1` sulla tabella del database `inventory retail` nell' AWS account `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "INSERT" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

Questa autorizzazione si applica solo ai dati in Amazon S3 e non ai dati in altri archivi di dati come Amazon RDS.

SELECT

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
SELECT	• TABLE	(Non sono necessarie autorizzazioni IAM aggiuntive se la posizione è registrata.)

Un principale con questa autorizzazione può visualizzare una tabella nel catalogo dati e interrogare i dati sottostanti in Amazon S3 nella posizione specificata dalla tabella. Il responsabile può visualizzare

la tabella nella console di Lake Formation e recuperare informazioni sulla tabella con l'AWS GlueAPI. Se il filtraggio delle colonne è stato applicato quando è stata concessa questa autorizzazione, il principale può visualizzare i metadati solo per le colonne incluse e può interrogare i dati solo dalle colonne incluse.

Note

È responsabilità del servizio di analisi integrato applicare il filtro delle colonne durante l'elaborazione di una query.

Example

L'esempio seguente concede l'SELECT autorizzazione all'utente `datalake_user1` sulla tabella del database `inventory retail` nell'AWS account `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "SELECT" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

Questa autorizzazione si applica solo ai dati in Amazon S3 e non ai dati in altri archivi di dati come Amazon RDS.

Puoi filtrare (limitare l'accesso a) colonne specifiche con un elenco di inclusione o un elenco di esclusione facoltativo. Un elenco di inclusione specifica le colonne a cui è possibile accedere. Un elenco di esclusione specifica le colonne a cui non è possibile accedere. In assenza di un elenco di inclusione o esclusione, tutte le colonne della tabella sono accessibili.

I risultati di `glue:GetTable` restituiscono solo le colonne che il chiamante è autorizzato a visualizzare. I servizi integrati come Amazon Athena e Amazon Redshift rispettano gli elenchi di inclusione ed esclusione delle colonne.

Example

L'esempio seguente concede SELECT all'utente presente nella tabella `l'inventario` utilizzo di `datalake_user1` un elenco di inclusione.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
```

```
permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName":"retail",
  "Name":"inventory", "ColumnNames": ["prodcode","location","period","withdrawals"]}]'
```

Example

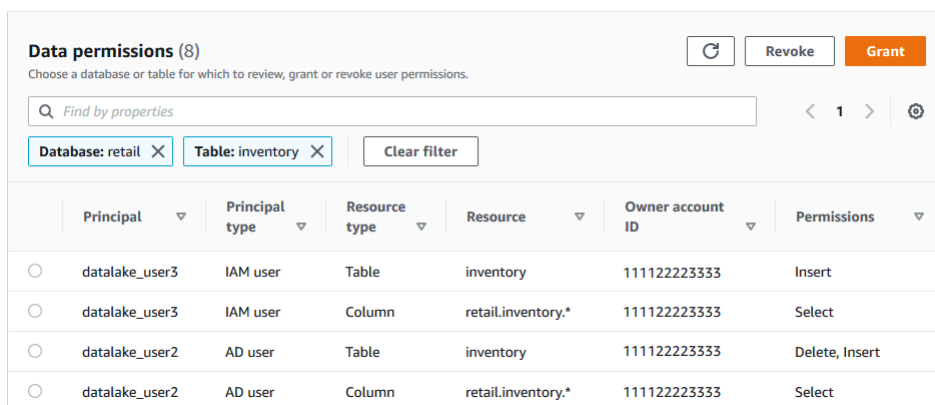
L'esempio successivo prevede l'utilizzo di SELECT un elenco di esclusione inventory nella tabella.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName":"retail",
  "Name":"inventory", "ColumnWildcard": {"ExcludedColumnNames": ["intkey",
  "prodcode"]}]}'
```

Le seguenti restrizioni si applicano all'SELECT autorizzazione:

- Quando si concede SELECT, non è possibile includere l'opzione di concessione se viene applicato il filtro delle colonne.
- Non è possibile limitare il controllo dell'accesso alle colonne che sono chiavi di partizione.
- A un principale con l'SELECT autorizzazione su un sottoinsieme di colonne di una tabella non può essere concessa l'INSERT autorizzazione ALTER, DROP, DELETE, o per quella tabella. Analogamente, a un'entità con l'INSERT autorizzazione ALTER, DROP, DELETE, o su una tabella non può essere concessa l'SELECT autorizzazione con il filtro delle colonne.

L'SELECT autorizzazione appare sempre nella pagina Autorizzazioni dati della console Lake Formation come una riga separata. L'immagine seguente mostra che SELECT è concesso agli utenti `datalake_user2` e `datalake_user3` su tutte le colonne della `inventory` tabella.



Data permissions (8)						
Choose a database or table for which to review, grant or revoke user permissions.						
Find by properties						
Database: retail		Table: inventory		Clear filter		
	Principal	Principal type	Resource type	Resource	Owner account ID	Permissions
☐	datalake_user3	IAM user	Table	inventory	111122223333	Insert
☐	datalake_user3	IAM user	Column	retail.inventory.*	111122223333	Select
☐	datalake_user2	AD user	Table	inventory	111122223333	Delete, Insert
☐	datalake_user2	AD user	Column	retail.inventory.*	111122223333	Select

Super

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
Super	DATABASE	glue:*Database*
Super	TABLE	glue:*Table*, glue:*Partition*

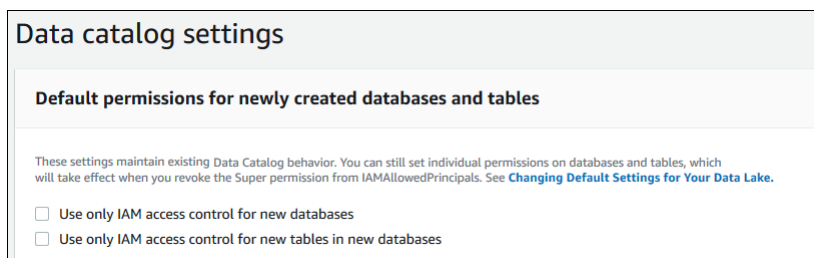
Questa autorizzazione consente a un principale di eseguire tutte le operazioni di Lake Formation supportate sul database o sulla tabella. Non puoi concedere Super un database a un account esterno.

Questa autorizzazione può coesistere con le altre autorizzazioni di Lake Formation. Ad esempio, puoi concedere le INSERT autorizzazioni SuperSELECT, e su una tabella di metadati. Il principale può quindi eseguire tutte le operazioni supportate sulla tabella. Quando si revoca Super, le INSERT autorizzazioni SELECT e rimangono invariate e il principale può eseguire solo operazioni di selezione e inserimento.

Invece di concedere Super a un singolo committente, puoi concederlo al gruppo.

IAMAllowedPrincipals Il IAMAllowedPrincipals gruppo viene creato automaticamente e include tutti gli utenti e i ruoli IAM a cui è consentito l'accesso alle risorse del Data Catalog dalle politiche IAM. Quando Super viene concesso a IAMAllowedPrincipals una risorsa Data Catalog, l'accesso alla risorsa è effettivamente controllato esclusivamente dalle politiche IAM.

Puoi avere l'Superautorizzazione da concedere automaticamente IAMAllowedPrincipals per nuove risorse del catalogo sfruttando le opzioni nella pagina Impostazioni della console Lake Formation.



- Per concedere Super a IAMAllowedPrincipals tutti i nuovi database, seleziona Usa solo il controllo di accesso IAM per i nuovi database.

- Per concedere Super a IAMAllowedPrincipals tutte le nuove tabelle nei nuovi database, seleziona Usa solo il controllo di accesso IAM per le nuove tabelle nei nuovi database.

Note

Questa opzione fa sì che la casella di controllo Usa solo il controllo di accesso IAM per le nuove tabelle in questo database nella finestra di dialogo Crea database sia selezionata per impostazione predefinita. Non fa altro che questo. È la casella di controllo nella finestra di dialogo Crea database che abilita la concessione di Super a IAMAllowedPrincipals.

Queste opzioni della pagina Impostazioni sono abilitate per impostazione predefinita. Per ulteriori informazioni, consulta gli argomenti seguenti:

- [the section called “Modifica delle impostazioni predefinite per il data lake”](#)
- [the section called “Aggiornamento delle autorizzazioni per AWS Glue i dati al modello Lake Formation”](#)

SUPER_USER

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
Super user	Catalog	glue:GetCatalog

È possibile concedere l'Super user autorizzazione solo a soggetti specifici sui cataloghi all'interno del Data Catalog predefinito. Non puoi concedere l'Super user autorizzazione sul catalogo predefinito o su altri tipi di risorse come database e tabelle o ai responsabili degli account esterni. L'Super user autorizzazione consente a un responsabile di eseguire tutte le operazioni di Lake Formation supportate sui database e sulle tabelle all'interno del catalogo concesso.

Con l'Super user autorizzazione, il principale (beneficiario) è in grado di eseguire le seguenti azioni sulle risorse (cataloghi, database e tabelle) all'interno del catalogo:

- CREATE_DATABASE, DESCRIBE autorizzazioni sul catalogo.

- DROP,ALTER,CREATE_TABLE, DESCRIBE (effettivamenteSUPER) autorizzazioni su tutti i database all'interno del catalogo.
- DROP,,ALTER,DESCRIBE, SELECTINSERT, DELETE (effettivamenteSUPER) autorizzazioni su tutte le tabelle all'interno di tutti i database all'interno del catalogo.
- All(effettivamente SUPER) autorizzazioni sui cataloghi all'interno del catalogo.
- Autorizzazioni concedibili (la possibilità di concedere queste autorizzazioni ad altri principali) su tutti i cataloghi, i database e le tabelle all'interno del catalogo.

Con l'Super user autorizzazione su una risorsa del catalogo, al beneficiario non è consentito eseguire o delegare alcuna azione sul catalogo. ALTER DROP

ASSOCIATE

Autorizzazione	Concesso su questa risorsa	Inoltre, il beneficiario ha bisogno
ASSOCIATE	LF-Tag	glue:GetDatabase glue:GetTable lakeformation:AddLFTagsToResource" lakeformation:RemoveLFTagsFromResource" lakeformation:GetResourceLFTags lakeformation:ListLFTags lakeformation:GetLFTag lakeformation:SearchTablesByLFTags lakeformation:SearchDatabasesByLFTags

Un principale con questa autorizzazione su un tag LF può assegnare il tag LF a una risorsa del catalogo dati. ASSOCIATEDESCRIBEConcedere concessioni implicite.

Example

Questo esempio concede all'utente l'ASSOCIATEautorizzazione per l'uso `dataLake_user1` del tag LF con la chiave. `module` Concede le autorizzazioni per visualizzare e assegnare tutti i valori per quella chiave, come indicato dall'asterisco (*).

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
dataLake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}]'
```

Integrazione di IAM Identity Center

Con AWS IAM Identity Center, puoi connetterti ai provider di identità (IdPs) e gestire centralmente l'accesso per utenti e gruppi tra i servizi di AWS analisi. È possibile integrare gestori di identità digitali come Okta, Ping e Microsoft Entra ID (in precedenza Azure Active Directory) con il Centro identità IAM per consentire agli utenti nell'organizzazione di accedere ai dati utilizzando un'esperienza di accesso singolo. Il Centro identità IAM supporta anche la connessione con altri gestori dell'identità digitali di terze parti.

Per ulteriori informazioni, consulta la sezione [Provider di identità supportati](#) nella Guida AWS IAM Identity Center per l'utente.

Puoi configurarlo AWS Lake Formation come applicazione abilitata in IAM Identity Center e gli amministratori del data lake possono concedere autorizzazioni granulari a utenti e gruppi autorizzati sulle risorse. AWS Glue Data Catalog

Gli utenti della tua organizzazione possono accedere a qualsiasi applicazione abilitata per Identity Center utilizzando il provider di identità dell'organizzazione e interrogare i set di dati utilizzando le autorizzazioni Lake Formation. Con questa integrazione, puoi gestire l'accesso ai AWS servizi, senza creare più ruoli IAM.

[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](#).

Per le limitazioni, consulta [Limitazioni dell'integrazione di IAM Identity Center](#).

Argomenti

- [Prerequisiti per l'integrazione di IAM Identity Center con Lake Formation](#)
- [Connessione di Lake Formation con IAM Identity Center](#)
- [Aggiornamento dell'integrazione con IAM Identity Center](#)
- [Eliminazione di una connessione Lake Formation con IAM Identity Center](#)
- [Concessione di autorizzazioni a utenti e gruppi](#)
- [Inclusione del contesto utente di IAM Identity Center nei log CloudTrail](#)

Prerequisiti per l'integrazione di IAM Identity Center con Lake Formation

Di seguito sono riportati i prerequisiti per l'integrazione di IAM Identity Center con Lake Formation.

1. Abilita IAM Identity Center: abilitare IAM Identity Center è un prerequisito per supportare l'autenticazione e la propagazione delle identità.
2. Scegli la tua fonte di identità: dopo aver abilitato IAM Identity Center, devi disporre di un provider di identità per gestire utenti e gruppi. È possibile utilizzare la directory Identity Center integrata come origine di identità o utilizzare un IdP esterno, ad esempio Microsoft Entra ID o Okta.

Per ulteriori informazioni, consulta [Manage your identity source](#) e [Connect to a un provider di identità esterno](#) nella Guida per l' AWS IAM Identity Center utente.

3. Crea un ruolo IAM: il ruolo che crea la connessione IAM Identity Center richiede le autorizzazioni per creare e modificare la configurazione dell'applicazione in Lake Formation e IAM Identity Center come nella seguente politica in linea.

È necessario aggiungere le autorizzazioni in base alle migliori pratiche IAM. Le autorizzazioni specifiche sono illustrate nelle procedure che seguono. Per ulteriori informazioni, consulta [Guida introduttiva a IAM Identity Center](#).

JSON

```
{
```

```

"Version": "2012-10-17",
"Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "lakeformation:CreateLakeFormationIdentityCenterConfiguration",
      "sso:CreateApplication",
      "sso:PutApplicationAssignmentConfiguration",
      "sso:PutApplicationAuthenticationMethod",
      "sso:PutApplicationGrant",
      "sso:PutApplicationAccessScope"
    ],
    "Resource": [
      "*"
    ]
  }
]
}

```

Se condividi risorse di Data Catalog con organizzazioni esterne Account AWS o esterne, devi disporre delle autorizzazioni AWS Resource Access Manager (AWS RAM) per creare condivisioni di risorse. Per ulteriori informazioni sulle autorizzazioni necessarie per condividere risorse, consulta [Prerequisiti per la condivisione dei dati tra account](#).

Le seguenti politiche in linea contengono autorizzazioni specifiche necessarie per visualizzare, aggiornare ed eliminare le proprietà dell'integrazione di Lake Formation con IAM Identity Center.

- Utilizza la seguente policy in linea per consentire a un ruolo IAM di visualizzare un'integrazione di Lake Formation con IAM Identity Center.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
        "sso:DescribeApplication"
      ],

```

```

        "Resource": [
            "*"
        ]
    }
]
}

```

- Utilizza la seguente policy in linea per consentire a un ruolo IAM di aggiornare un'integrazione di Lake Formation con IAM Identity Center. La policy include anche le autorizzazioni opzionali necessarie per condividere risorse con account esterni.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:UpdateLakeFormationIdentityCenterConfiguration",
        "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
        "sso:DescribeApplication",
        "sso:UpdateApplication"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}

```

- Utilizza la seguente policy in linea per consentire a un ruolo IAM di eliminare un'integrazione di Lake Formation con IAM Identity Center.

JSON

```

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

```

```

        "lakeformation:DeleteLakeFormationIdentityCenterConfiguration",
        "sso:DeleteApplication"
    ],
    "Resource": [
        "*"
    ]
}
]
}

```

- Per le autorizzazioni IAM necessarie per concedere o revocare le autorizzazioni del data lake per utenti e gruppi IAM Identity Center, consulta. [Autorizzazioni IAM necessarie per concedere o revocare le autorizzazioni di Lake Formation](#)

Descrizione delle autorizzazioni

- `lakeformation:CreateLakeFormationIdentityCenterConfiguration`— Crea la configurazione iDC di Lake Formation.
- `lakeformation:DescribeLakeFormationIdentityCenterConfiguration`— Descrive una configurazione iDC esistente.
- `lakeformation:DeleteLakeFormationIdentityCenterConfiguration`— Offre la possibilità di eliminare una configurazione iDC di Lake Formation esistente.
- `lakeformation:UpdateLakeFormationIdentityCenterConfiguration`— Usato per modificare una configurazione esistente di Lake Formation.
- `sso:CreateApplication`: utilizzato per creare un'applicazione Centro identità IAM.
- `sso:DeleteApplication`: utilizzato per eliminare un'applicazione Centro identità IAM.
- `sso:UpdateApplication`: utilizzato per aggiornare un'applicazione Centro identità IAM.
- `sso:PutApplicationGrant`: utilizzato per modificare le informazioni sull'emittente di token attendibile.
- `sso:PutApplicationAuthenticationMethod`— Garantisce l'accesso all'autenticazione di Lake Formation.
- `sso:GetApplicationGrant`: utilizzato per elencare informazioni sull'emittente di token attendibile.
- `sso:DeleteApplicationGrant`: elimina le informazioni sull'emittente di token attendibile.
- `sso:PutApplicationAccessScope`— Aggiunge o aggiorna l'elenco degli obiettivi autorizzati per un ambito di accesso di IAM Identity Center per un'applicazione.

- `sso:PutApplicationAssignmentConfiguration`— Utilizzato per configurare il modo in cui gli utenti accedono a un'applicazione.

Connessione di Lake Formation con IAM Identity Center

Prima di poter utilizzare IAM Identity Center per gestire le identità e concedere l'accesso alle risorse del Data Catalog utilizzando Lake Formation, devi completare i seguenti passaggi. Puoi creare l'integrazione con IAM Identity Center utilizzando la console Lake Formation oppure AWS CLI.

Console di gestione AWS

Per connettere Lake Formation con IAM Identity Center

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione a sinistra, seleziona Integrazione con IAM Identity Center.

Create IAM Identity Center Integration

Enable IAM Identity Center and then create Lake Formation - IAM Identity Center integration to manage identities from IAM Identity Center (external IdPs like Azure AD or Okta Universal Directory). [Learn more](#)

▼ How it works

Enable IAM Identity Center

Enable IAM Identity Center for your account or organization and select an identity provider.

Create Lake Formation integration

Integrate Lake Formation with IAM Identity Center to permit Lake Formation to access users from your selected identity provider.

Grant permissions

Grant permissions to users on Data Catalog databases and tables using fine-grained Lake Formation permissions.

Connect Lake Formation to IAM Identity Center



Connect to organization instance of IAM Identity Center

Manage access to Lake Formation by assigning users and groups from the Identity Center directory for your organization. [Learn more](#)

Recommended



Connect to account instance of IAM Identity Center

Manage access to Lake Formation by assigning existing or creating dedicated users and groups from your Identity Center directory. [Learn more](#)

instance of IAM Identity Center
Manage access to Lake Formation by assigning users and groups from your Identity Center directory.
 `arn:aws:sso:::instance/ssoins-6987513bf5410c2f`

Add AWS account and organization IDs

Add AWS accounts and organizations whose users need access to Lake Formation managed resources.

AWS Accounts and AWS organizations

Enter one or more AWS account IDs and AWS organization IDs. Press Enter after each ID.

 Choose AWS account, AWS organization ID

► Lake Formation application integration - optional

Connessione di Lake Formation con IAM Identity Center. Permette agli utenti di accedere a posizioni di dati registrate con Lake Formation on behalf of the user.

 After this step, you can't edit the connection. You can edit AWS accounts, organizations, and applications. If you want to modify the connection, delete it and create a new connection.

3. (Facoltativo) Inserisci una o più organizzazioni IDs o unità and/or organizzative valide Account AWS IDs IDs per consentire agli account esterni di accedere alle risorse del Data Catalog. Quando gli utenti o i gruppi di IAM Identity Center cercano di accedere alle risorse del Data Catalog gestite da Lake Formation, Lake Formation assume un ruolo IAM per autorizzare l'accesso ai metadati. Se il ruolo IAM appartiene a un account esterno che non dispone di una policy di AWS Glue risorse e di una condivisione di AWS RAM risorse, gli utenti e i gruppi di IAM Identity Center non saranno in grado di accedere alla risorsa anche se dispongono delle autorizzazioni Lake Formation.

Lake Formation utilizza il servizio AWS Resource Access Manager (AWS RAM) per condividere la risorsa con account e organizzazioni esterni. AWS RAM invia un invito all'account del beneficiario per accettare o rifiutare la condivisione delle risorse.

Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

 Note

Lake Formation consente ai ruoli IAM di account esterni di agire come portatori per conto degli utenti e dei gruppi di IAM Identity Center per l'accesso alle risorse del Data Catalog, ma le autorizzazioni possono essere concesse solo sulle risorse del Data Catalog all'interno dell'account proprietario. Se provi a concedere le autorizzazioni agli utenti e ai gruppi di IAM Identity Center sulle risorse di Data Catalog in un account esterno, Lake Formation genera il seguente errore: «Le sovvenzioni tra account non sono supportate per il principale».

4. (Facoltativo) Nella schermata di integrazione Create Lake Formation, specifica le applicazioni ARNs di terze parti che possono accedere ai dati nelle località Amazon S3 registrate con Lake Formation. Lake Formation fornisce credenziali temporanee limitate sotto forma di token alle sedi Amazon S3 registrate in base AWS STS alle autorizzazioni effettive, in modo che le applicazioni autorizzate possano accedere ai dati per conto degli utenti.
5. Selezionare Invia.

Dopo che l'amministratore di Lake Formation ha completato i passaggi e creato l'integrazione, le proprietà di IAM Identity Center vengono visualizzate nella console di Lake Formation. Il completamento di queste attività rende Lake Formation un'applicazione abilitata per IAM Identity Center. Le proprietà della console includono lo stato dell'integrazione. Lo

stato dell'integrazione indica Success quando è completata. Questo stato indica se la configurazione di IAM Identity Center è stata completata.

AWS CLI

- L'esempio seguente mostra come creare l'integrazione di Lake Formation con IAM Identity Center. Puoi anche specificare il Status (ENABLED,DISABLED) delle applicazioni.

```
aws lakeformation create-lake-formation-identity-center-configuration \  
  --catalog-id <123456789012> \  
  --instance-arn <arn:aws:sso:::instance/ssoins-112111f12ca1122p> \  
  --share-recipients '[{"DataLakePrincipalIdentifier": "<123456789012>"},  
                        {"DataLakePrincipalIdentifier": "<555555555555>"}]' \  
  --external-filtering '{"AuthorizedTargets": [<app arn1>", "<app arn2>"],  
                        "Status": "ENABLED"}'
```

- L'esempio seguente mostra come visualizzare un'integrazione di Lake Formation con IAM Identity Center.

```
aws lakeformation describe-lake-formation-identity-center-configuration  
  --catalog-id <123456789012>
```

Aggiornamento dell'integrazione con IAM Identity Center

Dopo aver creato la connessione, puoi aggiungere applicazioni di terze parti per l'integrazione di IAM Identity Center da integrare con Lake Formation e ottenere l'accesso ai dati di Amazon S3 per conto degli utenti. Puoi anche rimuovere le applicazioni esistenti dall'integrazione con IAM Identity Center. È possibile aggiungere o rimuovere applicazioni utilizzando la console Lake Formation e utilizzando [UpdateLakeFormationIdentityCenterConfiguration](#) operation. AWS CLI

Note

Dopo aver creato l'integrazione con IAM Identity Center, non puoi aggiornare l'istanzaARN.

Console di gestione AWS

Per aggiornare una connessione IAM Identity Center esistente con Lake Formation

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione a sinistra, seleziona Integrazione con IAM Identity Center.
3. Seleziona Aggiungi nella pagina di integrazione di IAM Identity Center.
4. Inserisci una o più organizzazioni IDs o unità and/or organizzative valide Account AWS IDs IDs per consentire agli account esterni di accedere alle risorse del Data Catalog.
5. Nella schermata Aggiungi applicazioni, inserisci l'applicazione IDs delle applicazioni di terze parti che desideri integrare con Lake Formation.
6. Selezionare Aggiungi.

AWS CLI

Puoi aggiungere o rimuovere applicazioni di terze parti per l'integrazione con IAM Identity Center eseguendo il AWS CLI comando seguente. Quando imposti lo stato del filtro esterno suENABLED, consente a IAM Identity Center di fornire la gestione delle identità per le applicazioni di terze parti per accedere ai dati gestiti da Lake Formation. Puoi anche abilitare o disabilitare l'integrazione di IAM Identity Center impostando lo stato dell'applicazione.

```
aws lakeformation update-lake-formation-identity-center-configuration \
  --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"], "Status": "ENABLED"}' \
  --share-recipients '[{"DataLakePrincipalIdentifier": "<444455556666>"} {"DataLakePrincipalIdentifier": "<777788889999>"}]' \
  --application-status ENABLED
```

Eliminazione di una connessione Lake Formation con IAM Identity Center

Se desideri eliminare un'integrazione esistente di IAM Identity Center, puoi farlo utilizzando la console o l'[DeleteLakeFormationIdentityCenterConfiguration](#) operazione Lake Formation. AWS CLI

Console di gestione AWS

Per eliminare una connessione IAM Identity Center esistente con Lake Formation

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione a sinistra, seleziona Integrazione con IAM Identity Center.
3. Seleziona Elimina nella pagina di integrazione di IAM Identity Center.
4. Nella schermata di conferma dell'integrazione, conferma l'azione e seleziona Elimina.

AWS CLI

Puoi eliminare l'integrazione con IAM Identity Center eseguendo il AWS CLI comando seguente.

```
aws lakeformation delete-lake-formation-identity-center-configuration \  
  --catalog-id <123456789012>
```

Concessione di autorizzazioni a utenti e gruppi

L'amministratore del data lake può concedere autorizzazioni agli utenti e ai gruppi di IAM Identity Center sulle risorse del Data Catalog (database, tabelle e viste) per consentire un facile accesso ai dati. Per concedere o revocare le autorizzazioni del data lake, il concedente richiede le autorizzazioni per le seguenti azioni di IAM Identity Center.

- [DescribeUser](#)
- [DescribeGroup](#)
- [DescribeInstance](#)

Puoi concedere le autorizzazioni utilizzando la console Lake Formation, l'API o il AWS CLI.

Per ulteriori informazioni sulla concessione delle autorizzazioni, consulta. [the section called "Concessione delle autorizzazioni relative ai dati"](#)

Note

Puoi concedere autorizzazioni solo sulle risorse del tuo account. Per assegnare a catena le autorizzazioni a utenti e gruppi sulle risorse condivise con te, devi utilizzare AWS RAM le condivisioni di risorse.

Console di gestione AWS

Per concedere autorizzazioni a utenti e gruppi

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Seleziona le autorizzazioni del data lake in Autorizzazioni nella console Lake Formation.
3. Seleziona Concedi.
4. Nella pagina Concedi le autorizzazioni del data lake, scegli utenti e gruppi di IAM Identity Center.
5. Seleziona Aggiungi per scegliere gli utenti e i gruppi a cui concedere le autorizzazioni.

Grant permissions

Principals
Choose the principals to grant permissions.

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **IAM Identity Center - new**
Users and groups configured in IAM Identity Center.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles
Add one or more IAM users or roles.

Choose IAM principals to add

AWSServiceRoleForRedshift ✕
Role

6. Nella schermata Assegna utenti e gruppi, scegli i gruppi di utenti and/or a cui concedere le autorizzazioni.

Seleziona Assegna.

Assign users and groups ✕

Users

user1 Remove

user2 Remove

Groups

DataStewards Remove

Manage groups [↗](#)

[Learn more about managing groups from IAM Identity Center](#) [↗](#)

Cancel Assign

7. Quindi, scegli il metodo per concedere le autorizzazioni.

Per istruzioni sulla concessione delle autorizzazioni utilizzando il metodo delle risorse denominate, vedere. [Concessione delle autorizzazioni per i dati utilizzando il metodo di risorsa denominato](#)

Per istruzioni su come concedere l'autorizzazione utilizzando i tag LF, vedere. [Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC](#)

8. Scegliete le risorse del Data Catalog per le quali desiderate concedere le autorizzazioni.
9. Scegli le autorizzazioni del Data Catalog da concedere.
10. Seleziona Concedi.

AWS CLI

L'esempio seguente mostra come concedere l'SELECT autorizzazione utente di IAM Identity Center su una tabella.

```
aws lakeformation grant-permissions \  
--principal DataLakePrincipalIdentifier=arn:aws:identitystore::user/<UserId> \  
--permissions "SELECT" \  
--resource '{ "Table": { "DatabaseName": "retail", "TableWildcard": {} } }'
```

Per eseguire il recupero UserId da IAM Identity Center, consulta il [GetUserId](#) funzionamento nel riferimento all'API IAM Identity Center.

Inclusione del contesto utente di IAM Identity Center nei log CloudTrail

Lake Formation utilizza la funzionalità di [vendita di credenziali](#) per fornire un accesso temporaneo ai dati di Amazon S3. Per impostazione predefinita, quando un utente di IAM Identity Center invia una query a un servizio di analisi integrato, CloudTrail i log includono solo il ruolo IAM assunto dal servizio per fornire un accesso a breve termine. Se utilizzi un ruolo definito dall'utente per registrare la posizione dei dati di Amazon S3 con Lake Formation, puoi scegliere di includere il contesto dell'utente di IAM Identity Center negli eventi e quindi tenere traccia CloudTrail degli utenti che accedono alle tue risorse.

Important

Per includere le richieste API Amazon S3 a livello di oggetto in, devi CloudTrail abilitare CloudTrail la registrazione degli eventi per i bucket e gli oggetti Amazon S3. Per ulteriori informazioni, consulta [Enabling CloudTrail event logging for Amazon S3 bucket and objects nella Amazon S3 User Guide](#).

Per abilitare il controllo della distribuzione delle credenziali nelle sedi di data lake registrate con ruoli definiti dall'utente

1. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nella barra di navigazione a sinistra, espandi Amministrazione e scegli le impostazioni del catalogo dati.

3. In Controllo avanzato, scegli Propagate il contesto fornito.
4. Scegli Save (Salva).

È inoltre possibile abilitare l'opzione di controllo avanzato impostando l'`Parameters` attributo nell'operazione. [PutDataLakeSettings](#) Per impostazione predefinita, il valore del `SET_CONTEXT` parametro è impostato su «true».

```
{
  "DataLakeSettings": {
    "Parameters": {"SET_CONTEXT": "true"},
  }
}
```

Di seguito è riportato un estratto di un CloudTrail evento con l'opzione di controllo avanzata. Questo log include sia il contesto della sessione dell'utente IAM Identity Center sia il ruolo IAM definito dall'utente assunto da Lake Formation per accedere alla posizione dati di Amazon S3. Vedi il `onBehalfOf` parametro nel seguente estratto.

```
{
  "eventVersion": "1.09",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "AR0AW7F7M0X40YE6FLIFN:access-grants-e653760c-4e8b-44fd-94d9-309e035b75ab",
    "arn": "arn:aws:sts::123456789012:assumed-role/accessGrantsTestRole/access-grants-e653760c-4e8b-44fd-94d9-309e035b75ab",
    "accountId": "123456789012",
    "accessKeyId": "ASIAW7F7M0X4CQLD4JIZN",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "AR0AW7F7M0X40YE6FLIFN",
        "arn": "arn:aws:iam::123456789012:role/accessGrantsTestRole",
        "accountId": "123456789012",
        "userName": "accessGrantsTestRole"
      },
      "attributes": {
        "creationDate": "2023-08-09T17:24:02Z",
        "mfaAuthenticated": "false"
      }
    }
  },
}
```



```
    "onBehalfOf": {
      "userId": "<identityStoreUserId>",
      "identityStoreArn": "arn:aws:identitystore::<restOfIdentityStoreArn>"
    },
    "eventTime": "2023-08-09T17:25:43Z",
    "eventSource": "s3.amazonaws.com",
    "eventName": "GetObject",
    ....
```

Aggiungere una posizione Amazon S3 al tuo data lake

Per aggiungere una posizione dati come spazio di archiviazione nel tuo data lake, devi registrare la posizione (Data lake location) con AWS Lake Formation. Puoi quindi utilizzare le autorizzazioni di Lake Formation per un controllo granulare degli accessi agli AWS Glue Data Catalog oggetti che puntano a questa posizione e ai dati sottostanti nella posizione.

Lake Formation consente inoltre di registrare una posizione dei dati in modalità di accesso ibrida e offre la flessibilità necessaria per abilitare selettivamente le autorizzazioni Lake Formation per database e tabelle nel Data Catalog. Con la modalità di accesso Hybrid, hai a disposizione un percorso incrementale che ti consente di impostare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti.

Per ulteriori informazioni sulla configurazione della modalità di accesso ibrida, vedere [Modalità di accesso ibrida](#)

Quando registri una posizione, vengono registrati quel percorso Amazon S3 e tutte le cartelle in quel percorso.

Ad esempio, supponiamo di avere un'organizzazione dei percorsi Amazon S3 come la seguente:

```
/mybucket/accounting/sales/
```

Se ti registri `S3://mybucket/accounting`, anche la `sales` cartella è registrata e sotto la gestione di Lake Formation.

Per ulteriori informazioni sulla registrazione delle sedi, vedere [Underlying data access control](#).

 Note

Le autorizzazioni Lake Formation sono consigliate per i dati strutturati (disposti in tabelle con righe e colonne). Se i tuoi dati contengono dati non strutturati basati su oggetti, prendi in considerazione l'utilizzo delle concessioni di accesso di Amazon S3 per gestire l'accesso ai dati.

Argomenti

- [Requisiti per i ruoli utilizzati per registrare le sedi](#)
- [Registrazione di una sede Amazon S3](#)
- [Registrazione di una posizione Amazon S3 crittografata](#)
- [Registrazione di una sede Amazon S3 in un altro account AWS](#)
- [Registrazione di una posizione Amazon S3 crittografata tra più account AWS](#)
- [Annullamento della registrazione di una sede Amazon S3](#)

Requisiti per i ruoli utilizzati per registrare le sedi

È necessario specificare un ruolo AWS Identity and Access Management (IAM) quando si registra una sede Amazon Simple Storage Service (Amazon S3). AWS Lake Formation assume quel ruolo quando accede ai dati in quella posizione.

È possibile utilizzare uno dei seguenti tipi di ruolo per registrare una posizione:

- Il ruolo legato ai servizi di Lake Formation. Questo ruolo concede le autorizzazioni necessarie sulla sede. L'utilizzo di questo ruolo è il modo più semplice per registrare la posizione. Per ulteriori informazioni, consultare [Utilizzo di ruoli collegati ai servizi per Lake Formation](#) e [Limitazioni dei ruoli legati ai servizi](#).
- Un ruolo definito dall'utente. Utilizza un ruolo definito dall'utente quando devi concedere più autorizzazioni rispetto a quelle fornite dal ruolo collegato al servizio.

È necessario utilizzare un ruolo definito dall'utente nelle seguenti circostanze:

- Quando si registra una sede in un altro account.

Per ulteriori informazioni, consultare [the section called “Registrazione di una sede Amazon S3 in un altro account AWS”](#) e [the section called “Registrazione di una posizione Amazon S3 crittografata tra più account AWS”](#).

- Se hai utilizzato una CMK (aws/s3) AWS gestita per crittografare la posizione Amazon S3.

Per ulteriori informazioni, consulta [Registrazione di una posizione Amazon S3 crittografata](#).

- Se prevedi di accedere alla posizione utilizzando Amazon EMR.

Se hai già registrato una sede con il ruolo collegato al servizio e desideri iniziare ad accedervi con Amazon EMR, devi annullare la registrazione della sede e registrarla nuovamente con un ruolo definito dall'utente. Per ulteriori informazioni, consulta [the section called “Annullamento della registrazione di una sede Amazon S3”](#).

Utilizzo di ruoli collegati ai servizi per Lake Formation

AWS Lake Formation utilizza un ruolo collegato al AWS Identity and Access Management servizio (IAM). Un ruolo collegato ai servizi è un tipo unico di ruolo IAM collegato direttamente a Lake Formation. Il ruolo collegato al servizio è predefinito da Lake Formation e include tutte le autorizzazioni richieste dal servizio per chiamare altri AWS servizi per tuo conto.

Un ruolo collegato al servizio semplifica la configurazione di Lake Formation perché non è necessario creare un ruolo e aggiungere manualmente le autorizzazioni necessarie. Lake Formation definisce le autorizzazioni del suo ruolo collegato ai servizi e, se non diversamente definito, solo Lake Formation 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.

Questo ruolo collegato al servizio si affida ai seguenti servizi per l'assunzione del ruolo:

- `lakeformation.amazonaws.com`

Quando utilizzi un ruolo collegato al servizio nell'account A per registrare una sede Amazon S3 di proprietà dell'account B, la policy del bucket Amazon S3 (una politica basata sulle risorse) nell'account B deve concedere le autorizzazioni di accesso al ruolo collegato al servizio nell'account A.

Per informazioni sull'utilizzo del ruolo collegato al servizio per registrare una posizione di dati, consulta. [Limitazioni dei ruoli legati ai servizi](#)

 Note

Le politiche di controllo del servizio (SCPs) non influiscono sui ruoli collegati ai servizi. Per ulteriori informazioni, consulta [Service control policies \(SCPs\) nella guida](#) per l'AWS Organizations utente.

Autorizzazioni di ruolo collegate al servizio per Lake Formation

Lake Formation utilizza il ruolo collegato al servizio denominato.

`AWSServiceRoleForLakeFormationDataAccess` Questo ruolo fornisce una serie di autorizzazioni Amazon Simple Storage Service (Amazon S3) che consentono al servizio integrato Lake Formation (ad esempio) di accedere alle sedi Amazon Athena registrate. Quando registri una posizione di data lake, devi fornire un ruolo con le read/write autorizzazioni Amazon S3 richieste in quella posizione. Invece di creare un ruolo con le autorizzazioni richieste di Amazon S3, puoi utilizzare questo ruolo collegato al servizio.

La prima volta che nomini il ruolo collegato al servizio come ruolo con cui registrare un percorso, il ruolo collegato al servizio e una nuova policy IAM vengono creati per tuo conto. Lake Formation aggiunge il percorso alla politica in linea e la associa al ruolo collegato ai servizi. Quando registri percorsi successivi con il ruolo collegato al servizio, Lake Formation aggiunge il percorso alla policy esistente.

Dopo aver effettuato l'accesso come amministratore del data lake, registra una posizione di data lake. Quindi, nella console IAM, cerca il ruolo `AWSServiceRoleForLakeFormationDataAccess` e visualizza le policy allegate.

Ad esempio, dopo aver registrato la località `s3://my-kinesis-test/logs`, Lake Formation crea la seguente politica in linea e la allega a `AWSServiceRoleForLakeFormationDataAccess`

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationDataAccessPermissionsForS3",
      "Effect": "Allow",
      "Action": [
```

```

        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject",
        "s3:AbortMultipartUpload",
        "s3:ListMultipartUploadParts"
    ],
    "Resource": [
        "arn:aws:s3:::my-kinesis-test/logs/*"
    ]
},
{
    "Sid": "LakeFormationDataAccessPermissionsForS3ListBucket",
    "Effect": "Allow",
    "Action": [
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads"
    ],
    "Resource": [
        "arn:aws:s3:::my-kinesis-test"
    ]
}
]
}

```

Creare un ruolo legato ai servizi per Lake Formation

Non hai bisogno di creare manualmente un ruolo collegato ai servizi. Quando registri una sede Amazon S3 con Lake Formation nell' AWS API Console di gestione AWS, Lake Formation crea automaticamente il ruolo collegato al servizio. AWS CLI

Important

Questo ruolo collegato ai servizi può apparire nell'account se è stata completata un'operazione in un altro servizio che utilizza le funzionalità supportate dal ruolo. Per ulteriori informazioni, consulta [Un nuovo ruolo è apparso nel mio account IAM](#).

Se elimini questo ruolo collegato al servizio, è possibile ricrearlo seguendo lo stesso processo utilizzato per ricreare il ruolo nell'account. Quando registri una sede Amazon S3 con Lake Formation, Lake Formation crea nuovamente il ruolo collegato al servizio per te.

Puoi anche utilizzare la console IAM per creare un ruolo collegato ai servizi con lo use case Lake Formation. Nella AWS CLI o nell' AWS API, crea un ruolo collegato al servizio con il nome del servizio. `lakeformation.amazonaws.com` Per ulteriori informazioni, consulta [Creazione di un ruolo collegato ai servizi](#) nella Guida per l'utente IAM. Se elimini il ruolo collegato ai servizi, è possibile utilizzare lo stesso processo per crearlo nuovamente.

Modifica di un ruolo collegato ai servizi per Lake Formation

Lake Formation non consente di modificare il ruolo `AWSServiceRoleForLakeFormationDataAccess` collegato al servizio. Dopo avere creato un ruolo collegato al servizio, non sarà possibile modificarne il nome perché varie entità potrebbero farvi riferimento. È possibile tuttavia modificarne la descrizione utilizzando IAM. Per ulteriori informazioni, consulta la sezione [Modifica di un ruolo collegato ai servizi](#) nella Guida per l'utente di IAM.

Eliminazione di un ruolo collegato al servizio per Lake Formation

Se non è più necessario utilizzare una funzionalità o un servizio che richiede un ruolo collegato al servizio, ti consigliamo di eliminare il ruolo. 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 al servizio prima di poterlo eliminare manualmente.

Note

Se il servizio Lake Formation utilizza il ruolo quando si tenta di eliminare le risorse, l'eliminazione potrebbe non riuscire. In questo caso, attendi alcuni minuti e quindi ripeti l'operazione.

Per eliminare le risorse di Lake Formation utilizzate da Lake Formation

- Se hai utilizzato il ruolo collegato al servizio per registrare le sedi Amazon S3 con Lake Formation, prima di eliminare il ruolo collegato al servizio, devi annullare la registrazione della posizione e registrarla nuovamente utilizzando un ruolo personalizzato.

Per eliminare manualmente il ruolo collegato ai servizi mediante IAM

Utilizza la console IAM, l'API o l'AWS CLI per eliminare il ruolo collegato al servizio. `AWSServiceRoleForLakeFormationDataAccess` Per ulteriori informazioni, consulta [Eliminazione del ruolo collegato al servizio](#) nella Guida per l'utente di IAM.

Di seguito sono riportati i requisiti per un ruolo definito dall'utente:

- Quando crei il nuovo ruolo, nella pagina Crea ruolo della console IAM, scegli AWS service, quindi in Choose a use case scegli Lake Formation.

Se crei il ruolo utilizzando un percorso diverso, assicurati che il ruolo abbia una relazione di fiducia con `lakeformation.amazonaws.com`. Per ulteriori informazioni, vedere [Modifica di una politica di attendibilità dei ruoli \(Console\)](#).

- Il ruolo deve avere una politica in linea che conceda le autorizzazioni di Amazon read/write S3 sulla posizione. Di seguito è riportata una politica tipica.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::awsexamplebucket/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::awsexamplebucket"
      ]
    }
  ]
}
```

- Aggiungi la seguente policy di fiducia al ruolo IAM per consentire al servizio Lake Formation di assumere il ruolo e fornire credenziali temporanee ai motori analitici integrati.

Per includere il contesto utente di IAM Identity Center nei CloudTrail log, la policy di fiducia deve disporre dell'autorizzazione per l'azione. `sts:SetContext`

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "DataCatalogViewDefinerAssumeRole1",
      "Effect": "Allow",
      "Principal": {
        "Service": [
          "lakeformation.amazonaws.com"
        ]
      },
      "Action": [
        "sts:AssumeRole",
        "sts:SetContext"
      ]
    }
  ]
}
```

- L'amministratore del data lake che registra la posizione deve disporre dell'`iam:PassRole` autorizzazione per il ruolo.

Di seguito è riportata una politica in linea che concede questa autorizzazione. Sostituiscilo `<account-id>` con un numero di AWS account valido e `<role-name>` sostituiscilo con il nome del ruolo.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "PassRolePermissions",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ]
    }
  ]
}
```



```

        "Resource": [
            "arn:aws:iam::111122223333:role/<role-name>"
        ]
    }
]
}

```

- Per consentire a Lake Formation di aggiungere log in CloudWatch Logs e pubblicare metriche, aggiungi la seguente politica in linea.

Note

La scrittura su CloudWatch Logs comporta un costo.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Sid1",
      "Effect": "Allow",
      "Action": [
        "logs:CreateLogStream",
        "logs:CreateLogGroup",
        "logs:PutLogEvents"
      ],
      "Resource": [
        "arn:aws:logs:us-east-1:111122223333:log-group:/aws-lakeformation-acceleration/*",
        "arn:aws:logs:us-east-1:111122223333:log-group:/aws-lakeformation-acceleration/*:log-stream:*"
      ]
    }
  ]
}

```

Registrazione di una sede Amazon S3

È necessario specificare un ruolo AWS Identity and Access Management (IAM) quando si registra una sede Amazon Simple Storage Service (Amazon S3). Lake Formation assume questo ruolo quando concede credenziali temporanee ai AWS servizi integrati che accedono ai dati in quella posizione.

Important

Evita di registrare un bucket Amazon S3 con Requester pay abilitato. Per i bucket registrati con Lake Formation, il ruolo utilizzato per registrare il bucket viene sempre visualizzato come richiedente. Se al bucket si accede da un altro AWS account, al proprietario del bucket viene addebitato l'accesso ai dati se il ruolo appartiene allo stesso account del proprietario del bucket.

Puoi utilizzare la AWS Lake Formation console, l'API Lake Formation o AWS Command Line Interface (AWS CLI) per registrare una sede Amazon S3.

Prima di iniziare

Rivedi i [requisiti per il ruolo utilizzato per registrare la sede](#).

Per registrare una posizione (console)

Important

Le seguenti procedure presuppongono che la posizione Amazon S3 si trovi nello stesso AWS account del Data Catalog e che i dati in essa contenuti non siano crittografati. Altre sezioni di questo capitolo riguardano la registrazione tra account e la registrazione di posizioni crittografate.

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come utente con l'autorizzazione `lakeformation:RegisterResource` IAM.
2. Nel riquadro di navigazione, in Amministrazione, seleziona Data lake locations.
3. Scegli Registra posizione, quindi scegli Sfoglia per selezionare un percorso Amazon Simple Storage Service (Amazon S3).

4. (Facoltativo, ma fortemente consigliato) Seleziona Rivedi le autorizzazioni della posizione per visualizzare un elenco di tutte le risorse esistenti nella posizione Amazon S3 selezionata e le relative autorizzazioni.

La registrazione della località selezionata potrebbe consentire agli utenti di Lake Formation di accedere ai dati già presenti in quella posizione. La visualizzazione di questo elenco ti aiuta a garantire che i dati esistenti rimangano sicuri.

5. Per il ruolo IAM, scegli il ruolo `AWSServiceRoleForLakeFormationDataAccess` collegato al servizio (predefinito) o un ruolo IAM personalizzato che soddisfi i requisiti di [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#)

Puoi aggiornare una posizione registrata o altri dettagli solo quando la registri utilizzando un ruolo IAM personalizzato. Per modificare una sede registrata utilizzando un ruolo collegato al servizio, è necessario annullare la registrazione della posizione e registrarla nuovamente.

6. Scegli l'opzione Enable Data Catalog Federation per consentire a Lake Formation di assumere un ruolo e vendere credenziali temporanee ai AWS servizi integrati per accedere alle tabelle nei database federati. Se una posizione è registrata con Lake Formation e desideri utilizzare la stessa posizione per una tabella in un database federato, devi registrare la stessa posizione con l'opzione Enable Data Catalog Federation.
7. Scegli la modalità di accesso ibrida per non abilitare le autorizzazioni di Lake Formation per impostazione predefinita. Quando registri una posizione Amazon S3 in modalità di accesso ibrido, puoi abilitare le autorizzazioni Lake Formation optando per i principali database e tabelle in quella posizione.

Per ulteriori informazioni sulla configurazione della modalità di accesso ibrida, consulta [Modalità di accesso ibrida](#)

8. Seleziona Registra posizione.

Per registrare una sede (AWS CLI)

1. Registra una nuova sede con Lake Formation

Questo esempio utilizza un ruolo collegato al servizio per registrare la posizione. È possibile utilizzare l' `--role-arn` argomento invece per fornire il proprio ruolo.

Sostituiscilo `<s3-path>` con un percorso Amazon S3 valido, un numero di account con un AWS account valido e `<s3-access-role>` con un ruolo IAM che dispone delle autorizzazioni per registrare una posizione dati.

 Note

Non puoi modificare le proprietà di una posizione registrata se è registrata utilizzando un ruolo collegato al servizio.

```
aws lakeformation register-resource \  
--resource-arn arn:aws:s3:::<s3-path> \  
--use-service-linked-role
```

L'esempio seguente utilizza un ruolo personalizzato per registrare la posizione.

```
aws lakeformation register-resource \  
--resource-arn arn:aws:s3:::<s3-path> \  
--role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>
```

2. Per aggiornare una località registrata con Lake Formation

Puoi modificare una sede registrata solo se è registrata utilizzando un ruolo IAM personalizzato. Per una sede registrata con un ruolo collegato al servizio, è necessario annullare la registrazione della posizione e registrarla nuovamente. Per ulteriori informazioni, consulta [the section called “Annullamento della registrazione di una sede Amazon S3”](#).

```
aws lakeformation update-resource \  
--role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>\  
--resource-arn arn:aws:s3:::<s3-path>
```

```
aws lakeformation update-resource \  
--resource-arn arn:aws:s3:::<s3-path> \  
--use-service-linked-role
```

3. Registra una posizione dati in modalità di accesso ibrida con federazione

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --hybrid-access-enabled
```

```
aws lakeformation register-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --with-federation
```

```
aws lakeformation update-resource \  
  --resource-arn arn:aws:s3:::<s3-path> \  
  --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \  
  --hybrid-access-enabled
```

Per ulteriori informazioni, consulta Funzionamento [RegisterResource](#) delle API.

Note

Dopo aver registrato una posizione Amazon S3, qualsiasi AWS Glue tabella che punta alla posizione (o a una delle sue sedi secondarie) restituirà il valore del `IsRegisteredWithLakeFormation` parametro come `true` nella chiamata `GetTable`. È noto che le operazioni dell'API Data Catalog, come `GetTables` e `SearchTables` non, aggiornano il valore del `IsRegisteredWithLakeFormation` parametro e restituiscono il valore predefinito, che è `false`. Si consiglia di utilizzare l'`GetTableAPI` per visualizzare il valore corretto del `IsRegisteredWithLakeFormation` parametro.

Registrazione di una posizione Amazon S3 crittografata

Lake Formation si integra con [AWS Key Management Service](#) (AWS KMS) per consentirti di configurare più facilmente altri servizi integrati per crittografare e decrittografare i dati nelle sedi Amazon Simple Storage Service (Amazon S3).

Entrambi sono gestiti dal cliente e sono supportati. AWS KMS keys Chiavi gestite da AWS
Attualmente, il lato client encryption/decryption è supportato solo con Athena.

È necessario specificare un ruolo AWS Identity and Access Management (IAM) quando si registra una sede Amazon S3. Per le sedi Amazon S3 crittografate, il ruolo deve disporre dell'autorizzazione per crittografare e decrittografare i dati con o la politica della AWS KMS key chiave KMS deve concedere le autorizzazioni sulla chiave del ruolo.

 Important

Evita di registrare un bucket Amazon S3 con Requester pay abilitato. Per i bucket registrati con Lake Formation, il ruolo utilizzato per registrare il bucket viene sempre visualizzato come richiedente. Se al bucket si accede da un altro AWS account, al proprietario del bucket viene addebitato l'accesso ai dati se il ruolo appartiene allo stesso account del proprietario del bucket.

Lake Formation utilizza un ruolo collegato al servizio per registrare le posizioni dei dati. [Tuttavia, questo ruolo presenta diverse limitazioni.](#) A causa di questi vincoli, consigliamo invece di creare e utilizzare un ruolo IAM personalizzato per maggiore flessibilità e controllo. Il ruolo personalizzato creato per registrare la sede deve soddisfare i requisiti specificati in. [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#)

 Important

Se hai utilizzato un Chiave gestita da AWS per crittografare la posizione Amazon S3, non puoi utilizzare il ruolo collegato al servizio Lake Formation. È necessario utilizzare un ruolo personalizzato e aggiungere le autorizzazioni IAM sulla chiave del ruolo. I dettagli sono forniti più avanti in questa sezione.

Le seguenti procedure spiegano come registrare una posizione Amazon S3 crittografata con una chiave gestita dal cliente o un. Chiave gestita da AWS

- [Registrazione di una posizione crittografata con una chiave gestita dal cliente](#)
- [Registrazione di una posizione crittografata con un Chiave gestita da AWS](#)

Prima di iniziare

Esamina i [requisiti per il ruolo utilizzato per registrare la sede](#).

Per registrare una posizione Amazon S3 crittografata con una chiave gestita dal cliente

 Note

Se la chiave KMS o la posizione Amazon S3 non si trovano AWS nello stesso account del Data Catalog, segui invece le istruzioni riportate [the section called “Registrazione di una posizione Amazon S3 crittografata tra più account AWS”](#) in.

1. Apri la AWS KMS console su <https://console.aws.amazon.com/kms> e accedi come utente amministrativo AWS Identity and Access Management (IAM) o come utente che può modificare la policy chiave della chiave KMS utilizzata per crittografare la posizione.
2. Nel riquadro di navigazione, scegli Customer managed keys, quindi scegli il nome della chiave KMS desiderata.
3. Nella pagina dei dettagli della chiave KMS, scegli la scheda Politica chiave, quindi esegui una delle seguenti operazioni per aggiungere il tuo ruolo personalizzato o il ruolo collegato al servizio Lake Formation come utente chiave KMS:
 - Se viene visualizzata la visualizzazione predefinita (con le sezioni Amministratori chiave, Eliminazione chiavi, Utenti chiave e Altri AWS account), nella sezione Utenti chiave, aggiungi il tuo ruolo personalizzato o il ruolo collegato al servizio Lake Formation. `AWSServiceRoleForLakeFormationDataAccess`
 - Se viene visualizzata la policy chiave (JSON), modifica la policy per aggiungere il ruolo personalizzato o il ruolo collegato `AWSServiceRoleForLakeFormationDataAccess` al servizio Lake Formation all'oggetto «Consenti l'uso della chiave», come mostrato nell'esempio seguente.

 Note

Se quell'oggetto manca, aggiungilo con le autorizzazioni mostrate nell'esempio. L'esempio utilizza il ruolo collegato al servizio.

```
...  
{
```

```

        "Sid": "Allow use of the key",
        "Effect": "Allow",
        "Principal": {
            "AWS": [
                "arn:aws:iam::111122223333:role/aws-service-role/
lakeformation.amazonaws.com/AWSServiceRoleForLakeFormationDataAccess",
                "arn:aws:iam::111122223333:user/keyuser"
            ]
        },
        "Action": [
            "kms:Encrypt",
            "kms:Decrypt",
            "kms:ReEncrypt*",
            "kms:GenerateDataKey*",
            "kms:DescribeKey"
        ],
        "Resource": "*"
    },
    ...

```

4. Apri la AWS Lake Formation console all'indirizzo. <https://console.aws.amazon.com/lakeformation/> Accedi come amministratore del data lake o come utente con l'autorizzazione `lakeformation:RegisterResource` IAM.
5. Nel riquadro di navigazione, in Amministrazione, scegli Data lake locations.
6. Scegli Registra posizione, quindi scegli Sfoglia per selezionare un percorso Amazon Simple Storage Service (Amazon S3).
7. (Facoltativo, ma fortemente consigliato) Scegli Rivedi le autorizzazioni della sede per visualizzare un elenco di tutte le risorse esistenti nella posizione Amazon S3 selezionata e le relative autorizzazioni.

La registrazione della località selezionata potrebbe consentire agli utenti di Lake Formation di accedere ai dati già presenti in quella posizione. La visualizzazione di questo elenco ti aiuta a garantire che i dati esistenti rimangano sicuri.

8. Per il ruolo IAM, scegli il ruolo `AWSServiceRoleForLakeFormationDataAccess` collegato al servizio (predefinito) o il ruolo personalizzato che soddisfa i. [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#)
9. Scegli Registra posizione.

Per ulteriori informazioni sul ruolo collegato al servizio, consulta [Autorizzazioni di ruolo collegate al servizio per Lake Formation](#).

Per registrare una posizione Amazon S3 crittografata con un Chiave gestita da AWS

⚠ Important

Se la sede Amazon S3 non si trova nello stesso AWS account del Data Catalog, segui invece le istruzioni riportate in [the section called “Registrazione di una posizione Amazon S3 crittografata tra più account AWS”](#).

1. Crea un ruolo IAM da utilizzare per registrare la posizione. Assicurati che soddisfi i requisiti elencati in [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#).
2. Aggiungi la seguente politica in linea al ruolo. Concede le autorizzazioni sulla chiave del ruolo. La Resource specifica deve indicare l'Amazon Resource Name (ARN) di Chiave gestita da AWS. È possibile ottenere l'ARN dalla AWS KMS console. Per ottenere l'ARN corretto, assicurati di accedere alla AWS KMS console con lo stesso AWS account e la stessa regione utilizzati per crittografare la Chiave gestita da AWS posizione.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "arn:aws:kms:us-east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
    }
  ]
}
```

Puoi utilizzare gli alias delle chiavi KMS anziché l'ID della chiave -
`arn:aws:kms:region:account-id:key/alias/your-key-alias`

Per ulteriori informazioni, consulta la sezione [Alias nella AWS KMS](#) sezione della Guida per gli AWS Key Management Service sviluppatori.

3. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come utente con l'autorizzazione `lakeformation:RegisterResource` IAM.
4. Nel riquadro di navigazione, in Amministrazione, scegli Data lake locations.
5. Scegli Registra posizione, quindi scegli Sfoglia per selezionare un percorso Amazon S3.
6. (Facoltativo, ma fortemente consigliato) Scegli Rivedi le autorizzazioni della sede per visualizzare un elenco di tutte le risorse esistenti nella posizione Amazon S3 selezionata e le relative autorizzazioni.

La registrazione della località selezionata potrebbe consentire agli utenti di Lake Formation di accedere ai dati già presenti in quella posizione. La visualizzazione di questo elenco ti aiuta a garantire che i dati esistenti rimangano sicuri.

7. Per il ruolo IAM, scegli il ruolo che hai creato nella fase 1.
8. Scegli la posizione di registrazione.

Registrazione di una sede Amazon S3 in un altro account AWS

AWS Lake Formation consente di registrare sedi Amazon Simple Storage Service (Amazon S3) su più account. AWS Ad esempio, se si AWS Glue Data Catalog trova nell'account A, un utente nell'account A può registrare un bucket Amazon S3 nell'account B.

La registrazione di un bucket Amazon S3 AWS nell'account B utilizzando AWS Identity and Access Management un ruolo (IAM) AWS nell'account A richiede le seguenti autorizzazioni:

- Il ruolo nell'account A deve concedere le autorizzazioni sul bucket dell'account B.
- La politica del bucket nell'account B deve concedere le autorizzazioni di accesso al ruolo nell'account A.

 Important

Evita di registrare un bucket Amazon S3 con Requester pay abilitato. Per i bucket registrati con Lake Formation, il ruolo utilizzato per registrare il bucket viene sempre visualizzato come richiedente. Se al bucket si accede da un altro AWS account, al proprietario del bucket viene addebitato l'accesso ai dati se il ruolo appartiene allo stesso account del proprietario del bucket.

Non puoi utilizzare il ruolo collegato al servizio Lake Formation per registrare una sede in un altro account. È invece necessario utilizzare un ruolo definito dall'utente. Il ruolo deve soddisfare i requisiti di [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#). Per ulteriori informazioni sul ruolo collegato al servizio, consulta [Autorizzazioni di ruolo collegate al servizio per Lake Formation](#).

Prima di iniziare

Esamina i [requisiti per il ruolo utilizzato per registrare la sede](#).

Per registrare una sede in un altro AWS account

 Note

Se la posizione è crittografata, segui [the section called “Registrazione di una posizione Amazon S3 crittografata tra più account AWS”](#) invece le istruzioni riportate in.

La procedura seguente presuppone che un titolare dell'account 1111-2222-3333, che contiene il Data Catalog, desideri registrare il `awsexamplebucket1` bucket Amazon S3, che si trova nell'account 1234-5678-9012.

1. Nell'account 1111-2222-3333, accedi e apri la console IAM all'indirizzo. Console di gestione AWS <https://console.aws.amazon.com/iam/>
2. Crea un nuovo ruolo o visualizza un ruolo esistente che soddisfa i requisiti in. [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#) Assicurati che il ruolo conceda le autorizzazioni di Amazon S3. `awsexamplebucket1`
3. Apri la console Amazon S3 all'indirizzo. <https://console.aws.amazon.com/s3/> Accedi con l'account 1234-5678-9012.
4. Nell'elenco dei nomi del bucket, scegli il nome del bucket, `awsexamplebucket1`

5. Seleziona Autorizzazioni.
6. Nella pagina Autorizzazioni, scegli Bucket Policy.
7. Nell'editor di policy Bucket, incolla la seguente politica. Sostituisci *<role-name>* con il nome del ruolo.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/<role-name>"
      },
      "Action": "s3:ListBucket",
      "Resource": "arn:aws:s3:::awsexamplebucket1"
    },
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/<role-name>"
      },
      "Action": [
        "s3:DeleteObject",
        "s3:GetObject",
        "s3:PutObject"
      ],
      "Resource": "arn:aws:s3:::awsexamplebucket1/*"
    }
  ]
}
```

8. Scegli Save (Salva).
9. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi all'account 1111-2222-3333 come amministratore del data lake o come utente con autorizzazioni sufficienti per registrare le sedi.
10. Nel riquadro di navigazione, in Amministrazione, scegli Posizioni Data lake.
11. Nella pagina delle sedi dei Data lake, scegli Registra posizione.

12. Nella pagina Registra posizione, per il percorso Amazon S3, inserisci il nome del bucket. `s3://awsexamplebucket1`

 Note

Devi digitare il nome del bucket perché i bucket tra account non vengono visualizzati nell'elenco quando scegli Sfoglia.

13. Per il ruolo IAM, scegli il tuo ruolo.
14. Scegli Registra la posizione.

Registrazione di una posizione Amazon S3 crittografata tra più account AWS

AWS Lake Formation si integra con [AWS Key Management Service](#) (AWS KMS) per consentirti di configurare più facilmente altri servizi integrati per crittografare e decrittografare i dati nelle sedi Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3).

Sono supportate entrambe le chiavi gestite dal cliente. Chiavi gestite da AWS Il lato client non encryption/decryption è supportato.

 Important

Evita di registrare un bucket Amazon S3 con Requester pay abilitato. Per i bucket registrati con Lake Formation, il ruolo utilizzato per registrare il bucket viene sempre visualizzato come richiedente. Se al bucket si accede da un altro AWS account, al proprietario del bucket viene addebitato l'accesso ai dati se il ruolo appartiene allo stesso account del proprietario del bucket.

Questa sezione spiega come registrare una sede Amazon S3 nelle seguenti circostanze:

- I dati nella posizione Amazon S3 sono crittografati con una chiave KMS creata in. AWS KMS
- La sede Amazon S3 non si trova nello stesso AWS account di. AWS Glue Data Catalog
- La chiave KMS è o non si trova nello stesso AWS account del Data Catalog.

La registrazione di un bucket Amazon S3 AWS KMS crittografato AWS nell'account B utilizzando AWS Identity and Access Management un ruolo (IAM) AWS nell'account A richiede le seguenti autorizzazioni:

- Il ruolo nell'account A deve concedere le autorizzazioni sul bucket dell'account B.
- La politica del bucket nell'account B deve concedere le autorizzazioni di accesso al ruolo nell'account A.
- Se la chiave KMS è nell'account B, la politica chiave deve concedere l'accesso al ruolo nell'account A e il ruolo nell'account A deve concedere le autorizzazioni sulla chiave KMS.

Nella procedura seguente, si crea un ruolo nell' AWS account che contiene il catalogo dati (l'account A nella discussione precedente). Quindi, si utilizza questo ruolo per registrare la posizione. Lake Formation assume questo ruolo quando accede ai dati sottostanti in Amazon S3. Il ruolo assunto dispone delle autorizzazioni richieste sulla chiave KMS. Di conseguenza, non è necessario concedere le autorizzazioni sulla chiave KMS ai responsabili che accedono ai dati sottostanti con lavori ETL o con servizi integrati come. Amazon Athena

Important

Non puoi utilizzare il ruolo collegato al servizio Lake Formation per registrare una sede in un altro account. È invece necessario utilizzare un ruolo definito dall'utente. Il ruolo deve soddisfare i requisiti di [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#). Per ulteriori informazioni sul ruolo collegato al servizio, consulta [Autorizzazioni di ruolo collegate al servizio per Lake Formation](#).

Prima di iniziare

Esamina i [requisiti per il ruolo utilizzato per registrare la sede](#).

Per registrare una posizione Amazon S3 crittografata tra più account AWS

1. Nello stesso AWS account del Data Catalog, accedi Console di gestione AWS e apri la console IAM all'indirizzo <https://console.aws.amazon.com/iam/>.
2. Crea un nuovo ruolo o visualizza un ruolo esistente che soddisfa i requisiti in [the section called “Requisiti per i ruoli utilizzati per registrare le sedi”](#). Assicurati che il ruolo includa una policy che conceda le autorizzazioni di Amazon S3 sulla location.

3. Se la chiave KMS non si trova nello stesso account del Data Catalog, aggiungi al ruolo una policy in linea che conceda le autorizzazioni richieste sulla chiave KMS. Di seguito è riportata una policy di esempio. Sostituisci la regione e l'ID dell'account con la regione e il numero di account della chiave KMS. Sostituisci *<key-id>* con l'ID della chiave.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "arn:aws:kms:us-east-1:111122223333:key/<key-id>"
    }
  ]
}
```

4. Sulla console Amazon S3, aggiungi una bucket policy che conceda le autorizzazioni Amazon S3 richieste per il ruolo. Di seguito è riportato un esempio di policy di bucket. Sostituisci l'ID dell'AWS account con il numero di account del Data Catalog, *<role-name>* con il nome del tuo ruolo e *<bucket-name>* con il nome del bucket.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/<role-name>"
      },
      "Action": "s3:ListBucket",
    }
  ]
}
```

```

        "Resource": "arn:aws:s3:::<bucket-name>"
    },
    {
        "Effect": "Allow",
        "Principal": {
            "AWS": "arn:aws:iam::111122223333:role/<role-name>"
        },
        "Action": [
            "s3:DeleteObject",
            "s3:GetObject",
            "s3:PutObject"
        ],
        "Resource": "arn:aws:s3:::<bucket-name>/*"
    }
]
}

```

5. In AWS KMS, aggiungi il ruolo come utente della chiave KMS.
 - a. Apri la AWS KMS console in <https://console.aws.amazon.com/kms>. Quindi, accedi come utente amministratore o come utente che può modificare la politica chiave della chiave KMS utilizzata per crittografare la posizione.
 - b. Nel riquadro di navigazione, scegli Customer managed keys, quindi scegli il nome della chiave KMS.
 - c. Nella pagina dei dettagli della chiave KMS, nella scheda Politica chiave, se la visualizzazione JSON della politica chiave non viene visualizzata, scegli Passa alla visualizzazione delle politiche.
 - d. Nella sezione Key policy, scegli Modifica e aggiungi l'Amazon Resource Name (ARN) del ruolo all'Allow use of the keyoggetto, come mostrato nell'esempio seguente.

 Note

Se quell'oggetto manca, aggiungilo con le autorizzazioni mostrate nell'esempio.

```

...
{
    "Sid": "Allow use of the key",
    "Effect": "Allow",
    "Principal": {

```



```

        "AWS": [
            "arn:aws:iam::<catalog-account-id>:role/<role-name>"
        ],
    },
    "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
    ],
    "Resource": "*"
},
...

```

Per ulteriori informazioni, consulta [Consentire agli utenti di altri account di utilizzare una chiave KMS nella Guida](#) per gli AWS Key Management Service sviluppatori.

6. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi all' AWS account Data Catalog come amministratore del data lake.
7. Nel riquadro di navigazione, in Amministrazione, scegli Posizioni Data lake.
8. Scegli Registra posizione.
9. Nella pagina Registra posizione, per il percorso Amazon S3, inserisci il percorso della posizione come. **s3://<bucket>/<prefix>** Sostituiscilo **<bucket>** con il nome del bucket e **<prefix>** con il resto del percorso della posizione.

Note

È necessario digitare il percorso perché i bucket tra account non vengono visualizzati nell'elenco quando si sceglie Sfoglia.

10. Per il ruolo IAM, scegli il ruolo dalla Fase 2.
11. Scegli Registra posizione.

Annullamento della registrazione di una sede Amazon S3

Puoi annullare la registrazione di una sede Amazon Simple Storage Service (Amazon S3) se non desideri più che venga gestita da Lake Formation. L'annullamento della registrazione di una sede

non influisce sulle autorizzazioni di localizzazione dei dati di Lake Formation concesse in quella località. È possibile registrare nuovamente una sede che è stata annullata e le autorizzazioni relative alla localizzazione dei dati restano valide. È possibile utilizzare un ruolo diverso per registrare nuovamente la posizione.

Per annullare la registrazione di una posizione (console)

1. Apri la AWS Lake Formation console all'indirizzo. <https://console.aws.amazon.com/lakeformation/> Accedi come amministratore del data lake o come utente con l'autorizzazione `lakeformation:RegisterResource` IAM.
2. Nel riquadro di navigazione, in Amministrazione, scegli Data lake locations.
3. Seleziona una posizione e nel menu Azioni scegli Rimuovi.
4. Quando viene richiesta la conferma, scegli Rimuovi.

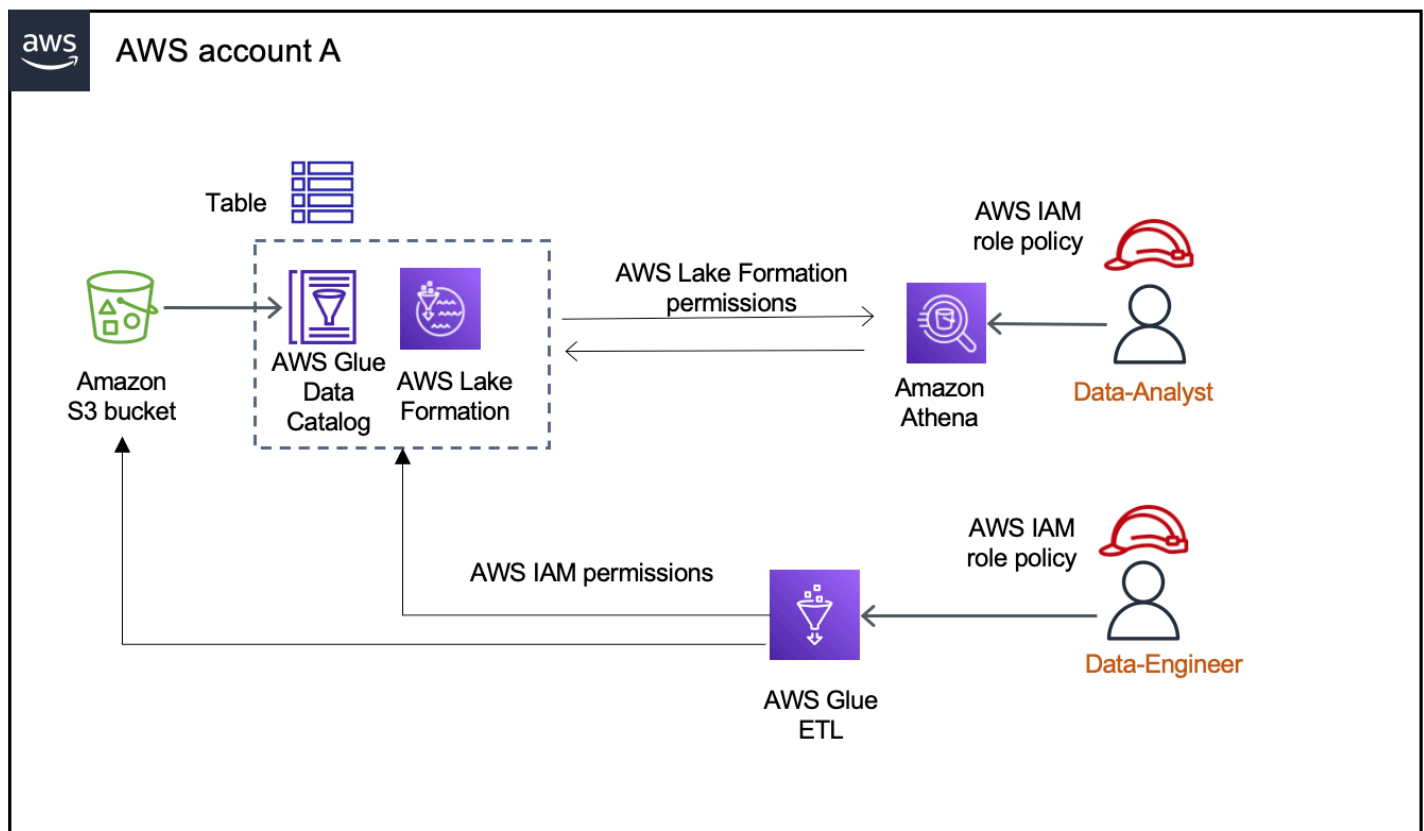
Modalità di accesso ibrida

AWS Lake Formation la modalità di accesso ibrida supporta due percorsi di autorizzazione per gli stessi oggetti. AWS Glue Data Catalog

Nel primo percorso, Lake Formation ti consente di selezionare principi specifici e concedere loro le autorizzazioni di Lake Formation per accedere a cataloghi, database, tabelle e visualizzazioni attivando il consenso. Il secondo percorso consente a tutti gli altri principi di accedere a queste risorse tramite le politiche e le azioni principali IAM predefinite per Amazon AWS Glue S3.

Quando registri una sede Amazon S3 con Lake Formation, hai la possibilità di applicare le autorizzazioni di Lake Formation per tutte le risorse in questa sede o utilizzare la modalità di accesso ibrida. Per impostazione predefinita, la modalità di accesso ibrida applica solo le `CREATE_TABLE` autorizzazioni. `CREATE_PARTITION` `UPDATE_TABLE` Quando una sede Amazon S3 è in modalità ibrida, puoi abilitare le autorizzazioni Lake Formation attivando i principi per gli oggetti Data Catalog in quella posizione. Significa che sia le autorizzazioni Lake Formation che le autorizzazioni IAM possono controllare l'accesso a quei dati. Ciò significa che i responsabili che hanno aderito richiederanno sia le autorizzazioni Lake Formation che le autorizzazioni IAM per accedere ai dati, mentre i non-opted-in principi continueranno ad accedere ai dati utilizzando solo le autorizzazioni IAM.

Pertanto, la modalità di accesso ibrido offre la flessibilità necessaria per abilitare selettivamente Lake Formation per cataloghi, database e tabelle nel tuo Data Catalog per un set specifico di utenti senza interrompere l'accesso per altri utenti o carichi di lavoro esistenti.



Per considerazioni e limitazioni, vedere [Considerazioni e limitazioni della modalità di accesso ibrido](#).

Termini e definizioni

Di seguito sono riportate le definizioni delle risorse di Data Catalog in base alla modalità di impostazione delle autorizzazioni di accesso:

Risorsa Lake Formation

Una risorsa registrata presso Lake Formation. Gli utenti richiedono le autorizzazioni di Lake Formation per accedere alla risorsa.

AWS Glue risorsa

Una risorsa che non è registrata presso Lake Formation. Gli utenti richiedono solo le autorizzazioni IAM per accedere alla risorsa perché dispone di autorizzazioni di IAMAllowedPrincipals gruppo. Le autorizzazioni di Lake Formation non vengono applicate.

Per ulteriori informazioni sulle autorizzazioni IAMAllowedPrincipals di gruppo, vedere.

[Autorizzazioni per i metadati](#)

Risorsa ibrida

Una risorsa registrata in modalità di accesso ibrida. In base agli utenti che accedono alla risorsa, la risorsa passa dinamicamente dall'essere una risorsa Lake Formation a una AWS Glue risorsa.

Casi d'uso comuni della modalità di accesso ibrido

È possibile utilizzare la modalità di accesso ibrido per fornire l'accesso in scenari di condivisione dei dati con account singolo e tra account:

Scenari con account singolo

- Convertire una AWS Glue risorsa in una risorsa ibrida: in questo scenario, attualmente non stai utilizzando Lake Formation ma desideri adottare le autorizzazioni Lake Formation per gli oggetti Data Catalog. Quando registri la posizione Amazon S3 in modalità di accesso ibrido, puoi concedere le autorizzazioni di Lake Formation agli utenti che optano per database e tabelle specifici che puntano a quella posizione.
- Convertire una risorsa Lake Formation in una risorsa ibrida: attualmente, utilizzi le autorizzazioni di Lake Formation per controllare l'accesso a un database Data Catalog, ma desideri fornire l'accesso a nuovi principali utilizzando le autorizzazioni IAM per Amazon S3 e AWS Glue senza interrompere le autorizzazioni Lake Formation esistenti.

Quando aggiorni la registrazione di una posizione dati alla modalità di accesso ibrida, i nuovi responsabili possono accedere al database Data Catalog puntando alla posizione Amazon S3 utilizzando le policy di autorizzazione IAM senza interrompere le autorizzazioni Lake Formation degli utenti esistenti.

Prima di aggiornare la registrazione della posizione dei dati per abilitare la modalità di accesso ibrida, devi prima attivare i principali che attualmente accedono alla risorsa con le autorizzazioni di Lake Formation.

Questo serve a prevenire potenziali interruzioni del flusso di lavoro corrente.

È inoltre necessario concedere al `IAMAllowedPrincipal` gruppo l'Superautorizzazione per le tabelle del database.

Scenari di condivisione dei dati tra account

- Condividi AWS Glue risorse utilizzando la modalità di accesso ibrido: in questo scenario, l'account produttore dispone di tabelle in un database che sono attualmente condivise con un account

consumer utilizzando le politiche di autorizzazione IAM per Amazon S3 AWS Glue e le azioni. La posizione dei dati del database non è registrata con Lake Formation.

Prima di registrare la posizione dei dati in modalità di accesso ibrido, è necessario aggiornare le impostazioni della versione dell'account Cross alla versione 4. La versione 4 fornisce le nuove politiche di AWS RAM autorizzazione necessarie per la condivisione tra account quando il `IAMAllowedPrincipal` gruppo dispone dell'Superautorizzazione sulla risorsa. Per le risorse con autorizzazioni di `IAMAllowedPrincipal` gruppo, puoi concedere le autorizzazioni di Lake Formation agli account esterni e consentire loro di utilizzare le autorizzazioni Lake Formation. L'amministratore del data lake nell'account del destinatario può concedere le autorizzazioni di Lake Formation ai responsabili dell'account e autorizzarli a far valere le autorizzazioni di Lake Formation.

- Condividi le risorse di Lake Formation utilizzando la modalità di accesso ibrida: attualmente, l'account produttore contiene tabelle in un database che vengono condivise con un account consumatore che applica le autorizzazioni di Lake Formation. La posizione dei dati del database è registrata presso Lake Formation.

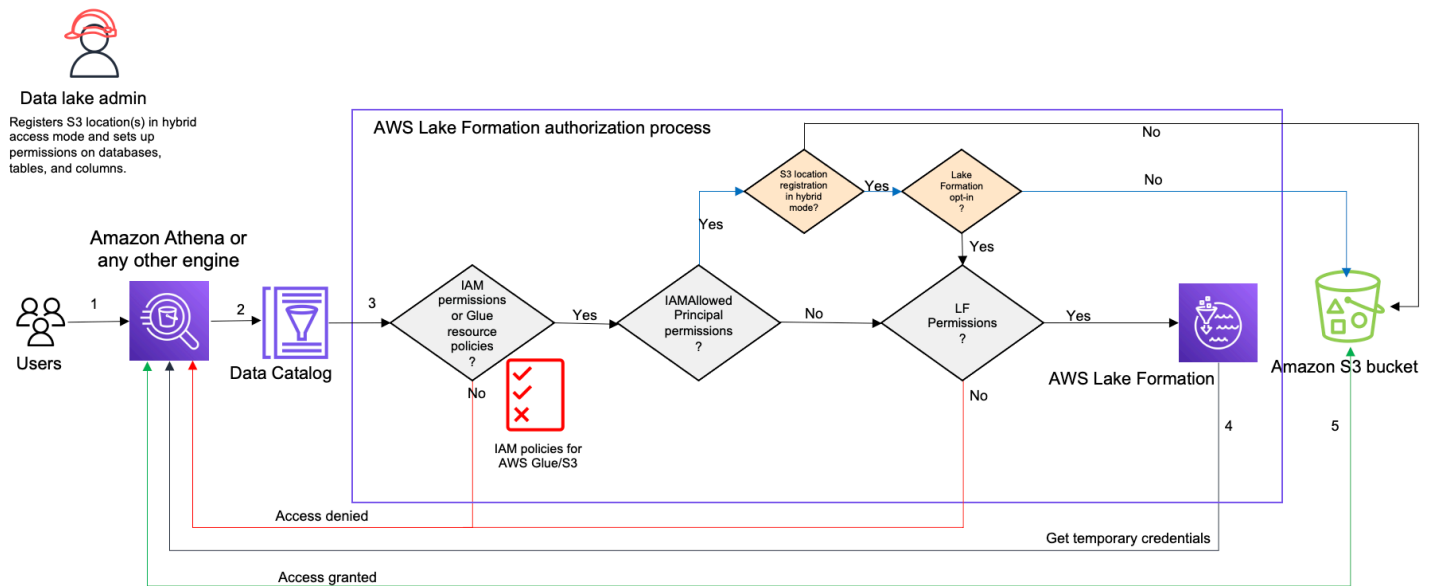
In questo caso, puoi aggiornare la registrazione della posizione Amazon S3 alla modalità di accesso ibrida e condividere i dati di Amazon S3 e i metadati di Data Catalog utilizzando le policy dei bucket di Amazon S3 e le politiche delle risorse di Data Catalog con i principali dell'account consumatore. È necessario concedere nuovamente le autorizzazioni esistenti di Lake Formation e attivare i principali prima di aggiornare la registrazione delle sedi Amazon S3. Inoltre, devi concedere al gruppo Super l'autorizzazione per le tabelle del database. `IAMAllowedPrincipals`

Argomenti

- [Come funziona la modalità di accesso ibrida](#)
- [Configurazione della modalità di accesso ibrida: scenari comuni](#)
- [Rimozione di principi e risorse dalla modalità di accesso ibrida](#)
- [Visualizzazione dei principali e delle risorse in modalità di accesso ibrida](#)
- [Risorse aggiuntive](#)

Come funziona la modalità di accesso ibrida

Il diagramma seguente mostra come funziona l'autorizzazione di Lake Formation in modalità di accesso ibrido quando si interrogano le risorse del Data Catalog.



Prima di accedere ai dati nel data lake, un amministratore del data lake o un utente con autorizzazioni amministrative configura le politiche utente delle singole tabelle di Data Catalog per consentire o negare l'accesso alle tabelle del Data Catalog. Quindi, un principale che dispone delle autorizzazioni per eseguire l'`RegisterResource` operazione registra la posizione Amazon S3 della tabella con Lake Formation in modalità di accesso ibrido. Se una posizione dati non è registrata con Lake Formation, l'amministratore concede le autorizzazioni di Lake Formation a utenti specifici sui database e sulle tabelle del Data Catalog e li autorizza a utilizzare le autorizzazioni Lake Formation per tali database e tabelle in modalità di accesso ibrido.

1. Invia una query: un principale invia una query o uno script ETL utilizzando un servizio integrato come Amazon Athena, Amazon EMR o AWS Glue Amazon Redshift Spectrum.
2. Richiede dati: il motore analitico integrato identifica la tabella richiesta e invia la richiesta di metadati al Data Catalog (,). `GetTable` `GetDatabase`
3. Verifica le autorizzazioni: il Data Catalog verifica le autorizzazioni di accesso del responsabile dell'interrogazione con Lake Formation.
 - a. Se alla tabella non sono allegate autorizzazioni di `IAMAllowedPrincipals` gruppo, vengono applicate le autorizzazioni di Lake Formation.
 - b. Se il principale ha scelto di utilizzare le autorizzazioni di Lake Formation nella modalità di accesso ibrida e alla tabella sono allegate le autorizzazioni di `IAMAllowedPrincipals` gruppo, le autorizzazioni di Lake Formation vengono applicate. Il motore di query applica i filtri ricevuti da Lake Formation e restituisce i dati all'utente.

- c. Se la posizione della tabella non è registrata con Lake Formation e il principale non ha scelto di utilizzare le autorizzazioni di Lake Formation in modalità di accesso ibrido, il Data Catalog verifica se alla tabella sono associate autorizzazioni di IAMAllowedPrincipals gruppo. Se questa autorizzazione esiste nella tabella, tutti i principali dell'account ottengono Super i All permessi sulla tabella.

La vendita di credenziali di Lake Formation non è disponibile, anche se è stato attivato, a meno che la posizione dei dati non sia registrata con Lake Formation.

4. Ottieni credenziali: il Data Catalog controlla e fa sapere al motore se la posizione della tabella è registrata o meno con Lake Formation. Se i dati sottostanti sono registrati con Lake Formation, il motore analitico richiede a Lake Formation le credenziali temporanee per accedere ai dati nel bucket Amazon S3.
5. Ottieni dati: se il responsabile è autorizzato ad accedere ai dati della tabella, Lake Formation fornisce l'accesso temporaneo al motore analitico integrato. Utilizzando l'accesso temporaneo, il motore analitico recupera i dati da Amazon S3 ed esegue i filtri necessari come il filtraggio di colonne, righe o celle. Quando il motore termina l'esecuzione del lavoro, restituisce i risultati all'utente. Questo processo è denominato distribuzione di credenziali. Per ulteriori informazioni, vedere [Integrazione con Lake Formation](#).

6.

Se la posizione dei dati della tabella non è registrata con Lake Formation, la seconda chiamata dal motore di analisi viene effettuata direttamente ad Amazon S3. La policy del bucket Amazon S3 interessata e la politica utente IAM vengono valutate per l'accesso ai dati. Ogni volta che si utilizzano le policy IAM, assicurati di seguire le best practice IAM. Per ulteriori informazioni, consulta le [best practice di sicurezza in IAM nella IAM User Guide](#).

Configurazione della modalità di accesso ibrida: scenari comuni

Come per le autorizzazioni di Lake Formation, in genere sono disponibili due tipi di scenari in cui è possibile utilizzare la modalità di accesso ibrida per gestire l'accesso ai dati: fornire l'accesso ai principali all'interno di uno Account AWS e fornire l'accesso a un principale Account AWS o esterno.

Questa sezione fornisce istruzioni per configurare la modalità di accesso ibrida nei seguenti scenari:

Gestisci le autorizzazioni in modalità di accesso ibrida all'interno di un'unica soluzione Account AWS

- [Conversione di una AWS Glue risorsa in una risorsa ibrida](#) — Attualmente stai fornendo l'accesso alle tabelle in un database per tutti i principali del tuo account utilizzando le autorizzazioni IAM per

Amazon S3 AWS Glue e desideri adottare Lake Formation per gestire le autorizzazioni in modo incrementale.

- [Conversione di una risorsa di Lake Formation in una risorsa ibrida](#) — Attualmente stai utilizzando Lake Formation per gestire l'accesso alle tabelle in un database per tutti i principali del tuo account, ma desideri utilizzare Lake Formation solo per i principali specifici. Desideri fornire l'accesso a nuovi principali utilizzando le autorizzazioni IAM per Amazon S3 sullo stesso database AWS Glue e sulle stesse tabelle.

Gestisci le autorizzazioni in modalità di accesso ibrida su s Account AWS

- [Condivisione di una AWS Glue risorsa utilizzando la modalità di accesso ibrida](#) — Al momento non stai utilizzando Lake Formation per gestire le autorizzazioni per una tabella, ma desideri applicare le autorizzazioni di Lake Formation per fornire l'accesso ai principali in un altro account.
- [Condivisione di una risorsa Lake Formation utilizzando la modalità di accesso ibrida](#) — Stai utilizzando Lake Formation per gestire l'accesso a una tabella ma desideri fornire l'accesso ai principali in un altro account utilizzando le autorizzazioni IAM per Amazon S3 sullo stesso database AWS Glue e sulle stesse tabelle.

Configurazione della modalità di accesso ibrida: passaggi di alto livello

1. Registra la posizione dei dati di Amazon S3 con Lake Formation selezionando la modalità di accesso ibrida.
2. I responsabili devono avere DATA_LOCATION l'autorizzazione su una posizione di data lake per creare tabelle o database di Data Catalog che puntano a quella posizione.
3. Imposta l'impostazione della versione per più account sulla versione 4.
4. Concedi autorizzazioni dettagliate a utenti o ruoli IAM specifici su database e tabelle. Allo stesso tempo, assicurati di impostare Super o All autorizzare il IAMAllowedPrincipals gruppo sul database e tutte le tabelle del database o solo alcune di esse.
5. Scegli i principi e le risorse. Gli altri responsabili dell'account possono continuare ad accedere ai database e alle tabelle utilizzando le policy di autorizzazione IAM AWS Glue e le azioni di Amazon S3.
6. Opzionalmente, pulisci le politiche di autorizzazione IAM per Amazon S3 per i principali che hanno scelto di utilizzare le autorizzazioni Lake Formation.

Prerequisiti per la configurazione della modalità di accesso ibrida

Di seguito sono riportati i prerequisiti per l'impostazione della modalità di accesso ibrida:

Note

Consigliamo a un amministratore di Lake Formation di registrare la posizione Amazon S3 in modalità di accesso ibrida e di attivare i principali e le risorse.

1. Concedi l'autorizzazione all'ubicazione dei dati (DATA_LOCATION_ACCESS) per creare risorse Data Catalog che puntino alle sedi Amazon S3. Le autorizzazioni di localizzazione dei dati controllano la capacità di creare cataloghi, database e tabelle di Data Catalog che puntano a particolari posizioni Amazon S3.
2. Per condividere le risorse di Data Catalog con un altro account in modalità di accesso ibrida (senza rimuovere le autorizzazioni di IAMAllowedPrincipals gruppo dalla risorsa), devi aggiornare le impostazioni della versione dell'account Cross alla versione 4. Per aggiornare la versione utilizzando la console Lake Formation, scegli Versione 4 nelle impostazioni della versione dell'account Cross nella pagina delle impostazioni del Data Catalog.

Puoi anche usare il `put-data-lake-settings` AWS CLI comando per impostare il `CROSS_ACCOUNT_VERSION` parametro sulla versione 4:

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
file://settings
{
  "DataLakeAdmins": [
    {
      "DataLakePrincipalIdentifier": "arn:aws:iam::<111122223333>:user/<user-name>"
    }
  ],
  "CreateDatabaseDefaultPermissions": [],
  "CreateTableDefaultPermissions": [],
  "Parameters": {
    "CROSS_ACCOUNT_VERSION": "4"
  }
}
```

3.

Per concedere le autorizzazioni su più account in modalità di accesso ibrida, il concedente deve disporre delle autorizzazioni IAM richieste per e i servizi. AWS Glue AWS RAM La policy AWS `AWSLakeFormationCrossAccountManager` gestita concede le autorizzazioni richieste. Per abilitare la condivisione dei dati tra account in modalità di accesso ibrido, abbiamo aggiornato la policy `AWSLakeFormationCrossAccountManager` gestita aggiungendo due nuove autorizzazioni IAM:

- ram: `ListResourceSharePermissions`
- ram: `AssociateResourceSharePermission`

Note

Se non utilizzi la politica AWS gestita per il ruolo di concedente, aggiungi le politiche precedenti alle politiche personalizzate.

Ubicazione del bucket Amazon S3 e accesso utente

Quando crei un catalogo, un database o una tabella in AWS Glue Data Catalog, puoi specificare la posizione del bucket Amazon S3 dei dati sottostanti e registrarli con Lake Formation. Le tabelle seguenti descrivono come funzionano le autorizzazioni per gli utenti AWS Glue e gli utenti di Lake Formation (responsabili) in base alla posizione dei dati in Amazon S3 della tabella o del database.

Sede Amazon S3 registrata presso Lake Formation

Ubicazione di un database in Amazon S3	AWS Glue utenti	Utenti di Lake Formation
Registrato a Lake Formation (in modalità di accesso ibrido o in modalità Lake Formation)	read/write Accedi alla posizione dati di Amazon S3 ereditando le autorizzazioni dal gruppo IAMAllowed Principal (super accesso).	Eredita le autorizzazioni per creare tabelle dall'autorizzazione CREATE TABLE concessa.
Nessuna sede Amazon S3 associata	Richiedi l'autorizzazione esplicita DATA LOCATION per eseguire le istruzioni CREATE TABLE e INSERT TABLE.	Richiedi l'autorizzazione esplicita DATA LOCATION per eseguire le istruzioni CREATE TABLE e INSERT TABLE.

IsRegisteredWithLakeFormation proprietà della tabella

La `IsRegisteredWithLakeFormation` proprietà di una tabella indica se la posizione dei dati della tabella è registrata con Lake Formation per il richiedente. Se la modalità di autorizzazione della posizione è registrata come Lake Formation, la `IsRegisteredWithLakeFormation` proprietà è valida `true` per tutti gli utenti che accedono alla posizione dei dati perché tutti gli utenti sono considerati iscritti a quella tabella. Se la posizione è registrata in modalità di accesso ibrida, il valore è impostato `true` solo per gli utenti che hanno optato per quella tabella.

Funzionamento di **IsRegisteredWithLakeFormation**

Modalità di autorizzazione	Utenti/ruoli	IsRegisteredWithLakeFormation	Description
Lake Formation	Tutti	True	Quando una posizione viene registrata con Lake Formation, la <code>IsRegisteredWithLakeFormation</code> proprietà verrà impostata su <code>true</code> per tutti gli utenti. Ciò significa che le autorizzazioni definite in Lake Formation si applicano alla sede registrata. La vendita di credenziali sarà effettuata da Lake Formation.
Modalità di accesso ibrida	Ha aderito	True	Per gli utenti che hanno scelto di utilizzare Lake Formation per

Modalità di autorizzazione	Utenti/ruoli	IsRegisteredWithLakeFormation	Description
			l'accesso e la governance dei dati per una tabella, la IsRegisteredWithLakeFormation proprietà verrà impostata su true per quella tabella. Sono soggetti alle politiche di autorizzazione definite in Lake Formation per la sede registrata.
Modalità di accesso ibrida	Non è stata attivata	False	Per gli utenti che non hanno scelto di utilizzare le autorizzazioni di Lake Formation , la IsRegisteredWithLakeFormation proprietà è impostata su. false Non sono soggetti alle politiche di autorizzazione definite in Lake Formation per la sede registrata. Gli utenti seguiranno invece le politiche di autorizzazione di Amazon S3.

Conversione di una AWS Glue risorsa in una risorsa ibrida

Segui questi passaggi per registrare una sede Amazon S3 in modalità di accesso ibrido e inserire nuovi utenti Lake Formation senza interrompere l'accesso ai dati degli utenti esistenti del Data Catalog.

Descrizione dello scenario: la posizione dei dati non è registrata con Lake Formation e l'accesso degli utenti al database e alle tabelle di Data Catalog è determinato dalle politiche di autorizzazione IAM per Amazon AWS Glue S3 e dalle azioni.

Per impostazione predefinita, il IAMAllowedPrincipals gruppo dispone di Super autorizzazioni su tutte le tabelle del database.

Per abilitare la modalità di accesso ibrido per una posizione dati non registrata con Lake Formation

1. Registra una posizione Amazon S3 abilitando la modalità di accesso ibrida.

Console

1. Accedi alla [console Lake Formation](#) come amministratore del data lake.
2. Nel riquadro di navigazione, scegli Posizioni dei data lake in Amministrazione.
3. Scegli Registra posizione.

Register location

Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

Amazon S3 path

Choose an Amazon S3 path for your data lake.

e.g.: s3://bucket/prefix/

Browse

Review location permissions - strongly recommended

Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

Review location permissions

IAM role

To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

AWSServiceRoleForLakeFormationDataAccess ▼

 Do not select the service linked role if you plan to use EMR.

☐ **Enable Data Catalog Federation**

Checking this box will allow Lake Formation to assume a role to access tables in a federated database.

Permission mode

Select the permission mode you want to use to manage access.

☒ **Hybrid access mode - new**

Lake Formation permissions can co-exist with IAM permission policies for AWS Glue and S3 actions to manage access. [Learn more](#) 

☐ **Lake Formation**

Only Lake Formation permissions are enforced.

Cancel

Register location

4. Nella finestra Registra posizione, scegli il percorso Amazon S3 che desideri registrare con Lake Formation.
5. Per il ruolo IAM, scegli il ruolo **AWSServiceRoleForLakeFormationDataAccess** collegato al servizio (predefinito) o un ruolo IAM personalizzato ruolo che soddisfa i requisiti di. [Requisiti per i ruoli utilizzati per registrare le sedi](#)

6. Scegli la modalità di accesso ibrida per applicare politiche di controllo degli accessi granulari di Lake Formation ai principali opt-in e ai database e alle tabelle di Data Catalog che puntano alla posizione registrata.

Scegli Lake Formation per consentire a Lake Formation di autorizzare le richieste di accesso alla posizione registrata.

7. Scegli Registra posizione.

AWS CLI

Di seguito è riportato un esempio di registrazione di una posizione dati con Lake Formation HybridAccessEnabled con:true/false. Il valore predefinito per il parametro è false.

HybridAccessEnabled Sostituisci il percorso, il nome del ruolo e l'ID AWS account di Amazon S3 con valori validi.

```
aws lakeformation register-resource --cli-input-json file:file path
json:
{
  "ResourceArn": "arn:aws:s3:::s3-path",
  "UseServiceLinkedRole": false,
  "RoleArn": "arn:aws:iam::<123456789012>:role/<role-name>",
  "HybridAccessEnabled": true
}
```

2. Concedi le autorizzazioni e attiva i principali per utilizzare le autorizzazioni di Lake Formation per le risorse in modalità di accesso ibrido

Prima di attivare i principali e le risorse in modalità di accesso ibrido, verifica che Super i All permessi di IAMAllowedPrincipals raggruppamento esistano sui database e sulle tabelle la cui posizione è registrata con Lake Formation in modalità di accesso ibrida.

Note

Non puoi concedere l'autorizzazione al IAMAllowedPrincipals gruppo All tables all'interno di un database. È necessario selezionare ogni tabella separatamente dal menu a discesa e concedere le autorizzazioni. Inoltre, quando crei nuove tabelle nel database, puoi scegliere di utilizzarle Use only IAM access control for new tables in new databases nelle Impostazioni del catalogo dati. Questa opzione

concede automaticamente l'Superautorizzazione al IAMAllowedPrincipals gruppo quando si creano nuove tabelle all'interno del database.

Console

1. Nella console Lake Formation, in Data Catalog, scegli Cataloghi, Database o Tabelle.
2. Seleziona un catalogo, un database o una tabella dall'elenco e scegli Concedi dal menu Azioni.
3. Scegliete i principali per concedere le autorizzazioni sul database, sulle tabelle e sulle colonne utilizzando il metodo della risorsa denominata o i tag LF.

In alternativa, scegli Autorizzazioni dati, seleziona i principali a cui concedere le autorizzazioni dall'elenco e scegli Concedi.

Per maggiori dettagli sulla concessione delle autorizzazioni per i dati, consulta.

[Concessione delle autorizzazioni per le risorse di Data Catalog](#)

Note

Se concedi l'autorizzazione Create table a un principale, devi anche concedere le autorizzazioni per la localizzazione dei dati (DATA_LOCATION_ACCESS) al principale. Questa autorizzazione non è necessaria per aggiornare le tabelle. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

4. Quando si utilizza il metodo Named resource per concedere le autorizzazioni, l'opzione per attivare i principali e le risorse è disponibile nella sezione inferiore della pagina Concedi l'autorizzazione ai dati.

Scegli Rendi immediatamente effettive le autorizzazioni di Lake Formation per abilitare le autorizzazioni di Lake Formation per i committenti e le risorse.

Hybrid access mode

In hybrid access mode, Lake Formation and IAM policies for AWS Glue and S3 work together.

- ☒ **Make Lake Formation permissions effective immediately**
Lake Formation permissions are enforced for databases, tables, and principals.

ⓘ You might get access denied.

If the checkbox is selected, your Lake Formation permissions are enforced. Make sure that you've completed the required setup for Lake Formation for the permissions to work. If the checkbox is clear, you can go to [hybrid access mode](#) [↗](#) to add resources and principals. [Learn more](#) [↗](#)

Cancel

Grant

5. Scegli Concessione.

Quando si attiva il principale A sulla tabella A che punta a una posizione dei dati, consente al principale A di accedere alla posizione di questa tabella utilizzando le autorizzazioni di Lake Formation se la posizione dei dati è registrata in modalità ibrida.

AWS CLI

Di seguito è riportato un esempio di attivazione di un principale e di una tabella in modalità di accesso ibrida. Sostituisci il nome del ruolo, l'id AWS dell'account, il nome del database e il nome della tabella con valori validi.

```
aws lakeformation create-lake-formation-opt-in --cli-input-json file://file path
json:
{
  "Principal": {
    "DataLakePrincipalIdentifier":
    "arn:aws:iam::<123456789012>:role/<hybrid-access-role>"
  },
  "Resource": {
    "Table": {
      "CatalogId": "<123456789012>",
      "DatabaseName": "<hybrid_test>",
      "Name": "<hybrid_test_table>"
    }
  }
}
```

- a. (Optional) Se scegli LF-Tags per concedere le autorizzazioni, puoi attivare i principali per utilizzare i permessi di Lake Formation in un passaggio separato. Puoi farlo scegliendo la modalità di accesso ibrido in Autorizzazioni nella barra di navigazione a sinistra.
- b. Nella sezione inferiore della pagina della modalità di accesso ibrida, scegli Aggiungi per aggiungere risorse e principali alla modalità di accesso ibrida.
- c. Nella pagina Aggiungi risorse e principali, scegli i cataloghi, i database e le tabelle registrati in modalità di accesso ibrido.

Puoi scegliere All tables all'interno di un database per concedere l'accesso.

Add resources and principals

Choose databases, tables, and principals to add in hybrid access mode. Lake Formation permissions will be enforced.

[Learn more](#) 

Resources

Catalogs

Databases

Select one or more databases.

Choose databases

testdb 

Tables - optional

Select one or more tables.

Choose tables

testtable 

- d. Scegli i titolari e accetta di utilizzare le autorizzazioni di Lake Formation in modalità di accesso ibrida.
 - Responsabili: puoi scegliere utenti e ruoli IAM nello stesso account o in un altro account. Puoi anche scegliere utenti e gruppi SAML.
 - Attributi: seleziona gli attributi per concedere le autorizzazioni in base agli attributi.

Principals

☐ **IAM users and roles**

Users or roles from this AWS account.

☐ **External accounts**

AWS account, AWS organization or IAM principal outside of this account.

☐ **SAML users and groups**

SAML users and group or QuickSight ARNs.

☒ **Attributes - new**

Define permissions using attributes (IAM session tags) with a Cedar policy expression.

Attribute expression

Add tag key-value pairs from IAM roles or [IAM session tags](#). Separate values with commas. **AND** logic applies between pairs, **OR** logic within keys. Verify tag entries for accuracy.

Key

Department

Value

Sales

Add key-value pair

Expression review

The attribute expression above will be written in the following Cedar policy. Use CLI/SDK to edit or write custom Cedar policy.

```
permit (principal, action, resource)
when {
  context.iam.principalTags.hasTag("Department")
  && context.iam.principalTags.getTag("Department") == "Sales"
};
```

Permission scope

☒ This account

☐ External account

- Inserisci la coppia chiave-valore per creare una concessione basata sugli attributi. Controlla l'espressione della policy Cedar sulla console. Per ulteriori informazioni su Cedar, vedi [Cos'è Cedar? | Cedar Policy Language Reference](#). GuideLink
- Scegliere Aggiungi.

A tutti gli IAM roles/users con attributi corrispondenti viene concesso l'accesso.

e. Scegliere Aggiungi.

Conversione di una risorsa di Lake Formation in una risorsa ibrida

Nei casi in cui attualmente utilizzi le autorizzazioni di Lake Formation per i database e le tabelle del Data Catalog, puoi modificare le proprietà di registrazione della posizione per abilitare la modalità di accesso ibrida. Ciò consente di fornire ai nuovi principali l'accesso alle stesse risorse utilizzando le politiche di autorizzazione IAM per Amazon S3 AWS Glue e le azioni senza interrompere le autorizzazioni esistenti di Lake Formation.

Descrizione dello scenario: i passaggi seguenti presuppongono che tu abbia una posizione dei dati registrata con Lake Formation e che tu abbia impostato le autorizzazioni per i principali su database, tabelle o colonne che puntano a quella posizione. Se la posizione è stata registrata con un ruolo collegato al servizio, non puoi aggiornare i parametri di posizione e abilitare la modalità di accesso ibrida. Per impostazione predefinita, il `IAMAllowedPrincipals` gruppo dispone di autorizzazioni Super sul database e su tutte le relative tabelle.

Important

Non aggiornate la registrazione di una posizione alla modalità di accesso ibrida senza attivare i principali che accedono ai dati in questa posizione.

Abilitazione della modalità di accesso ibrido per una posizione dati registrata con Lake Formation

1.

Warning

Non consigliamo di convertire una posizione dati gestita da Lake Formation in modalità di accesso ibrida per evitare di interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti.

Scegli i presidi esistenti che dispongono dei permessi di Lake Formation.

1. Elenca e rivedi le autorizzazioni che hai concesso ai responsabili su cataloghi, database e tabelle. Per ulteriori informazioni, consulta [Visualizzazione delle autorizzazioni per database e tabelle in Lake Formation](#).
2. Scegli la modalità di accesso ibrido in Autorizzazioni dalla barra di navigazione a sinistra e scegli Aggiungi.

3. Nella pagina Aggiungi principali e risorse, scegli i cataloghi, i database e le tabelle dalla posizione dati di Amazon S3 che desideri utilizzare in modalità di accesso ibrido. Scegli i presidi che dispongono già dei permessi di Lake Formation.
 4. Scegli Aggiungi per attivare i principali per utilizzare le autorizzazioni di Lake Formation in modalità di accesso ibrida.
2. Aggiorna la bucket/prefix registrazione Amazon S3 scegliendo l'opzione della modalità di accesso ibrida.

Console

1. Accedi alla console di Lake Formation come amministratore del data lake.
2. Nel riquadro di navigazione, in Register and Ingest, scegli Data lake locations.
3. Seleziona una posizione e nel menu Azioni scegli Modifica.
4. Scegli la modalità di accesso ibrida.
5. Scegli Save (Salva).
6. In Data Catalog, seleziona il database o la tabella e Super concedi All le autorizzazioni al gruppo virtuale chiamato IAMAllowedPrincipals.
7. Verifica che l'accesso degli utenti esistenti di Lake Formation non venga interrotto quando hai aggiornato le proprietà di registrazione della posizione. Accedi alla console Athena come principale di Lake Formation ed esegui una query di esempio su una tabella che punta alla posizione aggiornata.

Allo stesso modo, verifica l'accesso degli AWS Glue utenti che utilizzano le policy di autorizzazione IAM per accedere al database e alle tabelle.

AWS CLI

Di seguito è riportato un esempio di registrazione di una posizione dati con Lake Formation HybridAccessEnabled con:true/false. Il valore predefinito per il parametro è false. HybridAccessEnabled Sostituisci il percorso, il nome del ruolo e l'ID AWS account di Amazon S3 con valori validi.

```
aws lakeformation update-resource --cli-input-json file://file path
json:
{
  "ResourceArn": "arn:aws:s3:::<s3-path>",
```

```
"RoleArn": "arn:aws:iam::<123456789012>:role/<test>",  
"HybridAccessEnabled": true  
}
```

Condivisione di una AWS Glue risorsa utilizzando la modalità di accesso ibrida

Condividi i dati con un altro Account AWS o con un responsabile di un altro richiedente Account AWS applicando le autorizzazioni di Lake Formation senza interrompere l'accesso basato su IAM degli utenti di Data Catalog esistenti.

Descrizione dello scenario: l'account produttore dispone di un database Data Catalog il cui accesso è controllato tramite le politiche e AWS Glue le azioni principali di IAM per Amazon S3. La posizione dei dati del database non è registrata con Lake Formation. Il IAMAllowedPrincipals gruppo, per impostazione predefinita, dispone Super delle autorizzazioni per il database e tutte le relative tabelle.

Concessione di autorizzazioni Lake Formation per più account in modalità di accesso ibrida

1. Configurazione dell'account Producer

1. Accedi alla console di Lake Formation utilizzando un ruolo con autorizzazione `lakeformation:PutDataLakeSettings` IAM.
2. Vai alle impostazioni di Data Catalog e scegli Version 4 le impostazioni della versione dell'account Cross.

Se attualmente utilizzi la versione 1 o 2, consulta [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#) le istruzioni per l'aggiornamento alla versione 3.

Non sono necessarie modifiche alla politica di autorizzazione per l'aggiornamento dalla versione 3 alla 4.

3. Registra la posizione Amazon S3 del database o della tabella che intendi condividere in modalità di accesso ibrido.
4. Verifica che l'Superautorizzazione al IAMAllowedPrincipals gruppo esista sui database e sulle tabelle di cui hai registrato la posizione dei dati in modalità di accesso ibrida nel passaggio precedente.
5. Concedi le autorizzazioni di Lake Formation a AWS organizzazioni, unità organizzative (OUs) o direttamente con un responsabile IAM in un altro account.

6. Se concedi le autorizzazioni direttamente a un principale IAM, scegli l'indirizzo principale dall'account consumer per applicare le autorizzazioni di Lake Formation in modalità di accesso ibrido abilitando l'opzione Rendi le autorizzazioni di Lake Formation effettive immediatamente.

Se concedi autorizzazioni per più account a un altro AWS account, quando attivi l'account, le autorizzazioni di Lake Formation vengono applicate solo agli amministratori di quell'account. L'amministratore del data lake dell'account destinatario deve ripartire a cascata le autorizzazioni e attivare i principali dell'account per applicare le autorizzazioni di Lake Formation per le risorse condivise che si trovano in modalità di accesso ibrido.

Se scegli l'opzione Resources matched by LF-Tags per concedere le autorizzazioni su più account, devi prima completare la fase di concessione delle autorizzazioni. Puoi impostare i principali e le risorse per la modalità di accesso ibrido in un passaggio separato selezionando la modalità di accesso ibrida in Autorizzazioni nella barra di navigazione a sinistra della console Lake Formation. Quindi scegli Aggiungi per aggiungere le risorse e i presidi a cui desideri applicare le autorizzazioni di Lake Formation.

2. Configurazione dell'account consumatore

1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> come amministratore del data lake.
2. Vai alla https://console.aws.amazon.com/ram/home_page e accetta l'invito alla condivisione delle risorse. La scheda Condividi con me nella AWS RAM console mostra il database e le tabelle condivise con il tuo account.
3. Crea un link di risorsa alla and/or tabella del database condivisa in Lake Formation.
4. Concedi Describe l'autorizzazione sul link alla risorsa e l'Grant on target autorizzazione (sulla risorsa condivisa originale) ai principali IAM del tuo account (consumatore).
5. Concedi le autorizzazioni di Lake Formation sul database o sulla tabella condivisa con te ai responsabili del tuo account. Scegli i principi e le risorse per applicare le autorizzazioni di Lake Formation in modalità di accesso ibrida abilitando l'opzione Rendi immediatamente effettive le autorizzazioni di Lake Formation.
6. Verifica le autorizzazioni Lake Formation del preside eseguendo query Athena di esempio. Testa l'accesso esistente dei tuoi AWS Glue utenti con le policy principali di IAM per Amazon S3 e AWS Glue le azioni.

(Facoltativo) Rimuovi la policy del bucket di Amazon S3 per l'accesso ai dati e le politiche principali di IAM e l'accesso ai dati di Amazon S3 per AWS Glue i principali che hai configurato per utilizzare le autorizzazioni Lake Formation.

Condivisione di una risorsa Lake Formation utilizzando la modalità di accesso ibrida

Consenti ai nuovi utenti di Data Catalog in un account esterno di accedere ai database e alle tabelle di Data Catalog utilizzando policy basate su IAM senza interrompere le autorizzazioni di condivisione tra account esistenti di Lake Formation.

Descrizione dello scenario: l'account produttore dispone di database e tabelle gestiti da Lake Formation condivisi con un account esterno (consumatore) a livello di account o principale IAM. La posizione dei dati del database è registrata presso Lake Formation. Il `IAMAllowedPrincipals` gruppo non dispone `Super` delle autorizzazioni per il database e le relative tabelle.

Garantire l'accesso su più account ai nuovi utenti del Data Catalog tramite policy basate su IAM senza interrompere le autorizzazioni esistenti di Lake Formation

1. Configurazione dell'account Producer

1. Accedi alla console Lake Formation utilizzando un ruolo `cheLakeFormation:PutDataLakeSettings`.
2. In Impostazioni Data Catalog, scegli `Version 4` le impostazioni della versione dell'account Cross.

Se attualmente utilizzi la versione 1 o 2, consulta [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#) le istruzioni per l'aggiornamento alla versione 3.

Non sono necessarie modifiche alla politica di autorizzazione per l'aggiornamento dalla versione 3 alla 4.

3. Elenca le autorizzazioni che hai concesso ai principali su database e tabelle. Per ulteriori informazioni, consulta [Visualizzazione delle autorizzazioni per database e tabelle in Lake Formation](#).
4. Rigenera le autorizzazioni esistenti per più account di Lake Formation optando per i principali e le risorse.

 Note

Prima di aggiornare la registrazione di una posizione dati alla modalità di accesso ibrida per concedere le autorizzazioni tra più account, devi concedere nuovamente almeno una condivisione di dati tra account per account. Questo passaggio è necessario per aggiornare le autorizzazioni AWS RAM gestite allegate alla condivisione di risorse. AWS RAM

Nel luglio 2023, Lake Formation ha aggiornato le autorizzazioni AWS RAM gestite utilizzate per la condivisione di database e tabelle:

- `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueAllTablesReadWriteForDatabase`(politica di condivisione a livello di database)
- `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueTableReadWrite`(politica di condivisione a livello di tabella)

Le autorizzazioni per più account concesse prima di luglio 2023 non dispongono di queste autorizzazioni aggiornate. AWS RAM

Se hai concesso autorizzazioni per più account direttamente ai responsabili, devi concedere nuovamente tali autorizzazioni individualmente ai responsabili. Se salti questo passaggio, i principali che accedono alla risorsa condivisa potrebbero ricevere un errore di combinazione illegale.

5. [Vai a casa. https://console.aws.amazon.com/ram/](https://console.aws.amazon.com/ram/)

6. La scheda Condivisi da me nella AWS RAM console mostra i nomi di database e tabelle che hai condiviso con un account o un principale esterno.

Assicurati che le autorizzazioni allegate alla risorsa condivisa abbiano l'ARN corretto.

7. Verifica che lo stato delle risorse nella AWS RAM condivisione sia attivo. Associated Se lo stato è uguale aAssociating, attendi che entrino Associated nello stato. Se lo stato diventaFailed, fermati e contatta il team di assistenza di Lake Formation.

8. Scegli la modalità di accesso ibrido in Autorizzazioni dalla barra di navigazione a sinistra e scegli Aggiungi.

9. La pagina Aggiungi principi e risorse mostra i database, le and/or tabelle e i principali che hanno accesso. È possibile apportare gli aggiornamenti richiesti aggiungendo o rimuovendo i principali e le risorse.

10. Scegli i principali con autorizzazioni Lake Formation per il database e le tabelle che desideri modificare in modalità di accesso ibrida. Scegli i database e le tabelle.
11. Scegli Aggiungi per attivare i principali per applicare le autorizzazioni di Lake Formation in modalità di accesso ibrida.
12. Concedi Super l'autorizzazione al gruppo virtuale IAMAllowedPrincipals sul tuo database e sulle tabelle selezionate.
13. Modifica la registrazione della sede Amazon S3 Lake Formation in modalità di accesso ibrida.
14. Concedi le autorizzazioni agli AWS Glue utenti dell'account esterno (consumer) utilizzando le politiche di autorizzazione IAM per le azioni di Amazon AWS Glue S3.

2. Configurazione dell'account consumer

1. Accedi alla console di Lake Formation <https://console.aws.amazon.com/lakeformation/> come amministratore del data lake.
2. Vai alla https://console.aws.amazon.com/ram/home_page e accetta l'invito alla condivisione delle risorse. La scheda Risorse condivise con me della AWS RAM pagina mostra i nomi dei database e delle tabelle condivisi con il tuo account.

Per la AWS RAM condivisione, assicurati che l'autorizzazione allegata contenga l'ARN corretto dell'invito condiviso AWS RAM . Controlla se le risorse della AWS RAM condivisione sono in Associated stato. Se lo stato è uguale Associating, attendi che diventino Associated tali. Se lo stato diventa Failed, fermati e contatta il team di assistenza di Lake Formation.

3. Crea un link di risorsa alla and/or tabella del database condivisa in Lake Formation.
4. Concedi Describe l'autorizzazione sul link alla risorsa e l'Grant on target autorizzazione (sulla risorsa condivisa originale) ai principali IAM del tuo account (consumatore).
5. Successivamente, configura le autorizzazioni di Lake Formation per i principali del tuo account nel database o nella tabella condivisa.

Nella barra di navigazione a sinistra, in Autorizzazioni, scegli la modalità di accesso ibrido.

6. Scegli Aggiungi nella sezione inferiore della pagina della modalità di accesso ibrido per attivare i principali e il database o la tabella condivisi con te dall'account produttore.
7. Concedi le autorizzazioni agli AWS Glue utenti del tuo account utilizzando le politiche di autorizzazione IAM per le azioni di Amazon AWS Glue S3.
8. Verifica le autorizzazioni e AWS Glue le autorizzazioni di Lake Formation degli utenti eseguendo query di esempio separate sulla tabella utilizzando Athena

(Facoltativo) Pulisci le politiche di autorizzazione IAM per Amazon S3 per i principali che si trovano in modalità di accesso ibrido.

Rimozione di principi e risorse dalla modalità di accesso ibrida

Segui questi passaggi per rimuovere database, tabelle e principali dalla modalità di accesso ibrido.

Console

1. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. In Autorizzazioni, scegli la modalità di accesso ibrido.
3. Nella pagina della modalità di accesso ibrido, seleziona la casella di controllo accanto al nome del database o della tabella e scegli. Remove
4. Un messaggio di avviso richiede di confermare l'azione. Scegli Rimuovi.

Lake Formation non impone più le autorizzazioni per tali risorse e l'accesso a questa risorsa sarà controllato tramite IAM e AWS Glue le autorizzazioni. Ciò potrebbe far sì che l'utente non abbia più accesso a questa risorsa se non dispone delle autorizzazioni IAM appropriate.

AWS CLI

L'esempio seguente mostra come rimuovere risorse dalla modalità di accesso ibrida.

```
aws lakeformation delete-lake-formation-opt-in --cli-input-json file://file path

json:
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::<123456789012>:role/role name"
  },
  "Resource": {
    "Table": {
      "CatalogId": "<123456789012>",
      "DatabaseName": "<database name>",
      "Name": "<table name>"
    }
  }
}
```

Visualizzazione dei principali e delle risorse in modalità di accesso ibrida

Segui questi passaggi per visualizzare database, tabelle e principali in modalità di accesso ibrido.

Console

1. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. In Autorizzazioni, scegli la modalità di accesso ibrido.
3. La pagina della modalità di accesso ibrido mostra le risorse e i principali attualmente in modalità di accesso ibrida.

AWS CLI

L'esempio seguente mostra come elencare tutti i principali e le risorse di opt-in che si trovano in modalità di accesso ibrida.

```
aws lakeformation list-lake-formation-opt-ins
```

L'esempio seguente mostra come elencare gli opt-in per una specifica coppia principale-risorsa.

```
aws lakeformation list-lake-formation-opt-ins --cli-input-json file://file path

json:
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::<account-id>:role/<role name>"
  },
  "Resource": {
    "Table": {
      "CatalogId": "<account-id>",
      "DatabaseName": "<database name>",
      "Name": "<table name>"
    }
  }
}
```

```
}  
}
```

Risorse aggiuntive

Nel seguente post del blog, ti illustreremo le istruzioni per integrare le autorizzazioni di Lake Formation in modalità di accesso ibrida per utenti selezionati mentre il database è già accessibile ad altri utenti tramite le autorizzazioni IAM e Amazon S3. Esamineremo le istruzioni per configurare la modalità di accesso ibrida all'interno di un AWS account e tra due account.

- [Presentazione della modalità di accesso AWS Glue Data Catalog ibrida per l'accesso sicuro utilizzando le policy di Lake Formation e IAM e Amazon S3.](#)

Creazione di oggetti in AWS Glue Data Catalog

AWS Lake Formation utilizza AWS Glue Data Catalog (Data Catalog) per archiviare metadati su data lake, origini dati, trasformazioni e destinazioni. I metadati sono dati relativi ai dati sottostanti nel set di dati. Ogni AWS account dispone di un catalogo dati per regione. AWS

I metadati nel Data Catalog sono organizzati in una gerarchia di dati a tre livelli che comprende cataloghi, database e tabelle. Organizza i dati provenienti da varie fonti in contenitori logici chiamati cataloghi. Ogni catalogo rappresenta dati provenienti da fonti come data warehouse Amazon Redshift, Amazon DynamoDB database e fonti di dati di terze parti come Snowflake, MySQL e oltre 30 fonti di dati esterne, integrate tramite connettori federati. Puoi anche creare nuovi cataloghi nel Data Catalog per archiviare i dati in S3 Table Bucket o Redshift Managed Storage (RMS).

Le tabelle memorizzano informazioni sui dati sottostanti, tra cui informazioni sullo schema, sulle partizioni e sulla posizione dei dati. I database sono raccolte di tabelle. Il Data Catalog contiene anche collegamenti a risorse, che sono collegamenti a cataloghi, database e tabelle condivisi in account esterni e vengono utilizzati per l'accesso tra account diversi ai dati nel data lake.

Il Data Catalog è un oggetto di catalogo annidato che contiene cataloghi, database e tabelle. È referenziato dall' Account AWS ID ed è il catalogo predefinito in un account e in un. Regione AWS Il Data Catalog utilizza una gerarchia a tre livelli (catalogo.database.tabelle) per organizzare le tabelle.

- Catalogo: il livello più alto della gerarchia dei metadati a tre livelli di Data Catalog. È possibile aggiungere più cataloghi in un Data Catalog tramite la federazione.

- **Database:** il secondo livello della gerarchia dei metadati che comprende tabelle e viste. Un database viene anche definito schema in molti sistemi di dati come Amazon Redshift e Trino.
- **Tabella e visualizzazione:** il terzo livello della gerarchia dei dati a 3 livelli del Data Catalog.

Tutte le tabelle Iceberg in Amazon S3 sono archiviate nel Data Catalog predefinito con Catalog ID = ID Account AWS . Puoi creare cataloghi federati in AWS Glue Data Catalog cui archiviare le definizioni delle tabelle in Amazon Redshift, Amazon S3 Table Storage o altre fonti di dati di terze parti tramite la federazione.

Argomenti

- [Creazione di un catalogo](#)
- [Creazione di un database](#)
- [Creazione di tabelle](#)
- [Viste dell'edificio AWS Glue Data Catalog](#)

Creazione di un catalogo

I cataloghi rappresentano il livello più alto o più alto nella gerarchia dei metadati a tre livelli di AWS Glue Data Catalog. È possibile utilizzare diversi metodi per inserire i dati nel Data Catalog e creare cataloghi a più livelli.

Per ulteriori informazioni sulla creazione di cataloghi da fonti di dati esterne, consulta [Inserire i dati nel AWS Glue Data Catalog](#)

Per creare un catalogo utilizzando la console Lake Formation, devi accedere come amministratore del data lake o creatore di cataloghi. Un creatore di cataloghi è un preside a cui è stata concessa l'CREATE_CATALOG autorizzazione Lake Formation. Puoi vedere un elenco dei creatori di cataloghi nella pagina Ruoli e attività amministrative della console Lake Formation. Per visualizzare questo elenco, devi disporre dell'autorizzazione lakeformation:ListPermissions IAM ed essere registrato come amministratore del data lake o come creatore di cataloghi con l'opzione di concessione sull'CREATE_CATALOG autorizzazione.

Creazione di un database

Le tabelle di metadati nel Data Catalog sono archiviate all'interno di database. Puoi creare tutti i database di cui hai bisogno e puoi concedere diverse autorizzazioni Lake Formation su ciascun database.

I database possono avere una proprietà di localizzazione opzionale. Questa sede si trova in genere all'interno di una sede Amazon Simple Storage Service (Amazon S3) registrata presso Lake Formation. Quando specifichi una posizione, i responsabili non necessitano delle autorizzazioni per la localizzazione dei dati per creare tabelle Data Catalog che puntano a posizioni all'interno della posizione del database. Per ulteriori informazioni, consulta [Underlying data access control](#).

Per creare un database utilizzando la console Lake Formation, devi accedere come amministratore del data lake o creatore del database. Un creatore di database è un responsabile a cui è stata concessa l'`CREATE_DATABASE` autorizzazione Lake Formation. Puoi vedere un elenco dei creatori di database nella pagina Ruoli e attività amministrative della console Lake Formation. Per visualizzare questo elenco, è necessario disporre dell'autorizzazione `lakeformation:ListPermissions` IAM e accedere come amministratore del data lake o come creatore di database con l'opzione di concessione sull'`CREATE_DATABASE` autorizzazione.

Come creare un database

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake o creatore di database.
2. Nel riquadro di navigazione, in Catalogo dati, scegli Database.
3. Scegliere Crea database.
4. Nella finestra di dialogo Crea database, inserisci un nome di database, una posizione opzionale e una descrizione facoltativa.
5. Facoltativamente, seleziona Usa solo il controllo di accesso IAM per le nuove tabelle in questo database.

Per ulteriori informazioni su questa opzione, consulta [the section called “Modifica delle impostazioni predefinite per il data lake”](#).

6. Scegliere Crea database.

Creazione di tabelle

AWS Lake Formation le tabelle di metadati contengono informazioni sui dati nel data lake, incluse informazioni sullo schema, sulle partizioni e sulla posizione dei dati. Queste tabelle sono archiviate nel AWS Glue Data Catalog. Li usi per accedere ai dati sottostanti nel data lake e gestirli con le autorizzazioni di Lake Formation. Le tabelle vengono archiviate all'interno dei database del Data Catalog.

Esistono diversi modi per creare tabelle Data Catalog:

- Esegui un crawler. AWS Glue Vedi [Definizione dei crawler nella Guida per gli sviluppatori](#).AWS Glue
- Crea ed esegui un flusso di lavoro. Per informazioni, consulta [the section called “Importazione di dati tramite flussi di lavoro”](#).
- Crea una tabella manualmente utilizzando la console di Lake Formation, AWS Glue l'API o AWS Command Line Interface (AWS CLI).
- Crea una tabella utilizzando Amazon Athena.
- Crea un collegamento di risorsa a una tabella in un account esterno. Per informazioni, consulta [the section called “Creazione di collegamenti alle risorse”](#).

Creazione di tabelle Apache Iceberg

AWS Lake Formation supporta la creazione di tabelle Apache Iceberg che utilizzano il formato di dati Apache Parquet AWS Glue Data Catalog con dati che risiedono in Amazon S3. Una tabella nel catalogo dati è la definizione dei metadati che rappresenta i dati in un datastore. Per impostazione predefinita, Lake Formation crea tabelle Iceberg v2. Per la differenza tra le tabelle v1 e v2, consulta [Modifiche al tipo di formato](#) nella documentazione di Apache Iceberg.

[Apache Iceberg](#) è un formato a tabella aperta per set di dati analitici di dimensioni molto grandi. Iceberg consente di modificare facilmente lo schema, operazione nota anche come evoluzione dello schema, il che significa che gli utenti possono aggiungere, rinominare o rimuovere colonne da una tabella di dati senza pregiudicare i dati sottostanti. Iceberg fornisce anche supporto per il controllo delle versioni dei dati, che consente agli utenti di tenere traccia delle modifiche apportate ai dati nel tempo. Ciò abilita la funzionalità di query temporale, che consente agli utenti di accedere e interrogare le versioni storiche dei dati e analizzare le modifiche ai dati tra aggiornamenti ed eliminazioni.

Puoi utilizzare la console Lake Formation o l'`CreateTable` operazione nell' AWS Glue API per creare una tabella Iceberg nel Data Catalog. Per ulteriori informazioni, vedere [CreateTable action \(Python: create_table\)](#).

Quando si crea una tabella Iceberg nel catalogo dati, occorre specificare il formato della tabella e il percorso del file dei metadati in Amazon S3 per poter eseguire letture e scritture.

Puoi usare Lake Formation per proteggere la tua tabella Iceberg utilizzando autorizzazioni di controllo degli accessi granulari quando registri la posizione dati di Amazon S3 con. AWS Lake

Formation Per i dati di origine in Amazon S3 e i metadati non registrati con Lake Formation, l'accesso è determinato dalle politiche di autorizzazione IAM per Amazon S3 e dalle azioni. AWS Glue Per ulteriori informazioni, consulta [Gestione delle autorizzazioni di Lake Formation](#).

Note

Il catalogo dati non supporta la creazione di partizioni e l'aggiunta di proprietà delle tabelle Iceberg.

Argomenti

- [Prerequisiti](#)
- [Creazione di una tabella Iceberg](#)

Prerequisiti

Per creare tabelle Iceberg nel catalogo dati e configurare le autorizzazioni di accesso ai dati di Lake Formation, occorre soddisfare i seguenti requisiti:

1. Autorizzazioni richieste per creare tabelle Iceberg senza i dati registrati con Lake Formation.

Oltre alle autorizzazioni richieste per creare una tabella nel catalogo dati, il creatore della tabella richiede le seguenti autorizzazioni:

- `s3:PutObject` sulla risorsa `arn:aws:s3:::{bucketName}`
- `s3:GetObject` sulla risorsa `arn:aws:s3:::{bucketName}`
- `s3:DeleteObject` sulla risorsa `arn:aws:s3:::{bucketName}`

2. Autorizzazioni richieste per creare tabelle Iceberg con dati registrati con Lake Formation:

Per utilizzare Lake Formation per gestire e proteggere i dati nel data lake, registrare la posizione Amazon S3 che contiene i dati per le tabelle con Lake Formation. In questo modo Lake Formation può fornire credenziali a servizi di AWS analisi come Athena, Redshift Spectrum e Amazon EMR per accedere ai dati. Per ulteriori informazioni sulla registrazione di una sede Amazon S3, consulta. [Aggiungere una posizione Amazon S3 al tuo data lake](#)

Un principale che legge e scrive i dati sottostanti registrati con Lake Formation richiede le seguenti autorizzazioni:

- `lakeformation:GetDataAccess`
- `DATA_LOCATION_ACCESS`

Un principale che dispone di autorizzazioni di localizzazione dei dati in una sede dispone anche delle autorizzazioni di localizzazione su tutte le sedi secondarie.

Per ulteriori informazioni sulle autorizzazioni relative alla localizzazione dei dati, consulta [Controllo sottostante dell'accesso ai dati](#)

Per abilitare la compattazione, il servizio deve assumere un ruolo IAM con le autorizzazioni per aggiornare le tabelle nel catalogo dati. Per i dettagli, consulta [Prerequisiti per l'ottimizzazione delle tabelle](#).

Creazione di una tabella Iceberg

Puoi creare tabelle Iceberg v1 e v2 utilizzando la console Lake Formation o AWS Command Line Interface come documentato in questa pagina. Puoi anche creare tabelle Iceberg usando la console o. AWS Glue Crawler di AWS Glue Per ulteriori informazioni, consultare [Catalogo dati e crawler](#) nella Guida per gli sviluppatori di AWS Glue .

Per creare una tabella Iceberg

Console

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. In catalogo dati, scegliere Tabelle e utilizzare il pulsante Crea tabella per specificare i seguenti attributi:
 - Nome tabella: inserisci un nome per la tabella. Se si utilizza Athena per accedere alle tabelle, utilizzare questi [suggerimenti per la denominazione](#) della Guida per l'utente di Amazon Athena.
 - Database: scegli un database esistente o creane uno nuovo.
 - Descrizione: la descrizione della tabella. È possibile scrivere una descrizione per aiutare a comprendere i contenuti della tabella.
 - Formato tabella: per il formato tabella, scegli Apache Iceberg.

Table format [Info](#)

Data Catalog managed tables support data compaction for Apache Iceberg table type. [Learn more](#)

☐ **Standard AWS Glue table (default)**
Create a standard AWS Glue table.

☒ **Apache Iceberg table**
Create a table in Apache Iceberg table format.

Table optimization
Enable optimization capabilities for Apache Iceberg tables enhance query performance.

☒ **Compaction**
Combine small data files into larger, more efficient files to optimize table performance.

☒ **Snapshot retention - new**
Optimize table storage by removing old snapshots from metadata overhead and expiring files that are no longer needed.

☒ **Orphan file deletion - new**
Automatically clean up orphan files periodically.

- Ottimizzazione della tabella
 - Compattazione: I file di dati vengono uniti e riscritti per rimuovere i dati obsoleti e consolidare i dati frammentati in file più grandi ed efficienti.
 - Conservazione degli snapshot: gli snapshot sono versioni con data e ora di una tabella Iceberg. Le configurazioni di conservazione degli snapshot consentono ai clienti di stabilire per quanto tempo conservare gli snapshot e quanti snapshot conservare. La configurazione di un ottimizzatore di conservazione degli snapshot può aiutare a gestire il sovraccarico di archiviazione rimuovendo gli snapshot più vecchi e non necessari e i relativi file sottostanti.
 - Eliminazione di file orfani: i file orfani sono file a cui non fanno più riferimento i metadati della tabella Iceberg. Questi file possono accumularsi nel tempo, soprattutto dopo operazioni come l'eliminazione di tabelle o i processi ETL non riusciti. L'abilitazione dell'eliminazione dei file orfani AWS Glue consente di identificare e rimuovere periodicamente questi file non necessari, liberando spazio di archiviazione.

Per ulteriori informazioni, consulta [Ottimizzazione delle tabelle Iceberg](#).

- Ruolo IAM: per eseguire la compattazione, il servizio assume un ruolo IAM per conto dell'utente. Puoi scegliere un ruolo IAM utilizzando il menu a discesa. Assicurati che il ruolo disponga delle autorizzazioni necessarie per abilitare la compattazione.

Per ulteriori informazioni sulle autorizzazioni richieste, consulta [Prerequisiti per l'ottimizzazione delle tabelle](#).

- **Posizione:** specifica il percorso della cartella in Amazon S3 che memorizza la tabella dei metadati. Iceberg ha bisogno di un file di metadati e di una posizione nel catalogo dati per poter eseguire letture e scritture.
- **Schema:** scegli Aggiungi colonne per aggiungere colonne e tipi di dati delle colonne. Si dispone della possibilità di creare una tabella vuota e di aggiornare lo schema in un secondo momento. Il catalogo dati supporta i tipi di dati Hive. Per ulteriori informazioni, consultare [Tipi di dati Hive](#).

Iceberg consente di evolvere lo schema e la partizione dopo aver creato la tabella. È possibile usare le [query di Athena](#) per aggiornare lo schema della tabella e le [query Spark](#) per aggiornare le partizioni.

AWS CLI

```
aws glue create-table \  
  --database-name iceberg-db \  
  --region us-west-2 \  
  --open-table-format-input '{  
    "IcebergInput": {  
      "MetadataOperation": "CREATE",  
      "Version": "2"  
    }  
  }' \  
  --table-input '{"Name": "test-iceberg-input-demo",  
    "TableType": "EXTERNAL_TABLE",  
    "StorageDescriptor": {  
      "Columns": [  
        {"Name": "col1", "Type": "int"},  
        {"Name": "col2", "Type": "int"},  
        {"Name": "col3", "Type": "string"}  
      ],  
      "Location": "s3://DOC_EXAMPLE_BUCKET_ICEBERG/"  
    }  
  }'
```

Ottimizzazione delle tabelle Iceberg

Lake Formation supporta diverse opzioni di ottimizzazione delle tabelle per migliorare la gestione e le prestazioni delle tabelle Apache Iceberg utilizzate dai motori AWS analitici e dai job ETL. Questi ottimizzatori offrono un utilizzo efficiente dello storage, prestazioni di query migliorate e una gestione efficace dei dati. In Lake Formation sono disponibili tre tipi di ottimizzatori di tabelle:

- **Compattazione:** la compattazione dei dati compatta file di dati di piccole dimensioni per ridurre l'utilizzo dell'archiviazione e migliorare le prestazioni di lettura. I file di dati vengono uniti e riscritti per rimuovere i dati obsoleti e consolidare i dati frammentati in file più grandi ed efficienti. La compattazione può essere configurata per funzionare automaticamente o attivata manualmente secondo necessità.
- **Conservazione degli snapshot:** gli snapshot sono versioni con data e ora di una tabella Iceberg. Le configurazioni di conservazione degli snapshot consentono ai clienti di stabilire per quanto tempo conservare gli snapshot e quanti snapshot conservare. La configurazione di un ottimizzatore di conservazione degli snapshot può aiutare a gestire il sovraccarico di archiviazione rimuovendo gli snapshot più vecchi e non necessari e i relativi file sottostanti.
- **Eliminazione di file orfani:** i file orfani sono file a cui non fanno più riferimento i metadati della tabella Iceberg. Questi file possono accumularsi nel tempo, soprattutto dopo operazioni come l'eliminazione di tabelle o i processi ETL non riusciti. L'abilitazione dell'eliminazione dei file orfani consente di AWS Glue identificare e rimuovere periodicamente questi file non necessari, liberando spazio di archiviazione.

Puoi abilitare o disabilitare gli ottimizzatori di compattazione, conservazione delle istantanee e eliminazione dei file orfani per le singole tabelle Iceberg nel Data Catalog utilizzando la console o le AWS Glue operazioni API. AWS CLI AWS Glue

Per ulteriori informazioni, consulta [Optimizing Iceberg](#) tables nella Developer Guide. AWS Glue

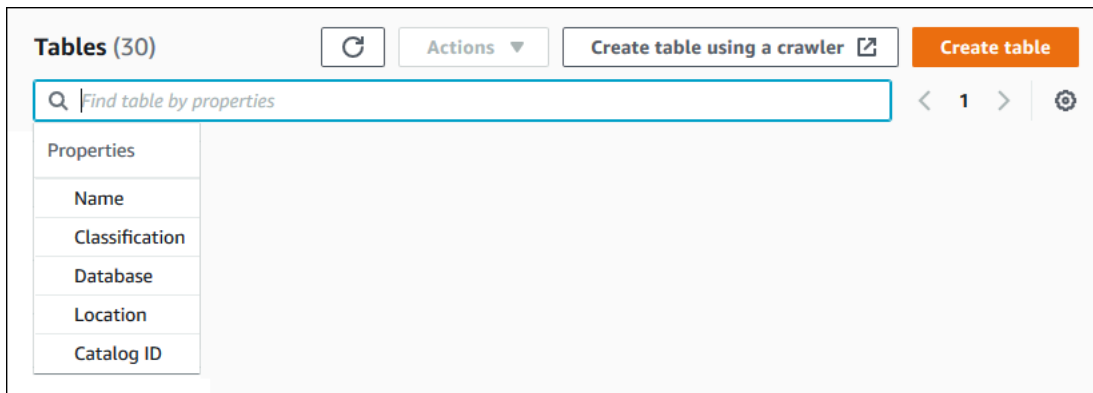
Ricerca di tabelle

Puoi usare la AWS Lake Formation console per cercare le tabelle del Data Catalog per nome, posizione, database contenente e altro. I risultati della ricerca mostrano solo le tabelle per le quali disponi delle autorizzazioni Lake Formation.

Per cercare tabelle (console)

1. Accedi a Console di gestione AWS e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel pannello di navigazione, seleziona Tabelle.
3. Posiziona il cursore nel campo di ricerca nella parte superiore della pagina. Il campo ha il testo segnaposto Trova tabella per proprietà.

Viene visualizzato il menu Proprietà, che mostra le varie proprietà della tabella in base alle quali effettuare la ricerca.



4. Esegui una delle seguenti operazioni:

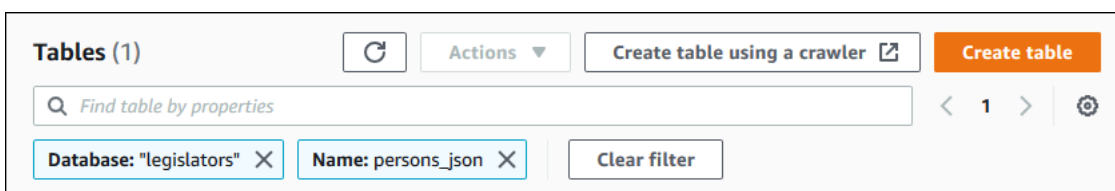
- Ricerca per database contenente.

1. Scegli Database dal menu Proprietà, quindi scegli un database dal menu Database visualizzato oppure digita il nome del database e premi Invio.

Vengono elencate le tabelle per le quali disponi delle autorizzazioni nel database.

2. (Facoltativo) Per restringere l'elenco a una singola tabella del database, posizionate nuovamente il cursore nel campo di ricerca, scegliete Nome dal menu Proprietà e scegliete il nome di una tabella dal menu Tabelle visualizzato oppure digitate un nome di tabella e premete Invio.

Viene elencata la singola tabella e sia il nome del database che il nome della tabella vengono visualizzati come riquadri sotto il campo di ricerca.



Per regolare il filtro, chiudi uno dei riquadri o scegli Cancella filtro.

- Cerca per altre proprietà.

1. Scegliete una proprietà di ricerca dal menu Proprietà.

Per eseguire la ricerca in base all'ID AWS dell'account, scegliete ID catalogo dal menu Proprietà, immettete un ID AWS account valido (ad esempio, 111122223333) e premete Invio.

Per effettuare una ricerca per località, scegli Posizione dal menu Proprietà e seleziona una località dal menu Posizioni visualizzato. Vengono restituite tutte le tabelle nella posizione principale della posizione selezionata (ad esempio, Amazon S3).

Ricerca nelle tabelle utilizzando AWS CLI

- L'esempio seguente mostra come eseguire una ricerca parziale. Il `--search-text` parametro consente di cercare tabelle che contengono il testo specificato nei relativi metadati. In questo caso, restituisce tutte le tabelle che contengono «cliente» nel nome, nella descrizione o in altri campi di metadati.

```
aws glue search-tables
  --search-text "customer"
  --region Regione AWS
  --max-results 10
  --sort-criteria "FieldName=Name,Sort=ASC"
```

Condivisione di tabelle e database del Data Catalog tra più account AWS

Puoi condividere le risorse del Data Catalog (database e tabelle) con AWS account esterni concedendo le autorizzazioni di Lake Formation sulle risorse agli account esterni. Gli utenti possono quindi eseguire query e lavori che uniscono e interrogano tabelle su più account. Con alcune restrizioni, quando condividi una risorsa del Catalogo dati con un altro account, i responsabili di quell'account possono utilizzare quella risorsa come se la risorsa fosse nel loro Catalogo dati.

Non condividi le risorse con responsabili specifici negli AWS account esterni, ma con un account o un'organizzazione. AWS Quando condividi una risorsa con un' AWS organizzazione, la condividi con tutti gli account a tutti i livelli dell'organizzazione. L'amministratore del data lake di ogni account

esterno deve quindi concedere le autorizzazioni sulle risorse condivise ai responsabili del proprio account.

Per ulteriori informazioni, consultare [Condivisione dei dati tra account in Lake Formation](#) e [Concessione delle autorizzazioni per le risorse di Data Catalog](#).

 Consulta anche:

- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)
- [Prerequisiti](#)

Viste dell'edificio AWS Glue Data Catalog

In AWS Glue Data Catalog, una vista è una tabella virtuale in cui i contenuti sono definiti da una query SQL che fa riferimento a una o più tabelle. Puoi creare una vista del catalogo dati che faccia riferimento a un massimo di 10 tabelle utilizzando editor SQL per Amazon Athena, Amazon Redshift o Apache Spark utilizzando EMR Serverless o la versione 5.0. AWS Glue Le tabelle di riferimento sottostanti per una vista possono appartenere allo stesso database o a database diversi all'interno dello stesso Data Catalog. Account AWS

Puoi fare riferimento a AWS Glue tabelle e tabelle standard in formati di tabelle aperte (OTF) come [Apache Hudi](#), Linux Foundation [Delta Lake](#) e [Apache Iceberg](#), con i dati sottostanti archiviati in posizioni Amazon S3 registrati con. AWS Lake Formation Inoltre, puoi creare viste da tabelle federate da condivisioni di dati Amazon Redshift condivise con Lake Formation.

Differenziazione delle viste del Data Catalog da altri tipi di visualizzazione

Le visualizzazioni del catalogo dati sono diverse da quelle di Apache Hive, Apache Spark e Amazon Athena. La visualizzazione Data Catalog è una funzionalità nativa di AWS Glue Data Catalog, ed è una visualizzazione multidialeale creata da un defintore. Puoi creare una vista del catalogo dati utilizzando uno dei servizi di analisi supportati, come Athena o Amazon Redshift Spectrum, e accedere alla stessa vista utilizzando altri servizi di analisi supportati. D'altra parte, le viste Apache Hive, Apache Spark e Athena vengono create indipendentemente in ogni servizio di analisi, come Athena e Amazon Redshift, e sono visibili e accessibili solo all'interno di tale servizio.

Cos'è una vista più definita?

Una vista definente è una vista SQL che funziona in base alle autorizzazioni del principale che l'ha creata. Il ruolo definer dispone delle autorizzazioni necessarie per accedere alle tabelle di riferimento ed esegue l'istruzione SQL che definisce la vista. Il definitore crea la vista e la condivide con altri utenti tramite il controllo granulare degli AWS Lake Formation accessi.

Quando un utente interroga la vista del definitore, il motore di query utilizza le autorizzazioni del ruolo definer per accedere alle tabelle di riferimento sottostanti. Questo approccio consente agli utenti di interagire con la vista senza richiedere l'accesso diretto alle tabelle di origine, migliorando la sicurezza e semplificando la gestione dell'accesso ai dati.

Per configurare una vista di definizione, il ruolo IAM definer può trovarsi all'interno dello stesso AWS account delle tabelle di base o in un account diverso utilizzando ruoli di definizione tra account. Per ulteriori informazioni sulle autorizzazioni richieste per il ruolo definer, vedere [Prerequisiti per la creazione di viste](#)

Un framework per visualizzazioni multidialettali

Il Data Catalog supporta la creazione di viste utilizzando più dialetti SQL (Structured Query Language). SQL è un linguaggio utilizzato per archiviare ed elaborare informazioni in un database relazionale e ogni motore AWS analitico utilizza la propria variante di SQL, o dialetto SQL.

È possibile creare una visualizzazione del catalogo dati in un dialetto SQL utilizzando uno dei motori di query di analisi supportati. Successivamente, puoi aggiornare la vista utilizzando l'ALTER VIEWistruzione in un dialetto SQL diverso all'interno di qualsiasi altro motore di analisi supportato. Tuttavia, ogni dialetto deve fare riferimento allo stesso set di tabelle, colonne e tipi di dati.

È possibile accedere ai diversi dialetti disponibili per la visualizzazione utilizzando l'GetTableAPI AWS CLI e AWS la console. Pertanto, la visualizzazione del catalogo dati è visibile e disponibile per le query su diversi motori di analisi supportati.

Definendo uno schema di visualizzazione e un oggetto di metadati comuni che è possibile interrogare da più motori, le viste del catalogo dati consentono di utilizzare viste uniformi su tutto il data lake.

Per maggiori dettagli su come viene risolto lo schema per ogni dialetto, consulta il [link al riferimento all'API](#). Per maggiori dettagli sulle regole di abbinamento per i diversi tipi, consulta il [collegamento alla sezione pertinente nel documento API](#).

Integrazione con le autorizzazioni di Lake Formation

Puoi utilizzare AWS Lake Formation per centralizzare la gestione delle autorizzazioni sulle AWS Glue Data Catalog visualizzazioni per gli utenti. È possibile concedere autorizzazioni dettagliate per le viste del Data Catalog utilizzando il metodo delle risorse denominato o i tag LF e condividerle tra organizzazioni e unità organizzative. Account AWS AWS Puoi anche condividere e accedere alle viste del Data Catalog utilizzando i link alle risorse. Regioni AWS Ciò consente agli utenti di fornire l'accesso ai dati senza duplicare l'origine dei dati e condividere le tabelle sottostanti.

L'istruzione CREATE VIEW DDL di una vista Data Catalog può fare riferimento alle AWS Glue tabelle e alle tabelle standard in formati di tabelle aperte (OTF) come Hudi, Delta Lake e Iceberg con dati sottostanti archiviati in posizioni Amazon S3 registrate con Lake Formation, nonché alle tabelle federate del datashare Amazon Redshift condivise con Lake Formation. Le tabelle possono essere di qualsiasi formato di file purché il motore utilizzato per interrogare la vista supporti tale formato. È inoltre possibile fare riferimento alle funzioni integrate del motore su cui viene eseguito, ma potrebbero non essere consentite altre risorse specifiche del motore. Per ulteriori dettagli, consulta [Visualizzazioni, considerazioni e limitazioni di Data Catalog](#).

Casi d'uso

Di seguito sono riportati i casi d'uso importanti per le visualizzazioni del catalogo dati:

- Crea e gestisci le autorizzazioni su uno schema a visualizzazione singola. Questo ti aiuta a evitare il rischio di autorizzazioni incoerenti su viste duplicate create in più motori.
- Concedi le autorizzazioni agli utenti per una vista che fa riferimento a più tabelle senza concedere le autorizzazioni direttamente sulle tabelle di riferimento sottostanti.
- Ottieni il filtraggio a livello di riga sulle tabelle utilizzando i tag LF (dove i tag LF si estendono a cascata solo fino al livello delle colonne) applicando i tag LF alle viste e concedendo autorizzazioni basate sui tag LF agli utenti.

Servizi di analisi supportati per le visualizzazioni AWS

I seguenti servizi AWS di analisi supportano la creazione di viste del catalogo dati:

- Amazon Redshift
- Amazon Athena versione 3
- Apache Spark su EMR Serverless
- Apache Spark nella versione 5.0 AWS Glue

Risorse aggiuntive

Puoi saperne di più sul Data Catalog in questa guida e utilizzare le seguenti risorse:

Il video seguente mostra come creare viste e interrogarle da Athena e Amazon Redshift.

Argomenti

- [Prerequisiti per la creazione di viste](#)
- [Creazione di viste del catalogo dati utilizzando istruzioni DDL](#)
- [Creazione di viste del catalogo dati utilizzando AWS Glue APIs](#)
- [Concessione delle autorizzazioni per le visualizzazioni del catalogo dati](#)
- [Viste materializzate](#)

Prerequisiti per la creazione di viste

- Per creare viste in Data Catalog, devi registrare le posizioni dei dati Amazon S3 sottostanti delle tabelle di riferimento con Lake Formation. Per i dettagli sulla registrazione dei dati con Lake Formation, vedere [Aggiungere una posizione Amazon S3 al tuo data lake](#).
- Solo i ruoli IAM possono creare viste del Data Catalog. Le altre identità IAM non possono creare viste del catalogo dati.
- Il ruolo IAM che definisce la vista deve disporre delle seguenti autorizzazioni:
 - SELECTAutorizzazione Lake Formation con Grantable opzione su tutte le tabelle di riferimento, incluse tutte le colonne.
 - CREATE_TABLEAutorizzazione Lake Formation sul database di destinazione in cui vengono create le viste.
 - Una politica di fiducia affinché la Lake Formation e AWS Glue i servizi assumano il ruolo.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "DataCatalogViewDefinerAssumeRole1",
      "Effect": "Allow",
      "Principal": {
        "Service": [
```

```

        "glue.amazonaws.com",
        "lakeformation.amazonaws.com"
    ],
    },
    "Action": "sts:AssumeRole"
}
]
}

```

- Lo scopo: PassRole autorizzazione per AWS Glue e Lake Formation.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "DataCatalogViewDefinerPassRole1",
      "Action": [
        "iam:PassRole"
      ],
      "Effect": "Allow",
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "iam:PassedToService": [
            "glue.amazonaws.com",
            "lakeformation.amazonaws.com"
          ]
        }
      }
    }
  ]
}

```

- AWS Glue e autorizzazioni Lake Formation.

JSON

```

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

```

```

        "Effect": "Allow",
        "Action": [
            "Glue:GetDatabase",
            "Glue:GetDatabases",
            "Glue:CreateTable",
            "Glue:GetTable",
            "Glue:GetTables",
            "Glue:BatchGetPartition",
            "Glue:GetPartitions",
            "Glue:GetPartition",
            "Glue:GetTableVersion",
            "Glue:GetTableVersions",
            "Glue:PassConnection",
            "lakeFormation:GetDataAccess"
        ],
        "Resource": "*"
    }
}

```

- Non è possibile creare viste in un database che dispone Super o ha ALL l'autorizzazione concessa al IAMAllowedPrincipals gruppo. Puoi revocare l'Superautorizzazione per il IAMAllowedPrincipals gruppo su un database [Passaggio 4: Passa i tuoi archivi dati al modello di autorizzazioni Lake Formation](#), vedere o creare un nuovo database con la casella Usa solo il controllo di accesso IAM per le nuove tabelle in questo database deselezionata in Autorizzazioni predefinite per le tabelle appena create.

Creazione di viste del catalogo dati utilizzando istruzioni DDL

Puoi creare AWS Glue Data Catalog viste utilizzando gli editor SQL per Athena, Amazon Redshift e utilizzando/. AWS Glue APIs AWS CLI

Per creare una vista del catalogo dati utilizzando gli editor SQL, scegli Athena o Redshift Spectrum e crea la vista utilizzando CREATE VIEW un'istruzione DDL (Data Definition Language). Dopo aver creato una vista nel dialetto del primo motore, puoi utilizzare un'istruzione ALTER VIEW DDL del secondo motore per aggiungere i dialetti aggiuntivi.

Quando si definiscono le viste, è importante considerare quanto segue:

- Definizione di viste multidialettali: quando si definisce una vista con più dialetti, gli schemi dei diversi dialetti devono corrispondere. Ogni dialetto SQL avrà una specifica sintattica leggermente

diversa. La sintassi della query che definisce la visualizzazione del catalogo dati dovrebbe corrispondere esattamente allo stesso elenco di colonne, inclusi sia i tipi che i nomi, in tutti i dialetti. Queste informazioni vengono memorizzate nella `StorageDescriptor` vista. I dialetti devono inoltre fare riferimento agli stessi oggetti della tabella sottostanti del Data Catalog.

Per aggiungere un altro dialetto a una vista utilizzando DDL, puoi usare l'istruzione `ALTER VIEW`. Se un'istruzione `ALTER VIEW` tenta di aggiornare la definizione della vista, ad esempio modificando il descrittore di archiviazione o le tabelle sottostanti della vista, l'istruzione restituisce un errore che dice «Input and existing storage descriptor mismatch». È possibile utilizzare le operazioni SQL cast per garantire che i tipi di colonne di visualizzazione corrispondano.

- Aggiornamento di una vista: per aggiornare la vista, puoi utilizzare l'UpdateTableAPI. Se aggiorni la vista senza corrispondere ai descrittori di archiviazione o alle tabelle di riferimento, puoi fornire il `FORCE` flag (consulta la documentazione SQL del motore per la sintassi). Dopo un aggiornamento forzato, la visualizzazione assumerà le tabelle forzate `StorageDescriptor` e di riferimento. Qualsiasi ulteriore `ALTER VIEW` DDL deve corrispondere ai valori modificati. Una vista che è stata aggiornata per includere dialetti incompatibili avrà lo stato «Stale». Lo stato della vista è visibile nella console di Lake Formation e utilizzando l'GetTableoperazione.
- Riferimento a un tipo di colonna varchar come stringa: non è possibile eseguire il cast di un tipo di colonna varchar di Redshift Spectrum su una stringa. Se viene creata una vista in Redshift Spectrum con un tipo di colonna varchar e un dialetto successivo tenta di fare riferimento a quel campo come una stringa, il Data Catalog la tratterà come stringa senza la necessità del flag. `FORCE`
- Trattamento di campi di tipo complesso: Amazon Redshift tratta tutti i tipi complessi come tipi [SUPER](#) mentre Athena specifica il tipo complesso. Se una vista ha un campo `SUPER` di tipo e un altro motore fa riferimento a quella colonna come a un particolare tipo complesso, ad esempio `struct(<street_address:struct<street_number:int, street_name:string, street_type:string>>)`, il Data Catalog presume che il campo sia il tipo complesso specifico e lo utilizza nel descrittore di archiviazione, senza richiedere il flag. `Force`

Per ulteriori informazioni sulla sintassi per la creazione e la gestione delle viste del catalogo dati, vedi:

- [Utilizzo delle AWS Glue Data Catalog visualizzazioni](#) nella Guida per l'utente di Amazon Athena.
- [Glue Data Catalog visualizza la sintassi delle query](#) nella Guida per l'utente di Amazon Athena.
- [Creazione di viste AWS Glue Data Catalog nella](#) Amazon Redshift Database Developer Guide.

Per ulteriori informazioni sui comandi SQL relativi alle viste nel Catalogo dati, consulta [CREATE EXTERNAL VIEW](#), [ALTER EXTERNAL VIEW](#) e [DROP EXTERNAL VIEW](#).

Dopo aver creato una vista Data Catalog, i dettagli della vista sono disponibili nella console Lake Formation.

1. Scegli Views in Data Catalog nella console Lake Formation.
2. Nella pagina delle visualizzazioni viene visualizzato un elenco delle viste disponibili.
3. Scegliete una vista dall'elenco e la pagina dei dettagli mostra gli attributi della vista.

[AWS Lake Formation](#) > [Views](#) > europe_players

europe_players

Version 1 (Current version) ▾ Actions ▾

Details

Name europe_players	Database views_demo_database	Definer role admin
Last updated November 22, 2023 at 10:41 PM UTC	Status Ready	Description -

Schema | **SQL definitions** | LF-Tags | Cross-account access | Underlying tables

SQL definitions (2)

List of available SQL definitions in different engines. Choose an engine from the list to add or edit the definition.

< 1 >

Engine name ▲	Version ▼	Status ▼	SQL statement	Edit definition
Athena	3	Ready	View	Amazon Athena
Redshift	1.0	Ready	View	Amazon Redshift

Schema

Scegliete una Column riga e selezionate Modifica tag LF per aggiornare i valori dei tag o assegnare nuovi tag LF.

Definizioni SQL

È possibile visualizzare un elenco di definizioni SQL disponibili. Seleziona **Aggiungi definizione SQL** e scegli un motore di query per aggiungere una definizione SQL. Scegli un motore di query (Athena o Amazon Redshift) sotto **Edit definition** la colonna per aggiornare una definizione SQL.

Tag LF

Scegliete **Modifica tag LF** per modificare i valori di un tag o assegnare nuovi tag. È possibile utilizzare i tag LF per concedere autorizzazioni sulle viste.

Accesso multi-account

Puoi vedere un elenco di Account AWS organizzazioni e unità organizzative (OUs) con cui hai condiviso la vista del catalogo dati.

Tabelle sottostanti

Le tabelle sottostanti a cui si fa riferimento nella definizione SQL utilizzata per creare la vista sono mostrate in questa scheda.

Creazione di viste del catalogo dati utilizzando AWS Glue APIs

È possibile utilizzare AWS Glue [CreateTable](#) e [UpdateTable](#) APIs creare e aggiornare le viste nel Data Catalog. Le `UpdateTable` operazioni `CreateTable` and hanno una nuova `TableInput` struttura `ViewDefinition`, mentre, `SearchTables`, `GetTable`, `GetTables` `GetTableVersion`, `GetTableVersions` le operazioni forniscono la `ViewDefinition` sintassi di output per le viste. Inoltre, c'è un nuovo `Status` campo nell'output dell'`GetTableAPI`.

Sono disponibili due nuove AWS Glue connessioni per la convalida del dialetto SQL per ogni motore di query supportato Amazon Athena e Amazon Redshift.

Le `CreateTable` e `UpdateTable` APIs sono asincrone se utilizzate con le viste. Quando questi APIs vengono chiamati con più dialetti SQL, la chiamata viene convalidata con ciascun motore per determinare se il dialetto può essere eseguito su quel motore e se lo schema risultante della vista di ciascun dialetto corrisponde. Il AWS Glue servizio utilizza queste connessioni per effettuare chiamate interne ai motori di analisi. Queste chiamate simulano ciò che fa il motore per convalidare se sul motore è stata eseguita una `DDL CREATE VIEW` o `ALTER VIEW SQL`.

Se l'SQL fornito è valido e gli schemi corrispondono tra i dialetti di visualizzazione, l' AWS Glue API esegue il commit atomico del risultato. Atomicity consente di creare o modificare viste con più dialetti senza tempi di inattività.

Argomenti

- [Creazione di connessioni per AWS Glue convalidare lo stato](#)
- [Convalida dello stato di generazione della vista](#)
- [Stati e operazioni asincroni](#)
- [Visualizza gli scenari di errore di creazione durante le operazioni asincrone](#)

Creazione di connessioni per AWS Glue convalidare lo stato

Per creare o aggiornare una AWS Glue Data Catalog vista utilizzando le UpdateTable operazioni CreateTable or, devi creare un nuovo tipo di AWS Glue connessione per la convalida e fornirlo al motore di analisi supportato. Queste connessioni sono necessarie per utilizzare le viste del catalogo dati con Athena o Amazon Redshift. Puoi creare queste connessioni solo utilizzando AWS CLI AWS SDKs, o. AWS Glue APIs Non è possibile utilizzare il Console di gestione AWS per creare la AWS Glue connessione.

Note

Se il ruolo view definer e il role calling CreateTable or UpdateTable sono diversi, entrambi richiedono l'glue:PassConnectionautorizzazione nella dichiarazione sulla politica IAM.

Per ulteriori informazioni, consulta la documentazione [create-connection](#) AWS CLI .

AWS CLI comando per creare una connessione

Di seguito è riportato un AWS CLI comando per creare una connessione:

```
aws glue create-connection --region us-east-1
--endpoint-url https://glue.us-east-1.amazonaws.com
--cli-input-json file:///root/path/to/create-connection.json
```

AWS CLI input JSON

Per Amazon Redshift:

```
{
  "CatalogId": "123456789012",
  "ConnectionInput": {
    "ConnectionType": "VIEW_VALIDATION_REDSHIFT",
    "Name": "views-preview-cluster-connection-2",
    "Description": "My first Amazon Redshift validation connection",
    "ConnectionProperties": {
      "DATABASE": "dev",
      "CLUSTER_IDENTIFIER": "glue-data-catalog-views-preview-cluster"
    }
  }
}
```

Per Amazon Athena:

```
{
  "CatalogId": "123456789012",
  "ConnectionInput": {
    "ConnectionType": "VIEW_VALIDATION_ATHENA",
    "Name": "views-preview-cluster-connection-3",
    "Description": "My first Amazon Athena validation connection",
    "ConnectionProperties": {
      "WORKGROUP_NAME": "workgroup-name"
    }
  }
}
```

Convalida dello stato di generazione della vista

Quando si eseguono le UpdateTable operazioni CreateTable or, il Status campo per l'output dell'GetTableAPI mostra i dettagli dello stato di creazione della vista. Per create le richieste in cui la tabella non esiste già, AWS Glue crea una tabella vuota per la durata del processo asincrono. Durante la chiamataGetTable, è possibile passare un flag booleano opzionaleIncludeStatusDetails, che mostra le informazioni diagnostiche sulla richiesta. In caso di errore, questo flag mostra un messaggio di errore con gli stati individuali di ogni dialetto.

Gli errori durante le operazioni di visualizzazione, creazione, lettura, aggiornamento ed eliminazione (CRUD) possono verificarsi durante l'elaborazione nel servizio AWS Glue/Lake Formation o durante la convalida SQL della visualizzazione in Amazon Redshift o Athena. Quando si verifica un errore

durante la convalida in un motore, il AWS Glue servizio fornisce il messaggio di errore restituito dal motore.

Campi di stato

I seguenti sono i campi di stato:

- **Status:** uno stato generico, che è agnostico per diversi tipi di lavori:
 - IN CODA
 - IN_PROGRESS
 - RIUSCITO
 - NON RIUSCITO
- **Azione:** indica quale azione è stata richiamata sulla tabella, attualmente solo CREATE o quali UPDATE operazioni sono disponibili.

La distinzione tra UPDATE e CREATE operazioni è importante quando si lavora con le viste. Il tipo di operazione determina come procedere con l'interrogazione delle tabelle.

Un'UPDATE operazione indica che la tabella esiste già nel Data Catalog. In questo caso, puoi continuare a interrogare la tabella creata in precedenza senza problemi. D'altra parte, un'CREATE operazione indica che la tabella non è mai stata creata correttamente prima. Se una tabella è contrassegnata per CREATE, il tentativo di interrogarla avrà esito negativo perché la tabella non esiste ancora nel sistema. Pertanto, è essenziale identificare il tipo di operazione (UPDATE o CREATE) prima di tentare di interrogare una tabella.

- **RequestedBy** — L'ARN dell'utente che ha richiesto la modifica asincrona.
- **UpdatedBy** — L'ARN dell'utente che per ultimo ha modificato manualmente il processo di modifica asincrono, ad esempio richiedendo una cancellazione o una modifica.
- **Errore:** questo campo viene visualizzato solo quando lo stato è FAILED. Si tratta di un messaggio di eccezione a livello principale. Potrebbero esserci errori diversi per ogni dialetto.
 - **ErrorCode** — Il tipo di eccezione.
 - **ErrorMessage** — una breve descrizione dell'eccezione.
- **RequestTime** — una stringa di data in formato ISO 8601 che indica l'ora in cui è stata avviata la modifica.
- **UpdateTime** — una stringa di data in formato ISO 8601 che indica l'ora dell'ultimo aggiornamento dello stato.

Stati e operazioni asincroni

Quando si esegue una `glue:CreateTable` richiesta, inizia la creazione asincrona della vista del catalogo dati. Nelle sezioni seguenti, questo documento descrive una `Status AWS Glue` vista disponibile in una risposta `glue:GetTable`. Per brevità, questa sezione omette la risposta completa.

```
{
  "Table": {
    ...
    "Status": {
      ...
      "Action": "CREATE",
      "State": "QUEUED",
    }
  }
}
```

Entrambi gli attributi precedenti rappresentano importanti informazioni diagnostiche che indicano lo stato dell'operazione asincrona e le azioni che possono essere eseguite su questa vista. Di seguito sono riportati i possibili valori che questi attributi possono assumere.

1. `Status.Action`

- a. CREATE
- b. UPDATE

2. `Status.State`

- a. IN CODA
- b. IN_PROGRESS
- c. RIUSCITO
- d. NON RIUSCITO

È inoltre importante notare che alcuni aggiornamenti su una visualizzazione del catalogo dati non richiedono un'operazione asincrona. Ad esempio, si potrebbe voler aggiornare l'`Description` attributo della tabella. Poiché ciò non richiede alcuna operazione asincrona, i metadati della tabella risultanti non ne `Status` avranno e l'attributo lo sarà. `NULL`

```
{
  "Table": {
```

```
    ...,
    "Description": "I changed this attribute!"
  }
}
```

Successivamente, questo argomento esplora in che modo le informazioni sullo stato di cui sopra possono influire sulle operazioni che possono essere eseguite su una vista. AWS Glue

colla: CreateTable

Non ci sono modifiche per questa API rispetto al modo in cui `glue:CreateTable` funziona per qualsiasi tabella Glue. `CreateTable` può essere chiamato per qualsiasi nome di tabella che non esista già.

colla: UpdateTable

Questa operazione non può essere eseguita su una AWS Glue vista che contiene le seguenti informazioni di stato:

1. Azione == CREATE e State == IN CODA
2. Azione == CREATE e State == IN_PROGRESS
3. Azione == CREATE e state == FALLITO
4. Azione == AGGIORNAMENTO e stato == IN CODA
5. Azione == AGGIORNAMENTO e stato == IN_PROGRESS

Per riassumere, è possibile aggiornare una vista del catalogo dati solo quando soddisfa i seguenti requisiti.

1. È stata creata con successo per la prima volta.
 - a. Azione == CREATE e State == SUCCESS
2. Ha raggiunto lo stato terminale dopo un'operazione di aggiornamento asincrona.
 - a. Azione == AGGIORNAMENTO e stato == SUCCESSO
 - b. Azione == AGGIORNAMENTO e stato == FALLITO
3. Ha un attributo di NULL stato come risultato di un aggiornamento sincrono.

colla: DeleteTable

Questa operazione non presenta modifiche rispetto al `glue:DeleteTable` funzionamento di qualsiasi AWS Glue tabella. È possibile eliminare una vista del catalogo dati indipendentemente dal suo stato.

colla: `GetTable`

Questa operazione non presenta modifiche rispetto al `glue:GetTable` funzionamento di qualsiasi AWS Glue tabella. Tuttavia, non è possibile interrogare una vista del catalogo dati dai motori di analisi finché non è stata creata correttamente per la prima volta. `Action == CREATE and State == SUCCESS`. Dopo aver creato correttamente una vista del catalogo dati per la prima volta, è possibile interrogare la vista indipendentemente dal suo stato.

Note

Tutte le informazioni contenute in questa sezione si applicano a tutte le tabelle `GetTable`, APIs ad esempio `GetTables`, `SearchTables`.

Visualizza gli scenari di errore di creazione durante le operazioni asincrone

Gli esempi seguenti sono rappresentativi dei tipi di errori che possono derivare `CreateTable` o `UpdateTable` visualizzare le chiamate API. Non sono esaustivi in quanto la superficie di errore degli errori delle query SQL è piuttosto ampia.

Scenario 1: errore di interrogazione su Amazon Redshift

La query fornita per Amazon Redshift include un nome di tabella con errori di ortografia che non può essere trovato nel Data Catalog durante la convalida. L'errore risultante viene mostrato nel `Status` campo della `GetTable` risposta della vista.

`GetTable` richiesta:

```
{
  "CatalogId": "123456789012",
  "DatabaseName": "async-view-test-db",
  "TableInput": {
    "Name": "view-athena-redshift-72",
    "Description": "This is an atomic operation",
    "StorageDescriptor": {
      "Columns": [
        { "Name": "col1", "Type": "int" },
```

```

        { "Name": "col2", "Type": "string" },
        { "Name": "col3", "Type": "double" }
    ]
},
"ViewDefinition": {
    "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
    "SubObjects": [ "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1" ],
    "Representations": [
        {
            "Dialect": "ATHENA",
            "DialectVersion": "3",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-playground-db\".
\"table_1\";",
            "ValidationConnection": "athena-connection"
        },
        {
            "Dialect": "REDSHIFT",
            "DialectVersion": "1.0",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-playground-external-
schema\".\"table__1\";",
            "ValidationConnection": "redshift-connection"
        }
    ]
}
}
}
}

```

GetTablerisposta:

```

IncludeStatusDetails = FALSE
{
    "Table": {
        "Name": "view-athena-redshift-72",
        "DatabaseName": "async-view-test-db",
        "Description": "",
        "CreateTime": "2024-07-11T11:39:19-07:00",
        "UpdateTime": "2024-07-11T11:39:19-07:00",
        "Retention": 0,
        "ViewOriginalText": "",
        "ViewExpandedText": "",
        "TableType": "",
        "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    }
}

```

```

    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:39:19-07:00",
      "UpdateTime": "2024-07-11T11:40:06-07:00",
      "Action": "CREATE",
      "State": "FAILED"
    }
  }
}

IncludeStatusDetails = TRUE
{
  "Table": {
    "Name": "view-athena-redshift-72",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:39:19-07:00",
    "UpdateTime": "2024-07-11T11:39:19-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:39:19-07:00",
      "UpdateTime": "2024-07-11T11:40:06-07:00",
      "Action": "CREATE",
      "State": "FAILED",
      "Error": {

```



```

      "ErrorCode": "QueryExecutionException",
      "ErrorMessage": "Error received during view SQL validation
using a connection: [Connection Name: redshift-connection | Query Execution
Id: ddb711d3-2415-4aa9-b251-6a76ab4f41b1 | Timestamp: Thu Jul 11 18:39:37 UTC
2024]: Redshift returned error for the statement: ERROR: AwsClientException:
EntityNotFoundException from glue - Entity Not Found"
    },
    "Details": {
      "RequestedChange": {
        "Name": "view-athena-redshift-72",
        "DatabaseName": "async-view-test-db",
        "Description": "This is an atomic operation",
        "Retention": 0,
        "StorageDescriptor": {
          "Columns": [
            {
              "Name": "col1",
              "Type": "int"
            },
            {
              "Name": "col2",
              "Type": "string"
            },
            {
              "Name": "col3",
              "Type": "double"
            }
          ],
          "Compressed": false,
          "NumberOfBuckets": 0,
          "SortColumns": [],
          "StoredAsSubDirectories": false
        },
        "TableType": "VIRTUAL_VIEW",
        "IsRegisteredWithLakeFormation": false,
        "CatalogId": "123456789012",
        "IsRowFilteringEnabled": false,
        "VersionId": "-1",
        "DatabaseId": "<databaseID>",
        "ViewDefinition": {
          "IsProtected": true,
          "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
          "SubObjects": [

```

```

        "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1"
    ],
    "Representations": [
        {
            "Dialect": "ATHENA",
            "DialectVersion": "3",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-
playground-db\".\"table_1\"",
            "IsStale": false
        },
        {
            "Dialect": "REDSHIFT",
            "DialectVersion": "1.0",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-
playground-external-schema\".\"table__1\";",
            "IsStale": false
        }
    ]
},
    "IsMultiDialectView": true
},
    "ViewValidations": [
        {
            "Dialect": "ATHENA",
            "DialectVersion": "3",
            "ViewValidationText": "SELECT * FROM \"gdc-view-playground-db
\".\"table_1\"",
            "UpdateTime": "2024-07-11T11:40:06-07:00",
            "State": "SUCCESS"
        },
        {
            "Dialect": "REDSHIFT",
            "DialectVersion": "1.0",
            "ViewValidationText": "SELECT * FROM \"gdc-view-playground-
external-schema\".\"table__1\";",
            "UpdateTime": "2024-07-11T11:39:37-07:00",
            "State": "FAILED",
            "Error": {
                "ErrorCode": "QueryExecutionException",
                "ErrorMessage": "Error received during view SQL validation
using a connection: [Connection Name: redshift-connection | Query Execution Id:
ddb711d3-2415-4aa9-b251-6a76ab4f41b1 | Timestamp: Thu

```

```
Jul 11 18:39:37 UTC 2024]: Redshift returned error for the statement: ERROR:  
AwsClientException: EntityNotFoundException from glue - Entity Not Found"  
    }  
  }  
]  
}  
}  
}
```

Scenario 2: connessione Amazon Redshift non valida

La connessione Amazon Redshift nell'esempio seguente non è valida perché fa riferimento a un database Amazon Redshift che non esiste nell'endpoint fornito. cluster/serverless Amazon Redshift non è in grado di convalidare la vista e il Status campo nella GetTable risposta mostra l'errore (da Amazon "State": "FAILED" Redshift).

GetTablerichiesta:

```
{
  "CatalogId": "123456789012",
  "DatabaseName": "async-view-test-db",
  "TableInput": {
    "Name": "view-athena-redshift-73",
    "Description": "This is an atomic operation",
    "StorageDescriptor": {
      "Columns": [
        { "Name": "col1", "Type": "int" },
        { "Name": "col2", "Type": "string" },
        { "Name": "col3", "Type": "double" }
      ]
    },
    "ViewDefinition": {
      "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
      "SubObjects": [ "arn:aws:glue:us-east-1:123456789012:table/gdc-view-playground-db/table_1" ],
      "Representations": [
        {
          "Dialect": "ATHENA",
          "DialectVersion": "3",
          "ViewOriginalText": "SELECT * FROM \"gdc-view-playground-db\".\"table_1\"",
          "ValidationConnection": "athena-connection"
        }
      ]
    }
  }
}
```

```

    },
    {
      "Dialect": "REDSHIFT",
      "DialectVersion": "1.0",
      "ViewOriginalText": "SELECT * FROM \"gdc-view-playground-external-
schema\".\"table_1\";",
      "ValidationConnection": "redshift-connection-malformed"
    }
  ]
}
}
}

```

GetTablerisposta:

```

IncludeStatusDetails = FALSE
{
  "Table": {
    "Name": "view-athena-redshift-73",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:43:27-07:00",
    "UpdateTime": "2024-07-11T11:43:27-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:43:27-07:00",
      "UpdateTime": "2024-07-11T11:43:40-07:00",
      "Action": "CREATE",
      "State": "FAILED"
    }
  }
}

```

```

}

IncludeStatusDetails = TRUE
{
  "Table": {
    "Name": "view-athena-redshift-73",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:43:27-07:00",
    "UpdateTime": "2024-07-11T11:43:27-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:43:27-07:00",
      "UpdateTime": "2024-07-11T11:43:40-07:00",
      "Action": "CREATE",
      "State": "FAILED",
      "Error": {
        "ErrorCode": "QueryExecutionException",
        "ErrorMessage": "Error received during view SQL validation using a
connection: [Connection Name: redshift-connection-malformed | Query Execution Id:
69bfafd4-3d51-4cb0-9320-7ce5404b1809 | Timestamp: Thu Jul 11 18:43:38 UTC 2024]:
Redshift returned error for the statement: FATAL: database \"devoov\" does not exist"
      },
      "Details": {
        "RequestedChange": {
          "Name": "view-athena-redshift-73",
          "DatabaseName": "async-view-test-db",
          "Description": "This is an atomic operation",
          "Retention": 0,
          "StorageDescriptor": {
            "Columns": [
              {

```

```

        "Name": "col1",
        "Type": "int"
    },
    {
        "Name": "col2",
        "Type": "string"
    },
    {
        "Name": "col3",
        "Type": "double"
    }
],
"Compressed": false,
"NumberOfBuckets": 0,
"SortColumns": [],
"StoredAsSubDirectories": false
},
"TableType": "VIRTUAL_VIEW",
"IsRegisteredWithLakeFormation": false,
"CatalogId": "123456789012",
"IsRowFilteringEnabled": false,
"VersionId": "-1",
"DatabaseId": "<databaseID>",
"ViewDefinition": {
    "IsProtected": true,
    "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
    "SubObjects": [
        "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1"
    ],
    "Representations": [
        {
            "Dialect": "ATHENA",
            "DialectVersion": "3",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-
playground-db\".\"table_1\"",
            "IsStale": false
        },
        {
            "Dialect": "REDSHIFT",
            "DialectVersion": "1.0",
            "ViewOriginalText": "SELECT * FROM \"gdc-view-
playground-external-schema\".\"table_1\";",
            "IsStale": false
        }
    ]
}

```

```

    }
  ],
  "IsMultiDialectView": true
},
"ViewValidations": [
  {
    "Dialect": "ATHENA",
    "DialectVersion": "3",
    "ViewValidationText": "SELECT * FROM \"gdc-view-playground-db\n\".\n\"table_1\n\"",
    "UpdateTime": "2024-07-11T11:43:40-07:00",
    "State": "SUCCESS"
  },
  {
    "Dialect": "REDSHIFT",
    "DialectVersion": "1.0",
    "ViewValidationText": "SELECT * FROM \"gdc-view-playground-external-schema\n\".\n\"table_1\n\";",
    "UpdateTime": "2024-07-11T11:43:38-07:00",
    "State": "FAILED",
    "Error": {
      "ErrorCode": "QueryExecutionException",
      "ErrorMessage": "Error received during view SQL validation using a connection: [Connection Name: redshift-connection-malformed | Query Execution Id: 69bfafd4-3d51-4cb0-9320-7ce5404b1809 | Time stamp: Thu Jul 11 18:43:38 UTC 2024]: Redshift returned error for the statement: FATAL: database \"devooo\" does not exist"
    }
  }
]
}
}
}
```

Scenario 3: interrogazione Athena non riuscita

L'SQL per Athena qui non è valido perché la query digita erroneamente il nome del database. La convalida della query Athena rileva questo problema e l'errore risultante viene visualizzato attraverso l'Statusoggetto in una chiamata. GetTable

GetTablerichiesta:

```
{
  "CatalogId": "123456789012",
  "DatabaseName": "async-view-test-db",
  "TableInput": {
    "Name": "view-athena-redshift-70",
    "Description": "This is an atomic operation",
    "StorageDescriptor": {
      "Columns": [
        { "Name": "col1", "Type": "int" },
        { "Name": "col2", "Type": "string" },
        { "Name": "col3", "Type": "double" }
      ]
    },
    "ViewDefinition": {
      "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
      "SubObjects": [ "arn:aws:glue:us-east-1:123456789012:table/gdc-view-playground-db/table_1" ],
      "Representations": [
        {
          "Dialect": "ATHENA",
          "DialectVersion": "3",
          "ViewOriginalText": "SELECT * FROM \"gdc--view-playground-db\".\"table_1\"",
          "ValidationConnection": "athena-connection"
        },
        {
          "Dialect": "REDSHIFT",
          "DialectVersion": "1.0",
          "ViewOriginalText": "SELECT * FROM \"gdc-view-playground-external-schema\".\"table_1\";",
          "ValidationConnection": "redshift-connection"
        }
      ]
    }
  }
}
```

GetTablerisposta:

```
IncludeStatusDetails = FALSE
{
```



```

"Table": {
  "Name": "view-athena-redshift-70",
  "DatabaseName": "async-view-test-db",
  "Description": "",
  "CreateTime": "2024-07-11T11:09:53-07:00",
  "UpdateTime": "2024-07-11T11:09:53-07:00",
  "Retention": 0,
  "ViewOriginalText": "",
  "ViewExpandedText": "",
  "TableType": "",
  "CreatedBy": "arn:aws:iam::123456789012:user/",
  "IsRegisteredWithLakeFormation": false,
  "CatalogId": "123456789012",
  "IsRowFilteringEnabled": false,
  "VersionId": "-1",
  "DatabaseId": "<databaseID>",
  "IsMultiDialectView": false,
  "Status": {
    "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "RequestTime": "2024-07-11T11:09:54-07:00",
    "UpdateTime": "2024-07-11T11:10:41-07:00",
    "Action": "CREATE",
    "State": "FAILED",
  }
}
}

```

```

IncludeStatusDetails = TRUE

```

```

{
  "Table": {
    "Name": "view-athena-redshift-70",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:09:53-07:00",
    "UpdateTime": "2024-07-11T11:09:53-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,

```

```

"VersionId": "-1",
"DatabaseId": "<databaseID>",
"IsMultiDialectView": false,
"Status": {
  "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
  "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
  "RequestTime": "2024-07-11T11:09:54-07:00",
  "UpdateTime": "2024-07-11T11:10:41-07:00",
  "Action": "CREATE",
  "State": "FAILED",
  "Error": {
    "ErrorCode": "QueryExecutionException",
    "ErrorMessage": "Error received during view SQL validation using
a connection: [Connection Name: athena-connection | Query Execution Id: d9bb1e6d-
ce26-4b35-8276-8a199af966aa | Timestamp: Thu Jul 11 18:10:
41 UTC 2024]: Athena validation FAILED: {ErrorCategory: 2,ErrorType: 1301,Retryable:
false,ErrorMessage: line 1:118: Schema 'gdc--view-playground-db' does not exist}"
  },
  "Details": {
    "RequestedChange": {
      "Name": "view-athena-redshift-70",
      "DatabaseName": "async-view-test-db",
      "Description": "This is an atomic operation",
      "Retention": 0,
      "StorageDescriptor": {
        "Columns": [
          {
            "Name": "col1",
            "Type": "int"
          },
          {
            "Name": "col2",
            "Type": "string"
          },
          {
            "Name": "col3",
            "Type": "double"
          }
        ],
        "Compressed": false,
        "NumberOfBuckets": 0,
        "SortColumns": [],
        "StoredAsSubDirectories": false
      }
    }
  }
},

```

```

        "TableType": "VIRTUAL_VIEW",
        "IsRegisteredWithLakeFormation": false,
        "CatalogId": "123456789012",
        "IsRowFilteringEnabled": false,
        "VersionId": "-1",
        "DatabaseId": "<databaseID>",
        "ViewDefinition": {
            "IsProtected": true,
            "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
            "SubObjects": [
                "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1"
            ],
            "Representations": [
                {
                    "Dialect": "ATHENA",
                    "DialectVersion": "3",
                    "ViewOriginalText": "SELECT * FROM \"gdc--view-
playground-db\".\"table_1\"",
                    "IsStale": false
                },
                {
                    "Dialect": "REDSHIFT",
                    "DialectVersion": "1.0",
                    "ViewOriginalText": "SELECT * FROM \"gdc-view-
playground-external-schema\".\"table_1\";",
                    "IsStale": false
                }
            ]
        },
        "IsMultiDialectView": true
    },
    "ViewValidations": [
        {
            "Dialect": "ATHENA",
            "DialectVersion": "3",
            "ViewValidationText": "SELECT * FROM \"gdc--view-playground-db
\".\"table_1\"",
            "UpdateTime": "2024-07-11T11:10:41-07:00",
            "State": "FAILED",
            "Error": {
                "ErrorCode": "QueryExecutionException",

```

```

        "ErrorMessage": "Error received during view SQL validation
using a connection: [Connection Name: athena-connection | Query Execution Id:
d9bb1e6d-ce26-4b35-8276-8a199af966aa | Timestamp: Thu J
ul 11 18:10:41 UTC 2024]: Athena validation FAILED: {ErrorCategory: 2,ErrorType:
1301,Retryable: false,ErrorMessage: line 1:118: Schema 'gdc--view-playground-db' does
not exist}"
    },
    {
        "Dialect": "REDSHIFT",
        "DialectVersion": "1.0",
        "ViewValidationText": "SELECT * FROM \"gdc-view-playground-
external-schema\".\"table_1\";",
        "UpdateTime": "2024-07-11T11:10:41-07:00",
        "State": "SUCCESS"
    }
]
}
}
}
}
}

```

Scenario 4: descrittori di archiviazione non corrispondenti

L'SQL fornito per il dialetto Athena seleziona col1, col2 mentre l'SQL per Redshift seleziona solo. col1 Ciò porta a un errore di mancata corrispondenza del descrittore di archiviazione.

GetTablerichiasta:

```

{
  "CatalogId": "123456789012",
  "DatabaseName": "async-view-test-db",
  "TableInput": {
    "Name": "view-athena-redshift-71",
    "Description": "This is an atomic operation",
    "StorageDescriptor": {
      "Columns": [
        { "Name": "col1", "Type": "int" },
        { "Name": "col2", "Type": "string" },
        { "Name": "col3", "Type": "double" }
      ]
    },
    "ViewDefinition": {

```

```

    "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
    "SubObjects": [ "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1" ],
    "Representations": [
      {
        "Dialect": "ATHENA",
        "DialectVersion": "3",
        "ViewOriginalText": "SELECT col1, col2 FROM \"gdc-view-playground-
db\".\"table_1\"",
        "ValidationConnection": "athena-connection"
      },
      {
        "Dialect": "REDSHIFT",
        "DialectVersion": "1.0",
        "ViewOriginalText": "SELECT col1 FROM \"gdc-view-playground-
external-schema\".\"table_1\";",
        "ValidationConnection": "redshift-connection"
      }
    ]
  }
}

```

GetTablerisposta:

```

IncludeStatusDetails = FALSE

{
  "Table": {
    "Name": "view-athena-redshift-71",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:22:02-07:00",
    "UpdateTime": "2024-07-11T11:22:02-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",

```

```

    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:22:02-07:00",
      "UpdateTime": "2024-07-11T11:23:19-07:00",
      "Action": "CREATE",
      "State": "FAILED"
    }
  }
}

IncludeStatusDetails = TRUE

{
  "Table": {
    "Name": "view-athena-redshift-71",
    "DatabaseName": "async-view-test-db",
    "Description": "",
    "CreateTime": "2024-07-11T11:22:02-07:00",
    "UpdateTime": "2024-07-11T11:22:02-07:00",
    "Retention": 0,
    "ViewOriginalText": "",
    "ViewExpandedText": "",
    "TableType": "",
    "CreatedBy": "arn:aws:iam::123456789012:user/zcaisse",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "IsMultiDialectView": false,
    "Status": {
      "RequestedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "UpdatedBy": "arn:aws:iam::123456789012:user/zcaisse",
      "RequestTime": "2024-07-11T11:22:02-07:00",
      "UpdateTime": "2024-07-11T11:23:19-07:00",
      "Action": "CREATE",
      "State": "FAILED",
      "Error": {
        "ErrorCode": "InvalidInputException",
        "ErrorMessage": "Engine and existing storage descriptor mismatch"
      }
    },
  },
}

```

```

"Details": {
  "RequestedChange": {
    "Name": "view-athena-redshift-71",
    "DatabaseName": "async-view-test-db",
    "Description": "This is an atomic operation",
    "Retention": 0,
    "StorageDescriptor": {
      "Columns": [
        {
          "Name": "col1",
          "Type": "int"
        },
        {
          "Name": "col2",
          "Type": "string"
        },
        {
          "Name": "col3",
          "Type": "double"
        }
      ],
      "Compressed": false,
      "NumberOfBuckets": 0,
      "SortColumns": [],
      "StoredAsSubDirectories": false
    },
    "TableType": "VIRTUAL_VIEW",
    "IsRegisteredWithLakeFormation": false,
    "CatalogId": "123456789012",
    "IsRowFilteringEnabled": false,
    "VersionId": "-1",
    "DatabaseId": "<databaseID>",
    "ViewDefinition": {
      "IsProtected": true,
      "Definer": "arn:aws:iam::123456789012:role/GDCViewDefiner",
      "SubObjects": [
        "arn:aws:glue:us-east-1:123456789012:table/gdc-view-
playground-db/table_1"
      ],
      "Representations": [
        {
          "Dialect": "ATHENA",
          "DialectVersion": "3",

```

```

        "ViewOriginalText": "SELECT col1, col2 FROM \"gdc-view-
playground-db\".\"table_1\"\"",
        "IsStale": false
    },
    {
        "Dialect": "REDSHIFT",
        "DialectVersion": "1.0",
        "ViewOriginalText": "SELECT col1 FROM \"gdc-view-
playground-external-schema\".\"table_1\";",
        "IsStale": false
    }
]
},
"IsMultiDialectView": true
},
"ViewValidations": [
    {
        "Dialect": "ATHENA",
        "DialectVersion": "3",
        "ViewValidationText": "SELECT col1, col2 FROM \"gdc-view-
playground-db\".\"table_1\"\"",
        "UpdateTime": "2024-07-11T11:23:19-07:00",
        "State": "FAILED",
        "Error": {
            "ErrorCode": "InvalidInputException",
            "ErrorMessage": "Engine and existing storage descriptor
mismatch"
        }
    },
    {
        "Dialect": "REDSHIFT",
        "DialectVersion": "1.0",
        "ViewValidationText": "SELECT col1 FROM \"gdc-view-playground-
external-schema\".\"table_1\";",
        "UpdateTime": "2024-07-11T11:22:49-07:00",
        "State": "FAILED",
        "Error": {
            "ErrorCode": "InvalidInputException",
            "ErrorMessage": "Engine and existing storage descriptor
mismatch"
        }
    }
]
}

```



```
}  
  }  
}
```

Concessione delle autorizzazioni per le visualizzazioni del catalogo dati

Dopo aver creato le viste in AWS Glue Data Catalog, puoi concedere le autorizzazioni per le visualizzazioni del data lake ai responsabili di tutte Account AWS le organizzazioni e le unità organizzative. È possibile concedere le autorizzazioni utilizzando i tag LF o il metodo di risorsa denominato. Per ulteriori informazioni sull'etichettatura delle risorse, consulta. [Controllo degli accessi basato su tag Lake Formation](#) Per ulteriori informazioni sulla concessione diretta delle autorizzazioni per le viste, consulta. [Concessione delle autorizzazioni sulle viste utilizzando il metodo di risorsa denominato](#)

Viste materializzate

Argomenti

- [Differenziazione delle viste materializzate da altri tipi di viste](#)
- [Casi d'uso](#)
- [Concetti chiave](#)
- [Autorizzazioni per le viste materializzate](#)
- [Creazione e gestione di viste materializzate](#)
- [Archiviazione e accesso ai dati](#)
- [AWS Lake Formation Integrazione con le autorizzazioni](#)
- [Monitoraggio e debug](#)
- [Gestione dei processi di aggiornamento](#)
- [Monitoraggio e risoluzione dei problemi](#)
- [Considerazioni e limitazioni](#)

Nel AWS Glue Data Catalog, una vista materializzata è una tabella gestita che memorizza il risultato precalcolato di una query SQL in formato Apache Iceberg. A differenza delle viste standard del Data Catalog che eseguono la query ogni volta che vi si accede, le viste materializzate archiviano fisicamente i risultati della query e li aggiornano man mano che le tabelle di origine sottostanti cambiano. Puoi creare viste materializzate utilizzando Apache Spark versione 3.5.6+ in Amazon Athena, Amazon EMR o. AWS Glue

Le viste materializzate fanno riferimento alle tabelle Apache Iceberg registrate nel AWS Glue Data Catalog, con dati precalcolati archiviati come tabelle Apache Iceberg in bucket Amazon S3 Tables o bucket Amazon S3 per uso generico, rendendole accessibili da più motori di query tra cui Amazon Athena, Amazon Redshift e motori di terze parti compatibili con Iceberg.

Differenziazione delle viste materializzate da altri tipi di viste

Le viste materializzate differiscono dalle viste AWS Glue Data Catalog, dalle viste Apache Spark e dalle viste Amazon Athena per alcuni aspetti fondamentali. Sebbene le viste del catalogo dati siano tabelle virtuali che eseguono la definizione della query SQL ogni volta che vi si accede, le viste materializzate archiviano fisicamente i risultati delle query precalcolate. Ciò elimina il calcolo ridondante e migliora significativamente le prestazioni delle query per le trasformazioni complesse a cui si accede di frequente.

Le viste materializzate differiscono anche dalle tradizionali pipeline di trasformazione dei dati create con AWS Glue ETL o job Spark personalizzati. Invece di scrivere codice personalizzato per gestire il rilevamento delle modifiche, gli aggiornamenti incrementali e l'orchestrazione del flusso di lavoro, definisci le viste materializzate utilizzando la sintassi SQL standard. Il AWS Glue Data Catalog monitora automaticamente le tabelle di origine, rileva le modifiche e aggiorna le viste materializzate utilizzando un'infrastruttura di elaborazione completamente gestita.

Casi d'uso

Di seguito sono riportati importanti casi d'uso per le viste materializzate:

- **Accelera le query analitiche complesse:** crea viste materializzate che precalcolano giunzioni, aggregazioni e funzioni di finestra costose. I motori Spark riscrivono automaticamente le query successive per utilizzare i risultati precalcolati, riducendo la latenza delle query e i costi di elaborazione.
- **Semplifica le pipeline di trasformazione dei dati:** sostituisci i complessi job ETL che gestiscono il rilevamento delle modifiche, gli aggiornamenti incrementali e l'orchestrazione del flusso di lavoro con semplici definizioni di viste materializzate basate su SQL. Il Data Catalog gestisce automaticamente tutta la complessità operativa. AWS Glue
- **Abilita l'analisi self-service con accesso regolamentato ai dati:** crea viste materializzate curate che trasformano i dati grezzi in set di dati pronti per l'uso. Concedi agli utenti l'accesso alle viste materializzate senza esporre le tabelle di origine sottostanti, semplificando la gestione della sicurezza e potenziando l'analisi self-service.
- **Ottimizza la progettazione delle funzionalità per l'apprendimento automatico:** definisci viste materializzate che implementano le trasformazioni delle funzionalità per i modelli ML. La

funzionalità di aggiornamento automatico assicura che gli archivi di funzionalità rimangano aggiornati man mano che i dati di origine si evolvono, mentre l'aggiornamento incrementale riduce al minimo i costi di elaborazione.

- Implementa una condivisione efficiente dei dati: crea viste materializzate che filtrano e trasformano i dati per consumatori specifici. Condividi le viste materializzate tra account e aree geografiche utilizzando AWS Lake Formation, eliminando la necessità di duplicazione dei dati e mantenendo al contempo una governance centralizzata.

Concetti chiave

Aggiornamento automatico

L'aggiornamento automatico è una funzionalità che monitora continuamente le tabelle di origine e aggiorna le viste materializzate in base a una pianificazione definita dall'utente. Quando si crea una vista materializzata, è possibile specificare una frequenza di aggiornamento utilizzando una pianificazione basata sul tempo con intervalli frequenti fino a un'ora. Il AWS Glue Data Catalog utilizza l'infrastruttura di elaborazione Spark gestita per eseguire operazioni di aggiornamento in background, gestendo in modo trasparente tutti gli aspetti del rilevamento delle modifiche e degli aggiornamenti incrementali.

Quando i dati di origine cambiano tra gli intervalli di aggiornamento, la vista materializzata diventa temporaneamente obsoleta. Le query che accedono direttamente alla vista materializzata possono restituire risultati obsoleti fino al completamento del successivo aggiornamento pianificato. Per gli scenari che richiedono l'accesso immediato ai dati più recenti, è possibile eseguire un aggiornamento manuale utilizzando il comando SQL. `REFRESH MATERIALIZED VIEW`

Aggiornamento incrementale

L'aggiornamento incrementale è una tecnica di ottimizzazione che elabora solo i dati modificati nelle tabelle di origine dall'ultimo aggiornamento, anziché ricalcolare l'intera vista materializzata. Il AWS Glue Data Catalog sfrutta il livello di metadati di Apache Iceberg per tracciare in modo efficiente le modifiche nelle tabelle di origine e determinare quali parti della vista materializzata richiedono aggiornamenti.

Questo approccio riduce in modo significativo i costi di elaborazione e la durata dell'aggiornamento rispetto alle operazioni di aggiornamento complete, in particolare per set di dati di grandi dimensioni in cui solo una piccola percentuale di dati cambia tra i cicli di aggiornamento. Il meccanismo di aggiornamento incrementale funziona automaticamente; non è necessario scrivere una logica personalizzata per rilevare o elaborare i dati modificati.

Riscrittura automatica delle interrogazioni

La riscrittura automatica delle query è una funzionalità di ottimizzazione delle query disponibile nei motori Spark di Amazon Athena, Amazon EMR e AWS Glue. Quando esegui una query su tabelle di base, l'ottimizzatore Spark analizza il tuo piano di query e determina automaticamente se le viste materializzate disponibili possono soddisfare la query in modo più efficiente. Se esiste una vista materializzata adatta, l'ottimizzatore riscrive in modo trasparente la query per utilizzare i risultati precalcolati anziché elaborare le tabelle di base.

Questa ottimizzazione avviene senza richiedere alcuna modifica al codice dell'applicazione o alle istruzioni di query. L'ottimizzatore Spark assicura che la riscrittura automatica delle query si applichi solo quando la vista materializzata è attuale e può produrre risultati accurati. Se una vista materializzata è obsoleta o non soddisfa completamente i requisiti della query, l'ottimizzatore esegue il piano di query originale sulla base delle tabelle di base, dando priorità alla correttezza rispetto alle prestazioni.

Visualizza il ruolo definito

Una vista materializzata funziona in base alle autorizzazioni del ruolo IAM che l'ha creata, noto come ruolo view definer. Il ruolo definer deve avere accesso in lettura a tutte le tabelle di base a cui si fa riferimento nella definizione della vista materializzata e creare i permessi per le tabelle sul database di destinazione. Quando il AWS Glue Data Catalog aggiorna una vista materializzata, assume il ruolo di definitore di accedere alle tabelle di origine e scrivere risultati aggiornati.

Questo modello di sicurezza consente di concedere agli utenti l'accesso alle viste materializzate senza concedere loro autorizzazioni dirette sulle tabelle di origine sottostanti. Se il ruolo View Definer perde l'accesso a qualsiasi tabella di base, le successive operazioni di aggiornamento avranno esito negativo fino al ripristino delle autorizzazioni.

Autorizzazioni per le viste materializzate

Per creare e gestire le viste materializzate, è necessario configurare le autorizzazioni. AWS Lake Formation Il ruolo IAM che crea la vista materializzata (il ruolo definer) richiede autorizzazioni specifiche sulle tabelle di origine e sui database di destinazione.

Autorizzazioni richieste per il ruolo di definitore

Il ruolo che definisce deve avere i seguenti permessi di Lake Formation:

- Nelle tabelle di origine: autorizzazioni SELECT o ALL senza filtri di riga, colonna o cella

- Sul database di destinazione: autorizzazione CREATE_TABLE
- Sul AWS Glue Data Catalog e sulle autorizzazioni API GetTable CreateTable

Quando si crea una vista materializzata, l'ARN del ruolo defintore viene memorizzato nella definizione della vista. Il AWS Glue Data Catalog assume questo ruolo durante l'esecuzione di operazioni di aggiornamento automatico. Se il ruolo del defintore perde l'accesso alle tabelle di origine, le operazioni di aggiornamento avranno esito negativo fino al ripristino delle autorizzazioni.

Autorizzazioni IAM per i lavori AWS Glue

Il ruolo IAM del tuo AWS Glue lavoro richiede le seguenti autorizzazioni:

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetCatalog",
        "glue:GetCatalogs",
        "glue:GetTable",
        "glue:GetTables",
        "glue:CreateTable",
        "glue:UpdateTable",
        "glue>DeleteTable",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "cloudwatch:PutMetricData"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetObject",
        "s3:PutObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket/*"
      ]
    }
  ],
}
```

```

    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "logs:CreateLogGroup",
        "logs:CreateLogStream",
        "logs:PutLogEvents"
      ],
      "Resource": [
        "arn:aws:logs:*:*:*:/aws-glue/*"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": "*"
    }
  ]
}

```

Il ruolo che usi per l'aggiornamento automatico di Materialized View deve avere l'PassRole autorizzazione iam: sul ruolo.

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [

```

```
        "arn:aws:iam::111122223333:role/materialized-view-role-name"
    ]
}
]
```

Affinché Glue aggiorni automaticamente la vista materializzata, il ruolo deve inoltre avere la seguente politica di attendibilità che consenta al servizio di assumere il ruolo.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::111122223333:role/materialized-view-role-name"
      ]
    }
  ]
}
```

Se la vista materializzata è archiviata in S3 Tables Bucket, devi anche aggiungere la seguente autorizzazione al ruolo.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3tables:PutTableMaintenanceConfiguration"
      ],
      "Resource": "arn:aws:s3tables*:123456789012:*"
    }
  ]
}
```

```
}
```

Concessione dell'accesso alle viste materializzate

Per concedere ad altri utenti l'accesso all'interrogazione di una vista materializzata, utilizzare questa opzione AWS Lake Formation per concedere l'autorizzazione SELECT sulla tabella delle viste materializzate. Gli utenti possono interrogare la vista materializzata senza richiedere l'accesso diretto alle tabelle di origine sottostanti.

Per informazioni dettagliate sulla configurazione delle autorizzazioni di Lake Formation, consulta [Concessione e revoca delle autorizzazioni sulle risorse di Data Catalog nella Developer Guide](#). AWS Lake Formation

Creazione e gestione di viste materializzate

Puoi creare viste materializzate utilizzando l'`CREATE MATERIALIZED VIEW` istruzione SQL nei motori Spark. La definizione della vista specifica la query SQL che definisce la logica di trasformazione, il database di destinazione e il nome della tabella e la configurazione di aggiornamento opzionale. È possibile definire trasformazioni complesse tra cui aggregazioni, join tra più tabelle, filtri e funzioni di finestra.

```
CREATE MATERIALIZED VIEW sales_summary
AS
SELECT
    region,
    product_category,
    SUM(sales_amount) as total_sales,
    COUNT(DISTINCT customer_id) as unique_customers
FROM sales_transactions
WHERE transaction_date >= current_date - interval '90' day
GROUP BY region, product_category;
```

Per configurare l'aggiornamento automatico, includi la pianificazione dell'aggiornamento nella definizione della vista:

```
CREATE MATERIALIZED VIEW sales_summary
SCHEDULE REFRESH EVERY 1 HOUR
AS
```



```
SELECT region, product_category, SUM(sales_amount) as total_sales
FROM sales_transactions
GROUP BY region, product_category;
```

È possibile aggiornare manualmente una vista materializzata in qualsiasi momento utilizzando il comando: `REFRESH MATERIALIZED VIEW`

```
REFRESH MATERIALIZED VIEW sales_summary;
```

Per modificare la pianificazione di aggiornamento di una vista materializzata esistente, utilizza l'istruzione: `ALTER MATERIALIZED VIEW`

```
ALTER MATERIALIZED VIEW sales_summary
ADD SCHEDULE REFRESH EVERY 2 HOURS;
```

Viste materializzate annidate

È possibile creare viste materializzate che fanno riferimento ad altre viste materializzate come tabelle di base, abilitando trasformazioni dei dati in più fasi. Quando si creano viste materializzate annidate, il AWS Glue Data Catalog tiene traccia delle dipendenze e propaga automaticamente gli aggiornamenti attraverso la gerarchia delle viste materializzate. Quando una vista materializzata di base viene aggiornata, tutte le viste materializzate a valle che dipendono da essa vengono aggiornate di conseguenza.

Questa funzionalità consente di scomporre trasformazioni complesse in fasi logiche, migliorando la manutenibilità e abilitando l'aggiornamento selettivo dei livelli di trasformazione in base ai requisiti di freschezza dei dati.

Archiviazione e accesso ai dati

Le viste materializzate archiviano i risultati precalcolati come tabelle Apache Iceberg in bucket S3 Tables o bucket S3 per uso generico all'interno del tuo account. AWS Il AWS Glue Data Catalog gestisce tutti gli aspetti della manutenzione delle tabelle Iceberg, tra cui la compattazione e la conservazione delle istantanee, attraverso le funzionalità di ottimizzazione automatizzata di S3 Tables.

Poiché le viste materializzate vengono archiviate come tabelle Iceberg, puoi leggerle direttamente da qualsiasi motore compatibile con Iceberg, tra cui Amazon Athena, Amazon Redshift e piattaforme di

analisi di terze parti. Questa accessibilità multi-motore garantisce che i dati precalcolati rimangano accessibili in tutto l'ecosistema di analisi senza duplicazione dei dati o conversione del formato.

AWS Lake Formation Integrazione con le autorizzazioni

Puoi utilizzarlo AWS Lake Formation per gestire autorizzazioni granulari sulle viste materializzate. Il creatore della vista diventa automaticamente il proprietario della vista materializzata e può concedere le autorizzazioni ad altri utenti o ruoli utilizzando AWS Lake Formation il metodo `named resource` o i tag LF.

Quando concedi l'`SELECT` autorizzazione a un utente per una vista materializzata, quest'ultimo può interrogare i risultati precalcolati senza richiedere l'accesso alle tabelle di origine sottostanti. Questo modello di sicurezza semplifica la gestione dell'accesso ai dati e consente di implementare il principio del privilegio minimo, fornendo agli utenti l'accesso solo alle trasformazioni di dati specifiche di cui hanno bisogno.

È possibile condividere viste materializzate tra AWS account, AWS organizzazioni e unità organizzative utilizzando le funzionalità di condivisione tra account AWS Lake Formation di cui dispone. È inoltre possibile accedere alle viste materializzate in tutte AWS le regioni utilizzando collegamenti alle risorse, abilitando una governance centralizzata dei dati con accesso distribuito ai dati.

Monitoraggio e debug

Il AWS Glue Data Catalog pubblica tutte le operazioni di aggiornamento delle viste materializzate e le metriche associate su Amazon. CloudWatch Puoi monitorare l'ora di inizio dell'aggiornamento, l'ora di fine, la durata, il volume di dati elaborati e lo stato dell'aggiornamento tramite metriche. CloudWatch Quando le operazioni di aggiornamento falliscono, i messaggi di errore e le informazioni di diagnostica vengono acquisiti nei registri. CloudWatch

È possibile impostare CloudWatch allarmi per ricevere notifiche quando i processi di aggiornamento superano la durata prevista o falliscono ripetutamente. Il AWS Glue Data Catalog pubblica anche eventi di modifica in caso di esecuzioni di aggiornamento riuscite e non riuscite, consentendoti di integrare le operazioni di visualizzazione materializzata in un'automazione più ampia del flusso di lavoro.

Per verificare lo stato corrente di una vista materializzata, utilizza il comando `DESCRIBE MATERIALIZED VIEW SQL`, che restituisce i metadati tra cui lo stato di inattività, il timestamp dell'ultimo aggiornamento e la configurazione della pianificazione dell'aggiornamento.

Gestione dei processi di aggiornamento

Avvio di un aggiornamento manuale

Attiva un aggiornamento immediato al di fuori dell'intervallo pianificato.

Autorizzazione richiesta: AWS le credenziali utilizzate per effettuare la chiamata API devono disporre dell'`glue:GetTable` autorizzazione per la vista materializzata.

Per S3 Tables Catalog:

```
aws glue start-materialized-view-refresh-task-run \  
  --catalog-id <ACCOUNT_ID>:s3tablescatalog/<CATALOG_NAME> \  
  --database-name <DATABASE_NAME> \  
  --table-name <MV_TABLE_NAME>
```

Per Root Catalog:

```
aws glue start-materialized-view-refresh-task-run \  
  --catalog-id <ACCOUNT_ID> \  
  --database-name <DATABASE_NAME> \  
  --table-name <MV_TABLE_NAME>
```

Verifica dello stato di aggiornamento

Ottieni lo stato di un processo di aggiornamento specifico:

```
aws glue get-materialized-view-refresh-task-run \  
  --catalog-id <CATALOG_ID> \  
  --materialized-view-refresh-task-run-id <TASK_RUN_ID>
```

Cronologia degli aggiornamenti degli elenchi

Visualizza tutti i lavori di aggiornamento per una vista materializzata:

```
aws glue list-materialized-view-refresh-task-runs \  
  --catalog-id <CATALOG_ID> \  
  --database-name <DATABASE_NAME> \  
  --table-name <MV_TABLE_NAME>
```

 Note

Utilizzare <ACCOUNT_ID>:s3tablescatalog/<CATALOG_NAME> per le tabelle S3 o <ACCOUNT_ID> per il catalogo principale.

Interruzione di un aggiornamento in esecuzione

Annullare un processo di aggiornamento in corso:

```
aws glue stop-materialized-view-refresh-task-run \  
  --catalog-id <CATALOG_ID> \  
  --database-name <DATABASE_NAME> \  
  --table-name <MV_TABLE_NAME>
```

Monitoraggio e risoluzione dei problemi

Esistono tre modi per monitorare i processi di aggiornamento delle viste materializzate:

CloudWatch Metriche

Visualizza le metriche aggregate per tutti i tuoi lavori di aggiornamento delle viste materializzate in: CloudWatch

Metriche disponibili:

- AWS/Namespce Glue con dimensioni:
 - CatalogId: il tuo identificatore del catalogo
 - DatabaseName: Database contenente la vista materializzata
 - TableName: nome della vista materializzata
 - TaskType: Imposta su "» MaterializedViewRefresh

Visualizzazione in console:

1. Vai a CloudWatch Console → Metriche
2. Seleziona lo spazio dei nomi AWS /Glue
3. Filtra per dimensioni: CatalogId,, DatabaseName TableName TaskType

4. Visualizza le metriche relative al successo, al fallimento e alla durata del lavoro

Esempio di interrogazione CloudWatch sulle metriche:

```
{AWS/Glue,CatalogId,DatabaseName,TableName,TaskType} MaterializedViewRefresh
```

Utilizzando AWS CLI:

```
aws cloudwatch get-metric-statistics \
  --namespace AWS/Glue \
  --metric-name <MetricName> \
  --dimensions Name=CatalogId,Value=<CATALOG_ID> \
    Name=DatabaseName,Value=<DATABASE_NAME> \
    Name=TableName,Value=<TABLE_NAME> \
    Name=TaskType,Value=MaterializedViewRefresh \
  --start-time <START_TIME> \
  --end-time <END_TIME> \
  --period 3600 \
  --statistics Sum \
  --region <REGION>
```

CloudWatch Registri

Visualizza i registri di esecuzione dettagliati per le singole operazioni di aggiornamento:

Gruppo di log: /aws-glue/materialized-views/<task_run_id>

<task_run_id> Dov'è un UUID (ad esempio, abc12345-def6-7890-ghij-klmnopqrstuv).

Visualizzazione dei registri:

```
# List log streams for a task run
aws logs describe-log-streams \
  --log-group-name /aws-glue/materialized-views/<TASK_RUN_ID> \
  --region <REGION>

# Get log events
aws logs get-log-events \
  --log-group-name /aws-glue/materialized-views/<TASK_RUN_ID> \
  --log-stream-name <LOG_STREAM_NAME> \
  --region <REGION>
```

Nella CloudWatch console:

1. Vai a CloudWatch → Gruppi di log
2. Cerca /aws-glue/materialized-views/
3. Seleziona il gruppo di log con il tuo ID di esecuzione dell'attività
4. Visualizza i log di esecuzione dettagliati, gli errori e l'output del job Spark

Notifications

Iscriviti agli eventi per ricevere notifiche in tempo reale sulle modifiche allo stato del processo di aggiornamento:

Tipi di eventi disponibili:

- Attività di aggiornamento di Glue Materialized View iniziata
- Attività di aggiornamento di Glue Materialized View riuscita
- Attività di aggiornamento di Glue Materialized View non riuscita
- Errore di invocazione dell'aggiornamento automatico di Glue Materialized View

Creazione di una regola:

```
aws events put-rule \  
  --name materialized-view-refresh-notifications \  
  --event-pattern '{  
    "source": ["aws.glue"],  
    "detail-type": [  
      "Glue Materialized View Refresh Task Started",  
      "Glue Materialized View Refresh Task Succeeded",  
      "Glue Materialized View Refresh Task Failed",  
      "Glue Materialized View Auto-Refresh Invocation Failure"  
    ]  
  }' \  
  --region <REGION>
```

Aggiungere una destinazione (ad esempio, argomento SNS):

```
aws events put-targets \  
  --rule materialized-view-refresh-notifications \  
  --targets 'arn:aws:sns:us-east-1:123456789012:my-topic'
```

```
--targets "Id"="1", "Arn"="arn:aws:sns:<REGION>:<ACCOUNT_ID>:<TOPIC_NAME>" \  
--region <REGION>
```

Visualizzazione dello stato di aggiornamento

Controlla lo stato dei tuoi lavori di aggiornamento delle viste materializzate utilizzando l'API: AWS Glue

```
aws glue get-materialized-view-refresh-task-run \  
  --catalog-id <CATALOG_ID> \  
  --materialized-view-refresh-task-run-id <TASK_RUN_ID> \  
  --region <REGION>
```

Oppure elenca tutte le esecuzioni di aggiornamento recenti:

```
aws glue list-materialized-view-refresh-task-runs \  
  --catalog-id <CATALOG_ID> \  
  --database-name <DATABASE_NAME> \  
  --table-name <MV_TABLE_NAME> \  
  --region <REGION>
```

Questo mostra:

- Ora dell'ultimo aggiornamento
- Stato di aggiornamento (RIUSCITO, FALLITO, IN ESECUZIONE, INTERROTTO)
- ID di esecuzione dell'attività
- Messaggi di errore (se falliti)

Stati di aggiornamento comuni:

- IN ESECUZIONE: il processo di aggiornamento è attualmente in esecuzione
- RIUSCITO: aggiornamento completato con successo
- FALLITO: l'aggiornamento ha rilevato un errore
- INTERROTTO: l'aggiornamento è stato annullato manualmente

Risoluzione dei problemi relativi agli aggiornamenti non riusciti:

Se un aggiornamento fallisce, controlla:

1. Autorizzazioni IAM: assicurati che il ruolo che definisce abbia accesso a tutte le tabelle di base e alla posizione della vista materializzata
2. Disponibilità della tabella di base: verifica che tutte le tabelle di riferimento esistano e siano accessibili
3. Validità della query: conferma che la query SQL è valida per il dialetto SQL di Spark
4. Limiti delle risorse: verifica se hai raggiunto i limiti di aggiornamento simultanei per il tuo account

Utilizza l' `GetMaterializedViewRefreshTaskRun` API per recuperare messaggi di errore dettagliati.

Considerazioni e limitazioni

- Le viste materializzate possono fare riferimento solo alle tabelle Apache Iceberg registrate nel AWS Glue Data Catalog come tabelle di base.
- La creazione di viste e la riscrittura automatica delle query sono disponibili solo dai motori Spark in Apache Spark versione 3.5.6 e successive su Amazon Athena, Amazon EMR e (versione 5.1). AWS Glue
- Le viste materializzate alla fine sono coerenti con le tabelle di base. Durante la finestra di aggiornamento, le query che accedono direttamente alla vista materializzata possono restituire dati obsoleti. Per un accesso immediato ai dati correnti, esegui un aggiornamento manuale.
- L'intervallo minimo di aggiornamento automatico è di un'ora. Per i casi d'uso che richiedono aggiornamenti più frequenti, esegui gli aggiornamenti manuali a livello di codice utilizzando il comando. `REFRESH MATERIALIZED VIEW`
- La riscrittura delle query dà priorità alla correttezza rispetto alle prestazioni. Se una vista materializzata è obsoleta o non è in grado di soddisfare con precisione i requisiti di query, i motori Spark eseguono la query originale sulle tabelle di base.

Importazione di dati utilizzando i flussi di lavoro in Lake Formation

Con AWS Lake Formation, puoi importare i tuoi dati utilizzando i flussi di lavoro. Un flusso di lavoro definisce l'origine dei dati e la pianificazione per l'importazione dei dati nel data lake. È un contenitore per AWS Glue crawler, job e trigger che viene utilizzato per orchestrare i processi per caricare e aggiornare il data lake.

Argomenti

- [Progetti e flussi di lavoro in Lake Formation](#)

- [Creazione di un flusso di lavoro](#)
- [Esecuzione di un workflow](#)

Progetti e flussi di lavoro in Lake Formation

Un flusso di lavoro racchiude una complessa attività di estrazione, trasformazione e caricamento (ETL) che coinvolge più processi. I flussi di lavoro generano AWS Glue crawler, job e trigger per orchestrare il caricamento e l'aggiornamento dei dati. Lake Formation esegue e traccia un flusso di lavoro come un'unica entità. È possibile configurare un flusso di lavoro in modo che venga eseguito su richiesta o in base a una pianificazione.

Note

Spark parquet writer non supporta caratteri speciali nei nomi delle colonne. Questa è una limitazione tecnica dello scrittore stesso, non un problema di configurazione.

I flussi di lavoro creati in Lake Formation sono visibili nella AWS Glue console come grafo aciclico diretto (DAG). Ogni nodo DAG è un job, un crawler o un trigger. Per monitorare l'avanzamento e risolvere i problemi, è possibile tenere traccia dello stato di ogni nodo del flusso di lavoro.

Una volta completato un flusso di lavoro Lake Formation, all'utente che lo ha eseguito viene concessa l'`SELECT` autorizzazione Lake Formation per le tabelle del Data Catalog create dal flusso di lavoro.

Puoi anche creare flussi di lavoro in AWS Glue. Tuttavia, poiché Lake Formation consente di creare un flusso di lavoro a partire da un progetto, la creazione di flussi di lavoro è molto più semplice e automatizzata in Lake Formation. Lake Formation fornisce i seguenti tipi di progetti:

- **Istantanea del database:** carica o ricarica i dati da tutte le tabelle nel data lake da una fonte JDBC. È possibile escludere alcuni dati dall'origine in base a uno schema di esclusione.
- **Database incrementale:** carica nel data lake solo nuovi dati da una fonte JDBC, in base ai segnalibri impostati in precedenza. Si specificano le singole tabelle da includere nel database di origine JDBC. Per ogni tabella, scegli le colonne dei segnalibri e l'ordinamento dei segnalibri per tenere traccia dei dati che sono stati caricati in precedenza. La prima volta che si esegue un blueprint di database incrementale su un set di tabelle, il flusso di lavoro carica tutti i dati dalle tabelle e imposta i segnalibri per la successiva esecuzione del blueprint di database incrementale.

È quindi possibile utilizzare un blueprint di database incrementale anziché il blueprint di snapshot del database per caricare tutti i dati, a condizione di specificare ogni tabella nell'origine dati come parametro.

- File di registro: carica in blocco i dati da fonti di file di registro AWS CloudTrail, inclusi i log di Elastic Load Balancing e i log di Application Load Balancer.

Utilizza la tabella seguente per decidere se utilizzare uno snapshot del database o un blueprint di database incrementale.

Usa lo snapshot del database quando...	Usa il database incrementale quando...
• L'evoluzione dello schema è flessibile. (Le colonne vengono rinominate, le colonne precedenti vengono eliminate e al loro posto vengono aggiunte nuove colonne.) • È necessaria una coerenza completa tra l'origine e la destinazione.	• L'evoluzione dello schema è incrementale. (Le colonne vengono aggiunte solo in successione). • Vengono aggiunte solo nuove righe; le righe precedenti non vengono aggiornate.

Note

Gli utenti non possono modificare i progetti e i flussi di lavoro creati da Lake Formation.

Creazione di un flusso di lavoro

Prima di iniziare, assicurati di aver concesso al ruolo le autorizzazioni necessarie per i dati e le autorizzazioni per la localizzazione dei dati. `LakeFormationWorkflowRole` In questo modo il flusso di lavoro può creare tabelle di metadati nel Data Catalog e scrivere dati nelle posizioni di destinazione in Amazon S3. Per ulteriori informazioni, consultare [\(Facoltativo\) Crea un ruolo IAM per i flussi di lavoro](#) e [Panoramica delle autorizzazioni di Lake Formation](#).

Note

Lake Formation utilizza `GetTemplateInstance` e `InstantiateTemplate` opera per creare flussi di lavoro a partire dai progetti. `GetTemplateInstances` Queste operazioni

non sono disponibili al pubblico e vengono utilizzate solo internamente per creare risorse per conto dell'utente. Ricevi CloudTrail eventi per la creazione di flussi di lavoro.

Per creare un flusso di lavoro da un blueprint

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come utente con autorizzazioni di data engineer. Per ulteriori informazioni, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).
2. Nel riquadro di navigazione, scegli Blueprint, quindi scegli Usa blueprint.
3. Nella pagina Usa un blueprint, scegli un riquadro per selezionare il tipo di blueprint.
4. In Origine di importazione, specifica l'origine dati.

Se state importando da una fonte JDBC, specificate quanto segue:

- Connessione al database: scegliere una connessione dall'elenco. Crea connessioni aggiuntive utilizzando la AWS Glue console. Il nome utente e la password JDBC nella connessione determinano gli oggetti del database a cui il flusso di lavoro ha accesso.
- Percorso dei dati di origine:
immettere `<database>/<schema>/<table>` o `<database>/<table>`, a seconda del prodotto del database. Oracle Database e MySQL non supportano lo schema nel percorso. È possibile sostituire o con il carattere percentuale (%). `<schema> <table>` Ad esempio, per un database Oracle con un identificatore di sistema (SID) di orcl, immettere orcl/% per importare tutte le tabelle a cui ha accesso l'utente indicato nella connessione.

 Important

Questo campo distingue tra maiuscole e minuscole. Il flusso di lavoro avrà esito negativo in caso di mancata corrispondenza tra maiuscole e minuscole per uno qualsiasi dei componenti.

Se si specifica un database MySQL AWS Glue, ETL utilizza il driver JDBC Mysql5 per impostazione predefinita, quindi My non è supportato nativamente. SQL8 È possibile modificare lo script di lavoro ETL per utilizzare un customJdbcDriverS3Path parametro

come descritto in [JDBC ConnectionType](#) Values nella AWS Glue Developer Guide per utilizzare un driver JDBC diverso che supporti My. SQL8

Se stai importando da un file di registro, assicurati che il ruolo specificato per il flusso di lavoro (il «ruolo del flusso di lavoro») disponga delle autorizzazioni IAM necessarie per accedere all'origine dati. Ad esempio, per importare AWS CloudTrail i log, l'utente deve disporre `cloudtrail:LookupEvents` delle autorizzazioni `cloudtrail:DescribeTrails` e per visualizzare l'elenco dei CloudTrail log durante la creazione del flusso di lavoro e il ruolo del flusso di lavoro deve disporre delle autorizzazioni sulla posizione in CloudTrail Amazon S3.

5. Esegui una di queste operazioni:

- Per il tipo di blueprint Database snapshot, identifica facoltativamente un sottoinsieme di dati da importare specificando uno o più modelli di esclusione. Questi modelli di esclusione sono modelli in stile Unix. `glob` Vengono memorizzati come proprietà delle tabelle create dal flusso di lavoro.

Per i dettagli sui modelli di esclusione disponibili, consulta [Includi ed escludi i modelli](#) nella Guida per gli AWS Glue sviluppatori.

- Per il tipo di blueprint del database incrementale, specificare i seguenti campi. Aggiungere una riga per ogni tabella da importare.

Nome tabella

Tabella da importare. Deve essere tutto in minuscolo.

Tasti per segnalibri

Elenco delimitato da virgole di nomi di colonne che definiscono le chiavi dei segnalibri.

Se vuoto, la chiave primaria viene utilizzata per determinare nuovi dati. Le maiuscole e le minuscole per ogni colonna devono corrispondere a quelle definite nell'origine dati.

 Note

La chiave primaria si qualifica come chiave predefinita per i segnalibri solo se è crescente o decrescente in sequenza (senza spazi vuoti). Se desideri utilizzare la chiave primaria come chiave del segnalibro e presenta degli spazi vuoti, devi denominare la colonna della chiave primaria come chiave del segnalibro.

Ordine dei segnalibri

Quando scegli Crescente, le righe con valori superiori ai valori contrassegnati dai segnalibri vengono identificate come nuove righe. Quando scegli Decrescente, le righe con valori inferiori ai valori contrassegnati dai segnalibri vengono identificate come nuove righe.

Schema di partizionamento

(Facoltativo) Elenco delle colonne chiave di partizionamento, delimitate da barre (/).

Esempio:.. year/month/day

Incremental data
Enter tables in the data source to import along with bookmark columns to determine previously imported data.

Table name	Bookmark keys	Bookmark order	Partitioning scheme - optional	
Enter a table name	Enter a bookmark Comma-delimited list of bookmark columns.	Choose a sort. ▼	Type partitioning	Remove
Add				

Per ulteriori informazioni, consulta [Tracciamento dei dati elaborati utilizzando i Job Bookmarks](#) nella AWS Glue Developer Guide.

6. In Import target, specifica il database di destinazione, la posizione Amazon S3 di destinazione e il formato dei dati.

Assicurati che il ruolo del workflow disponga delle autorizzazioni Lake Formation richieste sul database e sulla posizione di destinazione di Amazon S3.

Note

Attualmente, i blueprint non supportano la crittografia dei dati sulla destinazione.

7. Scegliete una frequenza di importazione.

È possibile specificare un'cronoespressione con l'opzione Personalizzata.

8. In Opzioni di importazione:

- a. Inserisci un nome per il flusso di lavoro.

- b. Per ruolo, scegli il ruolo `LakeFormationWorkflowRole` in cui hai creato [\(Facoltativo\) Crea un ruolo IAM per i flussi di lavoro](#).
 - c. Specificate facoltativamente un prefisso per la tabella. Il prefisso viene aggiunto ai nomi delle tabelle del catalogo dati create dal flusso di lavoro.
9. Scegli Crea e attendi che la console segnali che il flusso di lavoro è stato creato correttamente.

 Tip

Hai ricevuto il seguente messaggio di errore?

```
User: arn:aws:iam::<account-id>:user/<username> is not authorized
to perform: iam:PassRole on resource:arn:aws:iam::<account-
id>:role/<rolename>...
```

In tal caso, verifica di aver inserito `<account-id>` un numero di AWS account valido in tutte le politiche.

 Consulta anche:

- [Progetti e flussi di lavoro in Lake Formation](#)

Esecuzione di un workflow

Puoi eseguire un flusso di lavoro utilizzando la console Lake Formation, la AWS Glue console o la AWS Glue Command Line Interface (AWS CLI) o l'API.

Per eseguire un flusso di lavoro (console Lake Formation)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come utente con autorizzazioni di data engineer. Per ulteriori informazioni, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).
2. Nel pannello di navigazione scegli Blueprints (Progetti).
3. Nella pagina Blueprints, seleziona il flusso di lavoro. Quindi, nel menu Azioni, scegli Avvia.
4. Durante l'esecuzione del flusso di lavoro, visualizzane l'avanzamento nella colonna Stato dell'ultima esecuzione. Scegli il pulsante di aggiornamento di tanto in tanto.

Lo stato va da RUNNING, a Discovering, a Importing, a COMPLETED.

Quando il flusso di lavoro è completo:

- Il Data Catalog ha nuove tabelle di metadati.
- I tuoi dati vengono inseriti nel data lake.

Se il flusso di lavoro fallisce, procedi come segue:

- a. Seleziona un flusso di lavoro. Scegliete Azioni, quindi scegliete Visualizza grafico.

Il flusso di lavoro si apre nella AWS Glue console.

- b. Verifica che il flusso di lavoro sia selezionato e scegli la scheda History (Cronologia).
- c. In Cronologia, seleziona l'esecuzione più recente e scegli Visualizza i dettagli dell'esecuzione.
- d. Seleziona un processo o un crawler non riuscito nel grafico dinamico (di runtime) ed esamina il messaggio di errore. I nodi con errori sono rossi o gialli.

 Consulta anche:

- [Progetti e flussi di lavoro in Lake Formation](#)

Inserire i dati nel AWS Glue Data Catalog

Puoi creare cataloghi federati in AWS Glue Data Catalog (Data Catalog) e unificare i dati tra i data lake Amazon S3 e i data warehouse Amazon Redshift. Puoi anche integrare i dati dei tuoi database operativi come Amazon DynamoDB e fonti di dati di terze parti come PostgreSQL, BigQuery Google, MySQL, tra le altre. Il Data Catalog fornisce un archivio centralizzato di metadati che semplifica la gestione e la scoperta dei dati su sistemi diversi.

Il Data Catalog si integra con oltre 30 fonti di dati esterne tramite connettori federati. Con questa integrazione, puoi interrogare i dati da queste fonti esterne senza dover creare pipeline di dati in cui inserire prima i dati. AWS

Dopo aver catalogato i dati esterni, puoi utilizzarli AWS Lake Formation per gestire centralmente le autorizzazioni di accesso ai dati nel Data Catalog. Gli amministratori di Data Lake possono concedere autorizzazioni di accesso granulari ad altri principali IAM (utenti o ruoli) all'interno dello stesso account o tra più account utilizzando metodi di controllo degli accessi basati su tag (LF-Tags) e risorse denominate.

Utilizzando LF-Tags, gli amministratori di dati possono organizzare in modo logico le risorse in base ad attributi come dominio e livello di sensibilità, semplificando la gestione delle autorizzazioni e garantendo al contempo controlli di accesso coerenti tra i servizi di analisi e apprendimento automatico, tra cui Athena, Amazon EMR o Redshift Spectrum. AWS Glue

Il Data Catalog fornisce i seguenti metodi per gestire dati e autorizzazioni su set di dati esterni e metastore esterni:

- Integra i bucket Amazon S3 Table con Data Catalog: puoi pubblicare e catalogare le tabelle Amazon S3 come oggetti Data Catalog e registrare il catalogo come posizione dati di Lake Formation dalla console di Lake Formation o utilizzando le operazioni API. AWS Glue
- Importa i dati nei data warehouse di Amazon Redshift in AWS Glue Data Catalog: registra un namespace o un cluster [Amazon Redshift esistente](#) con Data Catalog e crea un catalogo federato a più livelli nel Data Catalog.

Puoi accedere ai tuoi dati utilizzando qualsiasi motore di query compatibile con le specifiche OpenAPI del catalogo REST di Apache Iceberg, come Amazon EMR Serverless e Amazon Athena.

- Federa Iceberg REST i cataloghi remoti su Data Catalog: unifica i Iceberg REST cataloghi remoti a Data Catalog e accedi in modo sicuro alle tabelle remote archiviate Iceberg in Amazon S3 utilizzando motori di analisi. AWS

- Entra nel Data Catalog da fonti di dati esterne: collega il Data Catalog a fonti di dati esterne utilizzando AWS Glue connessioni e crea cataloghi federati per gestire centralmente le autorizzazioni di accesso sui set di dati utilizzando Lake Formation. Non è necessaria alcuna migrazione dei metadati nel Data Catalog.
- Crea cataloghi per gestire le tabelle Amazon Redshift nel Data Catalog: potresti non avere un cluster di produttori Amazon Redshift o un datashare Amazon Redshift disponibile oggi, ma desideri creare e gestire tabelle Amazon Redshift utilizzando Data Catalog. Puoi iniziare creando un catalogo AWS Glue gestito utilizzando il funzionamento dell'`glue:CreateCatalogAPI` o la AWS Lake Formation console impostando il tipo di catalogo come Managed e Catalog source come Redshift.
- Pubblica le condivisioni di dati Amazon Redshift con Data Catalog: pubblica le condivisioni di dati Amazon Redshift [su](#) Data Catalog e usa Lake Formation per gestire centralmente l'accesso ai dati delle condivisioni di dati e limitare l'accesso degli utenti.

Puoi interrogare i tuoi dati utilizzando Amazon Redshift Spectrum.

- Connetti Data Catalog a metastore Hive esterni: collega il Data Catalog a metastore esterni per gestire le autorizzazioni di accesso ai set di dati in Amazon S3 utilizzando Lake Formation. Non è necessaria alcuna migrazione dei metadati nel Data Catalog.
- Integra Lake Formation con AWS Data Exchange: Lake Formation supporta la concessione di licenze di accesso ai tuoi dati tramite AWS Data Exchange. Se desideri concedere in licenza i tuoi dati di Lake Formation, consulta [Cosa c'è AWS Data Exchange](#) nella Guida AWS Data Exchange per l'utente.

Argomenti

- [Federazione dei cataloghi ai cataloghi Iceberg remoti](#)
- [Inserimento dei dati di Amazon Redshift nel AWS Glue Data Catalog](#)
- [La federazione in fonti di dati esterne in AWS Glue Data Catalog](#)
- [Creazione di un catalogo Amazon S3 Tables nel AWS Glue Data Catalog](#)
- [Creazione di un catalogo gestito di Amazon Redshift nel AWS Glue Data Catalog](#)
- [Gestione delle autorizzazioni per i dati in un datashare Amazon Redshift](#)
- [Gestione delle autorizzazioni sui set di dati che utilizzano metastore esterni](#)

Federazione dei cataloghi ai cataloghi Iceberg remoti

La federazione dei cataloghi AWS Glue fornisce un accesso diretto e sicuro alle tabelle Iceberg, archiviate in Amazon S3 e catalogate in cataloghi remoti, utilizzando motori di analisi. AWS La federazione dei cataloghi sincronizza i metadati tra Data Catalog e cataloghi remoti quando accedi a tabelle remote. È supportato da un'ampia varietà di motori di analisi, tra cui Amazon Redshift, Amazon EMR, Amazon Athena e motori di terze parti come Apache Spark e altri. AWS Glue

La federazione dei AWS Glue Data Catalog cataloghi comunica con i sistemi di catalogo remoti per scoprire le tabelle e Lake Formation per autorizzare l'accesso ai dati delle tabelle in Amazon S3. Quando esegui una query su una tabella federata, Data Catalog rileva le informazioni più recenti sulla tabella nel catalogo remoto al momento della query, ottenendo la posizione della tabella in Amazon S3, lo schema corrente e le informazioni sulla partizione. Il tuo motore di analisi (Amazon Athena, Amazon Redshift, Amazon EMR) utilizza quindi queste informazioni per accedere ai file di dati Iceberg direttamente da Amazon S3. Lake Formation gestisce l'accesso alle tabelle distribuendo credenziali mirate ai dati della tabella archiviati in Amazon S3, permettendo ai motori di applicare autorizzazioni granulari alle tabelle federate.

Caratteristiche di Catalog Federation

Governato utilizzando Lake Formation

I cataloghi Federated Iceberg in Data Catalog sono risorse registrate da Lake Formation, che consentono di concedere autorizzazioni granulari a livello di riga, colonna e cella alle tabelle Iceberg nei cataloghi Iceberg federati utilizzando le sovvenzioni Lake Formation. I cataloghi Federated Iceberg e gli oggetti associati possono essere condivisi in modo sicuro tra gli account. AWS I cataloghi Federated Iceberg funzionano anche con il controllo degli accessi basato su Lake Formation Tag, che consente di scalare la governance utilizzando i tag.

Configurazioni di rete

La federazione dei cataloghi supporta connessioni dirette a sorgenti di cataloghi remote utilizzando la connettività HTTPS standard. Supporta anche la connettività tramite Amazon VPC quando si desidera mantenere l'isolamento della rete e la connettività utilizzando il supporto proxy quando si desidera una comunicazione sicura attraverso i firewall dell'organizzazione.

Argomenti

- [Catalogo Federate to Snowflake Iceberg](#)
- [Federa a Databricks Unity Catalog](#)

Catalogo Federate to Snowflake Iceberg

AWS Glue Data Catalog si associa all'Snowflakeutilizzo delle OAuth2 credenziali di un responsabile del servizio. Snowflake È possibile utilizzare questa federazione per connettersi ai cataloghiSnowflake Horizon. Snowflake Polaris Questo meccanismo di autenticazione consente a Data Catalog di accedere ai metadati di vari oggetti (come cataloghi, database e tabelle) del catalogo, in base ai privilegi associati al principale del servizio. Per garantire l'accesso agli oggetti giusti, è essenziale concedere al responsabile del servizio le autorizzazioni necessarie per leggere Snowflake i metadati di questi oggetti.

Prerequisiti

Prima di creare un catalogo federato in Data Catalog governato da Lake Formation, assicurati di disporre delle seguenti autorizzazioni:

Il tuo principale IAM (utente o ruolo) deve disporre delle seguenti autorizzazioni:

- Autorizzazioni Lake Formation —, `lakeformation:RegisterResource`
`lakeformation:DescribeResource`
- AWS Glue autorizzazioni —`glue:CreateConnection`, `glue:CreateCatalog`
`glue:GetConnection` `glue:PassConnection`
- Autorizzazioni di Secrets Manager —, `secretsmanager:CreateSecret`
`secretsmanager:GetSecretValue`
- Autorizzazioni IAM —, `iam:CreateRole` `iam:AttachRolePolicy` `iam:PassRole`

Devi essere un amministratore del data lake di Lake Formation o avere l'`CREATE_CATALOG`autorizzazione per il Data Catalog.

Crea un catalogo federato

Utilizzo della console

1. Accedi alla console e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Scegli la AWS regione preferita nella sezione della pagina in alto a destra.
3. Nel riquadro di navigazione a sinistra, seleziona Cataloghi.
4. Scegli Crea catalogo per aprire il flusso di lavoro Crea catalogo.

5. Nel passaggio Scegli l'origine dei dati, Snowflake seleziona una delle opzioni disponibili.
6. Nel passaggio Imposta i dettagli del catalogo, fornisci tre informazioni: dettagli del catalogo, dettagli di connessione e dettagli di registrazione.
7. Nel contenitore dei dettagli del catalogo, fornisci un nome univoco al tuo catalogo AWS Glue federato e inserisci il nome del Snowflake catalogo esistente.
8. Nel contenitore dei dettagli delle connessioni, puoi scegliere tra una connessione esistente a cui hai accesso o fornire la configurazione per creare un nuovo connettore.
9. Le nuove configurazioni di connessione includono:
 - Nome connessione: nome univoco dell'oggetto di AWS Glue connessione.
 - URL dell'istanza: l'URL dell'endpoint dell'Snowflakeaccount esistente.
 - Autenticazione: specifica la configurazione di autenticazione AWS Glue utilizzata per la connessione al server di catalogo remoto. AWS Glue supporta sia OAuth2 l'autenticazione personalizzata.
 - URL del token: specifica l'URL del provider di identità del catalogo remoto.
 - OAuth2 ID client: specifica l'ID client della OAuth2 credenziale associata al catalogo remoto.
 - Segreto: archivia e utilizza OAuth2 il segreto del client utilizzando Gestione dei segreti AWS o immettendo il valore segreto nella casella di testo. Quando inserisci il segreto manualmente nella console, AWS Glue crea il segreto per tuo conto.
 - Ambito dell'URL del token: specifica l' OAuth ambito per l'autenticazione.
 - Catalog Casing Filter: scegli se portare oggetti in lettere minuscole o maiuscole dal catalogo remoto a Data Catalog.
10. Crea un ruolo IAM che AWS Glue i responsabili del servizio Lake Formation possano utilizzare per accedere rispettivamente alle posizioni segrete in Gestione dei segreti AWS e Amazon S3 delle tabelle Iceberg remote. Seleziona il ruolo IAM nel menu a discesa di registrazione. Fai riferimento ai passaggi 2 e 3 nella seguente sezione CLI per i dettagli delle policy IAM.
11. Seleziona Test Connection per verificare se le proprietà della connessione e l'accesso ai ruoli IAM sono configurati correttamente.
12. Seleziona Avanti per rivedere le impostazioni.
13. Seleziona Crea catalogo nella pagina di revisione.

Utilizzo della CLI

1. Crea un Gestione dei segreti AWS segreto

Il AWS Glue connettore supporta due tipi di autenticazione: OAuth2Personalizzata. Quando si utilizza OAuth2 l'opzione, utilizzare Gestione dei segreti AWS per memorizzare il segreto del client del principale Snowflake servizio. Successivamente utilizzerai questo segreto durante la creazione della AWS Glue connessione. Per l'autenticazione personalizzata, utilizzare Gestione dei segreti AWS per archiviare e recuperare il token di accesso.

Nell'esempio seguente, sostituisci *your-snowflake-secretclient_secret*, *region* con le tue informazioni.

```
aws secretsmanager create-secret \
--name your-snowflake-secret \
--description "Snowflake secret" \
--secret-string '{
"USER_MANAGED_CLIENT_APPLICATION_CLIENT_SECRET": "client_secret"
}' \
--region region
```

Note

USER_MANAGED_CLIENT_APPLICATION_CLIENT_SECRET è una parola chiave riservata che viene AWS Glue utilizzata per fare riferimento a un valore segreto del client nel segreto. Usa la stessa parola chiave anche quando crei il segreto nella console di Lake Formation.

2. Crea un ruolo IAM che dia accesso all'oggetto di AWS Glue connessione al segreto creato nel passaggio precedente

L'oggetto di AWS Glue connessione richiede l'accesso al Gestione dei segreti AWS segreto quando lo utilizzi Gestione dei segreti AWS per archiviare, recuperare e aggiornare il token OAuth segreto. L'oggetto di AWS Glue connessione richiede inoltre l'accesso per creare, descrivere e utilizzare le interfacce di rete Amazon VPC quando utilizzi un endpoint Amazon VPC per limitare la connettività al tuo account. Snowflake

Crea una policy IAM e collegala a un ruolo IAM. Aggiungi AWS Glue il service principal alla policy di fiducia.

Nell'esempio seguente *your-secrets-manager-ARN**your-vpc-id*, sostituisci e *your-subnet-id1* con le tue informazioni.

Example Politica IAM

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue",
      "secretsmanager:DescribeSecret",
      "secretsmanager:PutSecretValue"
    ],
    "Resource": [
      "your-secrets-manager-ARN"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "ec2:CreateNetworkInterface",
      "ec2:DeleteNetworkInterface",
      "ec2:DescribeNetworkInterfaces"
    ],
    "Resource": "*",
    "Condition": {
      "ArnEquals": {
        "ec2:Vpc": "arn:aws:ec2:region:account-id:vpc/your-vpc-id",
        "ec2:Subnet": ["arn:aws:ec2:region:account-id:subnet/your-  
subnet-id1"]
      }
    }
  }
]
}
```

Example Policy di attendibilità

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Service": "glue.amazonaws.com"
    }
  }
]
```

```

    },
    "Action": "sts:AssumeRole"
  ]
}

```

3. Crea una policy IAM che fornisca a Lake Formation l'accesso in lettura alla posizione Amazon S3 del catalogo

In qualità di proprietario di un catalogo federato in Data Catalog, utilizzi Lake Formation per concedere ai tuoi team di dati l'accesso granulare alle tabelle, l'accesso granulare a livello di colonna, a livello di riga e a livello di cella e l'accesso basato su tag. Lake Formation utilizza un ruolo IAM che gli consente di accedere alle posizioni Amazon S3 sottostanti delle tabelle Iceberg remote. Questo accesso consente a Lake Formation di fornire credenziali di accesso mirate ai motori di analisi che interrogano tabelle remote.

Crea una policy IAM e collegala a un ruolo IAM. Aggiungi il responsabile del servizio Lake Formation alla policy di trust del ruolo.

Nell'esempio seguente, sostituisci *amzn-s3-demo-bucketN* e *your-kms-key* con le tue informazioni.

Example Politica IAM

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::amzn-s3-demo-bucket1/*",
      "arn:aws:s3:::amzn-s3-demo-bucket2/*"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::amzn-s3-demo-bucket1",

```

```

        "arn:aws:s3:::amzn-s3-demo-bucket2"
    ],
    {
        "Effect": "Allow",
        "Action": [
            "kms:Decrypt",
            "kms:Encrypt"
        ],
        "Resource": [
            "your-kms-key"
        ]
    }
]
}

```

Example Policy di attendibilità

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Sid": "",
    "Effect": "Allow",
    "Principal": {
      "Service": "lakeformation.amazonaws.com"
    },
    "Action": "sts:AssumeRole"
  }]
}

```

Note

Quando utilizzi la console Lake Formation per creare un catalogo federato, la console utilizza un singolo ruolo IAM con entrambe le policy collegate per completare la configurazione.

4. Crea un oggetto di AWS Glue connessione

AWS Glue supporta per connettori OAuth2 e metodi di autenticazione personalizzati. L'oggetto di connessione Data Catalog supporta una CATALOG_CASING_FILTER configurazione che consente di inserire oggetti in lettere minuscole o maiuscole del catalogo remoto in Data Catalog.

L'esempio seguente utilizza la configurazione OAuth2 dell'autenticazione per creare una connessione. AWS Glue Sostituisci *highlighted sections* con le tue informazioni.

```
aws glue create-connection \
  --connection-input '{
    "Name": "your-glue-connection-to-snowflake-account",
    "ConnectionType": "SNOWFLAKEICEBERGRESTCATALOG",
    "ConnectionProperties": {
      "INSTANCE_URL": "your-snowflake-account-URL",
      "ROLE_ARN": "your-IAM-role-for-secrets-and-VPC-access",
      "CATALOG_CASING_FILTER": "LOWERCASE_ONLY"
    },
    "AuthenticationConfiguration": {
      "AuthenticationType": "OAUTH2",
      "OAuth2Properties": {
        "OAuth2GrantType": "CLIENT_CREDENTIALS",
        "TokenUrl": "your-internal-or-external-token-server-url",
        "OAuth2ClientApplication": {
          "UserManagedClientApplicationClientId": "our-client-id"
        },
        "TokenUrlParametersMap": {
          "scope": "all-apis"
        }
      },
      "SecretArn": "arn:aws:secretsmanager:your-aws-region:your-aws-account-id:secret:snowflake-secret"
    }
  }'
```

5. Registra AWS Glue la connessione come risorsa Lake Formation

Utilizzando l'oggetto di AWS Glue connessione (creato nella fase 4) e il ruolo IAM (creato nella fase 3), è ora possibile registrare l'oggetto di AWS Glue connessione come risorsa gestita da Lake Formation.

Sostituisci *your-glue-connector-arn* e *your-IAM-role-ARN-having-LF-access* con le tue informazioni.

```
aws lakeformation register-resource \
  --resource-arn your-glue-connector-arn \
  --role-arn your-IAM-role-ARN-having-LF-access \
  --with-federation \
```

```
--with-privileged-access
```

6. Crea un catalogo federato in Data Catalog

Dopo aver creato un oggetto di AWS Glue connessione e averlo registrato con Lake Formation, puoi creare un catalogo federato nel Data Catalog:

Fornisci al catalogo federato un nome univoco in *your-federated-catalog-name*, fai riferimento al catalogo in Snowflake at *catalog-name-in-Snowflake* e inserisci il nome della connessione creato in precedenza in. *your-glue-connection-name*

```
aws glue create-catalog \  
  --name your-federated-catalog-name \  
  --catalog-input '{  
    "FederatedCatalog": {  
      "Identifier": catalog-name-in-Snowflake,  
      "ConnectionName": your-glue-connection-name  
    },  
    "CreateTableDefaultPermissions": [],  
    "CreateDatabaseDefaultPermissions": []  
  }'
```

Considerazioni relative all'integrazione con Snowflake

- Quando inserisci risorse (come database e tabelle) Snowflake, Lake Formation non revoca automaticamente le autorizzazioni concesse su quella risorsa federata. Per rimuovere le autorizzazioni di accesso, è necessario revocare esplicitamente le autorizzazioni precedentemente concesse alla risorsa federata utilizzando Lake Formation.
- Quando si monta un catalogo remoto con `CATALOG_CASING_FILTER='UPPERCASE_ONLY'` configurazione, i database e le tabelle con identificatori maiuscoli sono federati ma gli oggetti con identificatori minuscoli non lo sono.
- È possibile riutilizzare la stessa connessione per creare più cataloghi federati. AWS Glue L'eliminazione di un catalogo non eliminerà la connessione associata. Per eliminare una connessione, utilizza il `aws glue delete-connection` comando AWS CLI e assicurati che tutti i cataloghi associati vengano prima eliminati.
- I namespace annidati nel catalogo non sono supportati. Polarice In altre parole, la federazione dei cataloghi può accedere a Iceberg tabelle remote che seguono una notazione in 3 parti. `catalog.database.table`

Federa a Databricks Unity Catalog

AWS Glue Data Catalog si associa all'utilizzo delle OAuth2 credenziali di un responsabile del servizio. Databricks Questo meccanismo di autenticazione consente di accedere AWS Glue Data Catalog ai metadati di vari oggetti (come cataloghi, database e tabelle) in Databricks Unity Catalog base ai privilegi associati all'entità del servizio. Per garantire l'accesso agli oggetti giusti, è essenziale concedere al responsabile del servizio le autorizzazioni necessarie Databricks per leggere i metadati di questi oggetti.

Successivamente, la federazione dei cataloghi consente l'individuazione e l'interrogazione delle Iceberg tabelle nel catalogo Databricks Unity. Per leggere le tabelle delta, assicurati che Iceberg i metadati siano disponibili per l'utilizzo di queste tabelle. Uniform Segui il Databricks tutorial e la documentazione per creare il principale servizio e i privilegi associati nel tuo Databricks spazio di lavoro.

Prerequisiti

Prima di creare un catalogo federato in Data Catalog governato da Lake Formation, assicurati di disporre delle seguenti autorizzazioni:

Il tuo principale IAM (utente o ruolo) deve disporre delle seguenti autorizzazioni:

- Autorizzazioni Lake Formation —, lakeformation:RegisterResource
lakeformation:DescribeResource
- AWS Glue autorizzazioni —glue:CreateConnection,, glue:CreateCatalog
glue:GetConnection glue:PassConnection
- Autorizzazioni di Secrets Manager —, secretsmanager:CreateSecret
secretsmanager:GetSecretValue
- Autorizzazioni IAM:, iam:CreateRole iam:AttachRolePolicy iam:PassRole

Devi essere un amministratore del data lake di Lake Formation o avere l'CREATE_CATALOGautorizzazione per il Data Catalog

Crea un catalogo federato

Utilizzo della console

1. Accedi alla console e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

2. Scegli la AWS regione preferita nella sezione della pagina in alto a destra.
3. Nel riquadro di navigazione a sinistra, seleziona Cataloghi.
4. Scegli Crea catalogo per aprire il flusso di lavoro Crea catalogo.
5. Nel passaggio Scegli l'origine dei dati, Databricks seleziona una delle opzioni disponibili.
6. Nel passaggio Imposta i dettagli del catalogo, fornisci tre informazioni: dettagli del catalogo, dettagli di connessione e dettagli di registrazione.
7. Nel contenitore dei dettagli del catalogo, fornisci un nome univoco al tuo catalogo AWS Glue federato e inserisci il nome del Databricks catalogo esistente.
8. Nel contenitore dei dettagli delle connessioni, puoi scegliere tra una connessione esistente a cui hai accesso o fornire la configurazione per creare un nuovo connettore.
9. Le nuove configurazioni di connessione includono:
 - Nome connessione: nome univoco dell'oggetto di AWS Glue connessione.
 - URL dell'area di lavoro: l'URL dell'endpoint dell'area di lavoro esistente Databricks.
 - Autenticazione: specifica la configurazione di autenticazione AWS Glue utilizzata per la connessione al server di catalogo remoto. AWS Glue supporta sia OAuth2 l'autenticazione personalizzata.
 - URL del token: specifica l'URL del provider di identità del catalogo remoto.
 - OAuth2 ID client: specifica l'ID client della OAuth2 credenziale associata al catalogo remoto.
 - Segreto: archivia e utilizza OAuth2 il segreto del client utilizzando Gestione dei segreti AWS o immettendo il valore segreto nella casella di testo. Quando inserisci il segreto manualmente nella console, AWS Glue crea il segreto per tuo conto.
 - Ambito dell'URL del token: specifica l' OAuth ambito per l'autenticazione.
10. Crea un ruolo IAM che AWS Glue i responsabili del servizio Lake Formation possano utilizzare per accedere rispettivamente alle posizioni segrete in e Amazon S3 delle tabelle Iceberg remote. Seleziona il ruolo IAM nel menu a discesa di registrazione. Fai riferimento ai passaggi 2 e 3 nella seguente sezione CLI per i dettagli delle policy IAM.
11. Seleziona Test Connection per verificare se le proprietà della connessione e l'accesso ai ruoli IAM sono configurati correttamente. La funzionalità di test di connessione non è disponibile quando ci si connette Databricks tramite Amazon VPC.
12. Seleziona Avanti per rivedere le impostazioni.
13. Seleziona Crea catalogo nella pagina di revisione.

Utilizzo della CLI

1. Crea un Gestione dei segreti AWS segreto

Il AWS Glue connettore supporta due tipi di autenticazione: OAuth2Personalizzata. Quando si utilizza OAuth2 l'opzione, utilizzare Gestione dei segreti AWS per memorizzare il segreto del client del principale Databricks servizio. Successivamente utilizzerai questo segreto durante la creazione della AWS Glue connessione. Per l'autenticazione personalizzata, utilizzare Gestione dei segreti AWS per archiviare e recuperare il token di accesso.

Nell'esempio seguente, sostituisci *<databricks-secret>*, *<client_secret>*, *<region>* con le tue informazioni.

```
aws secretsmanager create-secret \
--name <databricks-secret> \
--description "Databricks secret" \
--secret-string '{
"USER_MANAGED_CLIENT_APPLICATION_CLIENT_SECRET": "<client_secret>"
}' \
--region <region>
```

Note

USER_MANAGED_CLIENT_APPLICATION_CLIENT_SECRET è una parola chiave riservata che viene AWS Glue utilizzata per fare riferimento a un valore segreto del client nel segreto. Usa la stessa parola chiave anche quando crei il segreto nella console di Lake Formation.

2. Crea un ruolo IAM che dia accesso all'oggetto di AWS Glue connessione al segreto creato nel passaggio precedente

L'oggetto di AWS Glue connessione richiede l'accesso al Gestione dei segreti AWS segreto quando lo utilizzi Gestione dei segreti AWS per archiviare, recuperare e aggiornare il token OAuth segreto. L'oggetto di AWS Glue connessione richiede inoltre l'accesso per creare, descrivere e utilizzare le interfacce di rete Amazon VPC quando utilizzi un endpoint Amazon VPC per limitare la connettività al tuo spazio di lavoro. Databricks

Crea una policy IAM e collegala a un ruolo IAM. Aggiungi AWS Glue il service principal alla policy di fiducia.

Nell'esempio seguente *<your-secrets-manager-ARN><your-vpc-id>*, sostituisci e *<your-subnet-id1>* con le tue informazioni.

Example Politica IAM

```
{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "secretsmanager:GetSecretValue",
      "secretsmanager:DescribeSecret",
      "secretsmanager:PutSecretValue"
    ],
    "Resource": [
      "<your-secrets-manager-ARN>"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "ec2:CreateNetworkInterface",
      "ec2:DeleteNetworkInterface",
      "ec2:DescribeNetworkInterfaces"
    ],
    "Resource": "*",
    "Condition": {
      "ArnEquals": {
        "ec2:Vpc": "arn:aws:ec2:region:account-id:vpc/<your-vpc-id>",
        "ec2:Subnet": ["arn:aws:ec2:region:account-id:subnet/<your-
subnet-id1>"]
      }
    }
  }
]
```

Example Policy di attendibilità

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

```

    "Effect": "Allow",
    "Principal": {
        "Service": "glue.amazonaws.com"
    },
    "Action": "sts:AssumeRole"
  ]
}

```

3. Crea una policy IAM che fornisca a Lake Formation l'accesso in lettura alla posizione Amazon S3 del catalogo

In qualità di proprietario di un catalogo federato in Data Catalog, utilizzi Lake Formation per concedere ai tuoi team di dati l'accesso granulare alle tabelle, l'accesso granulare a livello di colonna, a livello di riga e a livello di cella e l'accesso basato su tag. Lake Formation utilizza un ruolo IAM che gli consente di accedere alle posizioni Amazon S3 sottostanti delle tabelle Iceberg remote. Questo accesso consente a Lake Formation di fornire credenziali di accesso mirate ai motori di analisi che interrogano tabelle remote.

Crea una policy IAM e collegala a un ruolo IAM. Aggiungi il responsabile del servizio Lake Formation alla policy di trust dei ruoli IAM.

Nell'esempio seguente, sostituisci *<your-s3-bucket-N>* e *<your-kms-key>* con le tue informazioni.

Example Politica IAM

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "s3:GetObject"
    ],
    "Resource": [
      "arn:aws:s3:::<your-s3-bucket-1>/*",
      "arn:aws:s3:::<your-s3-bucket-2>/*"
    ]
  },
  {
    "Effect": "Allow",
    "Action": [
      "s3:ListBucket"
    ]
  }
]
}

```

```

    ],
    "Resource": [
        "arn:aws:s3:::<your-s3-bucket-1>",
        "arn:aws:s3:::<your-s3-bucket-2>"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt",
        "kms:Encrypt"
    ],
    "Resource": [
        "<your-kms-key>"
    ]
}
]
}

```

Example Policy di attendibilità

```

{
    "Version": "2012-10-17",
    "Statement": [{
        "Sid": "",
        "Effect": "Allow",
        "Principal": {
            "Service": "lakeformation.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    }]
}

```

Note

Quando utilizzi la console Lake Formation per creare un catalogo federato, la console utilizza un singolo ruolo IAM con entrambe le policy collegate per completare la configurazione.

4. Crea un oggetto di AWS Glue connessione

Data Catalog supporta ConnectionTypeDATABRICKSICEBERGRESTCATALOG: per connettere Data Catalog a Databricks. Questo AWS Glue connettore supporta metodi OAuth2 di autenticazione personalizzati.

L'esempio seguente utilizza la configurazione OAuth2 dell'autenticazione per creare una AWS Glue connessione. Sostituisci *highlighted sections* con le tue informazioni.

```
aws glue create-connection \
  --connection-input '{
    "Name": "<your-glue-connection-to-databricks-unity-account>",
    "ConnectionType": "DATABRICKSICEBERGRESTCATALOG",
    "ConnectionProperties": {
      "INSTANCE_URL": "<your-databricks-workspace-catalog-URL>",
      "ROLE_ARN": "<your-IAM-role-for-secrets-and-VPC-access>"
    },
    "AuthenticationConfiguration": {
      "AuthenticationType": "OAUTH2",
      "OAuth2Properties": {
        "OAuth2GrantType": "CLIENT_CREDENTIALS",
        "TokenUrl": "<your-internal-or-external-token-server-url>",
        "OAuth2ClientApplication": {
          "UserManagedClientApplicationClientId": "<your-client-id>"
        },
        "TokenUrlParametersMap": {
          "scope": "all-apis"
        }
      },
      "SecretArn": "arn:aws:secretsmanager:<aws-region>:<your-aws-account-id>:secret:<databricks-secret>"
    }
  }'
```

5. Registra AWS Glue la connessione come risorsa Lake Formation

Utilizzando l'oggetto di AWS Glue connessione (creato nella fase 4) e il ruolo IAM (creato nella fase 3), ora è possibile registrare l'oggetto di AWS Glue connessione come risorsa gestita da Lake Formation.

Sostituisci *<your-glue-connector-arn>* e *<your-IAM-role-ARN-having-LF-access>* con le tue informazioni.

```
aws lakeformation register-resource \
  --resource-arn <your-glue-connector-arn> \
  --role-arn <your-IAM-role-ARN-having-LF-access> \
  --with-federation \
  --with-privileged-access
```

6. Crea un catalogo federato in Data Catalog

Dopo aver creato un oggetto di AWS Glue connessione e averlo registrato con Lake Formation, puoi creare un catalogo federato nel Data Catalog:

Fornisci al catalogo federato un nome univoco in *<your-federated-catalog-name>*, fai riferimento al catalogo in Databricks at "*<catalog-name-in-Databricks>*" e inserisci il nome della connessione creato in precedenza in. "*<your-glue-connection-name>*"

```
aws glue create-catalog \
  --name <your-federated-catalog-name> \
  --catalog-input '{
    "FederatedCatalog": {
      "Identifier": "<catalog-name-in-Databricks>",
      "ConnectionName": "<your-glue-connection-name>"
    },
    "CreateTableDefaultPermissions": [],
    "CreateDatabaseDefaultPermissions": []
  }'
```

Considerazioni relative all'integrazione con Databricks

- Quando lasci cadere risorse (come database e tabelle) in Databricks, Lake Formation non revoca automaticamente le autorizzazioni concesse su quella risorsa federata. Per rimuovere le autorizzazioni di accesso, è necessario revocare esplicitamente le autorizzazioni precedentemente concesse alla risorsa federata utilizzando Lake Formation.
- Puoi interrogare le tabelle Iceberg archiviate in Amazon S3 utilizzando questa integrazione. Quando utilizzi qualsiasi altro formato di tabella o storage di oggetti, puoi federare i metadati nei cataloghi remoti AWS Glue ed elencarli nei relativi database e tabelle, ma operazioni di query come quelle che eseguono la query `SELECT ColumnFoo from TableBar` falliranno e genereranno l'errore «Failed to read table». Apache Iceberg La posizione di archiviazione degli oggetti non è supportata. '

- È possibile riutilizzare la stessa AWS Glue connessione per creare più cataloghi federati. L'eliminazione di un catalogo non eliminerà l'oggetto di connessione associato. Per eliminare un oggetto di connessione, utilizza il `aws glue delete-connection` comando AWS CLI e assicurati che tutti i cataloghi associati vengano prima eliminati.

Inserimento dei dati di Amazon Redshift nel AWS Glue Data Catalog

Puoi gestire i dati analitici nei data warehouse Amazon Redshift in (Data Catalog) e unificare AWS Glue Data Catalog i data lake Amazon S3 e i data warehouse Amazon Redshift. Amazon Redshift è un servizio di data warehouse completamente gestito su scala di petabyte nel cloud. AWS Un data warehouse Amazon Redshift è costituito da un insieme di risorse di calcolo denominate nodi, strutturate in un gruppo denominato cluster. Ciascun cluster esegue un motore Amazon Redshift e contiene uno o più database.

In Amazon Redshift, puoi creare cluster e namespace serverless con provisioning di Amazon Redshift e registrarli nel Data Catalog. In questo modo, puoi unificare i dati nello storage gestito di Amazon Redshift (RMS) e nei bucket Amazon S3 e accedere ai dati da motori di analisi compatibili con Apache Iceberg.

Registrando namespace e cluster, puoi fornire l'accesso ai dati senza la necessità di copiarli o spostarli. Per ulteriori informazioni sulla registrazione di cluster e namespace in Amazon Redshift, consulta [Registrazione di cluster e namespace Amazon Redshift](#) su. AWS Glue Data Catalog

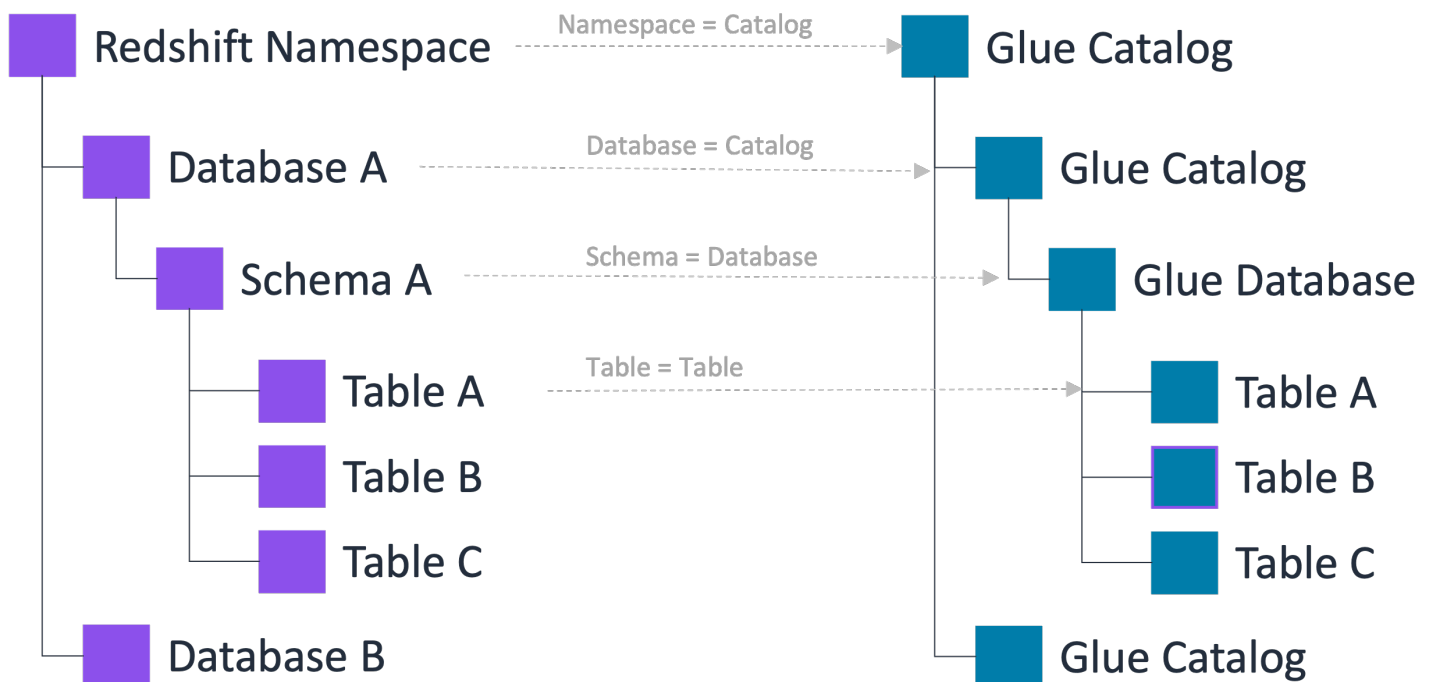
In Amazon Redshift, puoi eseguire la condivisione dei dati tramite datashare o registrando namespace e cluster con Data Catalog. Con le datashare, che operano a livello di singolo oggetto del database, devi abilitare la condivisione per ogni tabella o vista. Al contrario, la pubblicazione dei namespace funziona a livello di cluster o namespace. Quando si registra un cluster o uno spazio dei nomi con Data Catalog, tutti i database e le tabelle al suo interno vengono condivisi automaticamente, senza che sia necessario configurare la condivisione per i singoli oggetti.

Nel Data Catalog, puoi creare un catalogo federato per ogni namespace o cluster. Un catalogo viene definito catalogo federato quando punta a un'entità esterna al Data Catalog. Le tabelle e le viste nello spazio dei nomi Amazon Redshift sono elencate come singole tabelle nel Data Catalog. Puoi condividere database e tabelle nel catalogo federato con i principali IAM e gli utenti SAML selezionati all'interno dello stesso account o in un altro account con Lake Formation. Puoi anche

includere espressioni di filtro per righe e colonne per limitare l'accesso a determinati dati. Per ulteriori informazioni, consulta [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#).

Il Data Catalog supporta una gerarchia di metadati a tre livelli che comprende cataloghi, database e tabelle (e viste). Quando registri un namespace con Data Catalog, la gerarchia dei dati di Amazon Redshift viene mappata sulla gerarchia a 3 livelli del Data Catalog come segue:

- Lo spazio dei nomi Amazon Redshift diventa un catalogo multilivello nel Data Catalog.
- Il database Amazon Redshift associato è registrato come catalogo nel Data Catalog.
- Lo schema Amazon Redshift diventa un database nel Data Catalog.
- La tabella Amazon Redshift diventa una tabella nel Data Catalog.



Con questa gerarchia di metadati a tre livelli, puoi accedere alle tabelle di Amazon Redshift utilizzando la notazione in 3 parti: «catalog1/catalog2.database.table» nel Data Catalog. Inoltre, i team addetti ai dati possono mantenere la stessa organizzazione utilizzata da Amazon Redshift per organizzare le tabelle all'interno dell'account Data Catalog.

In Lake Formation, puoi gestire in modo sicuro i dati di Amazon Redshift utilizzando un controllo granulare degli accessi per le risorse del Data Catalog. Con questa integrazione, puoi gestire, proteggere e interrogare i dati analitici da un unico catalogo con un meccanismo di controllo degli accessi comune.

Per le limitazioni, consulta [Limitazioni per l'inserimento dei dati del data warehouse di Amazon Redshift nel AWS Glue Data Catalog](#).

Argomenti

- [Vantaggi principali](#)
- [Ruoli e responsabilità](#)
- [Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog](#)
- [Creazione di cataloghi federati Amazon Redshift](#)
- [Visualizzazione degli oggetti del catalogo](#)
- [Aggiornamento di un catalogo federato](#)
- [Accesso a un catalogo federato condiviso](#)
- [Eliminazione di un catalogo federato](#)
- [Interrogazione di cataloghi federati](#)
- [Risorse aggiuntive](#)

Vantaggi principali

La registrazione di cluster e namespace Amazon Redshift con e l' AWS Glue Data Catalog unificazione dei dati tra i data lake Amazon S3 e i data warehouse Amazon Redshift offre i seguenti vantaggi:

- Esperienza di interrogazione uniforme: interroga i dati e i dati gestiti di Amazon Redshift nei bucket Amazon S3 utilizzando qualsiasi motore di query compatibile con Apache Iceberg, come Amazon EMR Serverless e Amazon Athena, senza dover spostare o copiare i dati.
- Accesso coerente ai dati tra i servizi: non è necessario aggiornare i nomi di database e tabelle nelle pipeline di dati quando si accede alle stesse fonti di dati federate da diversi servizi di AWS analisi, poiché le fonti di dati sono registrate nel Catalogo dati.
- Controllo granulare degli accessi: puoi applicare le autorizzazioni di Lake Formation per gestire l'accesso alle fonti di dati federate utilizzando autorizzazioni di controllo degli accessi granulari.

Ruoli e responsabilità

Ruolo	Responsabilità

Produttore e amministratore del cluster Amazon Redshift	Registra il cluster o lo spazio dei nomi con il Data Catalog.
Amministratore del data lake Lake Formation	Accetta l'invito al cluster o al namespace, crea cataloghi federati e concede l'accesso ai cataloghi federati ad altri principali.
Amministratore di sola lettura di Lake Formation	Trova il catalogo federato, interroga le tabelle di Amazon Redshift nel catalogo federato.
Ruolo di trasferimento dei dati	Amazon Redshift si assume per tuo conto il trasferimento di dati da e verso il bucket Amazon S3.

Di seguito sono riportati i passaggi di alto livello per fornire agli utenti l'accesso a uno spazio dei nomi Amazon Redshift:

1. In Amazon Redshift, l'amministratore del cluster produttore registra un cluster o un namespace nel Data Catalog.
2. L'amministratore del data lake accetta l'invito allo spazio dei nomi dall'amministratore del cluster di produttori di Amazon Redshift e crea un catalogo federato nel Data Catalog.

Dopo aver completato questo passaggio, puoi gestire il catalogo dei namespace di Amazon Redshift all'interno del Data Catalog.

3. Concedi le autorizzazioni agli utenti su cataloghi, database e tabelle. Puoi condividere l'intero catalogo dei namespace o un sottoinsieme di tabelle con gli utenti dello stesso account o di un altro account.

Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog

1. Crea un amministratore del data lake: crea un ruolo IAM autorizzato ad accettare l'invito al namespace e crea AWS Glue Data Catalog gli oggetti (cataloghi, database, tabelle/viste) e concedi le autorizzazioni Lake Formation ad altri utenti.

Per step-by-step istruzioni sulla creazione di un amministratore di data lake, consulta. [Crea un amministratore del data lake](#)

2. Aggiorna le autorizzazioni di amministratore del data lake.

Oltre alle autorizzazioni di amministratore del data lake, l'amministratore del data lake richiede le seguenti autorizzazioni per accettare un invito allo spazio dei nomi Amazon Redshift in Lake Formation, creare o aggiornare le risorse del Data Catalog e abilitare l'accesso al data lake:

JSON

```
{
  "Version": "2012-10-17",
  "Id": "glue-enable-datalake-access",
  "Statement": [{
    "Effect": "Allow",
    "Action": [
      "redshift:AssociateDataShareConsumer",
      "redshift:DescribeDataSharesForConsumer",
      "redshift:DescribeDataShares",
      "redshift-serverless:CreateNamespace",
      "redshift-serverless:CreateWorkgroup",
      "redshift-serverless>DeleteNamespace",
      "redshift-serverless>DeleteWorkgroup",
      "ec2:DescribeAccountAttributes",
      "ec2:DescribeSubnets",
      "ec2:DescribeSecurityGroups",
      "ec2:DescribeAvailabilityZones",
      "s3:createBucket",
      "s3:deleteBucket",
      "s3:putBucketPolicy",
      "s3:putEncryptionConfiguration",
      "s3:putLifecycleConfiguration",
      "s3:putBucketVersioning",
      "iam:CreateRole"
    ],
    "Resource": "*"
  },
  {
    "Action": [
      "iam:PassRole"
    ],

```

```

    "Effect": "Allow",
    "Resource": "arn:aws:iam::*:role/data transfer role name",
    "Condition": {
      "StringLike": {
        "iam:PassedToService": [
          "glue.amazonaws.com"
        ]
      }
    }
  }
]
}

```

3. Se il ruolo IAM utilizzato per creare cataloghi federati non è un amministratore di data lake, devi concedere al ruolo l'autorizzazione. `Create catalog`

Per creare creatori di cataloghi

- a. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
- b. Scegli Ruoli e attività amministrative in Amministrazione.
- c. Scegli Concessione.
- d. Nella schermata Concedi le autorizzazioni, scegli un utente o un ruolo IAM.
- e. Seleziona Crea autorizzazione al catalogo.
- f. Facoltativamente, puoi anche concedere l'autorizzazione alla creazione di catalogo. L'autorizzazione concessa consente al creatore del catalogo di concedere l'`Create catalog` autorizzazione ad altri responsabili.
- g. Scegli Concessione.

AWS CLI esempio di concessione delle autorizzazioni per creare un catalogo federato.

```

aws lakeformation grant-permissions \
--cli-input-json \
'{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/Admin"
  },
  "Resource": {
    "Catalog": {
    }
  }
},

```



```

    "Permissions": [
      "CREATE_CATALOG",
      "DESCRIBE"
    ]
  }'

```

4. Crea un ruolo di amministratore di sola lettura per scoprire i cataloghi federati di Amazon Redshift nel Data Catalog da Amazon Redshift Query Editor v2.

Per interrogare le tabelle Amazon Redshift nel catalogo federato da Amazon Redshift Query Editor v2, assicurati che la policy del ruolo di amministratore di sola lettura contenga l'ARN per il ruolo collegato al servizio Amazon Redshift- `AWSServiceRoleForRedshift`

```

aws lakeformation put-data-lake-settings
  --region us-east-1 \
  --data-lake-settings \
  '{
    "DataLakeAdmins": [{"DataLakePrincipalIdentifier":"arn:aws:iam::123456789012:role/Admin"}],
    "ReadOnlyAdmins":[{"DataLakePrincipalIdentifier":"arn:aws:iam::123456789012:role/aws-service-role/redshift.amazonaws.com/AWSServiceRoleForRedshift"}],
    "CreateDatabaseDefaultPermissions":[],
    "CreateTableDefaultPermissions":[],
    "Parameters":{"CROSS_ACCOUNT_VERSION":"4","SET_CONTEXT":"TRUE"}
  }'

```

5. Crea un ruolo di trasferimento dati che Amazon Redshift può assumere per tuo conto per trasferire dati da e verso il bucket Amazon S3.

Quando abiliti l'accesso al data lake per motori di query compatibili con Apache Iceberg come Athena, Amazon EMR su EC2 Amazon per accedere alle risorse Amazon Redshift nel Data Catalog, devi creare un ruolo IAM con le autorizzazioni necessarie per eseguire il trasferimento di dati da e verso il bucket Amazon S3.

JSON

```

{
  "Version":"2012-10-17",
  "Id": "glue-enable-datalake-access",
  "Statement": [{

```

```

    "Sid": "DataTransferRolePolicy",
    "Effect": "Allow",
    "Action": [ "glue:GetCatalog",
                "glue:GetDatabase",
                "kms:GenerateDataKey",
                "kms:Decrypt"],
    "Resource": "*"
  }
]
}

```

6. Aggiungi la seguente policy di fiducia al ruolo di trasferimento dati per AWS Glue i servizi Amazon Redshift per assumere il ruolo di trasferimento dati da e verso il bucket Amazon S3.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [{
    "Effect": "Allow",
    "Principal": {
      "Service": [
        "redshift.amazonaws.com",
        "glue.amazonaws.com"
      ]
    },
    "Action": "sts:AssumeRole"
  }]
}

```

7. Aggiungi la seguente policy chiave alla chiave se utilizzi una AWS KMS chiave gestita dal cliente per crittografare i dati nel cluster/namespaces di Amazon Redshift. Sostituisci il numero di account con un numero di AWS account valido e specifica il nome del ruolo di trasferimento dati. Per impostazione predefinita, i dati nel cluster Amazon Redshift sono crittografati utilizzando una chiave KMS. Lake Formation offre un'opzione per creare una chiave KMS personalizzata per la crittografia. Se utilizzi una chiave gestita dal cliente, devi aggiungere politiche chiave specifiche alla chiave.

Per ulteriori informazioni relative alla gestione delle autorizzazioni di una chiave gestita dal cliente, consulta [Customer managed keys](#).

JSON

```
{
  "Version": "2012-10-17",
  "Id": "auto-redshift-3",
  "Statement": [{
    "Sid": "RedshiftAllowAccessPolicy",
    "Effect": "Allow",
    "Principal": {
      "AWS": "*"
    },
    "Action": [
      "kms:Encrypt",
      "kms:Decrypt",
      "kms:ReEncrypt*",
      "kms:GenerateDataKey*",
      "kms:CreateGrant",
      "kms:DescribeKey"
    ],
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "kms:CallerAccount": "111122223333",
        "kms:ViaService": "redshift.us-east-1.amazonaws.com"
      }
    }
  }],
  {
    "Sid": "RedshiftServerlessAllowAccessPolicy",
    "Effect": "Allow",
    "Principal": {
      "AWS": "*"
    },
    "Action": [
      "kms:Encrypt",
      "kms:Decrypt",
      "kms:ReEncrypt*",
      "kms:GenerateDataKey*",
      "kms:CreateGrant",
      "kms:DescribeKey"
    ]
  }
}
```

```

    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:CallerAccount": "111122223333",
            "kms:ViaService": "redshift-serverless.us-
east-1.amazonaws.com"
        }
    }
},
{
    "Sid": "DirectMetadataAccess",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
    },
    "Action": [
        "kms:Describe*",
        "kms:Get*",
        "kms:List*",
        "kms:RevokeGrant"
    ],
    "Resource": "*"
},
{
    "Sid": "GenerateDataKeyDecryptDataTransferRole",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/data-transfer-role-name"
    },
    "Action": [
        "kms:GenerateDataKey",
        "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:ViaService": "s3.us-east-1.amazonaws.com"
        }
    }
}
]
}

```

Creazione di cataloghi federati Amazon Redshift

Questo argomento descrive i passaggi da seguire per accettare un invito a un cluster o a un namespace, creare un catalogo federato a più livelli e concedere autorizzazioni ad altri principali. Puoi completare queste attività utilizzando la console Lake Formation, il AWS Command Line Interface (AWS CLI) o APIs/SDKs. Gli esempi in questo argomento mostrano il cluster/namespace del produttore, il Data Catalog e il data consumer nello stesso account.

Per ulteriori informazioni sulle funzionalità cross-account di Lake Formation, consulta [Condivisione dei dati tra account in Lake Formation](#).

Per gestire uno spazio dei nomi Amazon Redshift nel Data Catalog

1. Rivedi un invito al namespace e accettalo.

Console

1. Accedi alla console di Lake Formation come amministratore del data lake all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Vai alla pagina Cataloghi in Data Catalog.
2. Controlla l'invito al namespace a cui sei autorizzato ad accedere. La colonna Status indica il tuo attuale stato di partecipazione al namespace. Lo stato Non accettato indica che sei stato aggiunto al namespace, ma non l'hai ancora accettato o hai rifiutato l'invito.

Catalogs

▼ How it works

Create a catalog
Register Redshift databases as catalogs in the Data Catalog. [Learn more](#)

Manage catalog permissions
Manage permissions for specific catalogs, databases, tables and fine-grained data access. [Learn more](#)

Access from query editors
Access catalog objects from [Redshift Query Editor v2](#) and [Athena Console](#).

Create a federated catalog for your S3 Table Buckets. Enable S3 Table integration

Pending catalog invitations (4) Approve and create catalog Reject

View and manage Redshift namespace/cluster invitations in the AWS Glue Data Catalog.

Find invitations

Name	Source account ID	Received	Status
arn:aws:redshift-serverless:us-east-2:451785580005:namespace:c0381d75-3f21-49f2-b2c3-44d000803a71		November 20, 2024 at 10:16 PM UTC	Accepted, catalog not created
arn:aws:redshift-serverless:us-east-2:451785580005:namespace/4a798b4c-71d8-4df4-b77f-52cff8ac80a1		November 20, 2024 at 5:38 PM UTC	Accepted, catalog not created
arn:aws:redshift:us-east-2:451785580005:namespace:48a491a6-d5d8-415b-b5b2-3832a4affb08		November 26, 2024 at 3:45 PM UTC	Accepted, catalog not created
arn:aws:redshift:us-east-2:451785580005:namespace:a77f139c-5a19-4b53-a662-c2f50db2fc28		December 3, 2024 at 2:21 PM UTC	Accepted, catalog not created

Catalogs (1) Actions View Create catalog

A catalog is the top level in the Data Catalog's three-level data hierarchy and contains Data Catalog objects.

Find catalogs by name

Name	Type	Source	Owner account...	Shared resource	Shared resour...	Shared resource owner region
Default	Default catalog					

- Per rispondere a un invito al namespace o al cluster, seleziona il nome dell'invito e scegli Rivedi l'invito. In Accetta o rifiuta l'invito, controlla i dettagli dell'invito. Scegli Accetta per accettare l'invito o Rifiuta per rifiutarlo. Non avrai accesso al namespace se rifiuti l'invito.

AWS CLI

Gli esempi seguenti mostrano come visualizzare, accettare e registrare l'invito. Sostituisci l'Account AWS ID con un Account AWS ID valido. Sostituisci data-share-arn con l'effettivo Amazon Resource Name (ARN) che fa riferimento allo spazio dei nomi.

- Visualizza un invito in sospeso.

```
aws redshift describe-data-shares \
  --data-share-arn 'arn:aws:redshift:us-
east-1:123456789012:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
ds_internal_namespace' \
```

- Accetta un invito.

```
aws redshift associate-data-share-consumer \
  --data-share-arn 'arn:aws:redshift:us-
east-1:123456789012:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
ds_internal_namespace' \
```

```
--consumer-arn 'arn:aws:glue:us-east-1:123456789012:catalog'
```

3. Registra il cluster o lo spazio dei nomi nell'account Lake Formation. Utilizza l'operazione [RegisterResource](#) API per registrare il datashare in Lake Formation. DataShareArn è il parametro di input per. ResourceArn

 Note

Si tratta di un passaggio obbligatorio.

```
aws lakeformation register-resource \  
  --resource-arn 'arn:aws:redshift:us-  
east-1:123456789012:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/  
ds_internal_namespace'
```

2. Crea un catalogo federato.

Dopo aver accettato un invito, devi creare un catalogo federato nel Data Catalog che mappi gli oggetti nello spazio dei nomi Amazon Redshift al Data Catalog. Devi essere un amministratore del data lake o un utente o un ruolo che dispone delle autorizzazioni necessarie per creare un catalogo.

Console

1. Dopo aver accettato l'invito al namespace, viene visualizzata la pagina dei dettagli del catalogo Set.
2. Nella pagina dei dettagli del catalogo Set, inserisci un nome univoco per il catalogo. Usa lettere minuscole per i nomi dei cataloghi. I nomi dei cataloghi devono avere una lunghezza inferiore o uguale a 255 caratteri. Utilizzate questo identificatore per mappare lo spazio dei nomi internamente nella gerarchia dei metadati (CatalogID.DBName.Schema.Table).
3. Immettete una descrizione per il catalogo. La descrizione deve avere una lunghezza inferiore o uguale a 2048 caratteri.
4. Successivamente, seleziona la casella di controllo Accedi a questo catalogo dai motori compatibili con Iceberg per abilitare l'accesso alle risorse di Amazon Redshift utilizzando motori analitici compatibili con Apache Iceberg come Athena e Apache Spark su Amazon EMR.

Non è necessario abilitare l'accesso al data lake per accedere ai cataloghi federati tramite Amazon Redshift.

Catalog details

A catalog is the top level in the Data Catalog's three-level data hierarchy and contains Data Catalog objects.

Name

Catalog name is required, in lowercase characters, and no longer than 255 characters.

Type

Federated

Source

Redshift

Description - optional

Descriptions can be up to 2048 characters long.

Access from engines

You can access this catalog from open source engines as well as Amazon Redshift.

☒ Access this catalog from Iceberg compatible engines.
Choose this option to access the data catalog using with Apache Spark running on an EMR cluster.

IAM role

Role used by Redshift for loading data to and from S3 bucket that is created for the managed workgroup.

 [View](#) 

[Create an IAM role](#) 

Encryption options

Your data is encrypted by default with a key that AWS owns and manages for you. To choose a different key, customize your encryption settings.

☐ Customize encryption settings
To use the default key, clear this option.

[Cancel](#) [Skip to Review and create](#) [Next](#)

- Per consentire a questi motori di query di leggere e scrivere nei namespace di Amazon Redshift, crea un cluster Amazon Redshift AWS Glue gestito con le risorse di calcolo e storage necessarie per eseguire operazioni di lettura e scrittura senza influire sui carichi di lavoro del data warehouse di Amazon Redshift.

È inoltre necessario fornire a un ruolo IAM le autorizzazioni necessarie per trasferire dati da e verso il bucket Amazon S3.

- Per impostazione predefinita, i dati nel cluster Amazon Redshift sono crittografati utilizzando una chiave AWS gestita. Lake Formation offre un'opzione per creare una

chiave KMS personalizzata per la crittografia. Se utilizzi una chiave gestita dal cliente, devi aggiungere politiche chiave specifiche alla chiave.

Scegli Personalizza le impostazioni di crittografia se utilizzi una chiave gestita dal cliente per crittografare i dati nel cluster/namespaces di Amazon Redshift. Per utilizzare una chiave personalizzata, è necessario aggiungere una policy della chiave gestita personalizzata alla propria chiave KMS. Per ulteriori informazioni, consulta [Prerequisiti per la gestione dei namespaces di Amazon Redshift nel AWS Glue Data Catalog](#).

AWS CLI

Utilizza il seguente codice di esempio per creare un catalogo con i dati di Amazon Redshift pubblicati nel Data Catalog utilizzando il. AWS CLI

```
aws glue create-catalog
--cli-input-json \
'{
  "Name": "nscatalog",
  "CatalogInput": {
    "Description": "Redshift federated catalog",
    "CreateDatabaseDefaultPermissions" : [],
    "CreateTableDefaultPermissions": [],
    "FederatedCatalog": {
      "Identifier": "arn:aws:redshift:us-
east-1:123456789012:datashare:11524d7f-f56d-45fe-83f7-d7bb0a4d6d71/
ds_internal_namespace",
      "ConnectionName": "aws:redshift"
    },
    "CatalogProperties": {
      "DataLakeAccessProperties" : {
        "DataLakeAccess" : true,
        "DataTransferRole" :
"arn:aws:iam::123456789012:role/DataTransferRole"
      }
    }
  }
}'
```

3. Concedi le autorizzazioni agli utenti del tuo account o di account esterni.

Console di gestione AWS

1. Scegli Avanti per concedere le autorizzazioni ad altri utenti su cataloghi, database e tabelle condivisi.
2. Nella schermata Aggiungi autorizzazioni, scegli i principali e i tipi di autorizzazioni da concedere.

Add permissions

Principals

Choose the principals to grant permissions.

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

UserRole X
Role

Catalog permissions

Choose the permissions to grant on the catalog. Choosing Super user overwrites individual permissions, granting unrestricted administrative access.

☐ **Super user**
A super user has unrestricted administrative privileges to perform any operation on all resources within the catalog (databases, tables, and views).

Catalog permissions

Choose specific access permissions to grant.

☒ **Create database**
☐ **Describe**
☐ **Alter**

☐ **Drop**

☐ **Super**
This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions

Choose the permission that can be granted to others.

☒ **Create database**
☐ **Describe**
☐ **Alter**

☐ **Drop**

☐ **Super**
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel

Add

- a. Nella sezione Principali, scegli un tipo di principale e quindi specifica i principali per concedere le autorizzazioni.
- Utenti e ruoli IAM: scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

- Utenti e gruppi SAML: per SAML e Amazon Quick Suite utenti e gruppi, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire i comandi grant and ARNs revoke. AWS CLI AWS CLI

- Account esterni: per AWS, AWS organizzazione o responsabile IAM, inserisci uno o più AWS account IDs, organizzazione IDs IDs, unità organizzativa o ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID. Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole. L'ID di un'unità organizzativa inizia con «ou-» seguito da 4—32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.
- b. Nella sezione Autorizzazioni, seleziona le autorizzazioni e le autorizzazioni concedibili.

In Autorizzazioni del catalogo, seleziona una o più autorizzazioni da concedere.

In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere ad altri responsabili del proprio account. AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.

Scegli Super user per concedere all'utente autorizzazioni illimitate per le risorse (database, tabelle, viste) all'interno del catalogo.

3. Scegliere Aggiungi.

AWS CLI

Utilizzate gli esempi seguenti per concedere le autorizzazioni per cataloghi, database e tabelle utilizzando: AWS CLI

- L'esempio seguente mostra come concedere le autorizzazioni per il catalogo federato.

```
aws lakeformation grant-permissions
--cli-input-cli-json \
'{
    "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/
non-admin"
```

```

    },
    "Resource": {
      "Catalog": {
        "Id": "123456789012:nscatalog"
      }
    },
    "Permissions": [
      "DESCRIBE", "CREATE_CATALOG"
    ],
    "PermissionsWithGrantOption": [
    ]
  }'

```

- Utilizzate l'esempio seguente per concedere le autorizzazioni su un database.

```

aws lakeformation grant-permissions \
  --cli-input-json \
    '{
      "Principal": {

        "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/non-admin"
      },
      "Resource": {
        "Database": {
          "CatalogId": "123456789012:nscatalog/dev",
          "Name": "public"
        }
      },
      "Permissions": [
        "ALL"
      ]
    }'

```

- L'esempio seguente mostra come concedere le autorizzazioni su una tabella nel database Amazon Redshift.

```

aws lakeformation grant-permissions \
  --cli-input-json \
    '{
      "Principal": {

        "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/non-admin"
      },

```

```
"Resource": {
  "Table": {
    "CatalogId": "123456789012:nscatalog2/dev",
    "DatabaseName": "public",
    "TableWildcard" : {}
  },
  "Permissions": [
    "ALL"
  ]
}'
```

4. Scegli Avanti per esaminare i dettagli del catalogo e creare un catalogo federato. Il catalogo federato appena creato e gli oggetti del catalogo vengono visualizzati nella pagina Cataloghi.

Si fa riferimento a un catalogo federato Amazon Redshift con. catalogID =
123456789012:Redshift-federated catalog id

Visualizzazione degli oggetti del catalogo

Dopo aver creato il catalogo federato, puoi visualizzare gli oggetti nel catalogo utilizzando la console Lake Formation oppure AWS CLI.

Console di gestione AWS

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Scegli Catalogs in Data Catalog.
3. Scegli un catalogo federato dall'elenco nella pagina Cataloghi.
4. La pagina di riepilogo del catalogo mostra gli oggetti del catalogo (database e tabelle) per i quali disponi delle autorizzazioni. La scheda Autorizzazioni mostra i responsabili IAM a cui sono state concesse le autorizzazioni su questi oggetti.

AWS CLI

- L' AWS CLI esempio seguente mostra come richiedere il catalogo di primo livello.

```
aws glue get-catalog \  
--catalog-id 123456789012:nscatalog
```

Risposta

```
{
  "Catalog": {
    "CatalogId": "123456789012:nscatalog",
    "Name": "nscatalog",
    "ResourceArn": "arn:aws:glue:us-east-1:123456789012:catalog/nscatalog",
    "Description": "Redshift published Catalog",
    "CreateTime": "2024-09-05T14:49:16-07:00",
    "FederatedCatalog": {
      "Identifier": "arn:aws:redshift:us-east-1:123456789012:datashare:b1234589-e823-4a14-ad8e-077085540a50/ds_internal_namespace",
      "ConnectionName": "aws:redshift"
    },
    "CatalogProperties": {
      "DataLakeAccessProperties": {
        "DataLakeAccess": true,
        "DataTransferRole": "arn:aws:iam::123456789012:role/DataTransferRole",
        "KmsKey": "AWS_OWNED_KMS_KEY",
        "ManagedWorkgroupName": "123456789012:nscatalog",
        "ManagedWorkgroupStatus": "AVAILABLE",
        "RedshiftDatabaseName": "dev"
      }
    },
    "CatalogIdentifier": "e2309c2c2fb048f1a3069dfdc1c7883e",
    "CreateTableDefaultPermissions": [],
    "CreateDatabaseDefaultPermissions": []
  }
}
```

- L'esempio seguente mostra come richiedere tutti i cataloghi dell'account.

```
aws glue get-catalogs \
  --recursive
```

- La seguente richiesta di esempio mostra come ottenere un catalogo a livello di database Amazon Redshift.

```
aws glue get-catlog \
```

```
--catalog-id 123456789012:namespace catalog name/redshift database name
```

- La seguente richiesta di esempio mostra come ottenere i database nel catalogo a livello di database di Amazon Redshift.

```
aws glue get-databases \  
--catalog-id 123456789012:namespace catalog name/redshift database name
```

- La seguente richiesta di esempio mostra come inserire una tabella Amazon Redshift nel catalogo.

```
aws glue get-table \  
--catalog-id 123456789012:parent catalog name/redshift database \  
--database-name redshift schema name \  
--name table name
```

- L'esempio seguente mostra come ottenere tutte le tabelle dal database Amazon Redshift.

```
aws glue get-tables \  
--catalog-id 123456789012:namespace catalog name/redshift database name \  
--database-name RS schema name
```

Aggiornamento di un catalogo federato

Puoi aggiornare un catalogo federato Amazon Redshift nel Data Catalog utilizzando la console Lake Formation AWS CLI o l'[UpdateCatalog](#) operazione API.

Console di gestione AWS

Segui questi passaggi per aggiornare il tuo catalogo federato utilizzando la console Lake Formation.

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione a sinistra, scegli Cataloghi in Data Catalog.
3. Nella pagina Cataloghi, scegli il catalogo federato Amazon Redshift che desideri aggiornare.
4. In Azioni, scegli Modifica.

5. Nella schermata dei dettagli di Set Catalog, nella sezione Accesso dai motori, scegli Accedi a questo catalogo dai motori compatibili con Iceberg. La selezione di questa opzione abiliterà l'accesso al data lake per i motori di query compatibili con Apache Iceberg.
6. Successivamente, crea un nuovo ruolo IAM o scegli un ruolo IAM esistente con la policy che concede le autorizzazioni per eseguire il trasferimento di dati da e verso il bucket Amazon S3.

Per ulteriori informazioni sulle autorizzazioni, consulta. [Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog](#)

7. Per impostazione predefinita, i dati nel cluster Amazon Redshift sono crittografati utilizzando un. Chiave gestita da AWS Se scegli di crittografare i dati utilizzando una chiave gestita dal cliente, crea una chiave KMS o scegli una chiave esistente con le autorizzazioni definite nella sezione. [Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog](#)
8. Scegli Save (Salva).

Una volta completato con successo, la pagina dei dettagli del catalogo mostra il nome del gruppo di lavoro gestito con lo stato «Operazione completata».

AWS CLI

Di seguito è riportato un esempio di input update-catalog CLI con l'accesso al data lake disabilitato impostando il valore del DataLakeAccess parametro come. false

```
aws glue update-catalog --cli-input-json \  
'{  
  "Name": "nscatalog",  
  "CatalogInput": {  
    "Description": "Redshift published catalog",  
    "CreateDatabaseDefaultPermissions" : [],  
    "CreateTableDefaultPermissions": [],  
    "FederatedCatalog": {  
      "Identifier": "arn:aws:redshift:us-  
east-1:123456789012:datashare:11524d7f-f56d-45fe-83f7-d7bb0a4d6d71/  
ds_internal_namespace",  
      "ConnectionName": "aws:redshift"  
    },  
    "CatalogProperties": {  
      "DataLakeAccessProperties" : {  
        "DataLakeAccess" : false  
      }  
    }  
  }  
'
```

```
}  
  }  
}'
```

Accesso a un catalogo federato condiviso

AWS Lake Formation le funzionalità tra account consentono agli utenti di condividere in modo sicuro i data lake distribuiti tra più AWS aziende, organizzazioni o direttamente con Account AWS i responsabili IAM in un altro account, fornendo un accesso granulare ai metadati e ai dati sottostanti.

Lake Formation utilizza il servizio AWS Resource Access Manager (AWS RAM) per facilitare la condivisione delle risorse. Quando condividi una risorsa del catalogo con un altro account, AWS RAM invia un invito all'account del beneficiario per accettare o rifiutare la concessione della risorsa.

I servizi analitici integrati come Amazon Athena e Redshift Spectrum richiedono collegamenti alle risorse per poter includere risorse condivise nelle query. I responsabili devono creare un collegamento tra le risorse che rimandano a una risorsa AWS Glue Data Catalog condivisa da un altro Account AWS. Per ulteriori informazioni sui collegamenti alle risorse, vedi [Come funzionano i collegamenti alle risorse in Lake Formation](#).

Un contenitore di link Catalog è un oggetto Data Catalog, che fa riferimento a un catalogo a livello di database federato locale o interaccount proveniente da altri account. AWS È inoltre possibile creare link al database e link a tabelle all'interno di un contenitore di link al catalogo. Quando crei un link al database o un link a una tabella, devi specificare una risorsa di destinazione che si trova nello stesso catalogo di destinazione a livello di database Amazon Redshift (database Amazon Redshift).

Per creare un contenitore di link al catalogo, è necessario il Lake Formation `CREATE_CATALOG` o `glue:CreateCatalog` autorizzazione.

È necessario disporre delle impostazioni della versione 4 dell'account Cross per condividere database o tabelle nel catalogo federato tra Account AWS s.

Creazione di un contenitore di link al catalogo verso un catalogo federato tra più account

È possibile creare un contenitore di link al catalogo che punti a un catalogo federato a livello di database Redshift in qualsiasi AWS regione utilizzando la AWS Lake Formation console, AWS Glue `CreateCatalog` l'API o (). AWS Command Line Interface AWS CLI

Per creare un contenitore di link al catalogo verso un catalogo condiviso (console)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come preside con il CREATE_CATALOG permesso di Lake Formation.
2. Nel riquadro di navigazione, scegli Cataloghi, quindi scegli Crea catalogo.
3. Nella pagina Imposta i dettagli del catalogo, fornisci le seguenti informazioni:

Name

Inserisci un nome che rispetti le stesse regole del nome del catalogo. Il nome può essere lo stesso del catalogo condiviso di destinazione.

Tipo

Scegliete Catalog Link Container come tipo di catalogo.

Origine

Scegli Redshift.

Catalogo Target Redshift

Seleziona un catalogo federato a livello di database Redshift o scegli un catalogo locale (di proprietà) dall'elenco.

L'elenco contiene tutti i cataloghi condivisi con il tuo account. Nota che l'ID dell'account del proprietario del catalogo è riportato in ogni catalogo. Se non vedi un catalogo che sai è stato condiviso con il tuo account, controlla quanto segue:

- Se non sei un amministratore del data lake, verifica che l'amministratore del data lake ti abbia concesso le autorizzazioni Lake Formation sul catalogo.
- Se sei un amministratore del data lake e il tuo account non fa parte della stessa AWS organizzazione dell'account concedente, assicurati di aver accettato l'invito AWS Resource Access Manager (AWS RAM) alla condivisione delle risorse per il catalogo. Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

Note

Quando si crea un contenitore di link al catalogo tramite la console, il menu a discesa Target Redshift Catalog potrebbe essere visualizzato **No matches** quando si tenta di selezionare un catalogo Redshift per più account. Nonostante questa visualizzazione, puoi inserire manualmente l'ARN di destinazione del catalogo federato a livello di

database Amazon Redshift (database Amazon Redshift) nel campo di input e il modulo continuerà a funzionare correttamente. Ad esempio: `arn:aws:glue:us-east-1:123456789012:catalog/federated-catalog-redshift/dev`. Questo comportamento si verifica perché la console può cercare solo potenziali candidati all'interno dell'account attualmente connesso. Il menu a discesa è inteso come una funzionalità di completamento automatico, ma puoi comunque inserirlo manualmente ARNs per l'accesso su più account.

4. Per consentire ai motori di query Apache Iceberg di leggere e scrivere nei namespace di Amazon Redshift, crea AWS Glue un cluster Amazon Redshift gestito con le risorse di calcolo e storage necessarie per eseguire operazioni di lettura e scrittura senza influire sui carichi di lavoro del data warehouse di Amazon Redshift. È necessario fornire un ruolo IAM con le autorizzazioni necessarie per trasferire dati da e verso il bucket Amazon S3.
5. Scegli Next (Successivo).
6. (Facoltativo) Scegli Aggiungi autorizzazioni per concedere autorizzazioni ad altri principali.

Tuttavia, la concessione delle autorizzazioni su un contenitore di collegamenti al catalogo non concede le autorizzazioni sul catalogo di destinazione (collegato). È necessario concedere le autorizzazioni sul catalogo di destinazione separatamente affinché il link al catalogo sia visibile in Athena.

7. Successivamente, esaminate i dettagli del contenitore dei link al catalogo e scegliete Crea catalogo.

Puoi quindi visualizzare il nome del contenitore di link nella pagina Cataloghi.

Ora puoi creare link al database e link a tabelle nel contenitore di link del catalogo per consentire l'accesso dai motori di query.

Creare un esempio di CLI del contenitore di link al catalogo

- Nell'esempio seguente, l'`TargetRedshiftCatalogoggetto` specifica l'arn del catalogo federato a livello di database Amazon Redshift (database Amazon Redshift). `DataLakeAccessDeve` essere abilitato quando crei il contenitore di link al catalogo.

```
aws glue create-catalog \  
  --cli-input-json \  
  '{  
    "Name": "linkcontainer",
```

```

    "CatalogInput": {
      "TargetRedshiftCatalog": {
        "CatalogArn": "arn:aws:us-east-1:123456789012:catalog/nscatalog/dev"
      },
      "CatalogProperties": {
        "DataLakeAccessProperties": {
          "DataLakeAccess": true,
          "DataTransferRole": "arn:aws:iam::111122223333:role/
DataTransferRole"
        }
      }
    }
  }
}'

```

Creazione di collegamenti alle risorse nel contenitore dei link del catalogo

È possibile creare collegamenti a risorse a database e tabelle in un contenitore di collegamenti al catalogo. Quando crei link a risorse di database o link a risorse a tabelle, devi specificare una risorsa di destinazione che risieda nello stesso catalogo di destinazione a livello di database Amazon Redshift (database Amazon Redshift) a cui punta il contenitore di link.

Puoi creare un link di risorsa a un database o a una tabella Amazon Redshift condiviso utilizzando la AWS Lake Formation console, l'API o AWS Command Line Interface (AWS CLI).

- Per istruzioni dettagliate, vedi [Creazione di un collegamento di risorsa a un database Data Catalog condiviso](#).

Di seguito è riportato un AWS CLI esempio per creare un collegamento a una risorsa di database in un contenitore di link al catalogo.

```

aws glue create-database \
  --cli-input-json \
  '{
    "CatalogId": "111122223333:linkcontainer",
    "DatabaseInput": {
      "Name": "dblink",
      "TargetDatabase": {
        "CatalogId": "123456789012:nscatalog/dev",
        "DatabaseName": "schema1"
      }
    }
  }'

```

```
}
}'
```

- Per creare un collegamento alle risorse della tabella in un contenitore di collegamenti al catalogo, è necessario innanzitutto creare un AWS Glue database in locale che contenga il collegamento AWS Glue Data Catalog alla risorsa della tabella.

Per ulteriori informazioni sulla creazione di collegamenti di risorse a tabelle condivise, consulta [Creazione di un collegamento di risorsa a una tabella condivisa del Catalogo dati](#).

- Crea un database che contenga l'esempio di collegamento alle risorse della tabella

```
aws glue create-database \
  --cli-input-json \
  '{
    "CatalogId": "111122223333:linkcontainer",
    "DatabaseInput": {
      "Name": "db1",
      "Description": "creating parent database for table link"
    }
  }'
```

- Crea un esempio di collegamento alle risorse della tabella

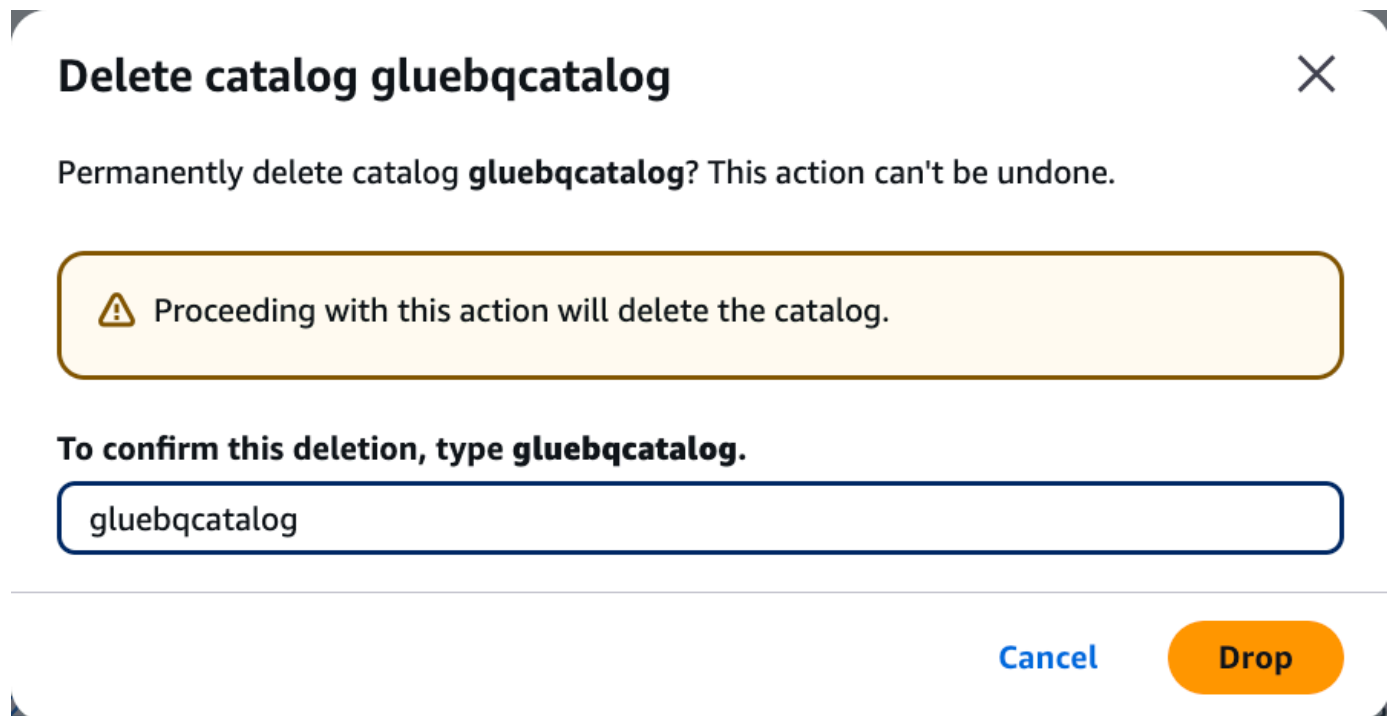
```
aws glue create-table \
  --cli-input-json \
  '{
    "CatalogId": "111122223333:linkcontainer",
    "DatabaseName": "db1",
    "TableInput": {
      "Name": "tablelink",
      "TargetTable": {
        "CatalogId": "123456789012:nscatalog/dev",
        "DatabaseName": "schema1",
        "Name": "table1"
      }
    }
  }'
```

Eliminazione di un catalogo federato

È possibile eliminare i cataloghi federati creati AWS Glue Data Catalog utilizzando l'operazione `glue:DeleteCatalog` o la console. AWS Lake Formation

Per eliminare un catalogo federato (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione, scegli Catalogs in Data Catalog.
3. Scegli il catalogo che desideri eliminare dall'elenco dei cataloghi.
4. Scegli Elimina da Azioni.
5. Scegli Elimina per confermare e il catalogo federato verrà eliminato dal Data Catalog.



Per eliminare un catalogo federato (CLI)

- ```
aws glue delete-catalog
--catalog-id 123456789012:catalog name
```

## Interrogazione di cataloghi federati

Dopo aver concesso le autorizzazioni ad altri responsabili, possono accedere e iniziare a interrogare le tabelle nei cataloghi federati accedendo agli strumenti SQL utilizzando Amazon Redshift, Amazon EMR ed ETL. Amazon Athena AWS Glue

[Per ulteriori informazioni sulla connessione all'endpoint di AWS Glue Data Catalog estensione Apache Iceberg Rest o all'applicazione Spark autonoma, consulta la sezione Accesso alla sezione della Guida per gli sviluppatori. AWS Glue Data Catalog AWS Glue](#)

Puoi utilizzare le query DDL (Data Definition Language) per creare e gestire tabelle nel database utilizzando Apache Spark su Amazon EMR. Per creare ed eliminare tabelle nel database Amazon Redshift, il principale deve disporre delle autorizzazioni Lake Formation `Create table`. `Drop`

Per ulteriori informazioni sulla concessione delle autorizzazioni di Data Catalog, consulta.

[Concessione delle autorizzazioni per le risorse di Data Catalog](#)

Per ulteriori informazioni sull'interrogazione delle risorse del catalogo Amazon Athena, consulta [Interrogazione Amazon Athena nella Guida per l'utente AWS Glue Data Catalog di Amazon Athena](#).

## Risorse aggiuntive

Puoi utilizzare [Amazon SageMaker Lakehouse](#) per ottenere un accesso unificato ai dati sia nei data warehouse che nei data lake. Tramite SageMaker Lakehouse, puoi utilizzare i motori preferiti di analisi, machine learning e business intelligence attraverso un'API REST Apache Iceberg aperta per garantire un accesso sicuro ai dati con controlli di accesso coerenti e dettagliati.

- [SageMaker Officina Amazon](#)
- [Semplifica l'accesso ai dati per la tua azienda con Amazon SageMaker Lakehouse](#)

## La federazione in fonti di dati esterne in AWS Glue Data Catalog

Puoi connettere AWS Glue Data Catalog (Data Catalog) a data warehouse come Amazon Redshift, Snowflake, database cloud come Amazon RDS Amazon DynamoDB e Oracle e servizi di streaming come Amazon MSK e sistemi locali come Teradata utilizzando connessioni. AWS Glue Queste connessioni vengono archiviate AWS Glue Data Catalog e registrate con AWS Lake Formation, consentendoti di creare un catalogo federato per ogni fonte di dati disponibile.



Un catalogo federato è un contenitore di primo livello che punta a un database in un sistema di dati esterno. Consente di interrogare i dati direttamente dal sistema di dati esterno senza il processo di estrazione, trasformazione e caricamento (ETL).

Per ulteriori informazioni sulle AWS Glue connessioni, consulta [Connecting to data](#) nella AWS Glue Developer Guide.

Gli amministratori di Data Lake possono creare cataloghi federati utilizzando [Amazon SageMaker Lakehouse](#) o [Amazon Athena](#).

Gli amministratori di Data Lake possono quindi concedere autorizzazioni dettagliate sugli oggetti all'interno del catalogo utilizzando Lake Formation, controllando l'accesso a vari livelli come catalogo, database, tabella, colonna, riga o cella. Gli analisti di dati possono scoprire e interrogare le fonti di dati catalogate utilizzando Athena, con Lake Formation che applica le politiche di accesso definite. Gli analisti possono unire i dati provenienti da più fonti in un'unica query senza dover connettersi a ciascuna fonte singolarmente.

## Argomenti

- [Flusso di lavoro](#)
- [Prerequisiti per connettere il Data Catalog a fonti di dati esterne](#)
- [Creazione di un catalogo federato utilizzando una connessione AWS Glue](#)
- [Visualizzazione degli oggetti del catalogo](#)
- [Eliminazione di un catalogo federato](#)
- [Interrogazione di cataloghi federati](#)
- [Risorse aggiuntive](#)

## Flusso di lavoro

Un amministratore del data lake o un utente con le autorizzazioni richieste completa i seguenti passaggi per connetterli AWS Glue Data Catalog a un'origine dati esterna.

1. Crea una AWS Glue connessione all'origine dati. Quando registri la connessione, il ruolo IAM utilizzato nella registrazione della connessione deve avere accesso alla funzione Lambda e alla posizione dello spill bucket di Amazon S3.
2. Registra la connessione con Lake Formation.

3. Crea un catalogo federato nel Data Catalog utilizzando una AWS Glue connessione per connettersi alle fonti di dati disponibili. I database, le tabelle e le viste vengono catalogati automaticamente nel Data Catalog e registrati con Lake Formation.
4. Concede l'accesso a cataloghi, database e tabelle specifici agli analisti di dati utilizzando le autorizzazioni di Lake Formation. È possibile definire politiche di controllo degli accessi granulari tra data lake, warehouse e fonti OLTP utilizzando Lake Formation, abilitando filtri di sicurezza a livello di riga e colonna.

Gli analisti di dati possono quindi accedere a tutti i dati tramite il Data Catalog utilizzando le query SQL in Athena, senza bisogno di connessioni separate o credenziali di origine dati. Gli analisti possono eseguire query SQL federate che analizzano i dati da più fonti, unendo i dati sul posto senza pipeline di dati complesse.

## Prerequisiti per connettere il Data Catalog a fonti di dati esterne

Per connetterli AWS Glue Data Catalog a fonti di dati esterne, registrare la connessione con Lake Formation e configurare cataloghi federati, devi soddisfare i seguenti requisiti:

### Note

Consigliamo a un amministratore del data lake di Lake Formation di creare le AWS Glue connessioni per connettersi a fonti di dati esterne e creare i cataloghi federati.

### 1. Creare ruoli IAM.

- Crea un ruolo con le autorizzazioni necessarie per distribuire le risorse (funzione Lambda, Amazon S3 spill bucket, ruolo IAM e AWS Glue connessione) necessarie per creare una connessione all'origine dati esterna.
- Crea un ruolo con le autorizzazioni minime necessarie per accedere alle proprietà della AWS Glue connessione (la funzione Lambda e lo spill bucket di Amazon S3). Questo è il ruolo che includerai quando registrerai la connessione con Lake Formation.

Per utilizzare Lake Formation per gestire e proteggere i dati nel tuo data lake, devi registrare la AWS Glue connessione con Lake Formation. In questo modo, Lake Formation può fornire credenziali ad Amazon Athena per interrogare le fonti di dati federate.

Il ruolo deve avere Select o Describe autorizzazioni sul bucket Amazon S3 e sulla funzione Lambda.

- s3: ListBucket
- s3: GetObject
- lambda: InvokeFunction

JSON

```
{
 "Version": "2012-10-17",
 "Statement": [
 {
 "Effect": "Allow",
 "Action": [
 "s3:*"
],
 "Resource": [
 "arn:aws:s3:::amzn-s3-demo-bucket1/object/*",
 "arn:aws:s3:::amzn-s3-demo-bucket1/object"
]
 },
 {
 "Sid": "lambdainvoke",
 "Effect": "Allow",
 "Action": "lambda:InvokeFunction",
 "Resource": "arn:aws:lambda:us-east-1:123456789012:function:example-lambda-function"
 },
 {
 "Sid": "gluepolicy",
 "Effect": "Allow",
 "Action": "glue:*",
 "Resource": "*"
 }
]
}
```

- Aggiungi la seguente politica di fiducia al ruolo IAM utilizzato nella registrazione della connessione:

## JSON

```
{
 "Version": "2012-10-17",
 "Statement": [
 {
 "Effect": "Allow",
 "Principal": {
 "Service": [
 "lakeformation.amazonaws.com"
]
 },
 "Action": "sts:AssumeRole"
 }
]
}
```

- L'amministratore del data lake che registra la connessione deve disporre dell'`iam:PassRole` autorizzazione per il ruolo.

Di seguito è riportata una politica in linea che concede questa autorizzazione. Sostituiscilo `<account-id>` con un numero di AWS account valido e `<role-name>` sostituiscilo con il nome del ruolo.

## JSON

```
{
 "Version": "2012-10-17",
 "Statement": [
 {
 "Sid": "PassRolePermissions",
 "Effect": "Allow",
 "Action": [
 "iam:PassRole"
],
 "Resource": [
 "arn:aws:iam::111122223333:role/example-role-name"
]
 }
]
}
```

- Per creare cataloghi federati in Data Catalog, assicurati che il ruolo IAM che stai utilizzando sia un amministratore del data lake di Lake Formation controllando le impostazioni del data lake (`aws lakeformation get-data-lake-settings`).

Se non sei un amministratore di data lake, hai bisogno dell'`CREATE_CATALOG` autorizzazione Lake Formation per creare un catalogo. L'esempio seguente mostra come concedere le autorizzazioni necessarie per creare cataloghi.

```
aws lakeformation grant-permissions \
--cli-input-json \
 '{
 "Principal": {
 "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/non-
admin"
 },
 "Resource": {
 "Catalog": {
 }
 },
 "Permissions": [
 "CREATE_CATALOG",
 "DESCRIBE"
]
 }'
```

2. Aggiungi la seguente politica chiave alla AWS KMS chiave se utilizzi una chiave gestita dal cliente per crittografare i dati nell'origine dati. Sostituisci il numero di account con un numero di AWS account valido e specifica il nome del ruolo. Per impostazione predefinita, i dati vengono crittografati utilizzando una chiave KMS. Lake Formation offre un'opzione per creare una chiave KMS personalizzata per la crittografia. Se utilizzi una chiave gestita dal cliente, devi aggiungere politiche chiave specifiche alla chiave.

Per ulteriori informazioni relative alla gestione delle autorizzazioni di una chiave gestita dal cliente, consulta [Customer managed keys](#).

JSON

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

```
"Statement": [
 {
 "Effect": "Allow",
 "Action": [
 "kms:Encrypt",
 "kms:Decrypt",
 "kms:ReEncrypt*",
 "kms:GenerateDataKey*",
 "kms:DescribeKey"
],
 "Resource": "arn:aws:kms:us-east-1:123456789012:key/key-1"
 }
]
```

## Creazione di un catalogo federato utilizzando una connessione AWS Glue

Per connetterla AWS Glue Data Catalog a fonti di dati esterne, è necessario utilizzare AWS Glue connessioni che consentano la comunicazione con le fonti di dati esterne. Puoi creare AWS Glue connessioni utilizzando la AWS Glue console, l'API [Create connection](#) e la console Amazon SageMaker Lakehouse.

Per istruzioni dettagliate sulla creazione di una AWS Glue connessione, consulta [Connessione ai dati](#) nella Guida per gli AWS Glue sviluppatori o [Creazione di connessioni in Amazon SageMaker Lakehouse](#).

Quando un utente esegue una query su tabelle federate, Lake Formation invia credenziali che richiamano una AWS Lambda funzione specificata nella AWS Glue connessione per recuperare oggetti di metadati dall'origine dati.

### Console di gestione AWS

Per creare un catalogo federato da un'origine dati esterna e impostare le autorizzazioni (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione, scegli Catalogs in Data Catalog.
3. Seleziona l'opzione Crea catalogo.
4. Nella pagina dei dettagli di Set Catalog, inserite le seguenti informazioni:

- Step 1  
 ● **Set catalog details**  
 Step 2 - optional  
 ● Grant permissions  
 Step 3  
 ● Review and create

## Set catalog details

Create a catalog in the Data Catalog.

### Catalog details

A catalog is the top level in the Data Catalog's three-level data hierarchy and contains Data Catalog objects.

#### Name

Catalog name is required, in lowercase characters, and no longer than 255 characters.

#### Type

#### Source

#### Connection

#### Description - optional

Descriptions can be up to 2048 characters long.

### Register Glue connection with Lake Formation

You can access this catalog from AWS Glue data connections.

#### IAM role

Choose a role that has permissions to invoke an AWS Glue connector.


[View](#)
[Create an IAM role](#)

☒ **Activate the connector and connect to the data source.**

A connector is a piece of code that runs on AWS Lambda that translates between the target data source and query engine (Athena).

#### Encryption options

Your data is encrypted by default with a key that AWS owns and manages for you. To choose a different key, customize your encryption settings.

☐ **Customize encryption settings**

To use the default key, clear this option.

[Cancel](#)
[Skip to Review and create](#)
[Next](#)

- Nome: un nome univoco per il tuo catalogo federato. Il nome non può essere modificato e deve essere scritto in minuscolo. Il nome può essere composto da un massimo di 255 caratteri. account.
  - Tipo: scegli il catalogo federato come tipo di catalogo.
  - Fonte: scegli una fonte di dati dal menu a discesa. Vengono visualizzate le origini dati per le quali sono state create connessioni. Per ulteriori informazioni sulla creazione di una AWS Glue connessione a un'origine dati esterna, consulta [Creazione di connessioni per connettori](#) nella Guida per gli AWS Glue sviluppatori o [Creazione di connessioni in Amazon SageMaker Lakehouse](#).
  - Connessione: scegli una AWS Glue connessione esistente all'origine dati.
  - Descrizione: immetti una descrizione per il catalogo creato dalla fonte di dati.
5. Scegli un ruolo IAM che Lake Formation assuma per vendere le credenziali al motore di interrogazione per accedere ai dati dalla fonte dati. Questo ruolo deve disporre delle

autorizzazioni necessarie per accedere alla AWS Glue connessione e richiamare la funzione Lambda per accedere ai dati dall'origine dati esterna.

Puoi anche creare un nuovo ruolo nella console IAM.

Consulta la [Prerequisiti per connettere il Data Catalog a fonti di dati esterne](#) sezione per le autorizzazioni richieste.

6. Seleziona l'opzione Attiva il connettore per connetterti all'origine dati per consentire ad Athena di eseguire query federate.

Per l'elenco dei connettori supportati, consulta [Registra la tua connessione](#) nella Guida per l'utente di Amazon Athena.

7. Opzioni di crittografia: scegli l'opzione Personalizza le impostazioni di crittografia se desideri utilizzare una chiave personalizzata per crittografare il catalogo. Per utilizzare una chiave personalizzata, è necessario aggiungere una policy della chiave gestita personalizzata alla propria chiave KMS.
8. Scegli Avanti per concedere le autorizzazioni ad altri principali.
9. Nella pagina Concedi autorizzazioni, scegli Aggiungi autorizzazioni.
10. Nella schermata Aggiungi autorizzazioni, scegli i principali e i tipi di autorizzazioni da concedere.



## Add permissions

### Principals

Choose the principals to grant permissions.

☒ **IAM users and roles**  
Users or roles from this AWS account.

☐ **SAML users and groups**  
SAML users and group or QuickSight ARNs.

☐ **External accounts**  
AWS account, AWS organization or IAM principal outside of this account

#### IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

UserRole X  
Role

### Catalog permissions

Choose the permissions to grant on the catalog. Choosing Super user overwrites individual permissions, granting unrestricted administrative access.

☐ **Super user**  
A super user has unrestricted administrative privileges to perform any operation on all resources within the catalog (databases, tables, and views).

#### Catalog permissions

Choose specific access permissions to grant.

☒ **Create database**

☐ Describe

☐ Alter

☐ Drop

☐ **Super**  
This permission is the union of all the individual permissions to the left, and supersedes them.

#### Grantable permissions

Choose the permission that can be granted to others.

☒ **Create database**

☐ Describe

☐ Alter

☐ Drop

☐ **Super**  
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel

Add

- Nella sezione Principali, scegli un tipo di principale e quindi specifica i principali per concedere le autorizzazioni.
- Utenti e ruoli IAM: scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

- Utenti e gruppi SAML: per SAML e Amazon Quick Suite utenti e gruppi, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.
- Nella sezione Autorizzazioni, seleziona le autorizzazioni e le autorizzazioni concedibili.

In Autorizzazioni del catalogo, seleziona una o più autorizzazioni da concedere.

Scegli Super utente per concedere autorizzazioni amministrative illimitate su tutte le risorse all'interno del catalogo.

In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere agli altri responsabili del proprio account. AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.

11. Scegli Avanti per esaminare le informazioni e creare il catalogo. L'elenco dei cataloghi mostra il nuovo catalogo federato.

L'elenco delle posizioni dei dati mostra la connessione federata appena registrata.

| Data lake locations (7)                                                                                 |                                                  |                                                   |                      |                    |                                   |
|---------------------------------------------------------------------------------------------------------|--------------------------------------------------|---------------------------------------------------|----------------------|--------------------|-----------------------------------|
| <input type="text" value="Find data lake storage"/> <span>Actions</span> <span>Register location</span> |                                                  |                                                   |                      |                    |                                   |
|                                                                                                         | Data lake location                               | IAM role                                          | Location Type        | Permission mode    | Last modified                     |
| <input type="radio"/>                                                                                   | <a href="#">ddb_ds_3</a>                         | <a href="#">SageMakerStudioQueryExecutionR...</a> | Federated connection | Lake Formation     | November 26, 2024 at 10:34 PM UTC |
| <input type="radio"/>                                                                                   | <a href="#">postgre_db2</a>                      | <a href="#">SageMakerStudioQueryExecutionR...</a> | Federated connection | Lake Formation     | November 24, 2024 at 11:12 AM UTC |
| <input type="radio"/>                                                                                   | <a href="#">sf_ds2</a>                           | <a href="#">SageMakerStudioQueryExecutionR...</a> | Federated connection | Lake Formation     | November 24, 2024 at 3:27 AM UTC  |
| <input type="radio"/>                                                                                   | <a href="#">s3://amazon-sagemaker-5390106...</a> | <a href="#">datazone_usr_role_50wwm8ts855...</a>  | Amazon S3            | Lake Formation     | November 24, 2024 at 3:10 AM UTC  |
| <input type="radio"/>                                                                                   | <a href="#">ddb_ds_2</a>                         | <a href="#">SageMakerStudioQueryExecutionR...</a> | Federated connection | Lake Formation     | November 24, 2024 at 3:05 AM UTC  |
| <input type="radio"/>                                                                                   | <a href="#">s3://amazon-sagemaker-5390106...</a> | <a href="#">datazone_usr_role_adtmv7d4lm98...</a> | Amazon S3            | Lake Formation     | November 23, 2024 at 9:15 PM UTC  |
| <input type="radio"/>                                                                                   | <a href="#">s3://data-lake-pk-us-east-2</a>      | <a href="#">AWSServiceRoleForLakeFormation...</a> | Amazon S3            | Hybrid access mode | November 21, 2024 at 7:40 PM UTC  |

## AWS CLI

Per creare un catalogo federato da un'origine dati esterna e impostare le autorizzazioni

1. L'esempio seguente mostra come creare una AWS Glue connessione.

```
aws glue create-connection
--connection-input \
 '{
 "Name": "DynamoDB connection",
```

```

 "ConnectionType": "DYNAMODB",
 "Description": "A connection created for DynamoDB",
 "ConnectionProperties": {},
 "AthenaProperties": {"spill_prefix": "your_spill_prefix",
 "lambda_function_arn": "Lambda_function_arn",
 "spill_bucket": "Your_Bucket_name",
 "AuthenticationConfiguration": {}
 }'

```

2. L'esempio seguente mostra come registrare una AWS Glue connessione con Lake Formation.

```

aws lakeformation register-resource
 --cli-input-json \
 '{"ResourceArn':"arn:aws:glue:us-east-1:123456789012:connection/
dynamo", "RoleArn": "arn:aws:iam::123456789012:role/
AdminTelemetry", "WithFederation": true}'

```

3. L'esempio seguente mostra come creare un catalogo federato.

```

aws glue create-catalog
 --cli-input-json \
 '{
 "Name": "ddbcatalog",
 "CatalogInput": {"CatalogProperties": {"DataLakeAccessProperties":
{"DataTransferRole": "arn:aws:iam::123456789012:role/role name"}},
 "CreateDatabaseDefaultPermissions": [],
 "CreateTableDefaultPermissions": [],
 "FederatedCatalog": {"ConnectionName": "dynamo", "Identifier": "dynamo"}
 }'

```

## Visualizzazione degli oggetti del catalogo

Per ogni fonte di dati disponibile, AWS Glue crea un catalogo corrispondente in AWS Glue Data Catalog. Dopo aver creato un catalogo, puoi visualizzare i database e le tabelle nel catalogo utilizzando la console Lake Formation o AWS CLI. In

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Scegli Catalogs in Data Catalog. La pagina dei cataloghi mostra i cataloghi per i quali disponi delle autorizzazioni.

**Catalogs (11)** Actions View Create catalog

A catalog is the top level in the Data Catalog's three-level data hierarchy and contains Data Catalog objects.

Find catalogs by name

|                       | Name                | Type      | Source                 | Owner account... | Shared resource | Shared resourc... | Shared resource owner region |
|-----------------------|---------------------|-----------|------------------------|------------------|-----------------|-------------------|------------------------------|
| <input type="radio"/> | 30005               | Default   | Default catalog        | 0005             | -               | -                 | -                            |
| <input type="radio"/> | bkaiyuan_nscatal... | Federated | Redshift               | 0005             | -               | -                 | -                            |
| <input type="radio"/> | bkaiyuan_test_ca... | Federated | -                      | 0005             | -               | -                 | -                            |
| <input type="radio"/> | linkcontainer-leon  | Managed   | Catalog Link container | 0005             | -               | -                 | -                            |
| <input type="radio"/> | mymulticatalog      | Federated | TPCDS                  | 0005             | -               | -                 | -                            |
| <input type="radio"/> | test-bug-share      | Federated | Redshift               | 0005             | -               | -                 | -                            |
| <input type="radio"/> | test-zetl           | Managed   | Redshift               | 0005             | -               | -                 | -                            |
| <input type="radio"/> | test-zetl-mwg       | Managed   | Redshift               | 0005             | -               | -                 | -                            |
| <input type="radio"/> | tpcdscatalog        | Federated | TPCDS                  | 0005             | -               | -                 | -                            |
| <input type="radio"/> | yansoncatalog2      | Federated | Redshift               | 0005             | -               | -                 | -                            |
| <input type="radio"/> | zetltest123         | Managed   | Redshift               | 0005             | -               | -                 | -                            |

3. Scegliete un catalogo dall'elenco per visualizzare i database e le tabelle in esso contenuti. L'elenco contiene i database presenti nell'account e i collegamenti alle risorse, che sono collegamenti a database e tabelle condivisi negli account esterni e vengono utilizzati per l'accesso tra account diversi ai dati nel data lake.

### Catalog summary

|                                                                                   |                             |                                      |
|-----------------------------------------------------------------------------------|-----------------------------|--------------------------------------|
| <b>Name</b><br>451785580005                                                       | <b>Data encryption</b><br>- | <b>IAM role</b><br>-                 |
| <b>Catalog ARN</b><br><a href="#">arn:aws:glue:us-west-2:451785580005:catalog</a> |                             | <b>KMS key for optimization</b><br>- |

Objects Permissions Table optimizations

### Databases (1/14)

Find databases

Actions
View

| Name                                                          | Owner account ID | Lake Fo... | Default ...  | Shared ...    | Shared ...  | Shared ... | Amazo...                  | Descri...    |
|---------------------------------------------------------------|------------------|------------|--------------|---------------|-------------|------------|---------------------------|--------------|
| <b>arfarajpostgresqldb</b>                                    |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> aws:cloudtrail                          |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> default                                 |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> gluedynamodb                            |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> mysnowflakedb                           |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> snowflakedb                             |                  | -          | Lake Form... | -             | -           | -          | -                         | -            |
| <input type="radio"/> test-db-0737fa687d584b2d9ab72fbd...     |                  | -          | Lake Form... | test-db-0...  | 45178558... | -          | -                         | -            |
| <input type="radio"/> test-db-1927b03560764a4b81a216e9...     |                  | -          | Lake Form... | test-db-1...  | 45178558... | -          | -                         | -            |
| <input type="radio"/> test-db-32ee54b6949b4 added85b8ff066... |                  | -          | Lake Form... | test-db-3...  | 45178558... | -          | -                         | -            |
| <input type="radio"/> test-db-5978c2e076aa4583a28df124f...    |                  | -          | Lake Form... | test-db-5...  | 45178558... | -          | -                         | -            |
| <input type="radio"/> test-db-5cebe417bf734eafbbeaf7d3d6...   |                  | -          | Lake Form... | test-db-5c... | 45178558... | -          | -                         | -            |
| <input type="radio"/> test-db-7fa7ca8de2b84232ae3e8dcf...     |                  | -          | Lake Form... | -             | -           | -          | <a href="#">http://db</a> | database ... |
| <input type="radio"/> test-db-bb19fb486dc14ad68813612...      |                  | -          | Lake Form... | -             | -           | -          | <a href="#">http://db</a> | database ... |
| <input type="radio"/> test-db-cdec6ecaa0f143b7b4d6f24a8...    |                  | -          | Lake Form... | test-db-cd... | 45178558... | -          | -                         | -            |

4. Scegli l'opzione Tabelle in Visualizza per visualizzare e gestire le tabelle nel database.

## AWS CLI esempi per la visualizzazione di cataloghi e database

L'esempio seguente mostra come visualizzare un catalogo utilizzando AWS CLI

```
aws glue get-catalog \
--catalog-id 123456789012:dynamodbcatalog
```

L'esempio seguente mostra come richiedere tutti i cataloghi dell'account.

```
aws glue get-catalogs \
 --recursive
```

La seguente richiesta di esempio mostra come ottenere i database nel catalogo.

```
aws glue get-database \
--catalog-id 123456789012:dynamodbcatalog
```

```
--database-name database name
```

## Eliminazione di un catalogo federato

È possibile eliminare i cataloghi federati creati AWS Glue Data Catalog utilizzando l'operazione `glue:DeleteCatalog` o la console. AWS Lake Formation

Per eliminare un catalogo federato (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione, scegli Catalogs in Data Catalog.
3. Scegli il catalogo che desideri eliminare dall'elenco dei cataloghi.
4. Scegli Elimina da Azioni.
5. Scegli Elimina per confermare e il catalogo federato verrà eliminato dal Data Catalog.

### Delete catalog gluebqcatalog



Permanently delete catalog **gluebqcatalog**? This action can't be undone.

 Proceeding with this action will delete the catalog.

To confirm this deletion, type **gluebqcatalog**.

Cancel

Drop

Per eliminare un catalogo federato (CLI)

- ```
aws glue delete-catalog  
--catalog-id 123456789012:catalog name
```

Interrogazione di cataloghi federati

Dopo aver concesso le autorizzazioni ad altri responsabili, questi possono accedere e iniziare a interrogare le tabelle nei cataloghi federati utilizzando Athena.

Per creare ed eliminare tabelle nel database federato, il principale deve disporre delle Drop autorizzazioni di Lake Formation `Create table`.

Per ulteriori informazioni sulla concessione delle autorizzazioni di Data Catalog, consulta.

[Concessione delle autorizzazioni per le risorse di Data Catalog](#)

Per ulteriori informazioni sull'interrogazione del catalogo dati da Amazon Athena, consulta [Interrogazione AWS Glue Data Catalog Amazon Athena](#) nella Guida per l'utente di Amazon Athena.

Risorse aggiuntive

In questo post del blog, mostriamo come gli analisti di dati possono ora accedere e interrogare in modo sicuro i dati archiviati all'esterno dei data lake S3, inclusi i data warehouse e i Amazon DynamoDB database di Amazon Redshift, il tutto attraverso un'unica esperienza unificata. Gli amministratori possono ora applicare controlli di accesso a diversi livelli di granularità per garantire che i dati sensibili rimangano protetti ampliando al contempo l'accesso ai dati. Ciò consente alle organizzazioni di accelerare le iniziative relative ai dati mantenendo al contempo la sicurezza e la conformità, portando a un processo decisionale più rapido e basato sui dati.

- [Cataloga e gestisci le query federate di Amazon Athena con Amazon Lakehouse SageMaker](#)

Creazione di un catalogo Amazon S3 Tables nel AWS Glue Data Catalog

[Le tabelle Amazon S3](#) forniscono uno storage S3 specificamente ottimizzato per i carichi di lavoro di analisi, migliorando le prestazioni delle query e riducendo i costi. I dati in Tabelle S3 sono archiviati in un nuovo tipo di bucket: un bucket di tabelle, che archivia le tabelle come sottorisorse. Le tabelle S3 hanno il supporto integrato per lo standard Apache Iceberg, che consente di interrogare facilmente i dati tabulari nei bucket di tabelle Amazon S3 utilizzando i motori di query più diffusi come Apache Spark.

Puoi integrare bucket e tabelle di Amazon S3 con AWS Glue Data Catalog (Data Catalog) e registrare il catalogo come posizione dati di Lake Formation dalla console Lake Formation o

utilizzando il servizio. APIs Quando la tua organizzazione gestisce i dati nel Data Catalog e registra la posizione dei dati con Lake Formation, puoi utilizzare Lake Formation per controllare l'accesso ai tuoi set di dati.

Puoi applicare le autorizzazioni di Lake Formation utilizzando il controllo degli accessi basato su tag e il metodo della risorsa denominata sui database federati e condividerli tra più Account AWS AWS organizzazioni e unità organizzative (). OUs Puoi anche condividere i database federati direttamente con i responsabili IAM di un altro account.

Per ulteriori informazioni, consulta [Using Amazon S3 Tables with AWS analytics services](#) nella Amazon Simple Storage Service User Guide.

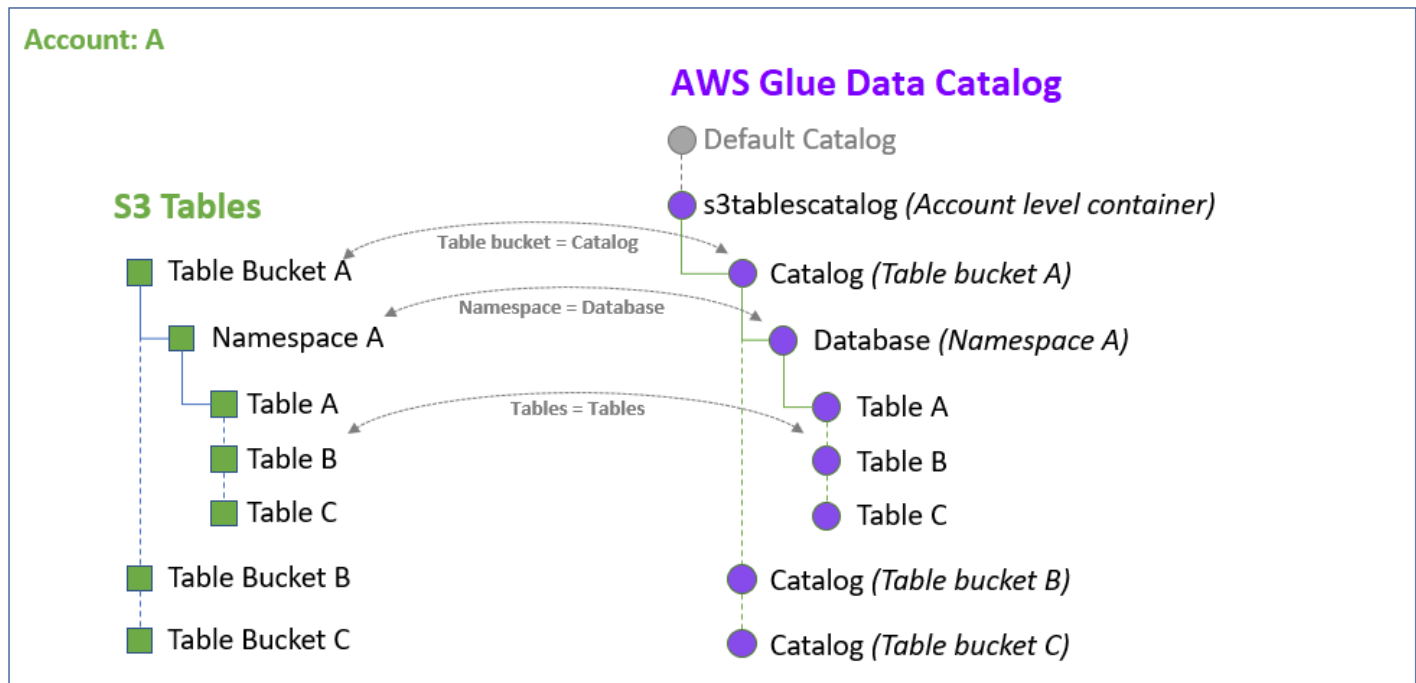
Argomenti

- [Come funziona l'integrazione tra Data Catalog e Lake Formation](#)
- [Prerequisiti per l'integrazione del catalogo di tabelle Amazon S3 con Data Catalog e Lake Formation](#)
- [Abilitazione dell'integrazione con Amazon S3 Tables](#)
- [Creazione di database e tabelle nel catalogo di tabelle S3](#)
- [Registrazione di un bucket da tavolo Amazon S3 in un altro account AWS](#)
- [Concessione di autorizzazioni](#)

Come funziona l'integrazione tra Data Catalog e Lake Formation

Quando integri il catalogo di tabelle S3 con Data Catalog e Lake Formation, il AWS Glue servizio crea un unico catalogo federato chiamato `s3tablescatalog` nel Data Catalog predefinito del tuo account specifico per il tuo. Regione AWS L'integrazione mappa tutte le risorse del table bucket di Amazon S3 nel tuo account e Regione AWS nel catalogo federato nel modo seguente:

- I bucket da tavolo Amazon S3 diventano un catalogo multilivello nel Data Catalog.
- Lo spazio dei nomi Amazon S3 associato è registrato come database nel Data Catalog.
- Le tabelle Amazon S3 nel bucket di tabelle diventano tabelle nel Data Catalog.



Dopo l'integrazione con Lake Formation, puoi creare tabelle Apache Iceberg nel catalogo di table buckets e accedervi tramite motori di AWS analisi integrati come Amazon Athena Amazon EMR e motori di analisi di terze parti.

Prerequisiti per l'integrazione del catalogo di tabelle Amazon S3 con Data Catalog e Lake Formation

Di seguito sono riportati i prerequisiti per abilitare l'integrazione AWS Glue Data Catalog delle tabelle Amazon S3 con e. AWS Lake Formation

1. Il processo di integrazione dei servizi di AWS analisi è stato aggiornato. Se l'integrazione è stata configurata con la versione di anteprima, è possibile continuare a utilizzare l'integrazione attuale. Tuttavia, il processo di integrazione aggiornato offre miglioramenti delle prestazioni. Per aggiornare l'integrazione:
 1. Innanzitutto, elimina il catalogo di tabelle S3 esistente in Lake Formation. Per eliminare il catalogo, seleziona il **s3tablescatalog** catalogo dall'elenco dei cataloghi e scegliete Elimina dalle azioni.
 2. Quindi, annulla la registrazione della posizione dei dati per. s3tablescatalog
 - a. Nella console Lake Formation, nella sezione Amministrazioni, scegli Data Locations.
 - b. Seleziona una posizione e dal menu Azioni scegli Rimuovi.

- c. Quando viene richiesta la conferma, scegli Rimuovi.

Per istruzioni dettagliate sull'annullamento della registrazione di una posizione dati, consulta la sezione. [Annullamento della registrazione di una sede Amazon S3](#)

- d. Quindi, segui i passaggi di integrazione aggiornati nella [Abilitazione dell'integrazione con Amazon S3 Tables](#) sezione.

2. Quando abiliti l'integrazione delle tabelle Amazon S3, Lake Formation registra automaticamente la posizione delle tabelle S3. Per registrare la posizione del table bucket con Lake Formation, è necessario un IAM role/user con `lakeformation:RegisterResource` e `lakeformation:CreateCatalog` autorizzazioni. `lakeformation:RegisterResourceWithPrivilegedAccess` Quando un utente non amministratore con queste autorizzazioni registra una posizione di catalogo, Lake Formation concede automaticamente l'`DATA_LOCATION_ACCESS` autorizzazione per quella posizione, concedendo al responsabile chiamante le autorizzazioni per eseguire tutte le operazioni di Lake Formation supportate sulla posizione dati registrata.
3. Quando abiliti l'integrazione delle tabelle S3, devi scegliere un ruolo IAM per Lake Formation per vendere le credenziali per consentire l'accesso ai dati. Crea un ruolo IAM per l'accesso ai dati di Lake Formation ai tuoi bucket di tabelle S3. Il ruolo IAM utilizzato durante la registrazione del table bucket con Lake Formation richiede le seguenti autorizzazioni:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationPermissionsForS3ListTableBucket",
      "Effect": "Allow",
      "Action": [
        "s3tables:ListTableBuckets"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "LakeFormationDataAccessPermissionsForS3TableBucket",
      "Effect": "Allow",
```

```

        "Action": [
            "s3tables:CreateTableBucket",
            "s3tables:GetTableBucket",
            "s3tables:CreateNamespace",
            "s3tables:GetNamespace",
            "s3tables:ListNamespaces",
            "s3tables>DeleteNamespace",
            "s3tables>DeleteTableBucket",
            "s3tables:CreateTable",
            "s3tables>DeleteTable",
            "s3tables:GetTable",
            "s3tables:ListTables",
            "s3tables:RenameTable",
            "s3tables:UpdateTableMetadataLocation",
            "s3tables:GetTableMetadataLocation",
            "s3tables:GetTableData",
            "s3tables:PutTableData"
        ],
        "Resource": [
            "arn:aws:s3tables:us-east-1:123456789012:bucket/*"
        ]
    }
}

```

Per ulteriori informazioni, consulta [Requisiti per i ruoli utilizzati per registrare le sedi](#).

4. Aggiungi la seguente policy di fiducia al ruolo IAM per consentire al servizio Lake Formation di assumere il ruolo e fornire credenziali temporanee ai motori analitici integrati.

```

{
  "Effect": "Allow",
  "Principal": {
    "Service": "lakeformation.amazonaws.com"
  },
  "Action": [
    "sts:AssumeRole",
    "sts:SetSourceIdentity",
    "sts:SetContext" # add action to trust relationship when using IAM Identity
                    center principals with Lake Formation
  ]
}

```

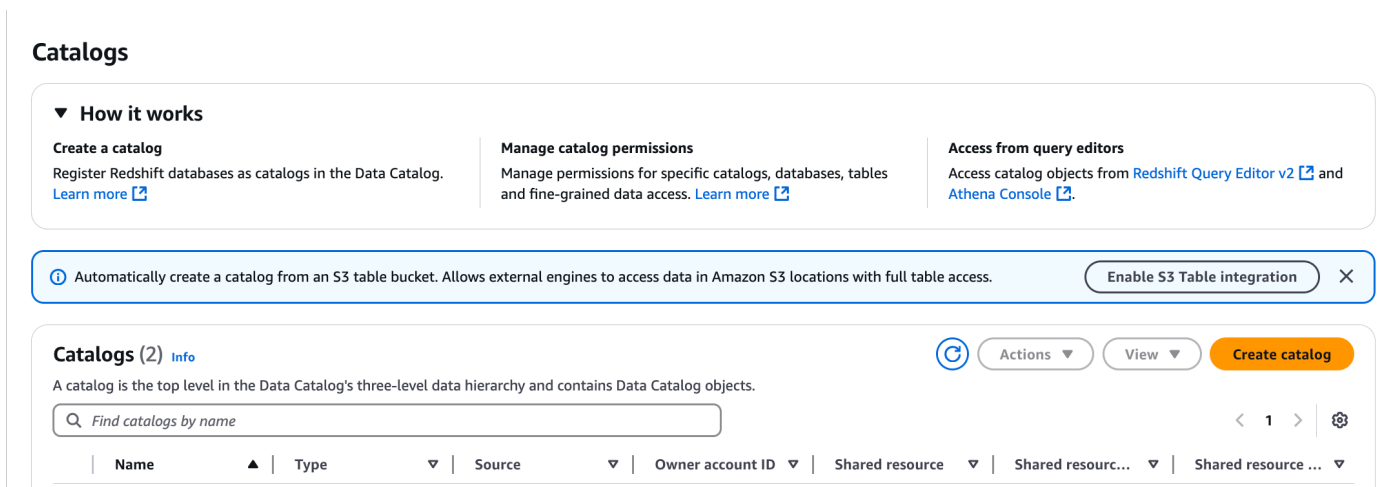
Abilitazione dell'integrazione con Amazon S3 Tables

Puoi creare bucket di tabelle Amazon S3 utilizzando la console Amazon S3 e integrarla con i servizi di analisi. AWS Per ulteriori informazioni, consulta [Usare le tabelle Amazon S3 con i servizi di AWS analisi](#).

Nel AWS Lake Formation, puoi abilitare l'integrazione di Amazon S3 Tables con AWS Glue Data Catalog e AWS Lake Formation utilizzando la console Lake Formation o utilizzarla. AWS CLI

Per integrare Amazon S3 Tables con Data Catalog e Lake Formation (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione, scegli Catalogs in Data Catalog.
3. Scegli Abilita l'integrazione con S3 Table nella pagina Cataloghi.



4. Scegli un ruolo IAM con le autorizzazioni richieste a Lake Formation da assumere per vendere le credenziali ai motori di query analitici. Per le autorizzazioni richieste per il ruolo di accesso ai dati, consulta la sezione sui [step3-permissions](#) prerequisiti.

Enable S3 Table integration

Once integration is enabled, every table bucket in this account and region will automatically be available under the s3tablescatalog catalog in AWS Data Catalog.

Select a principal to register

AWS LakeFormation needs to be able to call S3 Tables APIs on your behalf to retrieve S3 Table buckets, namespace, and tables. S3 tables must be registered with AWS LakeFormation with an IAM role that can be assumed.

Select a role to register

☐ Allow external engines to access data in Amazon S3 locations with full table access

Cancel Enable

i When you register an S3 tables catalog in Lake Formation, the registering principal automatically receives permissions to perform all supported Lake Formation operations on the registered data location.

5. Seleziona Consenti ai motori esterni di accedere ai dati nelle posizioni Amazon S3 con l'opzione di accesso completo alla tabella. Quando abiliti l'accesso completo alla tabella per i motori di terze parti, Lake Formation restituisce le credenziali direttamente al motore di terze parti senza eseguire la convalida dei tag di sessione IAM. Ciò significa che non è possibile applicare controlli di accesso granulari di Lake Formation alle tabelle a cui si accede.
6. Scegli Abilita . Il nuovo catalogo per S3 Tables viene aggiunto all'elenco dei cataloghi. Quando abiliti l'integrazione del catalogo di tabelle S3, il servizio registra la posizione dei dati del bucket di tabelle S3 con Lake Formation.
7. Scegli il catalogo per visualizzare gli oggetti del catalogo e concedere le autorizzazioni ad altri principali.

Success Successfully created catalog s3tablescatalog.

s3tablescatalog Actions

Catalog summary

Name	Permissions for newly created tables	Description
s3tablescatalog	-	-

Catalog connection details

Connect to data stored in lakes, warehouses, and other external data sources and publish to unified Iceberg data catalog.

Access for Open Source Engine	IAM role
-	-
Namespace register status	KMS key
Registered to AWS Data Catalog	-

Objects
Permissions

Data permissions for catalog s3tablescatalog (1)

View all permissions
Revoke
Grant

Filter permissions by property or value

☐	Principal	Princip...	Princip...	Resour...	Database	Table	Resource	Catalog	LF-Tag ex...	Permissions	Grantable	RAM F
☐	Admin	IAM role	arn:aws:ia...	Catalog	-	-	-	45178558...	-	All, Alter, ...	All, Alter, ...	-

Per creare cataloghi a più livelli, consulta la sezione [Creazione di un bucket da tabella](#) nella Guida per l'utente di Amazon Simple Storage Service.

Per integrare le tabelle Amazon S3 con Data Catalog e Lake Formation (CLI)

1. Registra il catalogo S3 Tables come posizione dati di Lake Formation.

```
aws lakeformation register-resource \
  --resource-arn 'arn:aws:s3tables:us-east-1:123456789012:bucket/*' \
  --role-arn 'arn:aws:iam::123456789012:role/LakeFormationDataAccessRole' \
  --with-federation
  --with-privileged-access
```

2. Crea un catalogo.

```
aws glue create-catalog --cli-input-json file://input.json

'{
  "Name": "s3tablescatalog",
  "CatalogInput" : {
    "FederatedCatalog": {
      "Identifier": "arn:aws:s3tables:us-east-1:123456789012:bucket/*",
      "ConnectionName": "aws:s3tables"
    },
  },
}
```

```
"CreateDatabaseDefaultPermissions": [],  
"CreateTableDefaultPermissions": []  
}  
'
```

Creazione di database e tabelle nel catalogo di tabelle S3

Puoi creare database per organizzare le tabelle Apache Iceberg e tabelle per definire lo schema e la posizione dei dati nel catalogo delle tabelle S3.

Crea un database (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore di data lake o creatore di database.
2. Nel riquadro di navigazione, scegli Database in Data Catalog.
3. Scegliere Crea database.
4. Nella pagina Crea database, scegli l'opzione Database e inserisci i seguenti dettagli:
 - Nome: un nome univoco per il database
 - Catalogo dati: scegli il catalogo delle tabelle S3. Il database risiederà in questo catalogo.
 - Descrizione: (Facoltativo) Aggiungi una descrizione e una posizione.
 - Controllo di accesso IAM per nuove tabelle: seleziona Facoltativamente Usa solo il controllo di accesso IAM per le nuove tabelle in questo database. Per informazioni su questa opzione, consulta la sezione [Modifica delle impostazioni predefinite per il tuo data lake](#).
 - Scegli Crea database. Puoi vedere il database creato nel catalogo delle tabelle S3.

Crea un database usando AWS CLI

Il seguente comando CLI mostra come creare un database nel catalogo di tabelle S3.

```
aws glue create-database  
--region us-east-1 \  
--catalog-id "123456789012:s3tablescatalog/test" \  
--database-input \  
'{ "Name": "testglueclidbcreation" }'
```

Crea una tabella ()Console di gestione AWS

Puoi creare tabelle di metadati Apache Iceberg nel catalogo delle tabelle S3 utilizzando la console Lake Formation o l'API. AWS Glue `CreateTable`

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake o utente con `CreateTable` autorizzazione.
2. Nel riquadro di navigazione, scegli Tabelle in Data Catalog.
3. Scegliere Create table (Crea tabella).
4. Nella pagina Crea tabella, inserisci i dettagli della tabella:

Create table [Info](#)

Table details

Create a table in the Data Catalog.

Name

If you plan to access the table from Amazon Athena, then the name should be under 256 characters and contain only lowercase letters (a-z), numbers (0-9), and underscore (_). For more information, see [Athena names](#).

Catalog

Table is contained within this catalog.



[Create a catalog](#)

Database

Table is contained within this database.



[Create database](#)

Table format

☒ Apache Iceberg Table

Create a table in the Apache Iceberg table format

Schema [Info](#)

View and manage table schema.

[Upload schema](#)[Delete](#)[Edit](#)[Add column](#)

< 1 >

No available schema

[Cancel](#)[Submit](#)

- Nome: inserisci un nome univoco per la tabella.
- Catalogo: scegli il catalogo delle tabelle S3 come catalogo.
- Database: scegli il database nel catalogo delle tabelle S3.

- **Descrizione:** inserisci una descrizione per la tabella.
- **Schema:** scegli **Aggiungi colonne** per aggiungere colonne e tipi di dati delle colonne. Si dispone della possibilità di creare una tabella vuota e di aggiornare lo schema in un secondo momento. Iceberg consente di evolvere lo schema e la partizione dopo aver creato la tabella. È possibile usare le query di Athena per aggiornare lo schema della tabella e le query Spark per aggiornare le partizioni.

5. Seleziona **Invia**.

Crea una tabella (AWS CLI)

```
aws glue create-table \  
--database-name "testglueclidbcreation" \  
--catalog-id "123456789012:s3tablescatalog/test" \  
--region us-east-1 \  
--table-input \  
'{ "Name": "testtablegluecli", "Parameters": { "format": "ICEBERG" },  
  "StorageDescriptor": { "Columns": [ {"Name": "x", "Type": "int", "Parameters":  
    {"required": "true"}} ] } }'
```

Registrazione di un bucket da tavolo Amazon S3 in un altro account AWS

Puoi registrare singoli table bucket Amazon S3 (formato ARN:arn:aws:s3tables:us-east-1:**account-id**:bucket/**bucket-name**) da un account con AWS Lake Formation in un altro account. Ad esempio, puoi registrare un table bucket dall'account A in Lake Formation dell'account B.

Prerequisiti

Prima di iniziare la registrazione del bucket tra account:

- Crea un bucket da tavolo nell'account A.
- Crea un ruolo IAM nell'account B con le autorizzazioni appropriate per la registrazione del bucket.

Per ulteriori informazioni sulle autorizzazioni necessarie per registrare un table bucket con Lake Formation, vedere. [Prerequisiti per l'integrazione del catalogo di tabelle Amazon S3 con Data Catalog e Lake Formation](#)

- Registra il table bucket nell'account in cui intendi creare il catalogo di tabelle S3.

- Per l'accesso tra account diversi, il ruolo specificato al momento della registrazione del table bucket deve essere un ruolo interno all'account con le autorizzazioni appropriate per accedere al bucket tra account. Il ruolo richiede le azioni IAM di S3 Tables necessarie per accedere alle risorse del bucket.

Registrazione di bucket di tabelle tra account

Nelle seguenti procedure, l'account A è l'account proprietario delle risorse e l'account B è il luogo in cui verrà registrato il table bucket per la gestione delle autorizzazioni di accesso.

1. Accedi all'account A. Console di gestione AWS

Apri la console Amazon S3 all'indirizzo. <https://console.aws.amazon.com/s3/>

2. Creare un bucket di tabelle. Per ulteriori informazioni, consulta [Creating a table bucket](#) nella Amazon S3 User Guide.
3. Registra il table bucket nell'account B.

Usa il AWS CLI per registrare il table bucket dall'account A con Lake Formation nell'account B.

```
aws lakeformation register-resource \  
--resource-arn 'arn:aws:s3tables:us-east-1:account-A-id:bucket/single-bucket-name' \  
\  
--role-arn arn:aws:iam::account-B-id:role/role-name \  
--region us-east-1
```

Sostituisci account-A-id single-bucket-name, account-B-id, e role-name con i tuoi valori specifici.

4. Quindi, crea un catalogo per il table bucket nell'account B.

Crea un catalogo utilizzando. AWS CLI

```
aws glue create-catalog --region us-east-1 \  
--cli-input-json \  
{  
  "Name": "catalog-name",  
  "CatalogInput" : {  
    "FederatedCatalog": {  
      "Identifier": "arn:aws:s3tables:us-east-1:account A:bucket/single-bucket-  
name",  
      "ConnectionName": "aws:s3tables"  
    },  
  },  
}
```

```

    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": []
  }
}'

```

Sostituisci catalog-name e bucket-name con i tuoi valori specifici. account-A-id

L'esempio CLI seguente mostra come visualizzare i dettagli del catalogo.

```

aws glue get-catalog \
  --catalog-id account-id:catalog-name \
  --region us-east-1

```

5. Successivamente, crea database e tabelle nel catalogo appena creato nell'account B.

Crea un database.

```

aws glue create-database \
  --region us-east-1 \
  --catalog-id "account-B-id:catalog-name" \
  --database-input \
  '{
    "Name": "database-name"
  }'

```

Creare una tabella.

```

aws glue create-table \
  --database-name database-name \
  --catalog-id account-B-id:catalog-name \
  --region us-east-1 \
  --table-input \
  '{
    "Name": "table-name",
    "Parameters": {
      "format": "ICEBERG"
    },
    "StorageDescriptor": {
      "Columns": [
        {"Name": "x", "Type": "int", "Parameters": {"required": "true"}}
      ]
    }
  }'

```

```
}  
}'
```

Sostituisci il nome del database `account-B-id`, il nome del catalogo e il nome della tabella con i tuoi valori specifici.

6. Gli esempi seguenti mostrano come visualizzare gli oggetti nel catalogo.

Visualizza i dettagli del database.

```
aws glue get-database \  
  --name database-name \  
  --catalog-id account-B-id:catalog-name \  
  --region us-east-1
```

Visualizza i dettagli della tabella.

```
aws glue get-table \  
  --name table-name \  
  --database-name database-name \  
  --catalog-id account-B-id:catalog-name \  
  --region us-east-1
```

Sostituisci il nome del database `account-B-id`, il nome del catalogo e il nome della tabella con i tuoi valori specifici.

Concessione di autorizzazioni

Dopo aver integrato le tabelle S3 con AWS Lake Formation, puoi concedere le autorizzazioni sul catalogo di tabelle S3 e sugli oggetti del catalogo (table bucket, database, tabelle) ad altri ruoli e utenti IAM del tuo account. Le autorizzazioni di Lake Formation consentono di definire i controlli di accesso a livello di tabella, colonna e riga per gli utenti di motori analitici integrati come Amazon Redshift Spectrum e Athena.

È possibile concedere le autorizzazioni utilizzando il metodo `named resource` o il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC). Prima di concedere le autorizzazioni utilizzando i tag LF e le espressioni dei tag LF, è necessario definirle e assegnarle agli oggetti del Data Catalog.

Per ulteriori informazioni, consulta [Gestione dei tag LF per il controllo dell'accesso ai metadati](#).

Puoi condividere database e tabelle con AWS account esterni concedendo le autorizzazioni di Lake Formation agli account esterni. Gli utenti possono quindi eseguire query e lavori che uniscono e interrogano tabelle su più account. Quando condividi una risorsa del catalogo con un altro account, i responsabili di quell'account possono utilizzare quella risorsa come se la risorsa si trovasse nel loro catalogo dati.

Quando condividi database e tabelle con account esterni, l'autorizzazione Super user non è disponibile.

Per istruzioni dettagliate sulla concessione delle autorizzazioni, consulta la [Gestione delle autorizzazioni di Lake Formation](#) sezione.

AWS CLI esempio di concessione delle autorizzazioni su una tabella Amazon S3

```
aws lakeformation grant-permissions \
--cli-input-json \
'{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:role/DataAnalystRole"
  },
  "Resource": {
    "Table": {
      "CatalogId": "111122223333:s3tablescatalog/amzn-s3-demo-bucket1",
      "DatabaseName": "S3 table bucket namespace <example_namespace>",
      "Name": "S3 table bucket table name <example_table>"
    }
  },
  "Permissions": [
    "SELECT"
  ]
}'
```

Di seguito sono riportati i parametri da includere nel comando:

- **DataLakePrincipalIdentifier** — ARN di utenti, ruoli o gruppi IAM per concedere le autorizzazioni
- **CatalogId** — ID dell' AWS account a 12 cifre che possiede il Data Catalog
- **DatabaseName** — Nome dello spazio dei nomi del bucket di tabella Amazon S3
- **Nome**: nome della tabella bucket di Amazon S3

- Autorizzazioni: autorizzazioni da concedere. Le opzioni includono: SELECT, INSERT, DELETE, DESCRIBE, ALTER, DROP, ALL e SUPER

Accesso a tabelle Amazon S3 condivise

Dopo aver concesso le autorizzazioni per più account su un database o una tabella nel catalogo delle tabelle S3, per accedere alle risorse, devi creare collegamenti alle risorse ai database e alle tabelle condivisi.

1. Nell'account di destinazione (l'account che riceve le risorse condivise), crea un link alle risorse del database. Per istruzioni dettagliate, vedi [Creazione di un collegamento di risorsa a un database Data Catalog condiviso](#).

Esempio CLI per la creazione di un collegamento a una risorsa del database

```
aws glue create-database
--region us-east-1
--catalog-id "111122223333"
--database-input \
'{
  "Name": "s3table_resourcelink",
  "TargetDatabase": {
    "CatalogId": "011426214932:s3tablescatalog/chmni-s3-table-bucket-011426214932",
    "DatabaseName": "s3_table_ns"
  },
  "CreateTableDefaultPermissions": []
}'
```

2. Concedi l'autorizzazione per più account sulla tabella.

Esempio CLI per la concessione di autorizzazioni tra account

```
aws lakeformation grant-permissions \
--region us-east-1 \
--cli-input-json \
'{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:role/
S3TablesTestExecRole"
  },
  "Resource": {
```

```
"Table": {
  "CatalogId": "011426214932:s3tablescatalog/chmni-s3-table-
bucket-011426214932",
  "DatabaseName": "s3_table_ns",
  "Name": "test_s3_iceberg_table"
},
"Permissions": [
  "ALL"
]
```

3. Concedi DESCRIBE l'autorizzazione a Lake Formation sul link alla risorsa.

Esempio CLI per la concessione dell'autorizzazione alla descrizione sul collegamento alla risorsa.

```
aws lakeformation grant-permissions \
  --principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/
S3TablesTestExecRole
  --resource Database='{CatalogId=111122223333;, Name=s3table_resourcelink}' \
  --permissions DESCRIBE
```

Creazione di un catalogo gestito di Amazon Redshift nel AWS Glue Data Catalog

Forse non disponi attualmente di un cluster di produttori Amazon Redshift o di un datashare Amazon Redshift, ma desideri creare e gestire tabelle Amazon Redshift utilizzando il AWS Glue Data Catalog. Puoi iniziare creando un catalogo AWS Glue gestito utilizzando l'`glue:CreateCatalogAPI` o la AWS Lake Formation console impostando il tipo di catalogo come Managed e Catalog source come Redshift. Questo passaggio esegue le seguenti operazioni:

- Crea un catalogo nel Data Catalog
- Registra il catalogo come posizione dati di Lake Formation
- crea un gruppo di lavoro serverless gestito da Amazon Redshift
- Collega il gruppo di lavoro serverless Amazon Redshift e il Data Catalog utilizzando un oggetto datashare

Per creare un catalogo gestito e configurare le autorizzazioni (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel riquadro di navigazione, scegli Catalogs in Data Catalog.
3. Seleziona l'opzione Crea catalogo.
4. Nella pagina dei dettagli di Set Catalog, inserite le seguenti informazioni:
 - Nome: un nome univoco per il catalogo gestito. Il nome non può essere modificato e deve essere scritto in minuscolo. Il nome può essere composto da un massimo di 255 caratteri account.
 - Tipo: scegli Managed catalog come tipo di catalogo.
 - Archiviazione: scegli Redshift l'archiviazione.
 - Descrizione: immetti una descrizione per il catalogo creato dalla fonte di dati.
5. Puoi utilizzare le applicazioni Apache Spark in esecuzione su Amazon EMR su Amazon per accedere ai database EC2 Amazon Redshift in. AWS Glue Data Catalog

Per consentire ad Apache Spark di leggere e scrivere sullo storage gestito di Amazon Redshift AWS Glue , crea un cluster Amazon Redshift gestito con le risorse di calcolo e storage necessarie per eseguire operazioni di lettura e scrittura senza influire sui carichi di lavoro del data warehouse di Amazon Redshift. È inoltre necessario fornire a un ruolo IAM le autorizzazioni necessarie per trasferire dati da e verso il bucket Amazon S3. Per le autorizzazioni richieste per il ruolo di trasferimento dei dati, consulta il passaggio 5 nella sezione. [Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog](#)

6. Per impostazione predefinita, i dati nel cluster Amazon Redshift sono crittografati utilizzando una chiave AWS gestita. Lake Formation offre un'opzione per creare una chiave KMS personalizzata per la crittografia. Se utilizzi una chiave gestita dal cliente, devi aggiungere politiche chiave specifiche alla chiave.
7. Scegli Personalizza le impostazioni di crittografia se utilizzi una chiave gestita dal cliente per crittografare i dati nello storage gestito di Amazon Redshift. Per utilizzare una chiave personalizzata, devi aggiungere una politica di chiave gestita personalizzata alla tua chiave KMS. Per ulteriori informazioni, consulta [Prerequisiti per la gestione dei namespace di Amazon Redshift nel AWS Glue Data Catalog](#).
8. Opzioni di crittografia: scegli l'opzione Personalizza le impostazioni di crittografia se desideri utilizzare una chiave personalizzata per crittografare il catalogo. Per utilizzare una chiave personalizzata, devi aggiungere una politica di chiave gestita personalizzata alla tua chiave KMS.

9. Scegli Avanti per concedere le autorizzazioni ad altri principali.
10. Nella pagina Concedi autorizzazioni, scegli Aggiungi autorizzazioni.
11. Nella schermata Aggiungi autorizzazioni, scegli i principali e i tipi di autorizzazioni da concedere.

Add permissions

Principals

Choose the principals to grant permissions.

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

UserRole X
Role

Catalog permissions

Choose the permissions to grant on the catalog. Choosing Super user overwrites individual permissions, granting unrestricted administrative access.

☐ **Super user**
A super user has unrestricted administrative privileges to perform any operation on all resources within the catalog (databases, tables, and views).

Catalog permissions

Choose specific access permissions to grant.

☒ **Create database**
☐ Describe
 ☐ Alter

☐ Drop

☐ **Super**
This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions

Choose the permission that can be granted to others.

☒ **Create database**
☐ Describe
 ☐ Alter

☐ Drop

☐ **Super**
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel

Add

- Nella sezione Principali, scegli un tipo di principale e quindi specifica i principali per concedere le autorizzazioni.
- Utenti e ruoli IAM: scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.
- Utenti e gruppi SAML: per SAML e Amazon Quick Suite utenti e gruppi, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire i comandi grant and ARNs revoke. AWS CLI AWS CLI

- Nella sezione Autorizzazioni, seleziona le autorizzazioni e le autorizzazioni concedibili.

In Autorizzazioni del catalogo, seleziona una o più autorizzazioni da concedere.

Scegli Super utente per concedere autorizzazioni amministrative illimitate su tutte le risorse all'interno del catalogo.

In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere agli altri responsabili del proprio account. AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.

12. Scegli Avanti per esaminare le informazioni e creare il catalogo. L'elenco dei cataloghi mostra il nuovo catalogo gestito.

Per creare un catalogo federato (CLI)

- L'esempio seguente mostra come creare un catalogo federato.

```
aws glue create-catalog --cli-input-json file://input.json

{
  "Name": "CatalogName",
  "CatalogInput": {
    "Description": "Redshift published Catalog",
    "CreateDatabaseDefaultPermissions" : [],
    "CreateTableDefaultPermissions": [],
    "CatalogProperties": {
      "DataLakeAccessProperties" : {
        "DataLakeAccess" : "true",
        "DataTransferRole" : "DTR arn",
        "KMSKey": "kms key arn", // Optional
        "CatalogType": "aws:redshift"
      }
    }
  }
}
```

```

    }
  }
}

```

Glue get-catalog response

```

aws glue get-catalog \
  --catalog-id account-id:catalog-name \
  --region us-east-1

Response:
{
  "Catalog": {
    "Name": "CatalogName",
    "Description": "Glue Catalog for Redshift z-etl use case",
    "CreateDatabaseDefaultPermissions" : [],
    "CreateTableDefaultPermissions": [],
    "CatalogProperties": {
      "DataLakeAccessProperties" : {
        "DataLakeAccess": "true",
        "DataTransferRole": "DTR arn",
        "KMSKey": "kms key arn",
        "ManagedWorkgroupName": "MWG name",
        "ManagedWorkgroupStatus": "MWG status",
        "RedshiftDatabaseName": "RS db name",
        "NamespaceArn": "namespace key arn",
        "CatalogType": "aws:redshift"
      }
    }
  }
}

```

Gestione delle autorizzazioni per i dati in un datashare Amazon Redshift

Con AWS Lake Formation, puoi gestire i dati in modo sicuro in un datashare di Amazon Redshift. Amazon Redshift è un servizio di data warehouse completamente gestito su scala di petabyte nel cloud. AWS Utilizzando la funzionalità di condivisione dei dati, Amazon Redshift ti aiuta a condividere

i dati tra di loro. Account AWS Per ulteriori informazioni sulla condivisione dei dati di Amazon Redshift, consulta [Panoramica sulla condivisione dei dati in Amazon Redshift](#).

In Amazon Redshift, l'amministratore del cluster di produttori crea un datashare e lo condivide con l'amministratore del data lake. Per step-by-step istruzioni sulla creazione di un amministratore di data lake, consulta. [Crea un amministratore del data lake](#)

Dopo che l'amministratore del data lake ha accettato il datashare, è necessario creare un AWS Glue Data Catalog database per il datashare specifico. In questo modo puoi controllarne l'accesso utilizzando le autorizzazioni di Lake Formation. Lake Formation mappa ogni datashare su un database Data Catalog corrispondente. Questi vengono visualizzati come database federati nel Data Catalog.

Un database viene definito database federato quando punta a un'entità esterna al Data Catalog. Le tabelle e le viste nel datashare Amazon Redshift sono elencate come singole tabelle nel Data Catalog. Puoi condividere il database federato con i principali IAM e gli utenti SAML selezionati all'interno dello stesso account o in un altro account con Lake Formation. Puoi anche includere espressioni di filtro per righe e colonne per limitare l'accesso a determinati dati. Per ulteriori informazioni, consulta [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#).

Per fornire agli utenti l'accesso a un datashare Amazon Redshift, devi fare quanto segue:

1. Aggiorna le impostazioni del Data Catalog per abilitare le autorizzazioni di Lake Formation.
2. Accetta l'invito al datashare dall'amministratore del cluster di produttori di Amazon Redshift e registra il datashare in Lake Formation.

Dopo aver completato questo passaggio, puoi gestire il datashare all'interno del Lake Formation Data Catalog.

3. Crea un database federato e definisci le autorizzazioni su quel database.
4. Concedi le autorizzazioni agli utenti su database e tabelle. Puoi condividere l'intero database o un sottoinsieme di tabelle con gli utenti dello stesso account o di un altro account.

Per le limitazioni, consulta [Limitazioni della condivisione dei dati di Amazon Redshift](#).

Argomenti

- [Prerequisiti per la configurazione delle autorizzazioni sulle condivisioni di dati Amazon Redshift](#)
- [Configurazione delle autorizzazioni per le condivisioni di dati Amazon Redshift](#)
- [Interrogazione di database federati](#)

Prerequisiti per la configurazione delle autorizzazioni sulle condivisioni di dati Amazon Redshift

Aggiorna le impostazioni predefinite del Data Catalog

Per abilitare le autorizzazioni Lake Formation per le risorse Data Catalog, ti consigliamo di disabilitare le impostazioni predefinite del Data Catalog in Lake Formation. Per ulteriori informazioni, consulta [Modifica il modello di autorizzazione predefinito o utilizza la modalità di accesso ibrida](#).

Aggiorna le autorizzazioni

Oltre alle autorizzazioni di amministratore del data lake (AWSLakeFormationDataAdmin), sono necessarie anche le seguenti autorizzazioni per accettare un datashare Amazon Redshift in Lake Formation:

- `glue:PassConnection on aws:redshift`
- `redshift:AssociateDataShareConsumer`
- `redshift:DescribeDataSharesForConsumer`
- `redshift:DescribeDataShares`

L'utente IAM dell'amministratore del data lake dispone implicitamente delle seguenti autorizzazioni.

- `data_location_access`
- `create_database`
- Lake Formation: registra la risorsa

Configurazione delle autorizzazioni per le condivisioni di dati Amazon Redshift

Questo argomento descrive i passaggi da seguire per accettare un invito alla condivisione di dati, creare un database federato e concedere le autorizzazioni. Puoi usare la console Lake Formation o il AWS Command Line Interface (AWS CLI). Gli esempi in questo argomento mostrano il cluster di produttori, il Data Catalog e il consumatore di dati nello stesso account.

Per ulteriori informazioni sulle funzionalità cross-account di Lake Formation, consulta [Condivisione dei dati tra account in Lake Formation](#).

Per configurare le autorizzazioni per un datashare

1. Rivedi un invito al datashare e accettalo.

Console

1. Accedi alla console di Lake Formation come amministratore del data lake all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Vai alla pagina di condivisione dei dati.
2. Controlla le condivisioni di dati a cui sei autorizzato ad accedere. La colonna Stato indica il tuo attuale stato di partecipazione al datashare. Lo stato In sospeso indica che sei stato aggiunto a un datashare, ma non l'hai ancora accettato o hai rifiutato l'invito.
3. Per rispondere a un invito al datashare, seleziona il nome del datashare e scegli Rivedi l'invito. In Accetta o rifiuta il datashare, esamina i dettagli dell'invito. Scegli Accetta per accettare l'invito o Rifiuta per rifiutarlo. Non avrai accesso al datashare se rifiuti l'invito.

AWS CLI

Gli esempi seguenti mostrano come visualizzare, accettare e registrare l'invito. Sostituisci l'Account AWS ID con un Account AWS ID valido. Sostituisci data-share-arn con l'effettivo Amazon Resource Name (ARN) che fa riferimento al datashare.

1. Visualizza un invito in sospeso.

```
aws redshift describe-data-shares \
  --data-share-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds' \
```

2. Accetta un datashare.

```
aws redshift associate-data-share-consumer \
  --data-share-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds' \
  --consumer-arn 'arn:aws:glue:us-east-1:111122223333:catalog
```

3. Registra il datashare nell'account Lake Formation. Utilizza l'operazione [RegisterResourceAPI](#) per registrare il datashare in Lake Formation. DataShareArn è il parametro di input per ResourceArn

 Note

Si tratta di un passaggio obbligatorio.

```
aws lakeformation register-resource \  
  --resource-arn 'arn:aws:redshift:us-  
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/  
federatedds'
```

2. Crea un database.

Dopo aver accettato un invito alla condivisione di dati, devi creare un database che punti al database Amazon Redshift associato al datashare. Devi essere un amministratore di data lake per creare un database.

Console

1. Seleziona il datashare dal riquadro Inviti e scegli Imposta i dettagli del database.
2. In Imposta i dettagli del database, inserisci un nome e un identificatore univoci per il datashare. Questo identificatore viene utilizzato per mappare il datashare internamente nella gerarchia dei metadati (dbname.schema.table).
3. Scegli Avanti per concedere le autorizzazioni ad altri utenti sul database e sulle tabelle condivisi.

AWS CLI

Utilizza il codice di esempio seguente per creare un database che punti al database Amazon Redshift condiviso con Lake Formation utilizzando il. AWS CLI

```
aws glue create-database --cli-input-json \  
{  
  "CatalogId": "111122223333",  
  "DatabaseInput": {  
    "Name": "tahoedb",  
    "FederatedDatabase": {
```



```
"Identifier": "arn:aws:redshift:us-east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/federatedds",
  "ConnectionName": "aws:redshift"
}
}
```

3. Concedi i permessi.

Dopo aver creato il database, puoi concedere le autorizzazioni agli utenti del tuo account o a organizzazioni esterne Account AWS . Non potrai concedere autorizzazioni per i dati di scrittura (inserimento, eliminazione) e per i metadati (modifica, eliminazione, creazione) sul database federato mappato su un datashare Amazon Redshift. Per ulteriori informazioni sulla concessione delle autorizzazioni, consulta. [Gestione delle autorizzazioni di Lake Formation](#)

Note

In qualità di amministratore del data lake, puoi visualizzare solo le tabelle nei database federati. Per eseguire qualsiasi altra azione, è necessario concedere ulteriori autorizzazioni su tali tabelle.

Console

1. Nella schermata Concedi autorizzazioni, seleziona gli utenti a cui concedere le autorizzazioni.
2. Scegli Concessione.

AWS CLI

Utilizza gli esempi seguenti per concedere le autorizzazioni per database e tabelle utilizzando: AWS CLI

```
aws lakeformation grant-permissions --input-cli-json file://input.json

{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/non-admin"
  },
```

```
"Resource": {
  "Database": {
    "CatalogId": "111122223333",
    "Name": "tahoedb"
  }
},
"Permissions": [
  "DESCRIBE"
],
"PermissionsWithGrantOption": [
]
}
```

```
aws lakeformation grant-permissions --input-cli-json file://input.json
```

```
{
  "Principal": {
    "DataLakePrincipalIdentifier":
"arn:aws:iam::111122223333:user/non-admin"
  },
  "Resource": {
    "Table": {
      "CatalogId": "111122223333",
      "DatabaseName": "tahoedb",
      "Name": "public.customer"
    }
  },
  "Permissions": [
    "SELECT"
  ],
  "PermissionsWithGrantOption": [
    "SELECT"
  ]
}
```

Interrogazione di database federati

Dopo aver concesso le autorizzazioni, gli utenti possono accedere e iniziare a interrogare il database federato utilizzando Amazon Redshift. Gli utenti possono ora utilizzare il nome del database locale per fare riferimento al datashare Amazon Redshift nelle query SQL. In Amazon Redshift, la tabella clienti nello schema pubblico condiviso tramite il datashare avrà una tabella corrispondente creata come `public.customer` nel Data Catalog.

1. Prima di interrogare il database federato utilizzando Amazon Redshift, l'amministratore del cluster crea un database dal database Data Catalog utilizzando il seguente comando:

```
CREATE DATABASE sharedcustomerdb FROM ARN
'arn:aws:glue:<region>:111122223333:database/tahoedb' WITH DATA CATALOG SCHEMA
tahoedb
```

2. L'amministratore del cluster concede le autorizzazioni di utilizzo sul database.

```
GRANT USAGE ON DATABASE sharedcustomerdb TO IAM:user;
```

3. L'utente federato può ora accedere agli strumenti SQL per interrogare la tabella.

```
Select * from sharedcustomerdb.public.customer limit 10;
```

Per ulteriori informazioni, consulta la sezione [Querying AWS Glue Data Catalog](#) in Amazon Redshift Management Guide.

Gestione delle autorizzazioni sui set di dati che utilizzano metastore esterni

Con la federazione dei AWS Glue Data Catalog metadati (Data Catalog federation), puoi connettere il Data Catalog a metastore esterni che archiviano i metadati per i tuoi dati Amazon S3 e gestire in modo sicuro le autorizzazioni di accesso ai dati utilizzando. AWS Lake Formation Non è necessario migrare i metadati dal metastore esterno al Data Catalog.

Il Data Catalog fornisce un archivio centralizzato di metadati che semplifica la gestione e la scoperta dei dati su sistemi diversi. Quando la tua organizzazione gestisce i dati nel catalogo dati, puoi utilizzarli AWS Lake Formation per controllare l'accesso ai tuoi set di dati in Amazon S3.

 Note

Attualmente, supportiamo solo la federazione dei metastore di Apache Hive (versione 3 e successive).

[Per configurare la federazione di Data Catalog, forniamo un'applicazione AWS Serverless Application Model \(AWS SAM\) chiamata GlueDataCatalogFederation - in. HiveMetastore](#) AWS Serverless Application Repository

L'implementazione di riferimento è fornita GitHub come progetto open source presso [AWS Glue Data Catalog Federation - Hive Metastore](#).

L' AWS SAM applicazione crea e distribuisce le seguenti risorse necessarie per connettere il Data Catalog al metastore Hive:

- Una AWS Lambda funzione: ospita l'implementazione del servizio federativo che comunica tra il Data Catalog e il metastore Hive. AWS Glue richiama questa funzione Lambda per recuperare oggetti di metadati dal metastore Hive.
- Amazon API Gateway— L'endpoint di connessione per il metastore Hive che funge da proxy per indirizzare tutte le chiamate alla funzione Lambda.
- Un ruolo IAM: un ruolo con le autorizzazioni necessarie per creare la connessione tra il Data Catalog e il metastore Hive.
- AWS Glue connessione: un Amazon API Gateway tipo di AWS Glue connessione che memorizza l' Amazon API Gateway endpoint e un ruolo IAM per richiamarlo.

Quando esegui una query sulle tabelle, il AWS Glue servizio effettua una chiamata di runtime al metastore Hive e recupera i metadati. La funzione Lambda funge da traduttore tra il metastore Hive e Data Catalog.

Dopo aver stabilito la connessione, per sincronizzare i metadati nel metastore Hive con il Data Catalog, è necessario creare un database federato nel Data Catalog utilizzando i dettagli di connessione del metastore Hive e mappare questo database al database Hive. Un database viene definito database federato quando punta a un'entità esterna al Data Catalog.

Puoi applicare le autorizzazioni di Lake Formation utilizzando il controllo degli accessi basato su tag e il metodo della risorsa denominata sul database federato e condividerlo tra più Account AWS unità

organizzative (). AWS Organizations OUs Puoi anche condividere il database federato direttamente con i responsabili IAM di un altro account.

Puoi definire autorizzazioni dettagliate a livello di colonna, riga e cella utilizzando i filtri dati di Lake Formation sulle tabelle Hive esterne. Puoi usare Amazon Athena, Amazon Redshift o Amazon EMR per interrogare le tabelle Hive esterne gestite da Lake Formation.

Per ulteriori informazioni sulla condivisione e il filtraggio dei dati tra account, consulta:

- [Condivisione dei dati tra account in Lake Formation](#)
- [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#)

Fasi di alto livello per la federazione dei metadati di Data Catalog

1. Crei utenti e ruoli IAM con le autorizzazioni appropriate per distribuire l' AWS SAM applicazione e creare database federati.
2. Registri la posizione dei dati di Amazon S3 con Lake Formation selezionando l'Enable Data Catalog federationopzione per i set di dati che utilizzano un metastore Hive esterno.
3. È possibile configurare le impostazioni AWS SAM dell'applicazione (nome della AWS Glue connessione, URL del metastore Hive e parametri della funzione Lambda) e distribuire l'applicazione. AWS SAM
4. L' AWS SAM applicazione distribuisce le risorse necessarie per connettere il metastore Hive esterno al Data Catalog.
5. Per applicare le autorizzazioni di Lake Formation al database e alle tabelle Hive, crei un database nel Data Catalog utilizzando i dettagli di connessione del metastore Hive e mappi questo database al database Hive.
6. Concedi le autorizzazioni sui database federati ai responsabili del tuo account o di un altro account.

Note

Puoi connettere il Data Catalog a un mestastore Hive esterno, creare database federati ed eseguire query e script ETL su database e tabelle Hive senza applicare le autorizzazioni di Lake Formation. Per i dati di origine in Amazon S3 che non sono registrati con Lake Formation, l'accesso è determinato dalle politiche di autorizzazione IAM per Amazon S3 e dalle azioni. AWS Glue

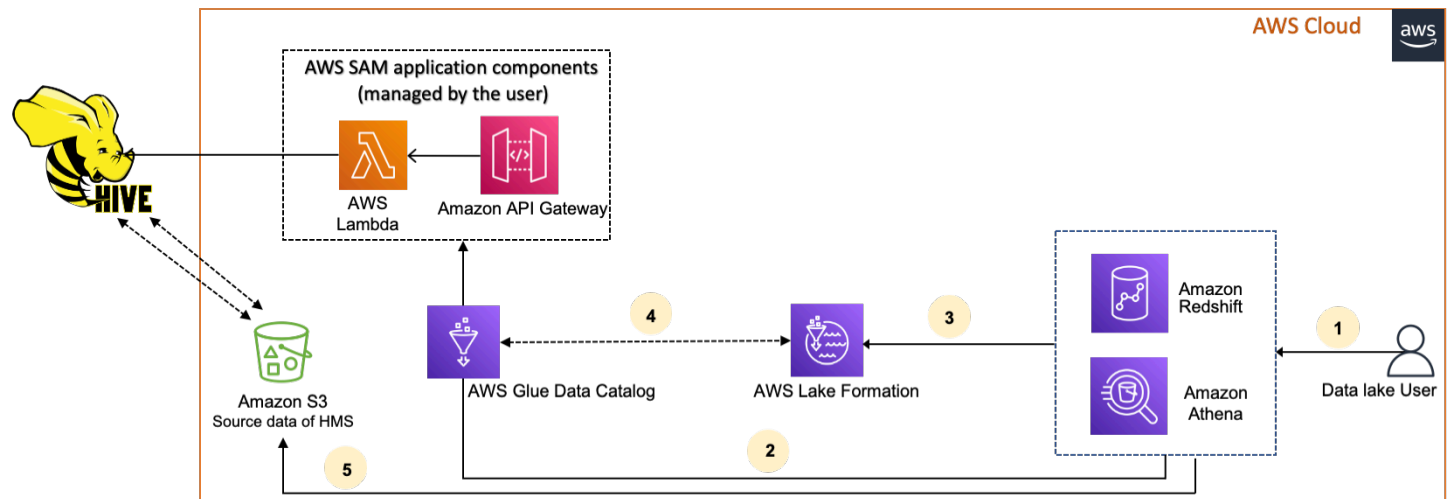
Per le limitazioni, consulta [I metadati di Hive archiviano: considerazioni e limitazioni sulla condivisione dei dati](#).

Argomenti

- [Flusso di lavoro](#)
- [Prerequisiti per connettere il Data Catalog al metastore Hive](#)
- [Connessione del Data Catalog a un metastore Hive esterno](#)
- [Risorse aggiuntive](#)

Flusso di lavoro

Il diagramma seguente mostra il flusso di lavoro per la connessione di un metastore AWS Glue Data Catalog Hive esterno.



1. Un principale invia una richiesta utilizzando un servizio integrato come Athena o Redshift Spectrum.
2. Il servizio integrato effettua una chiamata al Data Catalog per i metadati, che a sua volta richiama l'endpoint Hive metastore disponibile e riceve risposte alle Amazon API Gateway richieste di metadati.
3. Il servizio integrato invia la richiesta a Lake Formation per verificare le informazioni sulla tabella e le credenziali per accedere alla tabella.
4. Lake Formation autorizza la richiesta e invia credenziali temporanee all'applicazione integrata, che consente l'accesso ai dati.

5. Utilizzando le credenziali temporanee ricevute da Lake Formation, il servizio integrato legge i dati da Amazon S3 e condivide i risultati con il responsabile.

Prerequisiti per connettere il Data Catalog al metastore Hive

Per connetterlo AWS Glue Data Catalog a un metastore Apache Hive esterno e configurare le autorizzazioni di accesso ai dati, è necessario soddisfare i seguenti requisiti:

Note

Consigliamo che un amministratore di Lake Formation distribuisca l' AWS SAM applicazione e solo un utente privilegiato utilizzi la connessione metastore Hive per creare i database federati corrispondenti.

1. Creare ruoli IAM.

Per distribuire l'applicazione AWS SAM

- Crea un ruolo con le autorizzazioni necessarie per la distribuzione delle risorse (funzione Lambda Amazon API Gateway, ruolo IAM e AWS Glue connessione) necessarie per creare una connessione al metastore Hive.

Per creare database federati

Le seguenti autorizzazioni sono richieste per le risorse:

- `glue:CreateDatabase` on resource `arn:aws:glue:region:account-id:database/gluedatabasename`
- `glue:PassConnection` on resource `arn:aws:glue:region:account-id:connection/hms_connection`

2. Registra la sede Amazon S3 con Lake Formation.

Per utilizzare Lake Formation per gestire e proteggere i dati nel tuo data lake, devi registrare la posizione Amazon S3 che contiene i dati per le tabelle nel metastore Hive con Lake Formation. In questo modo, Lake Formation può fornire credenziali per servizi di AWS analisi come Athena, Redshift Spectrum e Amazon EMR.

Per ulteriori informazioni sulla registrazione di una sede Amazon S3, consulta. [Aggiungere una posizione Amazon S3 al tuo data lake](#)

Quando registri la posizione Amazon S3, seleziona la casella di controllo Enable Data Catalog Federation per consentire a Lake Formation di assumere un ruolo per accedere alle tabelle in un database federato.

[AWS Lake Formation](#) > [Data lake locations](#) > Register location

Register location

Amazon S3 location

Register an Amazon S3 path as the storage location for your data lake.

Amazon S3 path

Choose an Amazon S3 path for your data lake.

e.g.: s3://bucket/prefix/

Browse

Review location permissions - strongly recommended

Registering the selected location may result in your users gaining access to data already at that location. Before registering a location, we recommend that you review existing location permissions on resources in that location.

Review location permissions

IAM role

To add or update data, Lake Formation needs read/write access to the chosen Amazon S3 path. Choose a role that you know has permission to do this, or choose the **AWSServiceRoleForLakeFormationDataAccess** service-linked role. When you register the first Amazon S3 path, the service-linked role and a new inline policy are created on your behalf. Lake Formation adds the first path to the inline policy and attaches it to the service-linked role. When you register subsequent paths, Lake Formation adds the path to the existing policy.

AWSServiceRoleForLakeFormationDataAccess ▼

 Do not select the service linked role if you plan to use EMR.

☐ **Enable Data Catalog Federation**

Checking this box will allow Lake Formation to assume a role to access tables in a federated database.

Cancel

Register location

Per ulteriori informazioni sulla registrazione di una posizione dati con Lake Formation, vedere [Configura una posizione Amazon S3 per il tuo data lake](#).

3. Usa la versione corretta di Amazon EMR.

Per utilizzare Amazon EMR con i database metastore Hive federati, devi disporre della versione 3.x o successiva di Hive e della versione 6.x o successiva di Amazon EMR.

Connessione del Data Catalog a un metastore Hive esterno

[Per connetterlo AWS Glue Data Catalog a un metastore Hive, devi distribuire un'applicazione chiamata - AWS SAM GlueDataCatalogFederation HiveMetastore](#) Crea le risorse necessarie per connettere il metastore Hive esterno al Data Catalog. È possibile accedere all' AWS SAM applicazione in. AWS Serverless Application Repository

L' AWS SAM applicazione crea la connessione per il metastore Hive dietro Amazon API Gateway utilizzando una funzione Lambda. L' AWS SAM applicazione utilizza un URI (Uniform Resource Identifier) come input dell'utente e collega il metastore Hive esterno al Data Catalog. Quando un utente esegue una query sulle tabelle Hive, il Data Catalog chiama l'endpoint API Gateway. L'endpoint richiama la funzione Lambda per recuperare i metadati delle tabelle Hive.

Per connettere il Data Catalog al metastore Hive e configurare le autorizzazioni

1. Distribuisci l'applicazione. AWS SAM

1. Accedi a Console di gestione AWS e apri il AWS Serverless Application Repository.
2. Nel pannello di navigazione, scegli Available applications (Applicazioni disponibili).
3. Scegli Applicazioni pubbliche.
4. Seleziona l'opzione Visualizzare le app che creano ruoli IAM personalizzati o policy delle risorse.
5. Nella casella di ricerca, inserisci il nome GlueDataCatalogFederation- HiveMetastore.
6. Scegli l'HiveMetastoreapplicazione GlueDataCatalogFederation-.
7. In Impostazioni dell'applicazione, inserisci le seguenti impostazioni minime richieste per la tua funzione Lambda:
 - Nome dell'applicazione: un nome per l' AWS SAM applicazione.
 - GlueConnectionName- Un nome per la connessione.
 - HiveMetastoreURIs- L'URI del tuo host metastore Hive.
 - LambdaMemory- La quantità di memoria Lambda in MB è compresa tra 128 e 10240. Il valore di default è 1024.

- LambdaTimeout- La durata massima di invocazione Lambda in secondi. Il valore predefinito è 30.
 - VPCSecurityGroupIds VPCSubnetIds: informazioni per il VPC in cui esiste il metastore Hive.
8. Seleziona I acknowledge that this app creates custom IAM roles and resource policies (Sono consapevole che questa app crea ruoli IAM personalizzati e policy della risorsa). Per ulteriori informazioni, scegliere il link Info (Informazioni) .
 9. Nella parte inferiore destra della pagina Application settings (Impostazioni dell'applicazione), scegli Deploy (Implementa). Al termine dell'implementazione, la funzione Lambda viene visualizzata nella sezione Resources (Risorse) nella console Lambda.

L'applicazione viene distribuita su Lambda. Il suo nome è preceduto da serverlessrepo- per indicare che l'applicazione è stata distribuita da AWS Serverless Application Repository. Selezionando l'applicazione si accede alla pagina Risorse in cui sono elencate tutte le risorse dell'applicazione che sono state distribuite. Le risorse includono la funzione Lambda che consente la comunicazione tra il Data Catalog e il metastore Hive, la AWS Glue connessione e altre risorse necessarie per la federazione del database.

2. Crea un database federato nel Data Catalog.

Dopo aver creato una connessione al metastore Hive, puoi creare database federati nel Data Catalog che puntano ai database metastore Hive esterni. Devi creare un database corrispondente nel Data Catalog per ogni database di metastore Hive che stai connettendo al Data Catalog.

Lake Formation console

1. Nella pagina Condivisione dati, scegli la scheda Database condivisi, quindi scegli Crea database.
2. Per Nome della connessione, scegli il nome della tua connessione al metastore Hive dal menu a discesa.
3. Inserisci un nome di database univoco e l'identificatore di origine della federazione per il database. Questo è il nome che usi nelle istruzioni SQL quando esegui query sulle tabelle. Il nome può essere composto da un massimo di 255 caratteri e deve essere univoco all'interno del tuo account.
4. Scegliere Crea database.

AWS CLI

```
aws glue create-database \  
'{  
  "CatalogId": "<111122223333>",  
  "database-input": {  
    "Name": "<fed_glue_db>",  
    "FederatedDatabase": {  
      "Identifier": "<hive_db_on_emr>",  
      "ConnectionName": "<hms_connection>"  
    }  
  }  
'
```

3. Visualizza le tabelle nel database federato.

Dopo aver creato il database federato, puoi visualizzare l'elenco delle tabelle nel tuo metastore Hive utilizzando la console Lake Formation o il AWS CLI

Lake Formation console

1. Seleziona il nome del database dalla scheda Database condivisi.
2. Nella pagina Database, scegli Visualizza tabelle.

AWS CLI

Gli esempi seguenti mostrano come recuperare la definizione di connessione, il nome del database e alcune o tutte le tabelle del database. Sostituisci l'ID del catalogo dati con l'Account AWS ID valido utilizzato per creare il database. Sostituisci `hms_connection` con il nome della connessione.

```
aws glue get-connection \  
--name <hms_connection> \  
--catalog-id 111122223333
```

```
aws glue get-database \  
--name <fed_glu_db> \  
--catalog-id 111122223333
```

```
--catalog-id 111122223333
```

```
aws glue get-tables \  
--database-name <fed_glue_db> \  
--catalog-id 111122223333
```

```
aws glue get-table \  
--database-name <fed_glue_db> \  
--name <hive_table_name> \  
--catalog-id 111122223333
```

4. Concedi i permessi.

Dopo aver creato il database, puoi concedere le autorizzazioni ad altri utenti e ruoli IAM nel tuo account o a organizzazioni esterne Account AWS . Non sarai in grado di concedere autorizzazioni per i dati di scrittura (inserimento, eliminazione) e autorizzazioni per i metadati (modifica, eliminazione, creazione) sui database federati. Per ulteriori informazioni sulla concessione delle autorizzazioni, vedere. [Gestione delle autorizzazioni di Lake Formation](#)

5. Interroga i database federati.

Dopo aver concesso le autorizzazioni, gli utenti possono accedere e iniziare a interrogare il database federato utilizzando Athena e Amazon Redshift. Gli utenti possono ora utilizzare il nome del database locale per fare riferimento al database Hive nelle query SQL.

Sintassi di query di esempio Amazon Athena

Sostituisci `fed_glue_db` con il nome del database locale creato in precedenza.

```
Select * from fed_glue_db.customers limit 10;
```

Risorse aggiuntive

Il seguente post sul blog contiene istruzioni dettagliate per configurare i permessi di Lake Formation su un database e tabelle di metastore Hive e interrogarli utilizzando Athena. Illustriamo anche un caso d'uso della condivisione tra account, in cui un responsabile di Lake Formation nell'account produttore A condivide un database Hive federato e tabelle che utilizzano il tag LF con l'account consumatore B.

- [Interroga il tuo metastore Apache Hive con le autorizzazioni AWS Lake Formation](#)

Gestione delle autorizzazioni di Lake Formation

Lake Formation fornisce controlli di accesso centralizzati per i dati nel tuo data lake. Puoi definire regole basate su policy di sicurezza per i tuoi utenti e le tue applicazioni in base al ruolo in Lake Formation e l'integrazione con AWS Identity and Access Management autentica tali utenti e ruoli. Una volta definite le regole, Lake Formation applica i controlli di accesso con granularità a livello di tabella, colonna e ro per gli utenti di Amazon Redshift Spectrum e Amazon Athena.

Argomenti

- [Concessione delle autorizzazioni per la localizzazione dei dati](#)
- [Concessione delle autorizzazioni per le risorse di Data Catalog](#)
- [Scenario di esempio di autorizzazioni](#)
- [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#)
- [Visualizzazione delle autorizzazioni per database e tabelle in Lake Formation](#)
- [Revoca dell'autorizzazione utilizzando la console Lake Formation](#)
- [Condivisione dei dati tra account in Lake Formation](#)
- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)
- [Creazione di collegamenti alle risorse](#)
- [Accesso alle tabelle in tutte le regioni](#)

Concessione delle autorizzazioni per la localizzazione dei dati

Le autorizzazioni di localizzazione dei dati AWS Lake Formation consentono ai responsabili di creare e modificare risorse del catalogo dati che puntano a sedi Amazon S3 registrate designate. Le autorizzazioni di localizzazione dei dati funzionano in aggiunta alle autorizzazioni per i dati di Lake Formation per proteggere le informazioni nel tuo data lake.

Lake Formation non utilizza il servizio AWS Resource Access Manager (AWS RAM) per le concessioni di autorizzazioni alla localizzazione dei dati, quindi non è necessario accettare inviti alla condivisione delle risorse per le autorizzazioni di localizzazione dei dati.

Puoi concedere le autorizzazioni per la localizzazione dei dati utilizzando la console, l'API o AWS Command Line Interface (AWS CLI) di Lake Formation.

 Note

Affinché una sovvenzione abbia successo, devi prima registrare la posizione dei dati con Lake Formation.

 Consulta anche:

- [Underlying data access control](#)

Argomenti

- [Concessione delle autorizzazioni per la localizzazione dei dati \(stesso account\)](#)
- [Concessione delle autorizzazioni per la localizzazione dei dati \(account esterno\)](#)
- [Concessione delle autorizzazioni su una posizione di dati condivisa con l'account](#)

Concessione delle autorizzazioni per la localizzazione dei dati (stesso account)

Segui questi passaggi per concedere le autorizzazioni per la localizzazione dei dati ai responsabili del tuo account. AWS Puoi concedere le autorizzazioni utilizzando la console Lake Formation, l'API o AWS Command Line Interface (AWS CLI).

Console di gestione AWS

Per concedere le autorizzazioni per la localizzazione dei dati (stesso account)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come responsabile che ha concesso le autorizzazioni per la posizione dei dati desiderata.
2. Nel riquadro di navigazione, in Autorizzazioni, scegli Posizioni dei dati.
3. Scegli Concessione.
4. Nella finestra di dialogo Concedi autorizzazioni, assicurati che il riquadro Il mio account sia selezionato. Fornisci quindi le seguenti informazioni:
 - Per gli utenti e i ruoli IAM, scegli uno o più principali.

- Per utenti e gruppi SAML e Amazon Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite.

Immettere un ARN alla volta e premere Invio dopo ogni ARN. Per informazioni su come costruire il ARNs, vedere. [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

- Per le posizioni di archiviazione, scegli Browse e scegli una posizione di storage Amazon Simple Storage Service (Amazon S3). La sede deve essere registrata presso Lake Formation. Scegli nuovamente Sfoglia per aggiungere un'altra località. Puoi anche digitare la posizione, ma assicurati di farla precedere da `s3://`.
- Per Ubicazione dell'account registrato, inserisci l'ID AWS dell'account in cui è registrata la sede. L'impostazione predefinita è l'ID del tuo account. In uno scenario con più account, gli amministratori del data lake di un account destinatario possono specificare qui l'account del proprietario quando concedono l'autorizzazione alla localizzazione dei dati ad altri responsabili dell'account del destinatario.
- (Facoltativo) Per consentire ai principali selezionati di concedere le autorizzazioni per la localizzazione dei dati nella posizione selezionata, seleziona Grantable.

Grant permissions ✕

Add access permissions for specific storage locations.

☒ **My account**
User or role from this AWS account.

☐ **External account**
AWS account or AWS organization outside of my account.

IAM users and roles
Add one or more IAM users or roles.

Choose IAM principals to add

datalake_user ✕
User

SAML and Amazon QuickSight users and groups
Enter a SAML user or group ARN or Amazon QuickSight ARN. Press Enter to add additional ARNs.

Ex: `arn:aws:iam::<AccountId>:saml-provider/<SamlProviderName>`

Storage locations
Choose one or more data lake locations.

s3://retail/transactions/2020q1 Browse

Registered account location
The account where this storage location is registered in AWS Lake Formation.

123456789012

☐ Grantable

Cancel Grant

5. Scegli Concessione.

AWS CLI

Per concedere le autorizzazioni per la localizzazione dei dati (stesso account)

- Esegui un `grant-permissions` comando e `DATA_LOCATION_ACCESS` concedi al principale, specificando il percorso Amazon S3 come risorsa.

Example

L'esempio seguente concede all'utente le autorizzazioni di localizzazione dei dati. `s3://retail datalake_user1`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1
```



```
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":  
{"ResourceArn":"arn:aws:s3:::retail"}}'
```

Example

L'esempio seguente concede le autorizzazioni per la localizzazione dei dati a un gruppo.
`s3://retail ALLIAMPrincipals`

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=111122223333:IAMPrincipals --  
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":  
{"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "111122223333"}}'
```

 Consulta anche:

- [Riferimento alle autorizzazioni di Lake Formation](#)

Concessione delle autorizzazioni per la localizzazione dei dati (account esterno)

Segui questi passaggi per concedere le autorizzazioni per la localizzazione dei dati a un AWS account o a un'organizzazione esterni.

Puoi concedere le autorizzazioni utilizzando la console Lake Formation, l'API o AWS Command Line Interface (AWS CLI).

Prima di iniziare

Assicurati che tutti i prerequisiti di accesso tra account siano soddisfatti. Per ulteriori informazioni, consulta [Prerequisiti](#).

Console di gestione AWS

Per concedere le autorizzazioni per la localizzazione dei dati (account esterno, console)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake.

2. Nel riquadro di navigazione, in Autorizzazioni, scegli Posizioni dei dati, quindi scegli Concedi.
3. Nella finestra di dialogo Concedi autorizzazioni, scegli il riquadro Account esterno.
4. Inserisci le informazioni che seguono:
 - Per ID AWS account o ID AWS organizzazione, inserisci numeri di AWS conto IDs, organizzazione o unità IDs organizzativa validi.

Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa è composto da «ou-» seguito da 4 a 32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo «-» (trattino) e da 8 a 32 lettere o cifre minuscole aggiuntive.

- In Luoghi di archiviazione, scegli Browse e scegli una posizione di storage Amazon Simple Storage Service (Amazon S3). La sede deve essere registrata presso Lake Formation.

Grant permissions X

Add access permissions for specific storage locations.

☐ My account
User or role from this AWS account.

☒ External account
AWS account or AWS organization outside of my account.

AWS account ID or AWS organization ID

Q Enter AWS account ID or AWS organization ID

111122223333 X
Account

Enter one or more AWS account IDs or AWS organization IDs. Press Enter after each ID.

Storage locations
Choose one or more data lake locations.

s3://retail/transactions/2020q1 Browse

☒ Grantable

Cancel Grant

5. Seleziona Grantable.
6. Scegli Concessione.

AWS CLI

Per concedere le autorizzazioni per la localizzazione dei dati (account esterno,) AWS CLI

- Per concedere le autorizzazioni a un AWS account esterno, inserisci un comando simile al seguente.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions
"DATA_LOCATION_ACCESS" --permissions-with-grant-option
"DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"CatalogId":"123456789012", "ResourceArn":"arn:aws:s3::retail/
transactions/2020q1"}}'
```

Questo comando concede l'opzione di concessione DATA_LOCATION_ACCESS all'account 1111-2222-3333 nella posizione s3://retail/transactions/2020q1 Amazon S3, che è di proprietà dell'account 1234-5678-9012.

Per concedere le autorizzazioni a un'organizzazione, inserisci un comando simile al seguente.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/
o-abcdefghijkl --permissions "DATA_LOCATION_ACCESS" --permissions-
with-grant-option "DATA_LOCATION_ACCESS" --resource '{"DataLocation":
{"CatalogId":"123456789012", "ResourceArn":"arn:aws:s3::retail/
transactions/2020q1"}}'
```

Questo comando concede un'opzione DATA_LOCATION_ACCESS di concessione all'organizzazione o-abcdefghijkl nella s3://retail/transactions/2020q1 posizione Amazon S3, di proprietà dell'account 1234-5678-9012.

Per concedere le autorizzazioni a un principale in un AWS account esterno, inserisci un comando simile al seguente.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3::retail/transactions/2020q1", "CatalogId":
"123456789012"}}'
```

Questo comando concede DATA_LOCATION_ACCESS a un principale nell'account 1111-2222-3333 nella posizione s3://retail/transactions/2020q1 Amazon S3, che è di proprietà dell'account 1234-5678-9012.

Example

L'esempio seguente concede le autorizzazioni di localizzazione dei dati per il raggruppamento in un account esterno. s3://retail ALLIAMPrincipals

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "123456789012"} }'
```

 Consulta anche:

- [Riferimento alle autorizzazioni di Lake Formation](#)

Concessione delle autorizzazioni su una posizione di dati condivisa con l'account

Dopo aver condiviso una risorsa Data Catalog con il tuo AWS account, in qualità di amministratore del data lake, puoi concedere le autorizzazioni sulla risorsa ad altri responsabili del tuo account. Se l'ALTERautorizzazione viene concessa su una tabella condivisa e la tabella punta a una posizione Amazon S3 registrata, devi anche concedere le autorizzazioni per la posizione dei dati sulla posizione. Allo stesso modo, se l'ALTERautorizzazione CREATE_TABLE o è concessa a un database condiviso e il database ha una proprietà di location che punta a una posizione registrata, devi anche concedere le autorizzazioni per la localizzazione dei dati sulla posizione.

Per concedere le autorizzazioni di localizzazione dei dati su una posizione condivisa a un responsabile del tuo account, al tuo account deve essere stata concessa l'DATA_LOCATION_ACCESSautorizzazione sulla posizione con l'opzione di concessione. Quando poi concedi DATA_LOCATION_ACCESS a un altro titolare del tuo account, devi includere l'ID Data Catalog (ID dell'AWS account) dell'account proprietario. L'account del proprietario è l'account che ha registrato la sede.

Puoi utilizzare la AWS Lake Formation console, l'API o AWS Command Line Interface (AWS CLI) per concedere le autorizzazioni per la localizzazione dei dati.

Per concedere le autorizzazioni su una posizione dati condivisa con il tuo account (console)

- Segui la procedura riportata in [Concessione delle autorizzazioni per la localizzazione dei dati \(stesso account\)](#).

Per le posizioni di archiviazione, è necessario digitare le posizioni. Per Ubicazione dell'account registrato, inserisci l' AWS ID dell'account del proprietario.

Per concedere autorizzazioni su una posizione dati condivisa con il tuo account (AWS CLI)

- Inserisci uno dei seguenti comandi per concedere le autorizzazioni a un utente o a un ruolo.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"CatalogId":"<owner-account-ID>", "ResourceArn":"arn:aws:s3:::<s3-location>"} }'
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
--permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
{"CatalogId":"<owner-account-ID>", "ResourceArn":"arn:aws:s3:::<s3-location>"} }'
```

Concessione delle autorizzazioni per le risorse di Data Catalog

Puoi concedere le autorizzazioni relative ai dati ai responsabili in AWS Lake Formation modo che i responsabili possano creare e gestire le risorse del Data Catalog e possano accedere ai dati sottostanti. È possibile concedere autorizzazioni Data Lake su cataloghi, database, tabelle e viste. Quando concedi le autorizzazioni per le tabelle, puoi limitare l'accesso a colonne o righe specifiche della tabella per un controllo degli accessi ancora più preciso.

È possibile concedere autorizzazioni su singoli cataloghi, database, tabelle e viste oppure con un'unica operazione di concessione, è possibile concedere autorizzazioni su tutti i database, le tabelle e le viste di un catalogo o di un database. Se concedi le autorizzazioni su tutte le tabelle di un database ai responsabili IAM, concedi implicitamente l'autorizzazione sul database. DESCRIBE Il database viene quindi visualizzato nella pagina Database della console e viene restituito dall'operazione API. GetDatabases Lo stesso principio si applica a livello di catalogo: quando si

ricevono le autorizzazioni per i database all'interno di un catalogo, si ottengono anche DESCRIBE le autorizzazioni per quel catalogo.

Important

L'DESCRIBE autorizzazione implicita si applica solo quando si concedono le autorizzazioni ai responsabili IAM all'interno dello stesso account. AWS Per le risorse tra più account, devi concedere esplicitamente le autorizzazioni. DESCRIBE La concessione automatica delle DESCRIBE autorizzazioni non si applica quando si utilizza il controllo degli accessi basato sugli attributi (ABAC). Quando si concedono le autorizzazioni su tutte le tabelle di un database utilizzando gli attributi, Lake Formation non concede implicitamente DESCRIBE l'autorizzazione al database.

È possibile concedere le autorizzazioni utilizzando il metodo named resource o il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC).

È possibile concedere le autorizzazioni ai responsabili dello stesso account o a organizzazioni esterne. Account AWS Quando concedi ad account o organizzazioni esterni, condividi oggetti Data Catalog di tua proprietà con tali account o organizzazioni. I responsabili di tali account o organizzazioni possono quindi accedere agli oggetti di Data Catalog di cui sei proprietario e ai dati sottostanti.

Note

Attualmente, il metodo LF-TBAC supporta la concessione di autorizzazioni su più account a dirigenti, Account AWS organizzazioni e unità organizzative IAM (). OUs

Quando concedi autorizzazioni ad account o organizzazioni esterni, devi includere l'opzione di concessione. Solo l'amministratore del data lake dell'account esterno può accedere agli oggetti condivisi finché l'amministratore non concede le autorizzazioni sugli oggetti condivisi ad altri responsabili dell'account esterno.

Puoi concedere le autorizzazioni di Data Catalog utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

 Note

Quando si elimina un oggetto Data Catalog, tutte le autorizzazioni associate all'oggetto diventano non valide. Ricreando la stessa risorsa con lo stesso nome, non verranno recuperate le autorizzazioni di Lake Formation. Gli utenti dovranno configurare nuovamente le nuove autorizzazioni.

 Consulta anche:

- [Condivisione di tabelle e database del Data Catalog tra più account AWS](#)
- [Controllo dell'accesso ai metadati](#)
- [Riferimento alle autorizzazioni di Lake Formation](#)

Autorizzazioni IAM necessarie per concedere o revocare le autorizzazioni di Lake Formation

Tutti i responsabili, incluso l'amministratore del data lake, necessitano delle seguenti autorizzazioni AWS Identity and Access Management (IAM) per concedere o revocare le autorizzazioni AWS Lake Formation Data Catalog o le autorizzazioni di localizzazione dei dati con l'API Lake Formation o il: AWS CLI

- `lakeformation:GrantPermissions`
- `lakeformation:BatchGrantPermissions`
- `lakeformation:RevokePermissions`
- `lakeformation:BatchRevokePermissions`
- `glue:GetTable`, `glue:GetDatabase`, o `glue:GetCatalog` per una tabella, un database o un catalogo a cui stai concedendo le autorizzazioni utilizzando il metodo di risorsa denominato.

 Note

Gli amministratori di Data Lake dispongono delle autorizzazioni implicite di Lake Formation per concedere e revocare le autorizzazioni di Lake Formation. Ma hanno comunque bisogno delle autorizzazioni IAM sulle operazioni API di concessione e revoca di Lake Formation.

I ruoli IAM con policy `AWSLakeFormationDataAdmin` AWS gestita non possono aggiungere nuovi amministratori di data lake perché questa policy contiene un rifiuto esplicito per il funzionamento dell'API Lake Formation, `PutDataLakeSetting`.

La seguente policy IAM è consigliata ai responsabili che non sono amministratori di data lake e che desiderano concedere o revocare le autorizzazioni utilizzando la console Lake Formation.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:ListPermissions",
        "lakeformation:GrantPermissions",
        "lakeformation:BatchGrantPermissions",
        "lakeformation:RevokePermissions",
        "lakeformation:BatchRevokePermissions",
        "glue:GetCatalogs",
        "glue:GetDatabases",
        "glue:SearchTables",
        "glue:GetTables",
        "glue:GetCatalog",
        "glue:GetDatabase",
        "glue:GetTable",
        "iam:ListUsers",
        "iam:ListRoles",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeGroup",
        "sso:DescribeInstance"
      ],
      "Resource": "*"
    }
  ]
}
```


Tutte le autorizzazioni `glue:` e le `iam:` autorizzazioni di questa policy sono disponibili nella policy gestita. `AWS AWSTheGlueConsoleFullAccess`

Per concedere le autorizzazioni utilizzando il controllo degli accessi basato su tag Lake Formation (LF-TBAC), i principali necessitano di autorizzazioni IAM aggiuntive. Per ulteriori informazioni, consultare [Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation](#) e [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

Autorizzazioni tra account

Gli utenti che desiderano concedere autorizzazioni Lake Formation su più account utilizzando il metodo della risorsa denominata devono disporre anche delle autorizzazioni nella `AWSLakeFormationCrossAccountManager` AWS politica gestita.

Gli amministratori di Data Lake necessitano delle stesse autorizzazioni per concedere autorizzazioni su più account, oltre all'autorizzazione AWS Resource Access Manager (AWS RAM) per consentire la concessione di autorizzazioni alle organizzazioni. Per ulteriori informazioni, consulta [Autorizzazioni di amministratore di Data Lake](#).

L'utente amministrativo

Un preside con autorizzazioni amministrative, ad esempio con la policy `AdministratorAccess` AWS gestita, dispone delle autorizzazioni per concedere le autorizzazioni di Lake Formation e creare amministratori di data lake. Per negare a un utente o a un ruolo l'accesso alle operazioni di amministratore di Lake Formation, allega o aggiungi alla sua policy una Deny dichiarazione per le operazioni dell'API dell'amministratore.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "lakeformation:GetDataLakeSettings",
        "lakeformation:PutDataLakeSettings"
      ],
      "Effect": "Deny",
      "Resource": [
        "*"
      ]
    }
  ]
}
```

```
}  
  }  
}  
}
```

Important

Per impedire agli utenti di aggiungersi come amministratori con uno script di estrazione, trasformazione e caricamento (ETL), assicurati che a tutti gli utenti e i ruoli non amministratori sia negato l'accesso a queste operazioni API. La policy `AWSLakeFormationDataAdmin` AWS gestita contiene una negazione esplicita del funzionamento dell'API `Lake Formation`, `PutDataLakeSetting` che impedisce agli utenti di aggiungere nuovi amministratori di data lake.

Concessione delle autorizzazioni per i dati utilizzando il metodo di risorsa denominato

Il metodo di risorse denominato `Data Catalog` è un modo per concedere autorizzazioni a `AWS Glue Data Catalog` oggetti, come cataloghi, database, tabelle, colonne e viste, utilizzando un approccio centralizzato. Consente di definire policy basate sulle risorse che controllano l'accesso a risorse specifiche all'interno del data lake.

Quando si utilizza il metodo della risorsa denominata per concedere le autorizzazioni, è possibile specificare il tipo di risorsa e le autorizzazioni che si desidera concedere o revocare per quella risorsa. È inoltre possibile revocare l'autorizzazione in un secondo momento, se necessario, rimuovendo così le autorizzazioni dalle risorse associate.

È possibile concedere le autorizzazioni utilizzando la `AWS Lake Formation console` o `()`. APIs `AWS Command Line Interface AWS CLI`

Argomenti

- [Concessione delle autorizzazioni di catalogo utilizzando il metodo di risorsa denominato](#)
- [Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#)
- [Concessione delle autorizzazioni per le tabelle utilizzando il metodo di risorsa denominato](#)
- [Concessione delle autorizzazioni sulle viste utilizzando il metodo di risorsa denominato](#)

Concessione delle autorizzazioni di catalogo utilizzando il metodo di risorsa denominato

I passaggi seguenti spiegano come concedere le autorizzazioni di catalogo utilizzando il metodo della risorsa denominata.

Console

Usa la pagina Concedi le autorizzazioni sulla console Lake Formation. La pagina è suddivisa nelle seguenti sezioni:

- **Tipo principale:** è possibile concedere autorizzazioni a destinatari specifici o utilizzare tag di attributo.
- **Principali:** utenti, ruoli, utenti e gruppi IAM Identity Center, utenti e gruppi SAML, AWS account, organizzazioni o unità organizzative a cui concedere le autorizzazioni.

Principal per attributi: aggiungi coppie chiave-valore di tag dai IAMroles o tag di sessione IAM. I principali con attributi corrispondenti ricevono l'accesso alla risorsa specificata.

- **Tag LF o risorse del catalogo:** cataloghi, database, tabelle, viste o collegamenti alle risorse per cui concedere le autorizzazioni.
- **Autorizzazioni:** autorizzazioni da concedere a The Lake Formation.

Note

Per concedere le autorizzazioni su un collegamento a una risorsa del database, vedere. [Concessione delle autorizzazioni per i collegamenti alle risorse](#)

1. Apri la pagina Concedi le autorizzazioni.

Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake, creatore del catalogo o utente IAM con autorizzazioni Grantable sul catalogo.

Esegui una delle seguenti operazioni:

- Nel pannello di navigazione, in Autorizzazioni, scegli Autorizzazioni dati. Quindi scegli Concedi.

- Nel riquadro di navigazione, scegli Cataloghi in Data Catalog. Quindi, nella pagina Cataloghi, scegli un catalogo e dal menu Azioni, in Autorizzazioni, scegli Concedi.

Note

Puoi concedere le autorizzazioni su un catalogo tramite il relativo link alla risorsa. A tale scopo, nella pagina Cataloghi, scegli un contenitore di link al catalogo e, nel menu Azioni, scegli Concedi sulla destinazione. Per ulteriori informazioni, consulta [Come funzionano i link alle risorse in Lake Formation](#).

2. Successivamente, nella sezione Tipo principale, scegli i principali o specifica gli attributi associati ai principali.

Grant permissions

Principal type

Grant access by IAM principal or attribute tags.

☒ Principals

Grant permissions directly to specific IAM users, roles, or Identity Center principals.

☐ Principals by attributes

Grant permissions using attribute tags with Cedar policy expressions.

Principals

Choose the principals to grant permissions.

☒ IAM users and roles

Users or roles from this AWS account.

☐ IAM Identity Center - *new*

Users and groups configured in IAM Identity Center.

☐ SAML users and groups

SAML users and group or QuickSight ARNs.

☐ External accounts

AWS account, AWS organization or IAM principal outside of this account

IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

lf-data-analyst 
User

Specificare i principali

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Centro identità IAM

Scegli uno o più utenti o gruppi dall'elenco Utenti e gruppi. Seleziona Aggiungi per aggiungere altri utenti o gruppi.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per Account AWS, AWS organizzazione o IAM Principal inserisci uno o più AWS account IDs, organizzazione IDs IDs, unità organizzativa o ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4—32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

Principi per attributi

Attributes

Aggiungi le coppie chiave-valore del tag IAM dal ruolo IAM.

Ambito di autorizzazione

Specificate se state concedendo le autorizzazioni ai responsabili con attributi corrispondenti nello stesso account o in un altro account.

3. Nella sezione LF-Tags o risorse del catalogo, scegliete Risorse del catalogo dati denominate.

LF-Tags or catalog resources

Choose a method to grant permissions.

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named Data Catalog resources
Manage permissions for specific databases or tables, in addition to fine-grained data access.

Catalogs

Choose catalogs

mycatalogdemo X

Databases

Select one or more databases.

Choose databases

tpcds1 X
mycatalogdemo

Tables - optional

Select one or more tables.

Choose tables

All tables X
mycatalogdemo

Views - optional

Select one or more views.

Choose views

Data filters - optional

Select one or more data filters.

Choose data filters

Load
more

Create new

[Manage data filters](#)

4. Scegli uno o più cataloghi dall'elenco Cataloghi. Puoi anche scegliere uno o più database, tabelle, filtri di and/or dati.
5. Nella sezione Autorizzazioni del catalogo, seleziona Autorizzazioni e autorizzazioni concedibili. In Autorizzazioni del catalogo, seleziona una o più autorizzazioni da concedere.

Catalog permissions

Choose the permissions to grant on the catalog. Choosing Super user overwrites individual permissions, granting unrestricted administrative access.

☐ Super user

A super user has unrestricted administrative privileges to perform any operation on all resources within the catalog (databases, tables, and views).

Catalog permissions

Choose specific access permissions to grant.

☐ Create database ☐ Describe ☐ Alter

☐ Drop

Grantable permissions

Choose the permission that can be granted to others.

☐ Create database ☐ Describe ☐ Alter

☐ Drop

☐ Super

This permission is the union of all the individual permissions to the left, and supersedes them.

☐ Super

This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel

Grant

Scegli Super user per concedere privilegi amministrativi illimitati per eseguire qualsiasi operazione su tutte le risorse all'interno del catalogo (database, tabelle e viste).

Note

Dopo aver concesso Create database o Alter su un catalogo che ha una proprietà di posizione che punta a una posizione registrata, assicuratevi di concedere anche ai responsabili le autorizzazioni di localizzazione dei dati sulla posizione. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

6. (Facoltativo) In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere agli altri responsabili del proprio account. AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.
7. Scegli Concessione.

La pagina Autorizzazioni relative ai dati mostra i dettagli delle autorizzazioni. Se hai utilizzato l'opzione Principali per attributo per concedere le autorizzazioni, puoi visualizzare l'autorizzazione concessa ALLPrincipals nell'elenco.

AWS CLI

Per concedere le autorizzazioni del catalogo utilizzando, consulta. AWS CLI [Creazione di cataloghi federati Amazon Redshift](#)

Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata

I passaggi seguenti spiegano come concedere le autorizzazioni al database utilizzando il metodo della risorsa denominata.

Console

Usa la pagina Concedi le autorizzazioni sulla console Lake Formation. La pagina è suddivisa nelle seguenti sezioni:

- Tipo principale: la sezione Principal include gli utenti, i ruoli, gli utenti e i gruppi IAM Identity Center, gli utenti e i gruppi SAML, gli AWS account, le organizzazioni o le unità organizzative a cui concedere le autorizzazioni. Nella sezione Principal by attributes, puoi specificare la chiave e i valori per gli attributi associati ai ruoli IAM.
- LF-Tag o risorse di catalogo: database, tabelle, viste o collegamenti alle risorse su cui concedere le autorizzazioni.
- Autorizzazioni: autorizzazioni da concedere a The Lake Formation.

Note

Per concedere le autorizzazioni su un collegamento a una risorsa del database, vedere. [Concessione delle autorizzazioni per i collegamenti alle risorse](#)

1. Apri la pagina Concedi le autorizzazioni.

Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake, creatore del database o utente IAM con autorizzazioni Grantable sul database.

Esegui una delle seguenti operazioni:

- Nel pannello di navigazione, in Autorizzazioni, scegli Autorizzazioni dati. Quindi scegli Concedi.
- Nel riquadro di navigazione, scegli Database in Data Catalog. Quindi, nella pagina Database, scegli un database e dal menu Azioni, in Autorizzazioni, scegli Concedi.

Note

Puoi concedere le autorizzazioni su un database tramite il relativo link alla risorsa. A tale scopo, nella pagina Database, scegli un collegamento a una risorsa e nel menu Azioni scegli Concedi alla destinazione. Per ulteriori informazioni, consulta [Come funzionano i link alle risorse in Lake Formation](#).

2. Nella sezione Tipo principale, specifica i principali o concedi le autorizzazioni ai principali utilizzando gli attributi.

Grant permissions

Principals

Choose the principals to grant permissions.

☒ IAM users and roles

Users or roles from this AWS account.

☐ IAM Identity Center

- *new*

Users and groups configured in IAM Identity Center.

☐ SAML users and groups

SAML users and group or QuickSight ARNs.

☐ External accounts

AWS account, AWS organization or IAM principal outside of this account

IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add

AWSServiceRoleForRedshift

Role

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Centro identità IAM

Scegli uno o più utenti o gruppi dall'elenco Utenti e gruppi. Seleziona Aggiungi per aggiungere altri utenti o gruppi.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per Account AWS, AWS organizzazione o IAM Principal inserisci uno o più AWS account IDs, organizzazione IDs IDs, unità organizzativa o ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4—32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

Principi per attributi

Specificare la chiave e i valori dell'attributo. Se scegli più di un valore, stai creando un'espressione di attributo con un operatore OR. Ciò significa che se uno qualsiasi dei

valori dei tag degli attributi assegnati a un ruolo o utente IAM corrisponde, role/user ottiene i permessi di accesso alla risorsa.

Scegli l'ambito delle autorizzazioni specificando se stai concedendo le autorizzazioni ai responsabili con attributi corrispondenti nello stesso account o in un altro account.

3. Nella sezione LF-Tags o risorse del catalogo, scegliete Risorse del catalogo dati denominate.

LF-Tags or catalog resources

Choose a method to grant permissions.

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named Data Catalog resources
Manage permissions for specific databases or tables, in addition to fine-grained data access.

Catalogs

Choose catalogs

Default catalog

Databases

Select one or more databases.

Choose databases

sales

Tables - optional

Select one or more tables.

Choose tables

Views - optional

Select one or more views.

Choose views

Data filters - optional

Select one or more data filters.

Choose data filters

Load more

Create new

[Manage data filters](#)

4. Scegli uno o più database dall'elenco Database. Puoi anche scegliere uno o più filtri Tables and/or Data.
5. Nella sezione Autorizzazioni, seleziona le autorizzazioni e le autorizzazioni concedibili. In Autorizzazioni del database, seleziona una o più autorizzazioni da concedere.

Database permissions

Database permissions

Choose specific access permissions to grant.

☐ Create table
 ☐ Alter
 ☐ Drop

☐ Describe

☐ Super

This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions

Choose the permission that may be granted to others.

☐ Create table
 ☐ Alter
 ☐ Drop

☐ Describe

☐ Super

This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Note

Dopo aver concesso `Create Table` o `Alter` su un database con una proprietà `location` che punta a una posizione registrata, assicurati di concedere anche ai responsabili le autorizzazioni per la localizzazione dei dati sulla posizione. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

- (Facoltativo) In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere ad altri responsabili del proprio account. AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.
- Scegli Concessione.

AWS CLI

È possibile concedere le autorizzazioni per il database utilizzando il metodo della risorsa denominato `awscli`. AWS Command Line Interface AWS CLI

Per concedere le autorizzazioni al database utilizzando AWS CLI

- Eseguite un `grant-permissions` comando e specificate un database o il Data Catalog come risorsa, a seconda dell'autorizzazione concessa.

Negli esempi seguenti, sostituiscilo `<account-id>` con un ID AWS account valido.

Example — Concedi la creazione di un database

Questo esempio concede `CREATE_DATABASE` all'utente `datalake_user1`. Poiché la risorsa su cui viene concessa questa autorizzazione è il Data Catalog, il comando specifica una `CatalogResource` struttura vuota come parametro. `resource`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {} }'
```

Example — Concedi la creazione di tabelle in un database designato

L'esempio successivo concede l'`CREATE_TABLE` accesso al database `retail` all'utente `datalake_user1`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_TABLE" --resource '{ "Database": { "Name": "retail" } }'
```

Example — Concedi a un AWS account esterno con l'opzione Grant

L'esempio successivo concede `CREATE_TABLE` con l'opzione `grant` sul database `retail` all'account esterno `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "CREATE_TABLE"
--permissions-with-grant-option "CREATE_TABLE" --resource '{ "Database":
{ "Name": "retail" } }'
```

Example — Concessione a un'organizzazione

L'esempio successivo concede all'organizzazione `ALTER o-abcdefghijkl` con l'opzione di concessione sul database `issues`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/
```

```
o-abcdefghijkl --permissions "ALTER" --permissions-with-grant-option "ALTER" --  
resource '{ "Database": {"Name":"issues"}}'
```

Example - Concedi a **ALLIAMPrincipals** nello stesso account

L'esempio successivo concede l'`CREATE_TABLE` autorizzazione sul database `retail` a tutti i responsabili dello stesso account. Questa opzione consente a tutti i principali dell'account di creare una tabella nel database e creare un collegamento alle risorse della tabella che consente ai motori di query integrati di accedere a database e tabelle condivisi. Questa opzione è particolarmente utile quando un principale riceve una sovvenzione tra account e non dispone dell'autorizzazione per creare collegamenti alle risorse. In questo scenario, l'amministratore del data lake può creare un database segnaposto e concedere l'`CREATE_TABLE` autorizzazione al `ALLIAMPrincipal` gruppo, consentendo a ogni principale IAM dell'account di creare collegamenti alle risorse nel database placeholder.

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=111122223333:IAMPrincipals  
--permissions "CREATE_TABLE" --resource '{ "Database":  
{"Name":"temp","CatalogId":"111122223333"}}'
```

Example - Concedi a **ALLIAMPrincipals** in un account esterno

L'esempio successivo concede l'`CREATE_TABLE` accesso al database `retail` a tutti i principali di un account esterno. Questa opzione consente a ogni principale dell'account di creare una tabella nel database.

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=111122223333:IAMPrincipals  
--permissions "CREATE_TABLE" --resource '{ "Database":  
{"Name":"retail","CatalogId":"123456789012"}}'
```

Note

Dopo aver concesso `CREATE_TABLE` o `ALTER` su un database con una proprietà `location` che punta a una posizione registrata, assicurati di concedere anche ai responsabili le autorizzazioni di localizzazione dei dati sulla posizione. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

 Consulta anche

- [Riferimento alle autorizzazioni di Lake Formation](#)
- [Concessione delle autorizzazioni su un database o una tabella condivisa con il tuo account](#)
- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)

Concessione delle autorizzazioni per le tabelle utilizzando il metodo di risorsa denominato

Puoi utilizzare la console Lake Formation o AWS CLI concedere le autorizzazioni di Lake Formation sulle tabelle di Data Catalog. È possibile concedere autorizzazioni su singole tabelle oppure con una singola operazione di concessione, è possibile concedere autorizzazioni su tutte le tabelle di un database.

Se concedi le autorizzazioni per tutte le tabelle di un database, concedi implicitamente le autorizzazioni per il DESCRIBE database. Il database viene quindi visualizzato nella pagina Database della console e viene restituito dall'GetDatabasesoperazione API. Questa concessione automatica di DESCRIBE autorizzazione non si applica quando si utilizza il controllo degli accessi basato sugli attributi (ABAC). Quando si concedono le autorizzazioni su tutte le tabelle di un database utilizzando gli attributi, Lake Formation non concede implicitamente DESCRIBE l'autorizzazione al database.

Quando scegli SELECT come autorizzazione da concedere, hai la possibilità di applicare un filtro di colonna, un filtro di riga o un filtro di cella.

Console

I passaggi seguenti spiegano come concedere le autorizzazioni alle tabelle utilizzando il metodo della risorsa denominata e la pagina Grant data lake permissions sulla console Lake Formation. La pagina è suddivisa nelle seguenti sezioni:

- Tipi principali: utenti, ruoli, AWS account, organizzazioni o unità organizzative a cui concedere le autorizzazioni. Puoi anche concedere autorizzazioni ai responsabili con attributi corrispondenti.
- Tag LF o risorse del catalogo: i database, le tabelle o i collegamenti alle risorse a cui concedere le autorizzazioni.
- Autorizzazioni: autorizzazioni da concedere a The Lake Formation.

 Note

Per concedere le autorizzazioni su un collegamento a una risorsa della tabella, vedere. [Concessione delle autorizzazioni per i collegamenti alle risorse](#)

1. Apri la pagina Concedi le autorizzazioni.

Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake, creatore della tabella o utente a cui sono state concesse le autorizzazioni sulla tabella con l'opzione di concessione.

Esegui una delle seguenti operazioni:

- Nel riquadro di navigazione, scegli Autorizzazioni dati in Autorizzazioni. Quindi scegli Concedi.
- Nel pannello di navigazione, seleziona Tabelle. Quindi, nella pagina Tabelle, scegli una tabella e nel menu Azioni, in Autorizzazioni, scegli Concedi.

 Note

Puoi concedere le autorizzazioni su una tabella tramite il relativo link alla risorsa. A tale scopo, nella pagina Tabelle, scegli un collegamento a una risorsa e nel menu Azioni scegli Concedi all'obiettivo. Per ulteriori informazioni, consulta [Come funzionano i link alle risorse in Lake Formation](#).

2. Successivamente, nella sezione Tipi principali, specifica i principali o i principali con attributi corrispondenti per concedere le autorizzazioni.

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Centro identità IAM

Scegli uno o più utenti o gruppi dall'elenco Utenti e gruppi.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o ARNs per utenti o gruppi Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per Account AWS , AWS organizzazione o IAM Principal, inserisci una o più organizzazioni Account AWS IDs IDs IDs, unità organizzative o l'ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4—32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo carattere «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

Principi per attributi

Specificare la chiave e i valori dell'attributo. Se scegli più di un valore, stai creando un'espressione di attributo con un operatore OR. Ciò significa che se uno qualsiasi dei valori dei tag degli attributi assegnati a un ruolo o utente IAM corrisponde, role/user ottiene i permessi di accesso alla risorsa

Scegli l'ambito delle autorizzazioni specificando se stai concedendo le autorizzazioni ai responsabili con attributi corrispondenti nello stesso account o in un altro account.

3. Nella sezione LF-Tags o risorse del catalogo, scegli un database. Quindi scegliete una o più tabelle o Tutte le tabelle.

LF-Tags or catalog resources

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named data catalog resources
Manager permissions for specific databases or tables, in addition to fine-grained data access.

Databases
Select one or more databases.

Choose databases ▼

retail ✕

Tables - optional
Select one or more tables.

Choose tables ▼

Load more

inventory ✕
No description available

4. Specificate le autorizzazioni senza filtraggio dei dati.

Nella sezione Autorizzazioni, seleziona la tabella autorizzazioni da concedere e, facoltativamente, seleziona le autorizzazioni concedibili.

Table and column permissions

Table permissions
Choose specific access permissions to grant.

☒ Alter	☒ Insert	☐ Drop	☐ Super
☐ Delete	☒ Select	☒ Describe	This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions
Choose the permission that may be granted to others.

☐ Alter	☐ Insert	☐ Drop	☐ Super
☐ Delete	☐ Select	☐ Describe	This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Se concedi Select, la sezione Autorizzazioni ai dati viene visualizzata sotto la sezione Autorizzazioni per tabelle e colonne, con l'opzione di accesso a tutti i dati selezionata per impostazione predefinita. Accetta l'impostazione predefinita.

Data permissions

☒ **All data access**
Grant access to all data without any restrictions.

☐ **Simple column-based access**
Grant data access to specific columns only.

☐ **Advanced cell-level filters**
Grant access to specific columns and/or rows with data filters.

5. Scegli Concessione.
6. Specificare l'autorizzazione Seleziona con filtraggio dei dati

Seleziona l'autorizzazione Seleziona. Non selezionare altre autorizzazioni.

La sezione Autorizzazioni per i dati viene visualizzata sotto la sezione Autorizzazioni per tabelle e colonne.

7. Esegui una delle seguenti operazioni:
 - Applica solo un semplice filtraggio delle colonne.
 1. Scegli Accesso semplice basato su colonne.

Table and column permissions

Table permissions

Choose specific access permissions to grant.

☐ Alter
 ☐ Insert
 ☐ Drop

☐ Delete
 ☒ Select
 ☐ Describe

☐ Super

This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions

Choose the permission that may be granted to others.

☐ Alter
 ☐ Insert
 ☐ Drop

☐ Delete
 ☒ Select
 ☐ Describe

☐ Super

This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Data permissions

☐ All data access

Grant access to all data without any restrictions.

☒ Simple column-based access

Grant data access to specific columns only.

☐ Advanced cell-level filters

Grant access to specific columns and/or rows with data filters.

Choose permission filter

Choose whether to include or exclude columns.

☒ Include columns

Grant permissions to access specific columns.

☐ Exclude columns

Grant permissions to access all but specific columns.

Select columns

Grantable permissions

Choose the permission that may be granted to others.

☐ Select

- Scegli se includere o escludere le colonne, quindi scegli le colonne da includere o escludere.

Quando si concedono le autorizzazioni a un AWS account o a un'organizzazione esterni, sono supportati solo gli elenchi di inclusione.

- (Facoltativo) In Autorizzazioni concedibili, attiva l'opzione di concessione per l'autorizzazione Seleziona.

Se includi l'opzione di concessione, il destinatario della concessione può concedere le autorizzazioni solo nelle colonne che gli concedi.

Note

Puoi anche applicare il filtro delle colonne solo creando un filtro dati che specifichi un filtro di colonna e specifichi tutte le righe come filtro di riga. Tuttavia, ciò richiede ulteriori passaggi.

- Applica il filtraggio di colonne, righe o celle.

1. Scegli filtri avanzati a livello di cella.

The screenshot shows the 'Data permissions' interface in AWS Lake Formation. At the top, there are three radio button options: 'All data access' (unselected), 'Simple column-based access' (unselected), and 'Advanced cell-level filters' (selected). Below these is a 'View existing permissions' link. The main section is titled 'Data filters to grant' and contains a search bar, a 'Manage filters' button, and a 'Create new filter' button. Below this is a table listing existing filters.

	Filter name	Table	Database	Table catalog ID
☐	restrict-pharma	orders	sales	111122223333
☐	no-pharma	orders	sales	111122223333

2. (Facoltativo) Espandi Visualizza le autorizzazioni esistenti.
3. (Facoltativo) Scegli Crea nuovo filtro.
4. (Facoltativo) Per visualizzare i dettagli dei filtri elencati o per creare nuovi filtri o eliminare filtri esistenti, scegli Gestisci filtri.

La pagina Filtri dati si apre in una nuova finestra del browser.

Al termine della pagina Filtri dati, torna alla pagina Concedi autorizzazioni e, se necessario, aggiorna la pagina per visualizzare i nuovi filtri di dati che hai creato.

5. Seleziona uno o più filtri di dati da applicare alla concessione.

 Note

Se non ci sono filtri di dati nell'elenco, significa che non sono stati creati filtri di dati per la tabella selezionata.

8. Scegli Concessione.

AWS CLI

È possibile concedere le autorizzazioni per la tabella utilizzando il metodo di risorsa denominato e il AWS Command Line Interface (AWS CLI).

Per concedere i permessi alle tabelle utilizzando il AWS CLI

- Eseguite un `grant-permissions` comando e specificate una tabella come risorsa.

Example — Concessione su una singola tabella, senza filtri

L'esempio seguente concede `SELECT` e `ALTER` all'utente `datalake_user1` nell' AWS account `1111-2222-3333` sulla tabella del database. `inventory retail`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
"Name":"inventory"}}'
```

 Note

Se concedi l'`ALTER` autorizzazione su una tabella i cui dati sottostanti si trovano in una posizione registrata, assicurati di concedere anche ai principali i permessi di localizzazione dei dati sulla posizione. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per la localizzazione dei dati](#).

Example — Concedi su tutte le tabelle con l'opzione `Grant`: nessun filtro

L'esempio successivo concede `SELECT` con l'opzione `grant` su tutte le tabelle del database. `retail`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --permissions-with-grant-option "SELECT" --resource '{ "Table":
{ "DatabaseName": "retail", "TableWildcard": {} } }'
```

Example — Grant con un semplice filtraggio delle colonne

Questo esempio successivo concede autorizzazioni SELECT su un sottoinsieme di colonne della tabella. `persons` Utilizza un semplice filtraggio delle colonne.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --resource '{ "TableWithColumns": {"DatabaseName": "hr",
"Name": "persons", "ColumnNames": ["family_name", "given_name", "gender"] } }'
```

Example — Concedi con un filtro dati

Questo esempio concede SELECT sulla `orders` tabella e applica il filtro `restrict-pharma` dati.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

Di seguito è riportato il contenuto del file `grant-params.json`.

```
{
  "Principal": {"DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["SELECT"],
  "PermissionsWithGrantOption": ["SELECT"]
}
```


 Consulta anche

- [Panoramica delle autorizzazioni di Lake Formation](#)
- [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#)
- [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#)
- [Concessione delle autorizzazioni per i collegamenti alle risorse](#)
- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)

Concessione delle autorizzazioni sulle viste utilizzando il metodo di risorsa denominato

I passaggi seguenti spiegano come concedere le autorizzazioni sulle viste utilizzando il metodo della risorsa denominata e la pagina Concedi le autorizzazioni. La pagina è suddivisa nelle seguenti sezioni:

- Tipi principali: utenti, ruoli, utenti e gruppi IAM Identity Center Account AWS, organizzazioni o unità organizzative a cui concedere le autorizzazioni. Puoi anche concedere autorizzazioni ai responsabili con attributi corrispondenti.
- Tag LF o risorse del catalogo: i database, le tabelle, le viste o i collegamenti alle risorse a cui concedere le autorizzazioni.
- Autorizzazioni: le autorizzazioni del data lake da concedere.

Apri la pagina Concedi le autorizzazioni

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake, creatore del database o utente IAM con autorizzazioni Grantable sul database.
2. Esegui una delle seguenti operazioni:
 - Nel pannello di navigazione, in Autorizzazioni, scegli Autorizzazioni dati. Quindi scegli Concedi.
 - Nel riquadro di navigazione, scegli Visualizzazioni in Data Catalog. Quindi, nella pagina Visualizzazioni, scegli una visualizzazione e dal menu Azioni, in Autorizzazioni, scegli Concedi.

 Note

Puoi concedere le autorizzazioni per una vista tramite il relativo link alla risorsa. A tale scopo, nella pagina Visualizzazioni, scegli un collegamento a una risorsa e nel menu Azioni scegli Concedi all'obiettivo. Per ulteriori informazioni, consulta [Come funzionano i link alle risorse in Lake Formation](#).

Specificate i tipi principali

Nella sezione Tipi principali, scegli Principali o Principali per attributi. Se scegli Principal, sono disponibili le seguenti opzioni:

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Centro identità IAM

Scegli uno o più utenti o gruppi dall'elenco Utenti e gruppi.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o per utenti o gruppi ARNs Amazon Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

 Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per Account AWS, AWS organizzazione o IAM Principal inserisci uno o più AWS account IDs, organizzazione IDs IDs, unità organizzativa o ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4—32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

Vedi anche

- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)

Principi per attributi

Specificare la chiave e i valori dell'attributo. Se scegli più di un valore, stai creando un'espressione di attributo con un operatore OR. Ciò significa che se uno qualsiasi dei valori dei tag degli attributi assegnati a un ruolo o utente IAM corrisponde, role/user ottiene i permessi di accesso alla risorsa

Scegli l'ambito delle autorizzazioni specificando se stai concedendo le autorizzazioni ai responsabili con attributi corrispondenti nello stesso account o in un altro account.

Specificate le visualizzazioni

Nella sezione LF-Tags o risorse del catalogo, scegliete una o più viste a cui concedere le autorizzazioni.

1. Scegliete Named data catalog resources.
2. Scegliete una o più viste dall'elenco Visualizzazioni. Puoi anche scegliere uno o più cataloghi, database, tabelle, filtri di and/or dati.

La concessione delle autorizzazioni per il data lake all'All tables interno di un database comporterà che il beneficiario disponga delle autorizzazioni su tutte le tabelle e le viste all'interno del database.

Specificare le autorizzazioni

Nella sezione Autorizzazioni, seleziona le autorizzazioni e le autorizzazioni concedibili.

View permissions

View permissions

Choose specific access permissions to grant.

☐ Select

☐ Describe

☐ Drop

☐ Super

This permission is the union of all the individual permissions to the left, and supersedes them.

Grantable permissions

Choose the permission that may be granted to others.

☐ Select

☐ Describe

☐ Drop

☐ Super

This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

Cancel

Grant

1. In Visualizza autorizzazioni, seleziona una o più autorizzazioni da concedere.
2. (Facoltativo) In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere ad altri responsabili del proprio territorio. Account AWS Questa opzione non è supportata quando si concedono autorizzazioni a un principale IAM da un account esterno.
3. Scegli Concessione.

Vedi anche

- [Riferimento alle autorizzazioni di Lake Formation](#)
- [Concessione delle autorizzazioni su un database o una tabella condivisa con il tuo account](#)

Controllo degli accessi basato su tag Lake Formation

Il controllo degli accessi basato su tag di Lake Formation (LF-TBAC) è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. In Lake Formation, questi attributi sono chiamati

LF-tag. È possibile allegare tag LF alle risorse del Data Catalog e concedere autorizzazioni ai responsabili di Lake Formation su tali risorse utilizzando questi tag LF. Lake Formation consente operazioni su tali risorse quando il principale ha concesso l'accesso a un valore di tag che corrisponde al valore del tag della risorsa.

LF-TBAC è utile in ambienti in rapida crescita e aiuta in situazioni in cui la gestione delle politiche diventa complicata.

LF-TBAC è il metodo consigliato da utilizzare per concedere le autorizzazioni di Lake Formation quando è presente un numero elevato di oggetti Data Catalog, inclusi cataloghi federati, database, tabelle e viste. Lake Formation supporta il controllo degli accessi basato su tag per cataloghi federati di tabelle Amazon S3, Amazon Redshift data warehouse e fonti di dati federate come Amazon DynamoDB SQL Server e Snowflake.

Note

I tag IAM non sono uguali ai tag LF. Questi tag non sono intercambiabili. I tag LF vengono utilizzati per concedere i permessi di Lake Formation e i tag IAM vengono utilizzati per definire le politiche IAM.

Come funziona il controllo degli accessi basato su tag Lake Formation

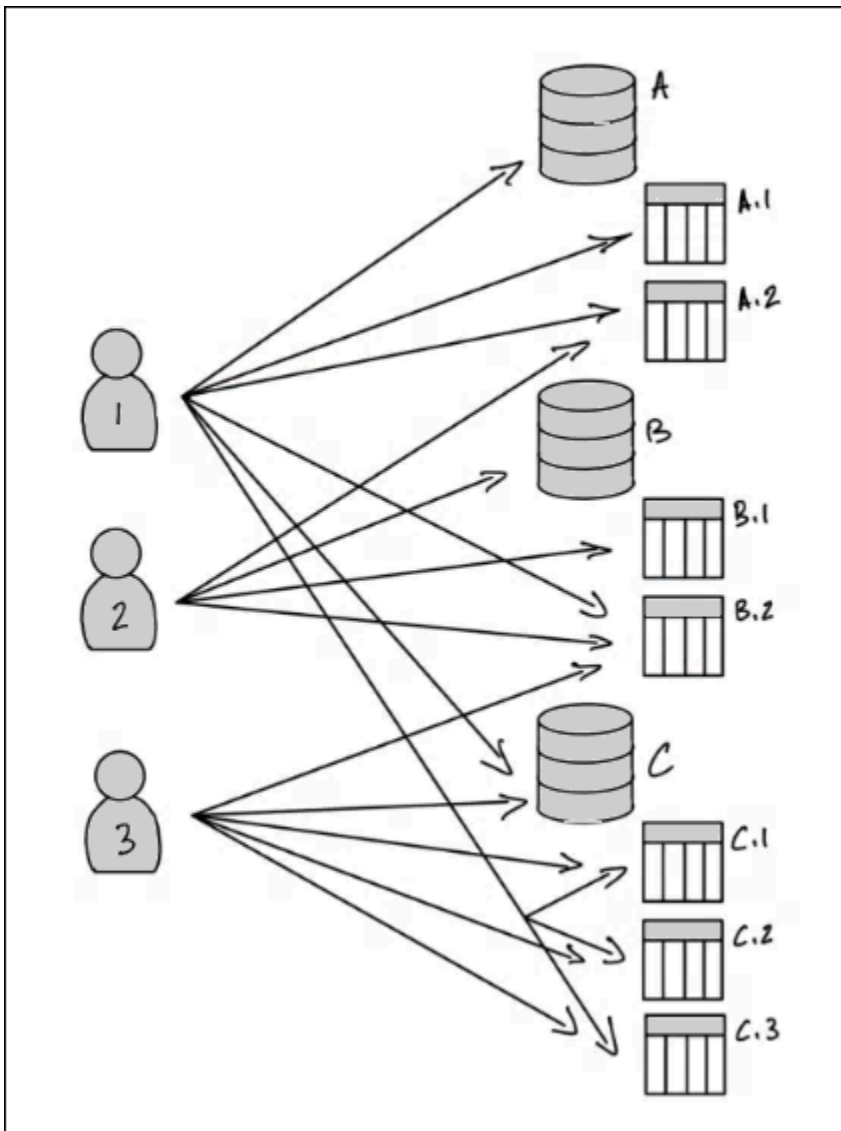
Ogni tag LF è una coppia chiave-valore, ad esempio o. `department=sales`
`classification=restricted` Una chiave può avere più valori definiti, ad esempio.
`department=sales,marketing,engineering,finance`

Per utilizzare il metodo LF-TBAC, gli amministratori del data lake e i data engineer eseguono le seguenti attività.

Operazione	Dettagli dell'attività
1. Definire le proprietà e le relazioni dei tag LF.	-
2. Crea i creatori di tag LF in Lake Formation.	Aggiungere creatori di tag LF
3. Crea il tag LF in Lake Formation.	Creazione di tag LF

Operazione	Dettagli dell'attività
4. Assegna tag LF alle risorse del Data Catalog.	Assegnazione di tag LF alle risorse del Data Catalog
5. Concedi le autorizzazioni ad altri principali per assegnare i tag LF alle risorse, opzionalmente con l'opzione grant.	Gestione dei permessi per i valori del tag LF
6. Concedi le espressioni dei tag LF ai principali, opzionalmente con l'opzione grant.	Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC
7. (Consigliato) Dopo aver verificato che i principali abbiano accesso alle risorse corrette tramite il metodo LF-TBAC, revocate le autorizzazioni concesse utilizzando il metodo della risorsa denominata.	-

Si consideri il caso in cui è necessario concedere le autorizzazioni a tre principali su tre database e sette tabelle.



Per ottenere le autorizzazioni indicate nel diagramma precedente utilizzando il metodo `named resource`, è necessario concedere 17 concessioni, come segue (in pseudo-codice).

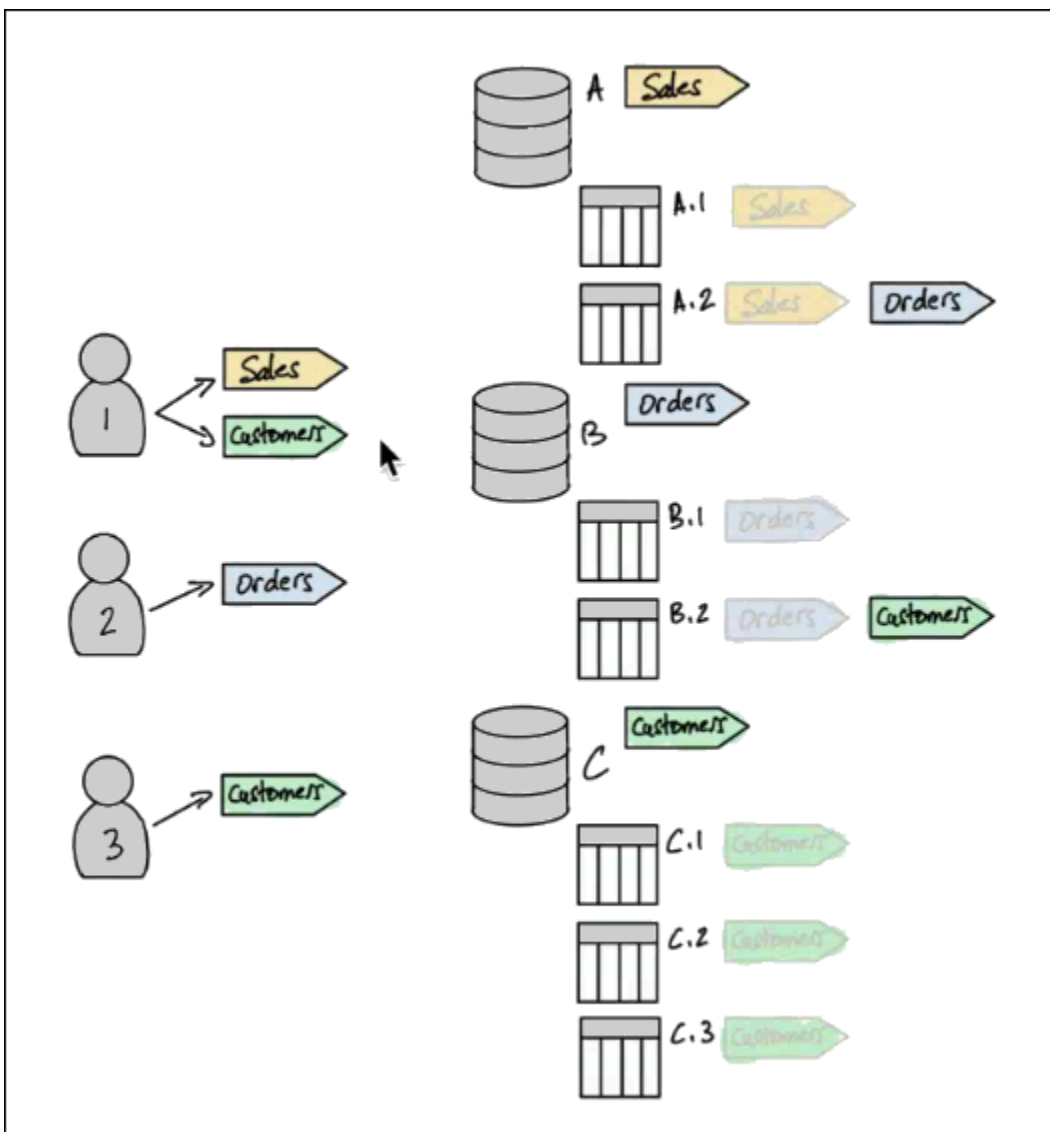
```
GRANT CREATE_TABLE ON Database A TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.1 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table B.2 TO PRINCIPAL 1
...
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 2
GRANT CREATE_TABLE ON Database B TO PRINCIPAL 2
...
GRANT SELECT, INSERT ON Table C.3 TO PRINCIPAL 3
```

Considerate ora come concedereste le autorizzazioni utilizzando LF-TBAC. Il diagramma seguente indica che sono stati assegnati tag LF a database e tabelle e che sono state concesse le autorizzazioni sui tag LF ai principali.

In questo esempio, i tag LF rappresentano aree del data lake che contengono analisi per diversi moduli di una suite di applicazioni ERP (Enterprise Resource Planning). È possibile utilizzarli per controllare l'accesso ai dati di analisi per i vari moduli. Tutti i tag LF hanno la chiave `module` e i valori `Sales`, `Orders`, `Customers`. Un esempio di tag LF ha il seguente aspetto:

```
module=Sales
```

Il diagramma mostra solo i valori del tag LF.



Assegnazione di tag alle risorse e all'ereditarietà di Data Catalog

Le tabelle ereditano i tag LF dai database e le colonne ereditano i tag LF dalle tabelle. I valori ereditati possono essere sovrascritti. Nel diagramma precedente, i tag LF oscurati vengono ereditati.

A causa dell'ereditarietà, l'amministratore del data lake deve effettuare solo le seguenti cinque assegnazioni di tag LF alle risorse (in pseudo-codice).

```
ASSIGN TAGS module=Sales TO database A
ASSIGN TAGS module=Orders TO table A.2
ASSIGN TAGS module=Orders TO database B
ASSIGN TAGS module=Customers TO table B.2
ASSIGN TAGS module=Customers TO database C
```

Contributi di tag ai committenti

Dopo aver assegnato i tag LF ai database e alle tabelle, l'amministratore del data lake deve concedere solo quattro tag LF ai principali, come segue (in pseudo-codice).

```
GRANT TAGS module=Sales TO Principal 1
GRANT TAGS module=Customers TO Principal 1
GRANT TAGS module=Orders TO Principal 2
GRANT TAGS module=Customers TO Principal 3
```

Ora, un principale con il tag module=Sales LF può accedere alle risorse del Data Catalog con il tag LF (ad esempio, il database A), un principale con il module=Sales tag LF può accedere alle risorse con il tag LF e così via. module=Customers module=Customers

I comandi di concessione precedenti sono incompleti. Questo perché, sebbene indichino tramite LF-Tags le risorse del Data Catalog su cui i responsabili hanno i permessi, non indicano esattamente quali SELECT permessi di Lake Formation (ad esempio ALTER) i principali hanno su tali risorse. Pertanto, i seguenti comandi in pseudo-codice sono una rappresentazione più accurata del modo in cui le autorizzazioni di Lake Formation vengono concesse alle risorse del Data Catalog tramite i tag LF.

```
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Sales TO Principal 1
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Sales TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 1
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Customers TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Orders TO Principal 2
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Orders TO Principal 2
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 3
```

```
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=Customers TO Principal 3
```

Mettere insieme: autorizzazioni risultanti sulle risorse

Dati i tag LF assegnati ai database e alle tabelle nel diagramma precedente e i tag LF concessi ai principali del diagramma, la tabella seguente elenca i permessi di Lake Formation che i principali hanno sui database e sulle tabelle.

Principale	Autorizzazioni concesse tramite LF-Tags
Principale 1	• CREATE_TABLE sul database A • SELECT, INSERT nella tabella A.1 • SELECT, INSERT nella tabella B.2 • CREATE_TABLE nel database C • SELECT, INSERT nella tabella C.1 • SELECT, INSERT nella tabella C.2 • SELECT, INSERT nella tabella C.3
Principal 2	• SELECT, INSERT nella tabella A.2 • CREATE_TABLE nel database B • SELECT, INSERT nella tabella B.1
Principal 3	• SELECT, INSERT nella tabella B.2 • CREATE_TABLE nel database C • SELECT, INSERT nella tabella C.1 • SELECT, INSERT nella tabella C.2 • SELECT, INSERT nella tabella C.3

Conclusione

In questo semplice esempio, utilizzando cinque operazioni di assegnazione e otto operazioni di concessione, l'amministratore del data lake è stato in grado di specificare 17 autorizzazioni. Quando ci sono decine di database e centinaia di tabelle, il vantaggio del metodo LF-TBAC rispetto al metodo delle risorse denominate diventa evidente. Nel caso ipotetico della necessità di concedere a ogni risorsa l'accesso principale, e $n(P)$ dov'è il numero di principali e il numero di risorse: $n(R)$

- Con il metodo Named Resource, il numero di sovvenzioni richieste è $\times n(P) n(R)$
- Con il metodo LF-TBAC, utilizzando un singolo tag LF, il totale del numero di sovvenzioni ai committenti e di assegnazioni alle risorse è $+ n(P) n(R)$

Consulta anche

- [Gestione dei tag LF per il controllo dell'accesso ai metadati](#)
- [Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC](#)

Argomenti

- [Gestione dei tag LF per il controllo dell'accesso ai metadati](#)
- [Gestione delle espressioni LF-Tag per il controllo dell'accesso ai metadati](#)
- [Gestione dei permessi per i valori del tag LF](#)

Gestione dei tag LF per il controllo dell'accesso ai metadati

Per utilizzare il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) per proteggere oggetti del Data Catalog come cataloghi, database, tabelle, viste e colonne, devi creare tag LF, assegnarli alle risorse e concedere le autorizzazioni LF-Tag ai principali.

Prima di poter assegnare LF-Tags agli oggetti del Data Catalog o concedere le autorizzazioni ai principali, è necessario definire i tag LF. Solo un amministratore del data lake o un responsabile con i permessi di creazione di LF-Tag può creare LF-Tag.

Creatori di tag LF

LF-Tag creator è un amministratore non amministratore che dispone delle autorizzazioni per creare e gestire i tag LF. Gli amministratori di Data Lake possono aggiungere creatori di tag LF utilizzando la console di Lake Formation o la CLI. I creatori di LF-Tag hanno i permessi impliciti di Lake Formation per aggiornare ed eliminare i tag LF, per assegnare i tag LF alle risorse e per concedere i permessi dei tag LF e i permessi per i valori dei tag LF ad altri principali.

Con i ruoli di creatore di LF-Tag, gli amministratori del data lake possono delegare attività di gestione dei tag come la creazione e l'aggiornamento di chiavi e valori dei tag a responsabili non amministratori. Gli amministratori di Data Lake possono anche concedere autorizzazioni valide ai

creatori di LF-Tag. `Create LF-Tag` Quindi, il creatore del tag LF può concedere il permesso di creare tag LF ad altri mandanti.

È possibile concedere due tipi di autorizzazioni sui tag LF:

- Autorizzazioni LF-Tag `-, e. Create LF-Tag Alter Drop` Queste autorizzazioni sono necessarie per creare, aggiornare ed eliminare i tag LF.

Gli amministratori di Data Lake e i creatori di LF-Tag dispongono implicitamente di queste autorizzazioni sui tag LF che creano e possono concedere queste autorizzazioni esplicitamente ai responsabili della gestione dei tag nel data lake.

- `Grant with LF-Tag expressions` Autorizzazioni `Assign della Describe` coppia chiave-valore LF-Tag `-, e.` Queste autorizzazioni sono necessarie per assegnare i tag LF agli oggetti del Data Catalog e per concedere le autorizzazioni sulle risorse ai responsabili che utilizzano il controllo degli accessi basato su tag Lake Formation. I creatori di LF-Tag ricevono implicitamente queste autorizzazioni durante la creazione di LF-Tag.

Dopo aver ricevuto l'`Create LF-Tag` autorizzazione e aver creato con successo i tag LF, il creatore di LF-Tag può assegnare i tag LF alle risorse e concedere le autorizzazioni LF-Tag (`Create LF-Tag, e`) ad altri soggetti non amministrativi per gestire i tag nel data lake. `Alter Drop` Puoi gestire i tag LF utilizzando la console Lake Formation, l'API o AWS Command Line Interface (AWS CLI).

Note

Gli amministratori di Data Lake dispongono delle autorizzazioni implicite di Lake Formation per creare, aggiornare ed eliminare i tag LF, assegnare i tag LF alle risorse e concedere i permessi LF-Tag ai principali.

Per le migliori pratiche e considerazioni, consulta [Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation](#)

Argomenti

- [Aggiungere creatori di tag LF](#)
- [Creazione di tag LF](#)
- [Aggiornamento dei tag LF](#)
- [Eliminazione dei tag LF](#)

- [Elenco dei tag LF](#)
- [Assegnazione di tag LF alle risorse del Data Catalog](#)
- [Visualizzazione dei tag LF assegnati a una risorsa](#)
- [Visualizzazione delle risorse a cui è assegnato un LF-Tag](#)
- [Ciclo di vita di un tag LF](#)
- [Confronto tra il controllo degli accessi basato su tag Lake Formation e il controllo degli accessi basato sugli attributi IAM](#)

Consulta anche

- [Gestione dei permessi per i valori del tag LF](#)
- [Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC](#)
- [Controllo degli accessi basato su tag Lake Formation](#)

Aggiungere creatori di tag LF

Per impostazione predefinita, gli amministratori del data lake possono creare, aggiornare ed eliminare i tag LF, assegnare tag agli oggetti del Data Catalog e concedere le autorizzazioni relative ai tag ai principali. Se desideri delegare le operazioni di creazione e gestione dei tag a responsabili non amministratori, l'amministratore del data lake può creare ruoli di creatore di tag LF e concedere l'autorizzazione a Lake Formation ai ruoli. Create LF-Tag Con l>Create LF-Tag autorizzazione concessa, i creatori di LF-Tag possono delegare le attività di creazione e manutenzione dei tag ad altri responsabili non amministrativi.

Affinché gli amministratori del data lake assegnino i tag LF alle risorse del Data Catalog, devono concedersi le autorizzazioni di associazione sui tag LF che non sono stati creati da loro.

Note

Le concessioni di autorizzazioni per più account possono includere solo autorizzazioni e. Describe Associate Non puoi concedere Create LF-Tag, DropAlter, e Grant with LFTag expressions autorizzazioni ai responsabili di un altro account.

Argomenti

- [Autorizzazioni IAM necessarie per creare tag LF](#)
- [Aggiungi creatori di tag LF](#)

 Consulta anche

- [Gestione dei permessi per i valori del tag LF](#)
- [Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC](#)
- [Controllo degli accessi basato su tag Lake Formation](#)

Autorizzazioni IAM necessarie per creare tag LF

È necessario configurare le autorizzazioni per consentire a un responsabile di Lake Formation di creare tag LF. Aggiungi la seguente dichiarazione alla politica delle autorizzazioni per il principale che deve essere un creatore di LF-Tag.

 Note

Sebbene gli amministratori dei data lake abbiano le autorizzazioni implicite di Lake Formation per creare, aggiornare ed eliminare i tag LF, assegnare i tag LF alle risorse e concedere i tag LF ai principali, gli amministratori del data lake necessitano anche delle seguenti autorizzazioni IAM.

Per ulteriori informazioni, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

```
{
  "Sid": "Transformational",
  "Effect": "Allow",
  "Action": [
    "lakeformation:AddLFTagsToResource",
    "lakeformation:RemoveLFTagsFromResource",
    "lakeformation:GetResourceLFTags",
    "lakeformation:ListLFTags",
    "lakeformation:CreateLFTag",
    "lakeformation:GetLFTag",
    "lakeformation:UpdateLFTag",
  ]
}
```

```
    "lakeformation:DeleteLFTag",  
    "lakeformation:SearchTablesByLFTags",  
    "lakeformation:SearchDatabasesByLFTags"  
  ]  
}
```

I principali che assegnano i tag LF alle risorse e concedono i tag LF ai principali devono avere le stesse autorizzazioni, ad eccezione delle autorizzazioni, e. CreateLFTag UpdateLFTag DeleteLFTag

Aggiungi creatori di tag LF

Un creatore di tag LF può creare un tag LF, aggiornare la chiave e i valori dei tag, eliminare tag, associare tag alle risorse del Data Catalog e concedere le autorizzazioni sulle risorse del Data Catalog ai principali utilizzando il metodo LF-TBAC. Il creatore di LF-Tag può anche concedere queste autorizzazioni ai principali.

È possibile creare ruoli di creatore di LF-Tag utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

console

Per aggiungere un creatore di tag LF

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore di datalake.

2. Nel pannello di navigazione, in Autorizzazioni, scegli LF-tags e permessi.

Nella pagina LF-Tag e permessi, scegli la sezione LF-Tag Creators e scegli Aggiungi creatori LF-Tag.

Add LF-Tag creators

LF-Tag creators can create and manage LF-Tags. [Learn more](#) 

LF-Tag creator details

IAM users and roles

Add IAM users or roles.

Choose IAM principals to add

If-developer 
User

Permission

Choose the permission to grant.

☒ Create LF-Tag

Grantable permission

Choose the permission that may be granted to others.

☒ Create LF-Tag

Cancel

Add

3. Nella pagina Aggiungi creatori di LF-Tag, scegli un ruolo o un utente IAM con le autorizzazioni necessarie per creare LF-Tag.
4. Abilita Create LF-Tag la casella di controllo delle autorizzazioni.
5. (Facoltativo) Per consentire ai responsabili selezionati di concedere l'Create LF-Tag autorizzazione ai mandanti, scegli Autorizzazione Create LF-Tag concedibile.
6. Scegliere Aggiungi.

AWS CLI

```
aws lakeformation grant-permissions --cli-input-json file://grantCreate
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:user/tag-manager"
  },
  "Resource": {
    "Catalog": {}
  },
  "Permissions": [
```



```

    "CreateLFTag"
  ],
  "PermissionsWithGrantOption": [
    "CreateLFTag"
  ]
}

```

Le seguenti sono le autorizzazioni disponibili per il ruolo di creatore di LF-Tag:

Autorizzazione	Description
Drop	Un principale con questa autorizzazione su un LF-Tag può eliminare un LF-Tag dal data lake. Il principale ottiene l' <code>Describe</code> autorizzazione implicita su tutti i valori dei tag di una risorsa LF-Tag.
Alter	Un principale con questa autorizzazione su un tag LF può aggiungere e o rimuovere un valore di tag da un tag LF. Il principale ottiene l' <code>Alter</code> autorizzazione implicita su tutti i valori dei tag di un tag LF.
Describe	Un principale con questa autorizzazione su un tag LF può visualizzare il tag LF e i suoi valori quando assegna tag LF alle risorse o concede autorizzazioni sui tag LF. È possibile concedere su tutti i valori chiave o su valori specifici. <code>Describe</code>
Associate	Un principale con questa autorizzazione su un tag LF può assegnare il tag LF a una risorsa del catalogo dati. <code>Associate</code> <code>Describe</code> Concedere concessioni implicite.
Grant with LF-Tag expression	Un principale con questa autorizzazione su un LF-Tag può concedere autorizzazioni sulle risorse di un Data Catalog utilizzando la chiave e i valori LF-Tag. <code>Grant with LF-Tag expression</code> La concessione implicita di concessioni. <code>Describe</code>

Queste autorizzazioni sono concesse. Un preside a cui sono state concesse queste autorizzazioni con l'opzione di concessione può concederle ad altri committenti.

Creazione di tag LF

Tutti i tag LF devono essere definiti in Lake Formation prima di poter essere utilizzati. Un tag LF è composto da una chiave e da uno o più valori possibili per la chiave.

Dopo che l'amministratore del data lake ha configurato le autorizzazioni IAM e le autorizzazioni Lake Formation richieste per il ruolo di creatore di LF-Tag, il responsabile può creare un LF-Tag. Il creatore del tag LF ottiene l'autorizzazione implicita ad aggiornare o rimuovere qualsiasi valore di tag dal tag LF e ad eliminare il tag LF.

È possibile creare tag LF utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per creare un tag LF

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come principale con i permessi di creazione di LF-Tag o come amministratore del data lake.

2. Nel pannello di navigazione, in Autorizzazioni, LF-Tag e autorizzazioni, scegli LF-Tags.

Viene visualizzata la pagina LF-Tags.

LF-Tags (2) Delete Edit Grant permissions Add LF-Tag

LF-Tags have a key and one or more values that can be associated with Data Catalog resources. [Learn more](#)

	Key	Values	Owner account ID	LF-Tag permissions
☐	test	111		View
☐	Department	Sales, Marketing		View

3. Scegli Aggiungi tag LF.
4. Nella finestra di dialogo Aggiungi LF-tag, inserite una chiave e uno o più valori.

Ogni chiave deve avere almeno un valore. Per inserire più valori, inserisci un elenco delimitato da virgole e premi Invio oppure inserisci un valore alla volta e scegli Aggiungi dopo ognuno di essi. Il numero massimo di valori consentito è 1000.

5. Seleziona Aggiungi tag.

AWS CLI

Per creare un tag LF

- Immettete un comando `create-lf-tag`.

L'esempio seguente crea un tag LF con chiave `module` e valori `Customers` e `Orders`

```
aws lakeformation create-lf-tag --tag-key module --tag-values Customers Orders
```

In qualità di creatore di tag, il principale ottiene l'`Alter` autorizzazione per questo tag LF e può aggiornare o rimuovere qualsiasi valore di tag da questo tag LF. Il responsabile del creatore di LF-Tag può anche concedere `Alter` il permesso a un altro principale di aggiornare e rimuovere i valori dei tag su questo LF-Tag.

Aggiornamento dei tag LF

Aggiorna un tag LF per il quale hai l'`Alter` autorizzazione aggiungendo o eliminando i valori chiave consentiti. Non è possibile modificare la chiave LF-Tag. Per cambiare la chiave, elimina il tag LF e aggiungine uno con la chiave richiesta. Oltre all'`Alter` autorizzazione, è necessaria anche l'autorizzazione `lakeformation:UpdateLFTag` IAM per aggiornare i valori.

Quando si elimina un valore del tag LF, non viene eseguito alcun controllo per verificare la presenza di tale valore in alcuna risorsa del Data Catalog. Se il valore del tag LF eliminato è associato a una risorsa, non è più visibile per la risorsa e tutti i principali a cui sono state concesse le autorizzazioni per quella coppia chiave-valore non dispongono più delle autorizzazioni.

Prima di eliminare un valore del tag LF, è possibile utilizzare facoltativamente il [remove-lf-tags-from-resource](#) comando per rimuovere il tag LF dalle risorse del Catalogo dati che hanno il valore che si desidera eliminare, quindi rietichettare la risorsa con i valori che si desidera conservare.

Solo gli amministratori del data lake, il creatore del tag LF e i responsabili che dispongono delle autorizzazioni per l'LF-Tag possono aggiornare un LF-Tag. `Alter`

È possibile aggiornare un LF-Tag utilizzando la console, l'API o il (). AWS Lake Formation AWS Command Line Interface AWS CLI

Console

Per aggiornare un LF-tag (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore di LF-Tag o principale con l'Altra autorizzazione per l'LF-Tag.

2. Nel pannello di navigazione, in Autorizzazioni, LF-Tag e autorizzazioni, scegli LF-Tags.
3. Nella pagina LF-Tags, selezionate un tag LF, quindi scegliete Modifica.
4. Nella finestra di dialogo Modifica LF-Tag, aggiungete o rimuovete i valori del tag LF.

Per aggiungere più valori, nel campo Valori, inserite un elenco delimitato da virgole e premete Invio, oppure inserite un valore alla volta o scegliete Aggiungi dopo ognuno di essi.

5. Scegli Save (Salva).

AWS CLI

Per aggiornare un tag LF (AWS CLI)

- Immettete un comando `update-lf-tag`. Fornite uno o entrambi i seguenti argomenti:
 - `--tag-values-to-add`
 - `--tag-values-to-delete`

Example

L'esempio seguente sostituisce il valore `vp` con il valore della chiave `vice-president` LF-Tag. `level`

```
aws lakeformation update-lf-tag --tag-key level --tag-values-to-add vice-president  
--tag-values-to-delete vp
```

Eliminazione dei tag LF

È possibile eliminare i tag LF che non sono più in uso. Non viene eseguito alcun controllo per la presenza del tag LF su una risorsa del Data Catalog. Se il tag LF eliminato è associato a una risorsa,

non è più visibile per la risorsa e tutti i principali a cui sono state concesse le autorizzazioni su quel tag LF non dispongono più delle autorizzazioni.

Prima di eliminare un tag LF, è possibile utilizzare facoltativamente il comando per rimuovere il tag LF da tutte le risorse. [remove-lf-tags-from-resource](#)

Solo gli amministratori del data lake, il creatore del tag LF o un principale che dispone Drop dell'autorizzazione per l'LF-Tag possono eliminare un LF-Tag. Oltre all'autorizzazione, il principale `lakeformation:DeleteLFTag` necessita anche dell'Dropautorizzazione IAM per eliminare un LF-Tag.

È possibile eliminare un tag LF utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per eliminare un LF-tag (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake.

2. Nel pannello di navigazione, in Autorizzazioni, LF-tags e autorizzazioni, scegli LF-tags.
3. Nella pagina LF-Tags, selezionate un LF-Tag, quindi scegliete Elimina.
4. Nell'ambiente Delete tag? nella finestra di dialogo, per confermare l'eliminazione, inserite il valore della chiave LF-Tag nel campo designato, quindi scegliete Elimina.

AWS CLI

Per eliminare un tag LF ()AWS CLI

- Immettete un comandodelete-lf-tag. Fornisci la chiave del tag LF da eliminare.

Example

L'esempio seguente elimina il tag LF con la chiave. region

```
aws lakeformation delete-lf-tag --tag-key region
```

Elenco dei tag LF

Puoi elencare i tag LF su cui hai i permessi o. `Describe` Associate I valori elencati con ogni chiave LF-Tag sono i valori per i quali avete i permessi.

LF-Tag creator ha i permessi impliciti per vedere i tag LF che ha creato.

Gli amministratori di Data Lake possono vedere tutti i tag LF definiti nell' AWS account locale e tutti i tag LF per i quali sono state concesse le Associate autorizzazioni `Describe` e all'account locale da account esterni. L'amministratore del data lake può vedere tutti i valori di tutti i tag LF.

È possibile elencare i tag LF utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per elencare i tag LF (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come creatore del tag LF, come amministratore del data lake o come responsabile a cui sono state concesse le autorizzazioni per i tag LF e che dispone dell'autorizzazione IAM. `lakeformation:ListLFTags`

2. Nel pannello di navigazione, sotto Autorizzazioni, LF-tags e permessi, scegliete LF-Tags.

Viene visualizzata la pagina LF-Tags.

LF-Tags (2) Delete Edit Grant permissions Add LF-Tag

LF-Tags have a key and one or more values that can be associated with Data Catalog resources. [Learn more](#)

	Key	Values	Owner account ID	LF-Tag permissions
☐	test	111		View
☐	Department	Sales, Marketing		View

Controlla la colonna Owner account ID per determinare i tag LF condivisi con il tuo account da un account esterno.

AWS CLI

Per elencare i tag LF ()AWS CLI

- Esegui il comando seguente come amministratore del data lake o come responsabile a cui sono state concesse le autorizzazioni sui tag LF e che dispone dell'autorizzazione IAM.

lakeformation:ListLFTags

```
aws lakeformation list-lf-tags
```

L'output è simile a quello riportato di seguito.

```
{
  "LFTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "level",
      "TagValues": [
        "director",
        "vp",
        "c-level"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "Orders",
        "Sales",
        "Customers"
      ]
    }
  ]
}
```

Per vedere anche i tag LF concessi da account esterni, includi l'opzione di comando. --resource-share-type ALL

```
aws lakeformation list-lf-tags --resource-share-type ALL
```

L'output è simile a quello riportato di seguito. Notate la `NextToken` chiave, che indica che ce ne sono altre da elencare.

```
{
  "LFTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "level",
      "TagValues": [
        "director",
        "vp",
        "c-level"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "Orders",
        "Sales",
        "Customers"
      ]
    }
  ],
  "NextToken": "eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ=="
}
```

Ripetete il comando e aggiungete l' `--next-token` argomento per visualizzare gli eventuali tag LF e tag LF locali rimanenti concessi da account esterni. I tag LF degli account esterni si trovano sempre su una pagina separata.

```
aws lakeformation list-lf-tags --resource-share-type ALL
--next-token eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ==
```

```
{
  "LFTags": [
    {
      "CatalogId": "123456789012",
      "TagKey": "region",
      "TagValues": [
```



```

        "central",
        "south"
    ]
}
]
}

```

API

Puoi utilizzare l'opzione SDKs available for Lake Formation per elencare i tag che il richiedente è autorizzato a visualizzare.

```

import boto3

client = boto3.client('lakeformation')
...

response = client.list_lf_tags(
    CatalogId='string',
    ResourceShareType='ALL',
    MaxResults=50'
)

```

Questo comando restituisce un dict oggetto con la seguente struttura:

```

{
  'LFTags': [
    {
      'CatalogId': 'string',
      'TagKey': 'string',
      'TagValues': [
        'string',
      ]
    },
  ],
  'NextToken': 'string'
}

```

Per ulteriori informazioni sulle autorizzazioni richieste, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

Assegnazione di tag LF alle risorse del Data Catalog

È possibile assegnare LF-tags alle risorse del Data Catalog (database, tabelle e colonne) per controllare l'accesso a tali risorse. Solo i principali a cui sono concessi i tag LF corrispondenti (e i principali a cui è concesso l'accesso con il metodo della risorsa denominata) possono accedere alle risorse.

Se una tabella eredita un tag LF da un database o una colonna eredita un tag LF da una tabella, è possibile sovrascrivere il valore ereditato assegnando un nuovo valore alla chiave LF-Tag.

Il numero massimo di LF-Tag che è possibile assegnare a una risorsa è 50.

Argomenti

- [Requisiti per la gestione dei tag assegnati alle risorse](#)
- [Assegnate i tag LF a una colonna della tabella](#)
- [Assegnate i tag LF a una risorsa del catalogo dati](#)
- [Aggiornamento dei tag LF per una risorsa](#)
- [Rimozione del tag LF da una risorsa](#)

Requisiti per la gestione dei tag assegnati alle risorse

Per assegnare un tag LF a una risorsa del catalogo dati, è necessario:

- Ottieni il ASSOCIATE permesso di Lake Formation sull'LF-Tag.
- Avere l'autorizzazione `IAMLakeformation:AddLFTagsToResource`.
- Have Glue: GetDatabase autorizzazione su un database Glue.
- Sii il proprietario della risorsa (creatore), disponi dell'autorizzazione Super Lake Formation sulla risorsa con l'GRANTopzione o disponi delle seguenti autorizzazioni con l'GRANTopzione:
 - Per i database nello stesso AWS account: DESCRIBE, CREATE_TABLEALTER, e DROP
 - Per i database in un account esterno: DESCRIBE, CREATE_TABLE e ALTER
 - Per le tabelle (e le colonne): DESCRIBEALTER,DROP,INSERT,SELECT, e DELETE

Inoltre, il tag LF e la risorsa a cui viene assegnato devono trovarsi nello stesso AWS account.

Per rimuovere un LF-tag da una risorsa del Data Catalog, è necessario soddisfare questi requisiti e disporre anche dell'autorizzazione IAM. `lakeformation:RemoveLFTagsFromResource`

Assegnate i tag LF a una colonna della tabella

Per assegnare tag LF a una colonna della tabella (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come utente che soddisfa i requisiti sopra elencati.

2. Nel pannello di navigazione, seleziona Tabelle.
3. Scegli un nome per la tabella (non il pulsante di opzione accanto al nome della tabella).
4. Nella pagina dei dettagli della tabella, nella sezione Schema, scegli Modifica schema.
5. Nella pagina Modifica schema, seleziona una o più colonne, quindi scegli Modifica tag LF.

Note

Se intendete aggiungere o eliminare colonne e salvare una nuova versione, fatelo prima. Quindi modificate i tag LF.

Viene visualizzata la finestra di dialogo Modifica LF-Tag, che mostra tutti i tag LF ereditati dalla tabella.

**Edit LF-Tags: product_id** [Learn More](#) 

LF-Tags

After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys	Values
 level	director (inherited) ▼
 module	Orders (inherited) ▼

Assign new LF-Tag

You can add 50 more tags.

Cancel

Save

6. (Facoltativo) Per l'elenco dei valori accanto a un campo Chiavi ereditate, scegliete un valore per sostituire il valore ereditato.
7. (Facoltativo) Scegliete Assegna nuovo tag LF. Quindi, per Tasti assegnati, scegliete una chiave e per Valori, scegliete un valore per la chiave.

Edit LF-Tags: product_id [Learn More](#) ✕

LF-Tags
After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys

Values

director (inherited) ▼

Orders (inherited) ▼

Assigned keys

✕

Assign new LF-Tag

Values

Production ▲

Production

Development

Remove

You can add 49 more tags.

Cancel

Save

8. (Facoltativo) Scegliete nuovamente Assegna nuovo LF-tag per aggiungere un altro LF-tag.
9. Scegli Save (Salva).

Assegnate i tag LF a una risorsa del catalogo dati

Console

Per assegnare LF-Tags a un database o a una tabella del Data Catalog

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come utente che soddisfa i requisiti elencati in precedenza.

2. Nel riquadro di navigazione, in Catalogo dati, esegui una delle seguenti operazioni:
 - Per assegnare tag LF ai database, scegliete Database.
 - Per assegnare i tag LF alle tabelle, scegliete Tabelle.
3. Scegliete un database o una tabella e nel menu Azioni, scegliete Modifica tag LF.

Viene visualizzata la finestra di dialogo Modifica tag LF: **resource-name**.

Se una tabella eredita i tag LF dal database che la contiene, la finestra mostra i tag LF ereditati. Altrimenti, visualizza il testo «Non ci sono tag LF ereditati associati alla risorsa».

Edit LF-Tags: inventory
[Learn More](#)
✕

LF-Tags

After they are associated with catalog resources, LF-Tags allow you to create scalable permissions.

Inherited keys

Values

Assigned keys

✕

You can add 49 more tags.

Values

4. (Facoltativo) Se una tabella ha ereditato i tag LF, per l'elenco Valori accanto a un campo Chiavi ereditate, è possibile scegliere un valore per sostituire il valore ereditato.
5. Per assegnare nuovi LF-Tag, effettuate le seguenti operazioni:
 - a. Scegliete Assegna nuovo LF-tag.
 - b. Nel campo Chiavi assegnate, scegliete una chiave LF-Tag e nel campo Valori scegliete un valore.

- c. (Facoltativo) Scegliete nuovamente Assegna nuovo LF-tag per assegnare un tag LF aggiuntivo.
6. Scegli Save (Salva).

AWS CLI

Per assegnare tag LF a una risorsa del Data Catalog

- Esegui il comando `add-lf-tags-to-resource`.

L'esempio seguente assegna il tag LF `module=orders` alla tabella del database. `orders erp` Per l'argomento viene utilizzata la sintassi della scorciatoia. `--lf-tags` La `CatalogID` proprietà `for` `--lf-tags` è facoltativa. Se non viene fornito, viene utilizzato l'ID di catalogo della risorsa (in questo caso, la tabella).

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":  
  {"DatabaseName":"erp", "Name":"orders"}}' --lf-tags  
CatalogId=111122223333,TagKey=module,TagValues=orders
```

Quanto segue è l'output se il comando ha esito positivo.

```
{  
  "Failures": []  
}
```

L'esempio successivo assegna due tag LF alla `sales` tabella e utilizza la sintassi JSON per l'argomento. `--lf-tags`

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":  
  {"DatabaseName":"erp", "Name":"sales"}}' --lf-tags '[{"TagKey":  
  "module","TagValues": ["sales"]}, {"TagKey": "environment","TagValues":  
  ["development"]}']
```

L'esempio successivo assegna il tag LF alla colonna della tabella `level=director. total sales`

```
aws lakeformation add-lf-tags-to-resource --resource '{ "TableWithColumns":
{"DatabaseName":"erp", "Name":"sales", "ColumnNames":["total"]}}' --lf-tags
TagKey=level,TagValues=director
```

Aggiornamento dei tag LF per una risorsa

Per aggiornare un tag LF per una risorsa del catalogo dati (AWS CLI)

- Utilizzate il `add-lf-tags-to-resource` comando, come descritto nella procedura precedente.

L'aggiunta di un tag LF con la stessa chiave di un tag LF esistente, ma con un valore diverso, aggiorna il valore esistente.

Rimozione del tag LF da una risorsa

Per rimuovere un tag LF per una risorsa del Data Catalog (AWS CLI)

- Esegui il comando `remove-lf-tags-from-resource`.

Se una tabella ha un valore di tag LF che sostituisce il valore ereditato dal database principale, la rimozione di tale tag LF dalla tabella ripristina il valore ereditato. Questo comportamento si applica anche a una colonna che sostituisce i valori chiave ereditati dalla tabella.

L'esempio seguente rimuove il tag LF `level=director` dalla colonna della `total` tabella. `sales` La `CatalogID` proprietà per `--lf-tags` è facoltativa. Se non viene fornito, viene utilizzato l'ID di catalogo della risorsa (in questo caso, la tabella).

```
aws lakeformation remove-lf-tags-from-resource
--resource '{ "TableWithColumns":
{ "DatabaseName": "erp", "Name": "sales", "ColumnNames": [ "total" ] } }'
--lf-tags CatalogId=111122223333,TagKey=level,TagValues=director
```

Visualizzazione dei tag LF assegnati a una risorsa

È possibile visualizzare i tag LF assegnati a una risorsa del Data Catalog. È necessario disporre dell'`ASSOCIATE` autorizzazione `DESCRIBE` o dell'autorizzazione su un LF-Tag per visualizzarlo.

Console

Per visualizzare i tag LF assegnati a una risorsa (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, proprietario della risorsa o utente a cui sono state concesse le autorizzazioni di Lake Formation sulla risorsa.

2. Nel riquadro di navigazione, sotto l'intestazione Catalogo dati, esegui una delle seguenti operazioni:
 - Per visualizzare i tag LF assegnati a un database, scegliete Database.
 - Per visualizzare i tag LF assegnati a una tabella, scegliete Tabelle.
3. Nella pagina Tabelle o Database, scegliete il nome del database o della tabella. Quindi, nella pagina dei dettagli, scorri verso il basso fino alla sezione LF-Tags.

La schermata seguente mostra i tag LF assegnati a una customers tabella, che è contenuta nel database. retail Il module tag LF viene ereditato dal database. Alla credit_limit colonna è assegnato il tag LFlevel=vp.

LF-Tags (3)

Edit tags

LF-Tags are key-value pairs that you can assign to data catalog resources, such as databases, tables, and columns. You can then grant permissions to principals based on these tags to control access to the resources. Table columns inherit all LF-Tags that are assigned to the table. [Learn More](#)

Find tags

< 1 > ⚙

Resource ▲	Key ▼	Value ▼	Inherited from
customers (table)	module	Customers	retail
customers (table)	environment	Production	-
credit_limit (column)	level	vp	-

AWS CLI

Per visualizzare i tag LF assegnati a una risorsa ()AWS CLI

- Utilizzare un comando simile al seguente:

```
aws lakeformation get-resource-lf-tags --show-assigned-lf-tags --
resource '{ "Table": {"CatalogId":"111122223333", "DatabaseName":"erp",
"Name":"sales"}}'
```

Questo comando restituisce il seguente output.

```
{
  "TableTags": [
    {
      "CatalogId": "111122223333",
      "TagKey": "module",
      "TagValues": [
        "sales"
      ]
    },
    {
      "CatalogId": "111122223333",
      "TagKey": "environment",
      "TagValues": [
        "development"
      ]
    }
  ],
  "ColumnTags": [
    {
      "Name": "total",
      "Tags": [
        {
          "CatalogId": "111122223333",
          "TagKey": "level",
          "TagValues": [
            "director"
          ]
        }
      ]
    }
  ]
}
```

```
}
```

Questo output mostra solo i tag LF assegnati in modo esplicito, non ereditati. Se volete vedere tutti i tag LF su tutte le colonne, compresi i tag LF ereditati, omettete l'opzione. -- show-assigned-lf-tags

Visualizzazione delle risorse a cui è assegnato un LF-Tag

È possibile visualizzare tutte le risorse del Data Catalog a cui è assegnata una particolare chiave LF-Tag. A tale scopo, sono necessarie le seguenti autorizzazioni di Lake Formation:

- Describeo Associate sul tag LF.
- Describeo qualsiasi altra autorizzazione di Lake Formation sulla risorsa.

Inoltre, sono necessarie le seguenti autorizzazioni AWS Identity and Access Management (IAM):

- lakeformation:SearchDatabasesByLFTags
- lakeformation:SearchTablesByLFTags

Console

Per visualizzare le risorse a cui è assegnato un tag LF (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake o come utente che soddisfa i requisiti elencati in precedenza.

2. Nel pannello di navigazione, in Autorizzazioni e LF-Tag e autorizzazioni, scegli LF-Tags.
3. Scegliete una chiave LF-Tag (non il pulsante di opzione accanto al nome della chiave).

La pagina dei dettagli del tag LF mostra un elenco di risorse a cui è stato assegnato il tag LF.

module

LF-Tag

DeleteEdit

Key

module

Values

Orders, Sales, Customers

Associated data catalog resources (12)

Find resource

Key	Values	Resource type	Resource
module	Customers	DATABASE	retail
module	Customers	TABLE	customers
module	Orders	TABLE	inventory
module	Customers	COLUMN	customers.cust_first_name
module	Customers	COLUMN	customers.work_phone_number
module	Customers	COLUMN	customers.company_name
module	Customers	COLUMN	customers.credit_limit

AWS CLI

Per visualizzare le risorse a cui è assegnato un tag LF

- Esegui un comando `search-tables-by-lf-tags` or. `search-databases-by-lf-tags`

Example

L'esempio seguente elenca le tabelle e le colonne a cui è assegnato il `level=vp` tag LF. Per ogni tabella e colonna elencate, vengono emessi tutti i tag LF assegnati alla tabella o alla colonna, non solo l'espressione di ricerca.

```
aws lakeformation search-tables-by-lf-tags --expression  
TagKey=level,TagValues=vp
```

Per ulteriori informazioni sulle autorizzazioni richieste, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

Ciclo di vita di un tag LF

1. Il creatore del tag LF Michael crea un tag LF. `module=Customers`
2. Michael concede l'LF-Tag all'Associateingegnere dei dati Eduardo. Concedere sovvenzioni implicitamente. `Associate Describe`
3. Michael concede `Super Custs` a Eduardo l'opzione `grant`, in modo che Eduardo possa assegnare dei tag LF al tavolo. Per ulteriori informazioni, consulta [Assegnazione di tag LF alle risorse del Data Catalog](#).
4. Eduardo assegna il tag LF alla tabella. `module=customers Custs`
5. Michael concede la seguente concessione all'ingegnere dei dati Sandra (in pseudo-codice).

```
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=customers TO Sandra WITH GRANT OPTION
```

6. Sandra concede la seguente concessione all'analista di dati Maria.

```
GRANT (SELECT ON TABLES) ON TAGS module=customers TO Maria
```

Maria ora può eseguire interrogazioni sul tavolo. `Custs`

Consulta anche

- [Controllo dell'accesso ai metadati](#)

Confronto tra il controllo degli accessi basato su tag Lake Formation e il controllo degli accessi basato sugli attributi IAM

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 alle risorse IAM, incluse le entità IAM (utenti o ruoli) e alle AWS risorse. È possibile creare una singola policy ABAC o un piccolo insieme di policy per i principali IAM. Queste policy ABAC possono essere definite affinché autorizzino le operazioni quando il tag dell'entità corrisponde al tag della risorsa. La strategia ABAC è utile in ambienti soggetti a una rapida crescita e aiuta in situazioni in cui la gestione delle policy diventa impegnativa.

I team di sicurezza e governance del cloud utilizzano IAM per definire le policy di accesso e le autorizzazioni di sicurezza per tutte le risorse, inclusi i bucket Amazon S3, le istanze EC2 Amazon e qualsiasi risorsa a cui puoi fare riferimento con un ARN. Le policy IAM definiscono autorizzazioni ampie (granulari) per le risorse del data lake, ad esempio per consentire o negare l'accesso a livello di bucket o prefisso Amazon S3 o a livello di database. [Per ulteriori informazioni su IAM ABAC, consulta *A cosa serve ABAC? AWS*](#) nella Guida per l'utente di IAM.

Ad esempio, è possibile creare tre ruoli associando a essi un tag con la chiave `project-access` e impostando il valore del tag del primo ruolo a `Dev`, del secondo a `Marketing` e del terzo a `Support`. Assegna tag con il valore appropriato alle risorse. È quindi possibile utilizzare una singola policy che consenta l'accesso quando il ruolo e la risorsa sono contrassegnati con lo stesso valore del tag `project-access`.

I team di governance dei dati utilizzano Lake Formation per definire autorizzazioni granulari per specifiche risorse di data lake. I tag LF sono assegnati alle risorse del Data Catalog (database, tabelle e colonne) e sono concessi ai principali. Un principale con tag LF che corrispondono ai tag LF di una risorsa può accedere a quella risorsa. Le autorizzazioni Lake Formation sono secondarie rispetto alle autorizzazioni IAM. Ad esempio, se le autorizzazioni IAM non consentono a un utente di accedere a un data lake, Lake Formation non concede l'accesso a nessuna risorsa all'interno di quel data lake a quell'utente, anche se il principale e la risorsa hanno tag LF corrispondenti.

Il controllo degli accessi basato su tag di Lake Formation (LF-TBAC) funziona con IAM ABAC per fornire livelli aggiuntivi di autorizzazioni per i dati e le risorse di Lake Formation.

- Le autorizzazioni TBAC di Lake Formation si adattano all'innovazione. Non è più necessario che un amministratore aggiorni le policy esistenti per consentire l'accesso a nuove risorse. Ad esempio, supponiamo di utilizzare una strategia IAM ABAC con il `project-access` tag per fornire l'accesso a database specifici all'interno di Lake Formation. Utilizzando LF-TBAC, il tag

LF Project=SuperApp viene assegnato a tabelle o colonne specifiche e lo stesso tag LF viene concesso a uno sviluppatore per quel progetto. Tramite IAM, lo sviluppatore può accedere al database e le autorizzazioni LF-TBAC garantiscono allo sviluppatore un ulteriore accesso a tabelle o colonne specifiche all'interno delle tabelle. Se viene aggiunta una nuova tabella al progetto, l'amministratore di Lake Formation deve solo assegnare il tag alla nuova tabella affinché lo sviluppatore possa accedere alla tabella.

- Lake Formation TBAC richiede meno policy IAM. Poiché utilizzi le policy IAM per garantire un accesso di alto livello alle risorse di Lake Formation e Lake Formation TBAC per gestire un accesso più preciso ai dati, crei meno policy IAM.
- Utilizzando Lake Formation TBAC, i team possono cambiare e crescere rapidamente. Questo perché le autorizzazioni per le nuove risorse vengono concesse automaticamente in base agli attributi. Ad esempio, se un nuovo sviluppatore si unisce al progetto, è facile concedere a tale sviluppatore l'accesso associando il ruolo IAM all'utente e quindi assegnandogli i tag LF richiesti. Non è necessario modificare la policy IAM per supportare un nuovo progetto o creare nuovi LF-tag.
- Autorizzazioni più dettagliate sono possibili utilizzando Lake Formation TBAC. Le policy IAM garantiscono l'accesso alle risorse di primo livello, come i database o le tabelle del Data Catalog. Utilizzando Lake Formation TBAC, puoi concedere l'accesso a tabelle o colonne specifiche che contengono valori di dati specifici.

Note

I tag IAM non sono gli stessi dei tag LF. Questi tag non sono intercambiabili. I tag LF vengono utilizzati per concedere i permessi di Lake Formation e i tag IAM vengono utilizzati per definire le politiche IAM.

Gestione delle espressioni LF-Tag per il controllo dell'accesso ai metadati

Le espressioni LF-tag sono espressioni logiche composte da uno o più tag LF (coppie chiave-valore) utilizzate per concedere autorizzazioni sulle risorse. AWS Glue Data Catalog Le espressioni LF-Tag consentono di definire regole che regolano l'accesso alle risorse di dati in base ai relativi tag di metadati. È possibile salvare queste espressioni e riutilizzarle in più autorizzazioni, garantendo la coerenza e semplificando la gestione delle modifiche all'ontologia dei tag nel tempo.

All'interno di una determinata espressione del tag LF, le chiavi dei tag vengono combinate utilizzando l'operazione AND, mentre i valori vengono combinati utilizzando l'operazione OR. Ad esempio,

l'espressione `tag content_type:Sales AND location:US` rappresenta risorse relative ai dati di vendita negli Stati Uniti.

È possibile creare fino a 1000 espressioni di tag LF in un Account AWS. Queste espressioni forniscono un modo flessibile e scalabile per gestire le autorizzazioni basate sui tag di metadati, garantendo che solo gli utenti o le applicazioni autorizzati possano accedere a risorse di dati specifiche in base alle regole di tag definite.

Le espressioni LF-Tag offrono i seguenti vantaggi:

- **Riusabilità:** definendo e salvando le espressioni dei tag LF, non è più necessario replicare manualmente le stesse espressioni quando si assegnano le autorizzazioni ad altre risorse o principali.
- **Coerenza:** il riutilizzo delle espressioni LF-Tag su più concessioni di autorizzazioni garantisce la coerenza nel modo in cui le autorizzazioni vengono concesse e gestite.
- **Gestione dell'ontologia dei tag:** le espressioni LF-Tag aiutano a gestire le modifiche all'ontologia dei tag nel tempo, poiché è possibile aggiornare le espressioni salvate anziché modificare le concessioni di autorizzazioni individuali.

Per ulteriori informazioni sul controllo degli accessi basato su tag, consulta [Controllo degli accessi basato su tag Lake Formation](#)

Creatori di espressioni con tag LF

Il creatore di espressioni LF-Tag è un preside che dispone delle autorizzazioni per creare e gestire espressioni LF-Tag. Gli amministratori di Data Lake possono aggiungere creatori di espressioni LF-Tag utilizzando la console, la CLI, l'API o l'SDK di Lake Formation. I creatori di espressioni LF-Tag hanno i permessi impliciti di Lake Formation per creare, aggiornare ed eliminare espressioni LF-Tag e per concedere i permessi di espressione LF-Tag ad altri principali.

I creatori di espressioni LF-Tag che non sono amministratori di Data Lake ricevono autorizzazioni implicite e solo per le espressioni che hanno creato. `Alter Drop Describe Grant with LF-Tag expression`

Gli amministratori di Data Lake possono anche concedere autorizzazioni valide ai creatori di espressioni LF-Tag. `Create LF-Tag expression` Quindi, il creatore di espressioni LF-Tag può concedere il permesso di creare espressioni LF-Tag ad altri responsabili.

Argomenti

- [Autorizzazioni IAM necessarie per creare espressioni con tag LF](#)
- [Aggiungi i creatori di espressioni LF-Tag](#)
- [Creazione di espressioni con tag LF](#)
- [Aggiornamento delle espressioni del tag LF](#)
- [Eliminazione delle espressioni dei tag LF](#)
- [Elenco delle espressioni dei tag LF](#)

 Consulta anche

- [Gestione dei permessi per i valori del tag LF](#)
- [Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC](#)
- [Controllo degli accessi basato su tag Lake Formation](#)

Autorizzazioni IAM necessarie per creare espressioni con tag LF

È necessario configurare le autorizzazioni per consentire a un principale di Lake Formation di creare espressioni di tag LF. Aggiungi la seguente dichiarazione alla politica di autorizzazione per il principale che deve essere un creatore di espressioni LF-Tag.

 Note

Sebbene gli amministratori dei data lake abbiano le autorizzazioni implicite di Lake Formation per creare, aggiornare ed eliminare i tag LF e le espressioni dei tag LF, per assegnare i tag LF alle risorse e per concedere l'autorizzazione alle espressioni LF-Tag e LF-Tag ai principali, gli amministratori del data lake necessitano anche delle seguenti autorizzazioni IAM.

Per ulteriori informazioni, consulta [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

```
{  
  "Sid": "Transformational",  
  "Effect": "Allow",  
  "Action": [  

```



```
    "lakeformation:AddLFTagsToResource",
    "lakeformation:RemoveLFTagsFromResource",
    "lakeformation:GetResourceLFTags",
    "lakeformation:ListLFTags",
    "lakeformation:CreateLFTag",
    "lakeformation:GetLFTag",
    "lakeformation:UpdateLFTag",
    "lakeformation>DeleteLFTag",
    "lakeformation:SearchTablesByLFTags",
    "lakeformation:SearchDatabasesByLFTags",
    "lakeformation:CreateLFTagExpression",
    "lakeformation>DeleteLFTagExpression",
    "lakeformation:UpdateLFTagExpression",
    "lakeformation:GetLFTagExpression",
    "lakeformation:ListLFTagExpressions",
    "lakeformation:GrantPermissions",
    "lakeformation:RevokePermissions",
    "lakeformation:BatchGrantPermissions",
    "lakeformation:BatchRevokePermissions"
  ]
}
```

Aggiungi i creatori di espressioni LF-Tag

I creatori di espressioni LF-Tag possono creare e salvare espressioni di tag LF riutilizzabili, aggiornare chiavi e valori dei tag, eliminare espressioni e concedere autorizzazioni sulle risorse del Data Catalog ai principali utilizzando il metodo LF-TBAC. Il creatore di espressioni LF-Tag può anche concedere queste autorizzazioni ai principali.

È possibile creare ruoli di creatore di espressioni LF-Tag utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

console

Per aggiungere un creatore di espressioni LF-Tag

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
Accedi come amministratore del data lake.
2. Nel riquadro di navigazione, in Autorizzazioni, scegli LF-tags e permessi.
3. Scegliete la scheda Espressioni LF-Tag.

4. Nella sezione Creatori di espressioni con tag LF, scegliete Aggiungi creatori di espressioni con tag LF.

Add LF-Tag expression creators

LF-Tag expression creators can create and manage LF-Tag expressions.

LF-Tag expression creator details

IAM users and roles
Add IAM users or roles.

Choose IAM principals to add

datalake_user X

User

Permission
Choose the permission to grant.

☒ Create LF-Tag expression

Grantable permission
Choose the permission that may be granted to others.

☒ Create LF-Tag expression

[Cancel](#)
[Add](#)

5. Nella pagina Aggiungi creatori di espressioni LF-Tag, scegli un ruolo o un utente IAM con le autorizzazioni necessarie per creare espressioni LF-Tag.
6. Seleziona Create LF-Tag expression la casella di controllo delle autorizzazioni.
7. (Facoltativo) Per consentire ai responsabili selezionati di concedere l>Create LF-Tag expression autorizzazione ai mandanti, scegli Autorizzazione Create LF-Tag expression concedibile.
8. Scegliere Aggiungi.

AWS CLI

```
aws lakeformation grant-permissions --cli-input-json file://grantCreate
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:user/tag-manager"
  },
  "Resource": {
    "Catalog": {}
  }
}
```

```

    },
    "Permissions": [
        "CreateLFTagExpression"
    ],
    "PermissionsWithGrantOption": [
        "CreateLFTagExpression"
    ]
}

```

Il ruolo di creatore di espressioni LF-Tag consente di creare, aggiornare o eliminare espressioni LF-Tag.

Autorizzazione	Description
Create	Un principale con questa autorizzazione può aggiungere espressioni di tag LF nel data lake.
Drop	Un principale con questa autorizzazione su un'espressione LF-Tag può eliminare un'espressione LF-Tag dal data lake.
Alter	Un principale con questa autorizzazione su un'espressione LF-Tag può aggiornare il corpo dell'espressione di un'espressione LF-Tag.
Describe	Un principale con questa autorizzazione su un'espressione LF-Tag può visualizzare il contenuto di un'espressione LF-Tag.
Grant with LF-Tag expression	Questa autorizzazione consente al destinatario di utilizzare l'espressione del tag come risorsa per concedere le autorizzazioni di accesso ai dati o ai metadati. Concedere concessioni Grant with LF-Tag expression implicite. Describe
Super	Per le espressioni LF-Tag, l'Superautorizzazione concede la possibilità di Describe AlterDrop, e concede le autorizzazioni sull'espressione del tag ad altri principali.

Queste autorizzazioni sono concesse. Un preside a cui sono state concesse queste autorizzazioni con l'opzione di concessione può concederle ad altri committenti.

Creazione di espressioni con tag LF

È necessario definire tutti i tag LF in Lake Formation e assegnarli alle risorse del Data Catalog prima che possano essere utilizzati per creare espressioni. Un'espressione LF-Tag è composta da un'altra chiave e uno o più valori possibili per ogni chiave.

Dopo che l'amministratore del data lake ha impostato le autorizzazioni IAM e le autorizzazioni Lake Formation richieste per il ruolo di creatore di espressioni LF-Tag, il responsabile può creare espressioni LF-Tag riutilizzabili. Il creatore dell'espressione LF-Tag ottiene le autorizzazioni implicite per aggiornare il corpo dell'espressione ed eliminare l'espressione LF-Tag.

È possibile creare espressioni LF-Tag utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per creare un'espressione LF-Tag

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come principale con i permessi di creazione di espressioni LF-Tag o come amministratore del data lake.

2. Nel pannello di navigazione, sotto Autorizzazioni, scegli LF-Tag e autorizzazioni.
3. Scegliete le espressioni LF-Tag. Viene visualizzata la pagina Aggiungi espressioni con tag LF.

Add LF-Tag Expression

LF-Tag expression creators can create and manage LF-Tag expressions

Expression details

Name

Enter a name that describes the expression. Expression name cannot be edited after creation.

Name must be less than 1000 characters.

Description - optional

Description can be up to 2048 characters.

Expression

Choose the keys and values for this expression. When multiple keys are specified, the keys are joined by an AND operator and when multiple values are specified, the values are joined by an OR operator.

Key

Values

Add LF-Tag key-value pair

You can add 49 more LF-Tags.

Expression review

The LF-Tag expression above will be interpreted in the following way.

Department = sales

► Grant permissions - optional

Grant permissions to allow others to manage or grant permissions with the expression.

Cancel

Add

4. Immetti le seguenti informazioni:

- Nome: immettere un nome univoco per l'espressione. Non è possibile aggiornare il nome dell'espressione.
- Descrizione: fornisce una descrizione facoltativa dell'espressione con i dettagli dell'espressione.
- Espressione: crea l'espressione specificando le chiavi dei tag e i valori associati. È possibile aggiungere fino a 50 chiavi per espressione. È necessario disporre

dell'autorizzazione Grant with LF-Tags Lake Formation per tutti i tag nel corpo dell'espressione.

Ogni chiave deve avere almeno un valore. Per inserire più valori, inserisci un elenco delimitato da virgole e premi Invio oppure inserisci un valore alla volta e scegli Aggiungi dopo ognuno di essi. Il numero massimo di valori consentiti per chiave è 1000.

Lake Formation utilizza la AND/OR logica per combinare più chiavi e valori in un'espressione. All'interno di una singola coppia (chiave: elenco di valori), i valori vengono combinati utilizzando l'operatore logico OR. Ad esempio, se la coppia è (Dipartimento: [Vendite, Marketing]), significa che il tag corrisponde se la risorsa ha il tag Department con valore Sales OR Marketing.

Quando si specificano più chiavi, le chiavi vengono unite da un operatore logico AND. Quindi, se l'espressione completa è (Department: [Sales, Marketing]) AND (Location: [US, Canada]), corrisponde alle risorse che hanno il tag Department con valore Sales OR Marketing e hanno anche il tag Location con valore US OR Canada. Di seguito è riportato un altro esempio con più chiavi e valori:

Espressione LF-tag: (ContentType : [Video, Audio]) AND (Regione: [Europa, Asia]) AND (Dipartimento: [Ingegneria, ProductManagement]).

Questa espressione corrisponderebbe a risorse che hanno: - Il ContentType tag con valore Video OR Audio AND - Il tag Region con valore Europe OR Asia AND - Il tag Department con valore Engineering OR. ProductManagement

È inoltre possibile salvare un'espressione di tag quando si concedono le autorizzazioni per il data lake utilizzando i tag LF. Scegliete le coppie chiave e valore e scegliete l'opzione Salva come nuova espressione. Immettete un nome che descriva l'espressione.

LF-Tags or catalog resources

Choose a method to grant permissions.

☒ **Resources matched by LF-Tags (recommended)**

Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☐ **Named Data Catalog resources**

Manage permissions for specific databases or tables, in addition to fine-grained data access.

☒ **LF-Tag key-value pairs**

☐ **Saved LF-Tag expressions - *new***

Key

Department ▼

Values

Choose LF-Tag values ▼

marketing ✕

sales ✕

Add LF-Tag key-value pair

You can add 49 more LF-Tags.

Expression review

The LF-Tag expression above will be interpreted in the following way.

Department = (marketing OR sales)

☒ **Save as new expression**

Use saved expressions to grant permissions. Create LF-Tag expression permissions are needed.

New LF-Tag expression name

Enter a name that describes the expression. Expression name cannot be edited after creation.

General access

Name must be less than 1000 characters.

5. (Facoltativo) Quindi, scegli gli utenti/ruoli e le autorizzazioni sull'espressione che desideri concedere loro nell'account. Puoi anche scegliere autorizzazioni concedibili che consentano agli utenti di concedere tali autorizzazioni ad altri utenti dell'account. Non puoi concedere autorizzazioni tra account per le espressioni dei tag.

▼ Grant permissions - optional

Grant permissions to allow others to manage or grant permissions with the expression.

IAM users and roles

Users or roles from this AWS account.

Choose IAM principals to add

Permissions

Choose the specific LF-Tag permissions to grant.

- ☐ **Describe**
See keys and values.
- ☐ **Alter**
Update or delete LF-Tag expressions.
- ☐ **Drop**
Delete LF-Tag expressions.
- ☐ **Grant with LF-Tag expression**
Allow principals to grant access permissions using LF-Tag expressions.
- ☐ **Super**
This permission supersedes the individual permissions set above.

Grantable permissions

Choose the permissions that the recipient can grant to other principals.

- ☐ **Describe**
See keys and values.
- ☐ **Alter**
Update or delete LF-Tag expressions.
- ☐ **Drop**
Delete LF-Tag expressions.
- ☐ **Grant with LF-Tag expression**
Allow principals to grant access permissions using LF-Tag expressions.
- ☐ **Super**
This permission supersedes the individual permissions set above.

Cancel

Add

6. Scegliere Aggiungi.

AWS CLI

Per creare un'espressione LF-tag

- Immettete un comando `create-lf-tag-expression`.

L'esempio seguente crea un'espressione LF-tag con il tag `Department` con valori `Sales` e `Marketing` il tag `Location` con il valore `US`

```
aws lakeformation create-lf-tag-expression \
```



```
-- name "my-tag-expression" \  
-- catalog-id "123456789012" \  
-- expression '{"Expression":[{"TagKey":"Department","TagValues":  
["Sales","Marketing"]}, {"TagKey":"Location","TagValues":["US"]}]}'
```

Questo comando CLI crea una nuova espressione LF-tag in AWS Glue Data Catalog. L'espressione può essere utilizzata per concedere autorizzazioni a risorse del catalogo dati come database, tabelle, viste o colonne in base ai tag associati. In questo esempio, l'espressione corrisponderà alle risorse che hanno la `Department` chiave con valori `Sales` o `Marketing` e la `Location` chiave con il valore `US`.

In qualità di creatore di espressioni di tag, il responsabile ottiene l'`Alter` autorizzazione per questa espressione di tag LF e può aggiornare o rimuovere l'espressione. Il responsabile del creatore dell'espressione LF-Tag può anche concedere `Alter` il permesso a un altro principale di aggiornare e rimuovere questa espressione.

Aggiornamento delle espressioni del tag LF

Solo gli amministratori del data lake, il creatore dell'espressione LF-Tag e i responsabili che dispongono `Alter` o hanno l'`Super` autorizzazione sull'espressione LF-Tag possono aggiornare un'espressione LF-Tag. Oltre all'`Alter` autorizzazione, è necessaria anche l'autorizzazione `lakeformation:UpdateLFTagExpression` IAM e l'`Grant with LF-Tag` autorizzazione su tutte le chiavi-valori sottostanti sul nuovo corpo dell'espressione per aggiornare le espressioni.

È possibile aggiornare un'espressione LF-Tag aggiornando la descrizione, il corpo dell'espressione e le autorizzazioni concesse all'espressione. Non è possibile modificare il nome dell'espressione LF-Tag. Per cambiare il nome, eliminate l'espressione del tag LF e aggiungetene una con i parametri richiesti.

È possibile aggiornare un'espressione LF-Tag utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per aggiornare un'espressione LF-Tag

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore del tag LF o principale con l'`Alter` autorizzazione per l'LF-Tag.

2. Nel pannello di navigazione, sotto autorizzazioni, scegli LF-Tag e autorizzazioni.
3. Scegliete la scheda Espressioni LF-Tag.
4. Nella sezione Espressioni di tag LF, selezionate un'espressione di tag LF, quindi scegliete Modifica.
5. Nella finestra di dialogo Modifica espressione LF-tag, aggiornate la descrizione e aggiorna il corpo dell'espressione aggiungendo o rimuovendo chiavi e valori.

Per aggiungere più valori, nel campo Valori, scegliete i valori dal menu a discesa.

6. Scegli Save (Salva).

AWS CLI

Il `update-lf-tag-expression` comando in Lake Formation consente di aggiornare un'espressione di tag LF esistente.

```
aws lakeformation update-lf-tag-expression \  
-- name expression_name \  
-- description new_description \  
-- catalog-id catalog_id \  
-- expression '{"Expression": [{"TagKey": "tag_key", "TagValues": ["tag_value1",  
"tag_value2", ...]}}'
```

Ecco cosa significano i parametri del comando fornito:

- `name`: il nome dell'espressione di tag denominata esistente che desideri aggiornare.
- `description` — Una nuova descrizione per l'espressione.

`catalog-id` — L'ID del Data Catalog in cui risiede l'espressione del tag denominata.

- `expression` — La nuova stringa di espressioni di tag con cui si desidera aggiornare l'espressione.

Eliminazione delle espressioni dei tag LF

È possibile eliminare espressioni di tag LF che non sono più in uso. Se hai concesso le autorizzazioni ai principali sulle risorse del Data Catalog utilizzando l'espressione LF-Tag, non avranno più le autorizzazioni.

Solo gli amministratori del data lake, il creatore dell'espressione LF-Tag o un responsabile con l'Dropautorizzazione sull'espressione LF-Tag possono eliminare un'espressione LF-Tag. Oltre all'autorizzazione, il principale necessita anche dell'Dropautorizzazione lakeformation:DeleteLFTagExpression IAM per eliminare un'espressione LF-Tag.

È possibile eliminare un'espressione LF-Tag utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

Per eliminare un'espressione LF-Tag (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore dell'espressione LF-Tag o principale che dispone delle autorizzazioni per eliminare l'espressione.

2. Nel pannello di navigazione, in Autorizzazioni, scegli LF-Tag e autorizzazioni.
3. Scegliete la scheda delle espressioni LF-Tag.
4. Nella sezione Espressioni di tag LF, selezionate un'espressione di tag LF, quindi scegliete Elimina.
5. Nell'espressione Elimina tag LF? nella finestra di dialogo, per confermare l'eliminazione, inserite il nome dell'espressione LF-Tag nel campo designato, quindi scegliete Elimina.

AWS CLI

Per eliminare un tag LF ()AWS CLI

- Immettete un comandodelete-lf-tag-expression. Fornite il nome dell'espressione e l'ID del catalogo da eliminare.

Example

L'esempio seguente elimina l'espressione LF-Tag con il nome my-tag-expression dal Data Catalog with ID. 123456789012 Il catalog-id parametro è facoltativo se utilizzi lo stesso account della configurazione. AWS CLI Dopo aver eliminato un'espressione LF-tag, Lake Formation pulisce i record di autorizzazione associati per quell'espressione. Ciò include sia i record di autorizzazioni individuali che i record di autorizzazioni aggregati che contengono l'espressione eliminata.

```
aws lakeformation delete-lf-tag-expression \  
--name "my-tag-expression" \  
--catalog-id "123456789012"
```

Elenco delle espressioni dei tag LF

Puoi elencare le espressioni del tag LF per le quali hai i permessi Descrivi. Gli amministratori di Data Lake, i creatori di espressioni LF-Tag e gli amministratori di sola lettura possono vedere implicitamente tutte le espressioni di tag nel proprio account.

È possibile elencare le espressioni LF-Tag utilizzando la console, l' AWS Lake Formation API o (). AWS Command Line Interface AWS CLI

Console

Per elencare le espressioni dei tag LF (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come creatore di espressioni LF-Tag, come amministratore del data lake o come responsabile a cui sono state concesse le autorizzazioni per le espressioni LF-Tag e che dispone dell'autorizzazione IAM. `lakeformation:ListLFTagExpressions`

2. Nel pannello di navigazione, sotto Autorizzazioni, LF-tags e permessi.
3. Scegliete la scheda Espressioni LF-Tag per vedere le espressioni. Questa sezione mostra le informazioni sulle espressioni di tag LF esistenti, incluso il nome dell'espressione, l'espressione stessa con i collegamenti ai tag inclusi e le opzioni per creare, modificare o eliminare le espressioni.

AWS CLI

Per elencare i tag LF ()AWS CLI

- Per elencare le espressioni dei tag LF usando il AWS CLI, puoi usare il comando. `list-lf-tag-expressions` La sintassi della richiesta è:

```
aws lakeformation list-lf-tag-expressions \  
-- catalog-id "123456789012" \  
-- max-items "100" \  

```

```
-- next-token "next-token"
```

Dove:

- `catalog-id` è l'ID dell'AWS account del Data Catalog per il quale desideri elencare le espressioni dei tag.
- `max-items` specifica il numero massimo di espressioni di tag da restituire. Se questo parametro non viene utilizzato, il valore predefinito è 100.
- `next-token` è un token di continuazione se i risultati sono stati troncati in una richiesta precedente.

La risposta includerà un elenco di espressioni di tag LF e un token successivo, se applicabile.

Gestione dei permessi per i valori del tag LF

È possibile concedere le `Drop` `Alter` autorizzazioni sui tag LF ai responsabili per gestire le espressioni di valore dei tag LF. È inoltre possibile concedere `Describe` `Grant with LF-Tag expressions` le autorizzazioni relative ai tag LF ai principali per visualizzare i tag LF e assegnarli alle risorse del Data Catalog (database, tabelle e colonne). Associate Quando i tag LF vengono assegnati alle risorse del Data Catalog, è possibile utilizzare il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) per proteggere tali risorse. Per ulteriori informazioni, consulta [Controllo degli accessi basato su tag Lake Formation](#).

È possibile concedere queste autorizzazioni con l'opzione di concessione in modo che altri responsabili possano concederle. Le Associate autorizzazioni `Grant with LF-Tag expressions` `Describe`, e sono spiegate in [Aggiungi creatori di tag LF](#)

È possibile concedere le Associate autorizzazioni `Describe` and su un tag LF a un account esterno. AWS Un amministratore di data lake in quell'account può quindi concedere tali autorizzazioni ad altri responsabili dell'account. I responsabili a cui l'amministratore del data lake dell'account esterno concede l'Associate autorizzazione possono quindi assegnare LF-Tags alle risorse del Data Catalog che hai condiviso con il loro account.

Quando si concede a un account esterno, è necessario includere l'opzione di concessione.

Puoi concedere le autorizzazioni sui tag LF utilizzando la console Lake Formation, l'API o il AWS Command Line Interface ().AWS CLI

Argomenti

- [Elencare le autorizzazioni LF-Tag utilizzando la console](#)
- [Concessione delle autorizzazioni LF-Tag tramite la console](#)
- [Gestione delle autorizzazioni LF-Tag utilizzando AWS CLI](#)

Per ulteriori informazioni, consulta [Gestione dei tag LF per il controllo dell'accesso ai metadati](#) e [Controllo degli accessi basato su tag Lake Formation](#).

Elencare le autorizzazioni LF-Tag utilizzando la console

Puoi usare la console Lake Formation per visualizzare le autorizzazioni concesse sui tag LF. Devi essere un creatore di LF-Tag, un amministratore di data lake o avere l'Associate autorizzazione Describe o l'autorizzazione su un LF-Tag per vederlo.

Per elencare le autorizzazioni LF-Tag (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come creatore di LF-tag, amministratore del data lake o come utente a cui sono state concesse le Drop, Alter Associate, o Describe le autorizzazioni per i tag LF.

2. Nel pannello di navigazione, in Autorizzazioni, scegli LF-Tag e permessi e scegli la sezione Autorizzazioni LF-Tag.

La sezione delle autorizzazioni LF-Tag mostra una tabella che contiene il principale, le chiavi dei tag, i valori e le autorizzazioni.

LF-Tags LF-Tag permissions LF-Tag creators - new																																																															
LF-Tag permissions (6) View and manage the permissions granted on LF-Tags. Learn more View Revoke Grant permissions Find permissions by LF-Tag key and value < 1 > ⚙ <table> <tr> <th></th><th>Principal ▲</th><th>Principal type ▼</th><th>Keys ▼</th><th>Values ▼</th><th>LF-Tag permissions ▼</th><th>LF-Tag value permissions ▼</th><th>Grantable ▼</th></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>module</td><td>All values</td><td>Alter, Drop</td><td>-</td><td>Alter, Drop</td></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>module</td><td>All values</td><td>-</td><td>Describe</td><td>Describe</td></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>module</td><td>All values</td><td>-</td><td>Associate</td><td>Associate</td></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>module</td><td>All values</td><td>-</td><td>Grant with LF-Tag expression</td><td>Grant with LF-Tag expression</td></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>LF-Test</td><td>All values</td><td>-</td><td>Describe</td><td>Describe</td></tr> <tr> <td>☐</td><td>arn:aws:iam::[redacted]:role/Admin</td><td>IAM role</td><td>LF-Test</td><td>All values</td><td>-</td><td>Associate</td><td>Associate</td></tr> </table>									Principal ▲	Principal type ▼	Keys ▼	Values ▼	LF-Tag permissions ▼	LF-Tag value permissions ▼	Grantable ▼	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	Alter, Drop	-	Alter, Drop	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Describe	Describe	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Associate	Associate	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Grant with LF-Tag expression	Grant with LF-Tag expression	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Describe	Describe	☐	arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Associate	Associate
	Principal ▲	Principal type ▼	Keys ▼	Values ▼	LF-Tag permissions ▼	LF-Tag value permissions ▼	Grantable ▼																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	Alter, Drop	-	Alter, Drop																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Describe	Describe																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Associate	Associate																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	module	All values	-	Grant with LF-Tag expression	Grant with LF-Tag expression																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Describe	Describe																																																								
☐	arn:aws:iam::[redacted]:role/Admin	IAM role	LF-Test	All values	-	Associate	Associate																																																								

Concessione delle autorizzazioni LF-Tag tramite la console

I passaggi seguenti spiegano come concedere le autorizzazioni sui tag LF utilizzando la pagina Grant LF-tag permissions sulla console Lake Formation. La pagina è suddivisa nelle seguenti sezioni:

- Tipi di autorizzazione: il tipo di autorizzazione da concedere.
- Principali: gli utenti o i ruoli IAM o gli utenti o i ruoli SAML a cui concedere le autorizzazioni.
- Autorizzazioni per le coppie chiave-valore LF-Tag: le coppie chiave-valore LF-Tag su cui concedere le autorizzazioni.
- Autorizzazioni LF-Tag — I tag LF su cui concedere le autorizzazioni.
- Autorizzazioni per le espressioni LF-Tag — Autorizzazioni per le espressioni LF-Tag su cui concedere le autorizzazioni.
- Autorizzazioni: le autorizzazioni da concedere.

Apri la pagina Concedi le autorizzazioni del tag LF

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come creatore di LF-Tag, amministratore di data lake o come utente. Le autorizzazioni LF-Tag o la coppia chiave-valore LF-Tag su LF-Tag sono state concesse con l'opzione Grant.

2. Nel pannello di navigazione, scegli LF-Tag e autorizzazioni, scegli la sezione Autorizzazioni LF-Tag.
3. Scegli Concedi autorizzazioni.

Specificate il tipo di autorizzazione

Nella sezione Tipo di autorizzazione, scegli un tipo di autorizzazione.

Autorizzazioni LF-Tag

Scegli le autorizzazioni LF-Tag per consentire ai mandanti di aggiornare i valori dei tag LF o eliminare i tag LF.

Autorizzazioni per la coppia chiave-valore LF-Tag

Scegliete i permessi della coppia chiave-valore LF-Tag per consentire ai mandanti di assegnare LF-Tags alle risorse del Data Catalog, visualizzare i tag e i valori LF-Tag e concedere ai principali autorizzazioni basate sui tag LF sulle risorse del Data Catalog.

Le opzioni disponibili nelle seguenti sezioni dipendono dal tipo di autorizzazione.

Autorizzazioni di espressione LF-tag

Scegliete i permessi di espressione LF-Tag per consentire ai principali di aggiornare le espressioni o eliminare le espressioni.

Specificate i principali

Note

Non puoi concedere autorizzazioni LF-Tag (AlterandDrop) ad account o responsabili esterni di un altro account.

Nella sezione Principali, scegli un tipo principale e specifica i principali a cui concedere le autorizzazioni.

Principals

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles
Add one or more IAM users or roles.

Choose IAM principals to add ▼

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o ARNs per utenti o gruppi Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

 Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per AWS account, inserisci uno o più AWS account validi IDs. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4 a 32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

Per il principale IAM, inserisci l'ARN per l'utente o il ruolo IAM.

Specificate i tag LF

Per concedere le autorizzazioni sui tag LF, nella sezione Autorizzazioni dei tag LF, specifica i tag LF su cui concedere le autorizzazioni.

LF-Tag permissions

LF-Tags

Choose the LF-Tags you want to grant permissions to.

Choose one or more LF-Tags ▼

Department X

Permissions

Choose the specific LF-Tag permissions to grant.

☒ **Alter**
Update or delete key values.

☒ **Drop**
Delete tag(s).

Grantable permissions

Choose the permissions that the grant recipient(s) can grant to other principals.

☒ **Alter**
Update or delete key values.

☒ **Drop**
Delete tag(s).

Cancel

Grant

- Scegli uno o più tag LF utilizzando il menu a discesa.

Specificate le coppie chiave-valore del tag LF

1. Per concedere i permessi sulle coppie chiave-valore LF-Tag, (devi prima scegliere Scegli i permessi della coppia chiave-valore LF-Tag come Tipo di autorizzazione) scegli Aggiungi coppia chiave-valore LF-Tag per visualizzare la prima riga di campi per specificare la chiave e i valori del tag LF.

LF-Tag key-value pair permissions

Key

Values

You can add 50 more LF-Tags.

Permissions

Choose the specific key-value pair permissions to grant.

☐ **Describe**
See keys and values.

☐ **Associate**
Assign LF-Tags to databases, tables, and columns.

☐ **Grant with LF-Tag expression**
Allow the principal(s) to grant access permissions using the LF-Tag(s).

Grantable permissions

Choose the permissions that the grant recipient(s) can grant to other principals.

☐ **Describe**
See keys and values.

☐ **Associate**
Assign LF-Tags to databases, tables, and columns.

☐ **Grant with LF-Tag expression**
Allow the principal(s) to grant access permissions using the LF-Tag(s).

2. Posizionate il cursore nel campo Chiave, facoltativamente iniziate a digitare per restringere l'elenco di selezione e selezionate una chiave LF-Tag.
3. Nell'elenco Valori, selezionate uno o più valori, quindi premete Tab oppure fate clic o toccate all'esterno del campo per salvare i valori selezionati.

Note

Se una delle righe dell'elenco Valori è attiva, premendo Invio si seleziona o deselecta la casella di controllo.

I valori selezionati vengono visualizzati sotto forma di riquadri sotto l'elenco Valori. Scegli  per rimuovere un valore. Scegli Rimuovi per rimuovere l'intero tag LF.

4. Per aggiungere un altro tag LF, scegliete nuovamente Aggiungi tag LF e ripetete i due passaggi precedenti.

Specificate le espressioni del tag LF

1. Per concedere i permessi sulle espressioni con tag LF, (devi prima scegliere i permessi per le espressioni con tag LF come tipo di autorizzazione).

Permission type

Choose the type of permission to grant. [Learn more](#)

☐ **LF-Tag permissions**
Grant permissions on LF-Tags to create, update, and delete LF-Tags.

☐ **LF-Tag key-value pair permissions**
Grant permissions on LF-Tag key-value pairs to assign LF-Tags to Data Catalog resources and grant permissions on the resources to principals.

☒ **LF-Tag expression permissions - new**
Grant permissions on LF-Tag expressions.

Principals

Choose the principals to grant permissions.

IAM users and roles
Add one or more IAM users or roles.

Choose IAM principals to add

datalake_user

User

2. Scegliete un'espressione LF-Tag.
3. Le espressioni selezionate appaiono come riquadri sotto l'elenco delle espressioni del tag LF. Scegliete ✕ per rimuovere un'espressione.
4. Per aggiungere un'altra espressione del tag LF, scegliete un'altra espressione.

Specificate le autorizzazioni

Questa sezione mostra le autorizzazioni del tag LF o le autorizzazioni con valore del tag LF in base al tipo di autorizzazione scelto nel passaggio precedente.

A seconda del tipo di autorizzazione che hai scelto di concedere, seleziona le autorizzazioni LF-Tag o la coppia chiave-valore LF-Tag, le autorizzazioni e le autorizzazioni concedibili.

1. In Autorizzazioni LF-Tag, seleziona le autorizzazioni da concedere.

La concessione di Drop and Alter implica la concessione implicita di Describe.

È necessario concedere le autorizzazioni Alter and Drop su tutti i valori dei tag.

2. In LT-Tag key-value permissions, seleziona le autorizzazioni da concedere.

Granting Associate concede implicitamente Describe. Scegliete Grant with LF-Tag expression per consentire al destinatario della concessione di concedere o revocare le autorizzazioni di accesso alle risorse del Data Catalog utilizzando il metodo LF-TBAC.

3. In Autorizzazioni di espressione LF-Tag, seleziona le autorizzazioni da concedere.

La concessione di Drop and Alter implica la concessione implicita di Describe.

Concedendo l'autorizzazione Super, vengono concesse tutte le autorizzazioni disponibili.

4. (Facoltativo) In Autorizzazioni concedibili, seleziona le autorizzazioni che il destinatario della sovvenzione può concedere agli altri responsabili del proprio account. AWS
5. Scegli Concessione.

Gestione delle autorizzazioni LF-Tag utilizzando AWS CLI

È possibile concedere, revocare ed elencare le autorizzazioni sui tag LF utilizzando (). AWS Command Line Interface AWS CLI

Per elencare le autorizzazioni LF-Tag ()AWS CLI

- Immettete un comando. `list-permissions` Devi essere il creatore del tag LF, un amministratore del data lake o avere l'Grant with LF-Tag permissionsautorizzazioneDrop,, Alter DescribeAssociate, per un tag LF per vederlo.

Il comando seguente richiede tutti i tag LF per i quali disponete delle autorizzazioni.

```
aws lakeformation list-permissions --resource-type LF_TAG
```

Di seguito è riportato un esempio di output per un amministratore di data lake, che vede tutti i tag LF concessi a tutti i principali. Gli utenti non amministrativi vedono solo i tag LF concessi loro.

Le autorizzazioni LF-Tag concesse da un account esterno vengono visualizzate in una pagina dei risultati separata. Per vederli, ripetete il comando e inserite nell'--next-token l'argomento il token restituito dalla precedente esecuzione del comando.

```
{
  "PrincipalResourcePermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_admin"
      },
      "Resource": {
        "LFTag": {
          "CatalogId": "111122223333",
          "TagKey": "environment",
          "TagValues": [
            "*"
          ]
        }
      },
      "Permissions": [
        "ASSOCIATE"
      ],
      "PermissionsWithGrantOption": [
        "ASSOCIATE"
      ]
    },
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
      },
      "Resource": {
        "LFTag": {
          "CatalogId": "111122223333",
          "TagKey": "module",
          "TagValues": [
            "Orders",
            "Sales"
          ]
        }
      },
      "Permissions": [
```

```

        "DESCRIBE"
      ],
      "PermissionsWithGrantOption": []
    },
    ...
  ],
  "NextToken": "eyJzaG91bGRRdWVy...Wlzc2lvdnMiOnRydWV9"
}

```

È possibile elencare tutte le concessioni per una chiave LF-Tag specifica. Il comando seguente restituisce tutte le autorizzazioni concesse per il tag LF. `module`

```
aws lakeformation list-permissions --resource-type LF_TAG --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

È inoltre possibile elencare i valori del tag LF concessi a un principale specifico per un tag LF specifico. Quando si fornisce l'--principalargomento, è necessario fornire l'argomento. --resource Pertanto, il comando può richiedere efficacemente solo i valori concessi a un principale specifico per una chiave LF-Tag specifica. Il comando seguente mostra come eseguire questa operazione per la chiave principale `datalake_user1` e per la chiave LF-Tag. `module`

```
aws lakeformation list-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --resource-type LF_TAG --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}'
```

Di seguito è riportato un output di esempio.

```

{
  "PrincipalResourcePermissions": [
    {
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
      },
      "Resource": {
        "LFTag": {
          "CatalogId": "111122223333",
          "TagKey": "module",
          "TagValues": [

```

```

        "Orders",
        "Sales"
    ]
    },
    "Permissions": [
        "ASSOCIATE"
    ],
    "PermissionsWithGrantOption": []
}
]
}

```

Per concedere i permessi sui tag LF (AWS CLI)

1. Utilizzare un comando simile al seguente: Questo esempio concede all'utente l'Associateautorizzazione per l'uso `datalake_user1` del tag LF con la chiave `module`. Concede le autorizzazioni per visualizzare e assegnare tutti i valori per quella chiave, come indicato dall'asterisco (*).

```

aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'

```

La concessione dell'autorizzazione implica la concessione implicita dell'Associateautorizzazione. `Describe`

L'esempio successivo concede Associate all' AWS account esterno 1234-5678-9012 sul tag LF con la chiave, con l'opzione `grant.module`. Concede le autorizzazioni per visualizzare e assegnare solo i valori `e. sales orders`

```

aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=123456789012 --permissions "ASSOCIATE"
--permissions-with-grant-option "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["sales", "orders"]}}'

```

2. La concessione dell'autorizzazione implica la concessione implicita `GrantWithLFTagExpression` dell'autorizzazione. `Describe`

L'esempio successivo concede `GrantWithLFTagExpression` a un utente il tag LF con la chiave `module`, con l'opzione `grant`. Concede le autorizzazioni per visualizzare e concedere le autorizzazioni sulle risorse del Data Catalog utilizzando solo i valori `e. sales orders`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "GrantWithLFTagExpression"
--permissions-with-grant-option "GrantWithLFTagExpression" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["sales", "orders"]}]'
```

3. L'esempio successivo concede `Drop` le autorizzazioni a un utente sul tag LF con la chiave, con l'opzione `grant`. `module` Concede le autorizzazioni per eliminare il tag LF. Per eliminare un tag LF, sono necessarie le autorizzazioni su tutti i valori di quella chiave.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "DROP"
--permissions-with-grant-option "DROP" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}]'
```

4. L'esempio successivo concede `Alter` le autorizzazioni all'utente sul tag LF con la chiave, con l'opzione `grant`. `module` Concede le autorizzazioni per eliminare il tag LF. Per aggiornare un LF-tag, sono necessarie le autorizzazioni su tutti i valori di quella chiave.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=111122223333 --permissions "ALTER"
--permissions-with-grant-option "ALTER" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}]'
```

Per revocare le autorizzazioni su LF-Tags ()AWS CLI

- Utilizzare un comando simile al seguente: Questo esempio revoca l'Associate autorizzazione sul tag LF con la chiave dell'utente. `module datalake_user1`

```
aws lakeformation revoke-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
{"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}]'
```

Concessione delle autorizzazioni per il data lake utilizzando il metodo LF-TBAC

Puoi concedere le autorizzazioni DESCRIBE e ASSOCIATE Lake Formation sui tag LF ai mandanti in modo che possano visualizzare i tag LF e assegnarli alle risorse del Data Catalog (database, tabelle, viste e colonne). Quando i tag LF vengono assegnati alle risorse del Data Catalog, è possibile utilizzare il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) per proteggere tali risorse. Per ulteriori informazioni, consulta [Controllo degli accessi basato su tag Lake Formation](#).

Inizialmente, solo l'amministratore del data lake può concedere queste autorizzazioni. Se l'amministratore del data lake concede queste autorizzazioni con l'opzione di concessione, altri responsabili possono concederle. Le ASSOCIATE autorizzazioni DESCRIBE e sono spiegate in [Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation](#)

È possibile concedere le ASSOCIATE autorizzazioni DESCRIBE and su un tag LF a un account esterno. AWS Un amministratore di data lake in quell'account può quindi concedere tali autorizzazioni ad altri responsabili dell'account. I responsabili a cui l'amministratore del data lake dell'account esterno concede l'ASSOCIATE autorizzazione possono quindi assegnare LF-Tags alle risorse del Data Catalog che hai condiviso con il loro account.

Quando si concede a un account esterno, è necessario includere l'opzione di concessione.

È possibile concedere le autorizzazioni sui tag LF utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Note

I seguenti passaggi non sono necessari per i cataloghi di S3 Tables. Puoi usare LF-Tags per concedere le autorizzazioni sui cataloghi S3 Tables esistenti senza eliminarli e ricrearli.

Abilitazione del supporto dei tag LF per i cataloghi federati esistenti che utilizzano le autorizzazioni Lake Formation

Segui questi passaggi, se disponi di cataloghi federati esistenti che utilizzano le autorizzazioni di Lake Formation, come Amazon Redshift o Amazon DynamoDB cataloghi creati prima che il supporto LF-Tags fosse disponibile per i cataloghi federati.

1. Elimina il catalogo esistente: richiama l'operazione `deleteCatalog` API per rimuovere il catalogo federato esistente che utilizza le autorizzazioni Lake Formation.
2. Crea un nuovo catalogo federato: crea un nuovo catalogo e indirizza il nuovo catalogo al tuo namespace/datashare esistente.

Usa un nuovo nome per il catalogo: questo processo aggiorna i cataloghi federati preesistenti per supportare la funzionalità LF-Tag. Se desideri utilizzare lo stesso nome di catalogo, contatta AWS il team di supporto per ricevere assistenza.

Argomenti

- [Concessione delle autorizzazioni per Data Catalog](#)

Consulta anche

- [Gestione dei permessi per i valori del tag LF](#)
- [Gestione dei tag LF per il controllo dell'accesso ai metadati](#)
- [Controllo degli accessi basato su tag Lake Formation](#)

Concessione delle autorizzazioni per Data Catalog

Usa la console Lake Formation o AWS CLI concedi le autorizzazioni di Lake Formation su database, tabelle, viste e colonne di Data Catalog utilizzando il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC).

Console

I passaggi seguenti spiegano come concedere le autorizzazioni utilizzando il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) e la pagina Grant data lake permissions sulla console Lake Formation. La pagina è suddivisa nelle seguenti sezioni:

- Principi: utenti, ruoli e autorizzazioni Account AWS a cui concedere le autorizzazioni.
- Tag LF o risorse del catalogo: database, tabelle o collegamenti a risorse a cui concedere le autorizzazioni.
- Autorizzazioni: autorizzazioni da concedere a The Lake Formation.

1. Apri la pagina Concedi le autorizzazioni del data lake.

Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake o come utente a cui sono state concesse le autorizzazioni Lake Formation sulle risorse del Data Catalog tramite LF-TBAC con l'opzione di concessione.

Nel pannello di navigazione, sotto Autorizzazioni, scegli Autorizzazioni Data Lake. Quindi scegli Concedi.

2. Specificate i principi.

Nella sezione Principali, scegli un tipo principale, quindi specifica i principali a cui concedere le autorizzazioni.

Grant permissions

Principals

Choose the principals to grant permissions.

☒ **IAM users and roles**
Users or roles from this AWS account.

☐ **IAM Identity Center - new**
Users and groups configured in IAM Identity Center.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☐ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

IAM users and roles

Add one or more IAM users or roles.

Choose IAM principals to add ▼

AWSServiceRoleForRedshift X

Role

Utenti e ruoli IAM

Scegli uno o più utenti o ruoli dall'elenco degli utenti e dei ruoli IAM.

Centro identità IAM

Scegli uno o più utenti o dall'elenco Utenti e gruppi.

Utenti e gruppi SAML

Per utenti e gruppi SAML e Quick Suite, inserisci uno o più Amazon Resource Names (ARNs) per utenti o gruppi federati tramite SAML o ARNs per utenti o gruppi Quick Suite. Premi Invio dopo ogni ARN.

Per informazioni su come costruire il, consulta. ARNs [Comandi di concessione e AWS CLI revoca di Lake Formation](#)

Note

L'integrazione di Lake Formation con Quick Suite è supportata solo per Quick Suite Enterprise Edition.

Account esterni

Per Account AWS, AWS organizzazione o responsabile IAM, inserisci una o più organizzazioni Account AWS IDs IDs IDs, unità organizzative o ARN validi per l'utente o il ruolo IAM. Premi Invio dopo ogni ID.

Un ID dell'organizzazione è composto da «o-» seguito da 10-32 lettere o cifre minuscole.

L'ID di un'unità organizzativa inizia con «ou-» seguito da 4 a 32 lettere o cifre minuscole (l'ID della radice che contiene l'unità organizzativa). Questa stringa è seguita da un secondo trattino «-» e da 8 a 32 lettere o cifre minuscole aggiuntive.

3. Specificate i tag LF.

Assicuratevi che sia selezionata l'opzione Risorse abbinate ai tag LF. Scegliete le coppie chiave-valore del tag LF o le espressioni di tag LF salvate.

1. Se scegliete l'opzione delle coppie chiave-valore LF-Tag, scegliete le chiavi e i valori.

Se scegliete più di un valore, state creando un'espressione LF-tag con un operatore. OR
Ciò significa che se uno qualsiasi dei valori del tag LF corrisponde a un tag LF assegnato a una risorsa del Data Catalog, vengono concesse le autorizzazioni sulla risorsa.

LF-Tags or catalog resources

Choose a method to grant permissions.

☒ **Resources matched by LF-Tags (recommended)**
 Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☐ **Named Data Catalog resources**
 Manage permissions for specific databases or tables, in addition to fine-grained data access.

☒ **LF-Tag key-value pairs**
☐ Saved LF-Tag expressions - *new*

Key	Values	
Location	Choose LF-Tag values	Remove
	US	
Department	Choose LF-Tag values	Remove
	marketing	
	sales	

[Add LF-Tag key-value pair](#)

You can add 48 more LF-Tags.

Expression review

The LF-Tag expression above will be interpreted in the following way.

```
Location = US
AND Department = (marketing OR sales)
```

☒ **Save as new expression**
 Use saved expressions to grant permissions. Create LF-Tag expression permissions are needed.

New LF-Tag expression name

Enter a name that describes the expression. Expression name cannot be edited after creation.

new-expression

Name must be less than 1000 characters.

2. (Facoltativo) Scegliete nuovamente Aggiungi coppia chiave-valore LF-tag per specificare un altro tag LF.

Se specificate più di un tag LF, state creando un'espressione di tag LF con un operatore. AND Al principale vengono concesse le autorizzazioni su una risorsa del Data Catalog solo se alla risorsa è stato assegnato un tag LF corrispondente per ogni tag LF nell'espressione del tag LF.

3. Scegliete l'opzione Salva come nuova espressione per riutilizzare l'espressione.

È necessario Create LF-Tag expression salvare le espressioni.

Per ulteriori informazioni sulle espressioni dei tag LF, vedere. [Gestione delle espressioni LF-Tag per il controllo dell'accesso ai metadati](#)

4. Specificare le autorizzazioni.

Specificate le autorizzazioni che desiderate concedere al principale per le risorse del Data Catalog corrispondenti. Le risorse corrispondenti sono quelle a cui sono stati assegnati tag LF che corrispondono a una delle espressioni LF-Tag concesse al principale.

È possibile specificare le autorizzazioni da concedere ai database corrispondenti, alle tabelle corrispondenti e alle viste corrispondenti.

The screenshot displays the 'Database permissions' and 'Table permissions' sections in the AWS Lake Formation console. Each section has a 'Database permissions' and 'Grantable permissions' subsection. The 'Database permissions' subsection includes checkboxes for 'Create table', 'Alter', 'Drop', 'Describe', and 'Super'. The 'Grantable permissions' subsection includes checkboxes for 'Create table', 'Alter', 'Drop', 'Describe', and 'Super'. The 'Table permissions' subsection includes checkboxes for 'Alter', 'Insert', 'Drop', 'Delete', 'Select', 'Describe', and 'Super'. The 'Super' checkbox is described as 'This permission is the union of all the individual permissions to the left, and supersedes them.' and 'This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.'

▼ Database permissions

Database permissions
Choose specific access permissions to grant.

☐ Create table ☐ Alter ☐ Drop ☐ Super
This permission is the union of all the individual permissions to the left, and supersedes them.

☐ Describe

Grantable permissions
Choose the permission that may be granted to others.

☐ Create table ☐ Alter ☐ Drop ☐ Super
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

☐ Describe

▼ Table permissions

Table permissions
Choose specific access permissions to grant.

☐ Alter ☐ Insert ☐ Drop ☐ Super
This permission is the union of all the individual permissions to the left, and supersedes them.

☐ Delete ☐ Select ☐ Describe

Grantable permissions
Choose the permission that may be granted to others.

☐ Alter ☐ Insert ☐ Drop ☐ Super
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

☐ Delete ☐ Select ☐ Describe

In Autorizzazioni database, seleziona le autorizzazioni del database da concedere al principale sui database corrispondenti.

In Autorizzazioni per le tabelle, seleziona le autorizzazioni per la tabella o la visualizzazione da concedere al principale per le tabelle e le viste corrispondenti.

Puoi anche scegliere tra Select le Describe Drop autorizzazioni della tabella e applicare alle viste le autorizzazioni.

5. Scegli Concessione.

AWS CLI

Puoi utilizzare il metodo AWS Command Line Interface (AWS CLI) e il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) per concedere le autorizzazioni di Lake Formation su database, tabelle e colonne di Data Catalog.

Concessione delle autorizzazioni per il data lake utilizzando il metodo e il metodo LF-TBAC AWS CLI

- Utilizza il comando `grant-permissions`.

Example

L'esempio seguente concede all'utente l'espressione LF-Tag `"module=*" (tutti i valori della chiave LF-Tag). module datalake_user1` Quell'utente avrà l'`CREATE_TABLE` autorizzazione su tutti i database corrispondenti, ovvero i database a cui è stato assegnato il tag LF con la chiave, con qualsiasi valore. `module`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "CREATE_TABLE" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333","ResourceType":"DATABASE","Expression":
[{"TagKey":"module","TagValues":["*"]}]]}'
```

Example

L'esempio successivo concede l'espressione LF-Tag `"" all'utente. (level=director) AND (region=west OR region=south) datalake_user1` Quell'utente disporrà delle `DROP` autorizzazioni `SELECT` `ALTER`, e con l'opzione `grant` sulle tabelle corrispondenti, tabelle a cui sono stati assegnati sia (che). `level=director region=west region=south`

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "SELECT" "ALTER" "DROP" --permissions-
with-grant-option "SELECT" "ALTER" "DROP" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333","ResourceType":"TABLE","Expression": [{"TagKey":
"level","TagValues": ["director"]},{ "TagKey": "region","TagValues": ["west",
"south"]}]]}'
```


Example

Il prossimo esempio concede l'espressione LF-tag "" all'account 1234-5678-9012. module=orders AWS L'amministratore del data lake di quell'account può quindi concedere l'espressione "" ai responsabili del proprio account. module=orders Tali responsabili avranno quindi l'CREATE_TABLE autorizzazione a far corrispondere i database di proprietà dell'account 1111-2222-3333 e condivisi con l'account 1234-5678-9012 utilizzando il metodo della risorsa denominata o il metodo LF-TBAC.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=123456789012 --permissions "CREATE_TABLE" --
permissions-with-grant-option "CREATE_TABLE" --resource '{ "LFTagPolicy":
{"CatalogId":"111122223333", "ResourceType":"DATABASE", "Expression":
[{"TagKey":"module", "TagValues":["orders"]}]}'
```

Controllo dell'accesso basato sugli attributi

In AWS Lake Formation, puoi concedere l'accesso a AWS Glue Data Catalog oggetti come cataloghi, database, tabelle e filtri di dati utilizzando attributi che sono tag IAM e tag di sessione associati a entità IAM come ruoli e utenti.

Per ulteriori informazioni sull'utilizzo dei tag di sessione, consulta [assume-role](#) nella guida per l'utente. AWS CLI

Il controllo degli accessi basato sugli attributi (ABAC) è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. AWS chiama questi attributi tag. È possibile utilizzare ABAC per concedere l'accesso ai principali all'interno dello stesso account o di un altro account sulle risorse del Data Catalog. Qualsiasi principale IAM con chiavi e valori del tag IAM o del tag di sessione corrispondenti ottiene l'accesso alla risorsa. È necessario disporre di autorizzazioni concedibili sulle risorse per effettuare queste sovvenzioni.

ABAC ti consente di concedere l'accesso a più utenti contemporaneamente. Quando nuovi utenti entrano a far parte dell'organizzazione, il loro accesso ai dati può essere determinato automaticamente in base ai loro attributi, come la funzione lavorativa o il reparto, senza richiedere agli amministratori di assegnare manualmente ruoli o autorizzazioni specifici. Utilizzando gli attributi anziché i ruoli, ABAC offre un modo più semplificato e gestibile per gestire l'accesso ai dati tra diversi sistemi e ambienti, migliorando in ultima analisi la governance e la conformità dei dati.

Per ulteriori informazioni sulla definizione degli attributi, vedere Definire le [autorizzazioni in base](#) agli attributi con l'autorizzazione ABAC.

Per informazioni su limitazioni, considerazioni e AWS regioni supportate, consulta. [Considerazioni, limitazioni e aree supportate sul controllo degli accessi basato sugli attributi](#)

Argomenti

- [Prerequisiti per la concessione di autorizzazioni tramite attributi](#)
- [Concessione delle autorizzazioni utilizzando il controllo degli accessi basato sugli attributi](#)

Prerequisiti per la concessione di autorizzazioni tramite attributi

Per concedere le autorizzazioni utilizzando il controllo degli accessi basato sugli attributi (ABAC), è necessario completare i seguenti prerequisiti:

- Aggiorna le impostazioni del Data Catalog per abilitare le autorizzazioni Lake Formation per gli oggetti Data Catalog. Per ulteriori informazioni, consulta la sezione [Modifica del modello di autorizzazione predefinito o utilizzo della modalità di accesso ibrida](#).
- Imposta le impostazioni della versione per più account su due o più.
- [Allega gli attributi](#) alle entità IAM che richiedono l'accesso.
- Solo un amministratore del data lake o un utente IAM con le autorizzazioni richieste può concedere l'accesso agli oggetti del Data Catalog. Per ulteriori informazioni sulle autorizzazioni richieste, consulta Autorizzazioni [IAM](#).

Concessione delle autorizzazioni utilizzando il controllo degli accessi basato sugli attributi

Questo argomento descrive i passaggi da seguire per concedere le autorizzazioni di accesso basate sugli attributi alle risorse del Data Catalog. Puoi usare la console Lake Formation o la AWS Command Line Interface (AWS CLI).

Concessione delle autorizzazioni tramite ABAC ()Console di gestione AWS

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/> e accedi come amministratore del data lake, creatore di risorse o utente IAM con autorizzazioni Grantable sulla risorsa.

2. Esegui una delle seguenti operazioni:

- Nel riquadro di navigazione, in Autorizzazioni, scegli Autorizzazioni Data Lake. Quindi scegli Concedi.
- Nel riquadro di navigazione, scegli Cataloghi in Data Catalog. Quindi, scegli un oggetto del catalogo (cataloghi, database, tabelle e filtri di dati) e dal menu Azioni in Autorizzazioni, scegli Concedi.

3. Nella pagina Concedi autorizzazioni, scegli Principali per attributo.

4. Specificate la chiave e i valori dell'attributo. Se scegli più di un valore, stai creando un'espressione di attributo con un OR operatore. Ciò significa che se uno qualsiasi dei valori dei tag degli attributi assegnati a un ruolo o utente IAM corrisponde, role/user ottiene i permessi di accesso alla risorsa.

Se specifichi più di un tag di attributo, stai creando un'espressione di attributo con un AND operatore. Al principale vengono concesse le autorizzazioni su una risorsa del Data Catalog solo se all'IAM role/user è stato assegnato un tag corrispondente per ogni tag di attributo nell'espressione dell'attributo.

Esamina l'espressione di policy Cedar risultante mostrata nella console.

Grant permissions

Principal type

Grant access by IAM principal or attribute tags.



Principals

Grant permissions directly to specific IAM users, roles, or Identity Center principals.



Principals by attributes

Grant permissions using attribute tags with Cedar policy expressions.

Attributes

Grant permissions using an attribute expression based on user, resource, and contextual attributes.

Attribute expression

Add tag key-value pairs from IAM roles or [IAM session tags](#). Separate values with commas. **AND** logic applies between pairs, **OR** logic within keys. Verify tag entries for accuracy.

Key

Value

Add key-value pair

Expression review

The attribute expression above will be written in the following Cedar policy. Use CLI/SDK to edit or write custom Cedar policy.

```
permit (principal, action, resource)
when {
  context.iam.principalTags.hasTag("Department")
  && context.iam.principalTags.getTag("Department") == "sales"
};
```

Permission scope



This account



External account

- Scegliete l'ambito delle autorizzazioni. Se i beneficiari appartengono a un account esterno, scegli Account esterno e inserisci l'ID dell' AWS account.
- Quindi, scegli l'account Data Catalog o gli account esterni. È necessario disporre delle autorizzazioni corrispondenti sulle risorse per completare correttamente le concessioni di autorizzazione.

7. Specificate le azioni che desiderate consentire ai principali (utenti o ruoli) con attributi corrispondenti. L'accesso è concesso alle entità IAM a cui sono stati assegnati tag e valori che corrispondono ad almeno una delle espressioni di attributo specificate. Esamina l'espressione della policy Cedar nella console. Per ulteriori informazioni su Cedar, vedi [Cos'è Cedar? | Cedar Policy](#) Language Reference. GuideLink
8. Quindi scegli le risorse del Data Catalog a cui concedere l'accesso. Puoi definire queste autorizzazioni per varie risorse del Data Catalog, inclusi cataloghi, database, tabelle e filtri di dati.
9. Scegli Concessione.

Questo approccio consente di controllare l'accesso in base agli attributi, garantendo che solo gli utenti o i ruoli con i tag appropriati possano eseguire azioni specifiche sulle risorse designate.

Concessione delle autorizzazioni tramite ABAC ()AWS CLI

L'esempio seguente mostra un'espressione di attributo che deve essere soddisfatta per ricevere tutte le autorizzazioni disponibili sulla risorsa. In alternativa, è possibile specificare autorizzazioni individuali come `SelectDescribe`, o. `Drop` L'espressione utilizza l'espressione politica Cedar. Per ulteriori informazioni su Cedar, vedere [Cos'è Cedar? | Cedar Policy](#) Language Reference. GuideLink

Questa condizione verifica se il principale IAM ha un `department` tag e il valore del `department` tag è uguale a `sales`.

```
aws lakeformation grant-permissions
--principal '{"DataLakePrincipalIdentifier": "111122223333:IAMPrincipals"}' \
--resource '{"Database": {"CatalogId": 111122223333, "Name": "abac-db"}}' \
--permissions ALL \
--condition '{"Expression": "context.iam.principalTags.hasTag(\"department\") \
&& context.iam.principalTags.getTag(\"department\") == \"sales\""}'
```

Scenario di esempio di autorizzazioni

Lo scenario seguente aiuta a dimostrare come è possibile impostare le autorizzazioni per proteggere l'accesso ai dati in AWS Lake Formation

Shirley è un'amministratore di dati. Vuole creare un data lake per la sua azienda, AnyCompany. Attualmente, tutti i dati sono archiviati in Amazon S3. John è un responsabile marketing e deve accedere per iscritto alle informazioni sugli acquisti dei clienti (contenute in `ins3://`

customerPurchases). Un analista di marketing, Diego, si unisce a John quest'estate. John deve poter concedere a Diego l'accesso per eseguire interrogazioni sui dati senza coinvolgere Shirley.

Mateo, del reparto finanziario, ha bisogno di accedere ai dati contabili interrogativi (ad esempio, `s3://transactions`). Vuole interrogare i dati delle transazioni nelle tabelle di un database (Finance_DB) utilizzato dal team finanziario. Il suo manager, Arnav, può dargli accesso a Finance_DB. Sebbene non dovrebbe essere in grado di modificare i dati contabili, ha bisogno della capacità di convertire i dati in un formato (schema) adatto alla previsione. Questi dati verranno archiviati in un bucket separato (`s3://financeForecasts`) che potrà modificare.

Per riassumere:

- Shirley è l'amministratore del data lake.
- John richiede CREATE_DATABASE CREATE_TABLE l'autorizzazione per creare nuovi database e tabelle nel Data Catalog.
- John richiede SELECT anche le autorizzazioni e DELETE le autorizzazioni per le tabelle che crea. INSERT
- Diego richiede SELECT l'autorizzazione sul tavolo per eseguire le interrogazioni.

I dipendenti di AnyCompany eseguono le seguenti azioni per impostare le autorizzazioni. Le operazioni API mostrate in questo scenario mostrano una sintassi semplificata per motivi di chiarezza.

1. Shirley registra il percorso Amazon S3 contenente le informazioni sugli acquisti dei clienti con Lake Formation.

```
RegisterResource(ResourcePath("s3://customerPurchases"), false, Role_ARN )
```

2. Shirley concede a John l'accesso al percorso Amazon S3 contenente le informazioni di acquisto dei clienti.

```
GrantPermissions(John, S3Location("s3://customerPurchases"),  
[DATA_LOCATION_ACCESS]) )
```

3. Shirley concede a John il permesso di creare database.

```
GrantPermissions(John, catalog, [CREATE_DATABASE])
```

- John crea il database. John_DB John ha automaticamente CREATE_TABLE l'autorizzazione per quel database perché lo ha creato.

```
CreateDatabase(John_DB)
```

- John crea la tabella John_Table che punta a. s3://customerPurchases Poiché ha creato la tabella, dispone di tutte le autorizzazioni e può concedere le autorizzazioni relative alla tabella.

```
CreateTable(John_DB, John_Table)
```

- John consente al suo analista, Diego, di accedere alla tabella. John_Table

```
GrantPermissions(Diego, John_Table, [SELECT])
```

- John consente al suo analista, Diego, di accedere a. s3://customerPurchases/London/ Poiché Shirley è già registrata s3://customerPurchases, le sue sottocartelle sono registrate con Lake Formation.

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, [DATA_LOCATION_ACCESS], [], S3Location("s3://customerPurchases/London/") )
```

- John consente al suo analista, Diego, di creare tabelle nel database. John_DB

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, John_DB, [CREATE_TABLE], [] )
```

- Diego crea una tabella John_DB in s3://customerPurchases/London/ e ottiene automaticamente ALTER,, DROP SELECT INSERT, e le DELETE autorizzazioni.

```
CreateTable( 123456789012/datalake, John_DB, Diego_Table )
```

Filtraggio dei dati e sicurezza a livello di cella in Lake Formation

Quando concedi le autorizzazioni di Lake Formation su una tabella Data Catalog, puoi includere specifiche di filtraggio dei dati per limitare l'accesso a determinati dati nei risultati delle query e nei motori integrati con Lake Formation. Lake Formation utilizza il filtraggio dei dati per ottenere la sicurezza a livello di colonna, la sicurezza a livello di riga e la sicurezza a livello di cella. Puoi definire e applicare filtri di dati sulle colonne nidificate se i dati di origine contengono strutture nidificate.

Con le funzionalità di filtraggio dei dati di Lake Formation, puoi implementare i seguenti livelli di sicurezza dei dati.

Sicurezza a livello di colonna

La concessione delle autorizzazioni su una tabella del Catalogo dati con sicurezza a livello di colonna (filtro delle colonne) consente agli utenti di visualizzare solo colonne specifiche e colonne nidificate a cui hanno accesso nella tabella. Prendiamo in considerazione una `persons` tabella utilizzata in più applicazioni per una grande società di comunicazioni multiregionale. La concessione di autorizzazioni per le tabelle di Data Catalog con filtro a colonne può impedire agli utenti che non lavorano nel reparto risorse umane di visualizzare informazioni di identificazione personale (PII) come il numero di previdenza sociale o la data di nascita. È inoltre possibile definire politiche di sicurezza e concedere l'accesso solo a sottostrutture parziali di colonne annidate.

Sicurezza a livello di riga

La concessione delle autorizzazioni su una tabella del Catalogo dati con sicurezza a livello di riga (filtro di riga) consente agli utenti di visualizzare solo righe di dati specifiche a cui hanno accesso nella tabella. Il filtraggio si basa sui valori di una o più colonne. È possibile includere strutture di colonne annidate quando si definiscono espressioni di filtro di riga. Ad esempio, se diversi uffici regionali della società di comunicazioni hanno i propri reparti delle risorse umane, è possibile limitare i record relativi alle persone che i dipendenti delle risorse umane possono vedere ai soli record relativi ai dipendenti della propria regione.

Sicurezza a livello di cella

La sicurezza a livello di cella combina il filtraggio di righe e il filtraggio di colonne per un modello di autorizzazioni altamente flessibile. Se si visualizzano le righe e le colonne di una tabella come griglia, utilizzando la sicurezza a livello di cella, è possibile limitare l'accesso ai singoli elementi (celle) della griglia in qualsiasi punto delle due dimensioni. In altre parole, è possibile limitare l'accesso a diverse colonne a seconda della riga. Ciò è illustrato dal diagramma seguente, in cui le colonne con restrizioni sono ombreggiate.

	Col1	Col2	Col3	Col4	Col5	Col6
Row1						
Row2						
Row3						
Row4						
Row5						

Continuando l'esempio della tabella delle persone, puoi creare un filtro dati a livello di cella che limita l'accesso alla colonna dell'indirizzo se la riga ha la colonna del paese impostata su «UK», ma consente l'accesso alla colonna dell'indirizzo se la riga ha la colonna del paese impostata su «US».

I filtri si applicano solo alle operazioni di lettura. Pertanto, puoi concedere solo l'autorizzazione SELECT Lake Formation con filtri.

Sicurezza a livello di cella sulle colonne annidate

Lake Formation consente di definire e applicare filtri di dati con sicurezza a livello di cella su colonne annidate. Tuttavia, i motori analitici integrati come Amazon Athena, Amazon EMR e Amazon Redshift Spectrum supportano l'esecuzione di query su tabelle nidificate gestite da Lake Formation con sicurezza a livello di riga e colonna.

Per le limitazioni, consulta [Limitazioni del filtraggio dei dati](#).

Argomenti

- [Filtri di dati in Lake Formation](#)
- [Supporto PartiQL nelle espressioni di filtro di riga](#)
- [Autorizzazioni necessarie per eseguire interrogazioni su tabelle con filtraggio a livello di cella](#)
- [Gestione dei filtri di dati](#)

Filtri di dati in Lake Formation

Puoi implementare la sicurezza a livello di colonna, riga e cella tramite la creazione di filtri di dati. Seleziona un filtro dati quando concedi l'autorizzazione SELECT Lake Formation sulle tabelle. Se la tabella contiene strutture di colonne nidificate, è possibile definire un filtro dati includendo o escludendo le colonne secondarie e definire espressioni di filtro a livello di riga sugli attributi nidificati.

Ogni filtro di dati appartiene a una tabella specifica del Data Catalog. Un filtro dati include le seguenti informazioni:

- Nome del filtro
- Il catalogo IDs della tabella associata al filtro
- Nome tabella
- Nome del database che contiene la tabella

- Specificazione delle colonne: un elenco di colonne e colonne annidate (con `struct` tipi di dati) da includere o escludere nei risultati della query.
- Espressione di filtro di riga: un'espressione che specifica le righe da includere nei risultati della query. Con alcune restrizioni, l'espressione ha la sintassi di una `WHERE` clausola nel linguaggio PartiQL. Per specificare tutte le righe, scegli Accesso a tutte le righe in Accesso a livello di riga nella console o Utilizza nelle chiamate API. `AllRowsWildcard`

Per ulteriori informazioni su ciò che è supportato nelle espressioni di filtro di riga, consulta.

[Supporto PartiQL nelle espressioni di filtro di riga](#)

Il livello di filtraggio ottenuto dipende da come si popola il filtro dati.

- Quando specifichi il carattere jolly "tutte le colonne" e fornisci un'espressione di filtro di riga, stabilisci solo la sicurezza a livello di riga (filtraggio di riga).
- Quando si includono o si escludono colonne specifiche e colonne nidificate e si specifica «tutte le righe» utilizzando il carattere jolly per tutte le righe, si stabilisce solo la sicurezza a livello di colonna (filtraggio delle colonne).
- Quando includi o escludi colonne specifiche e fornisci anche un'espressione di filtro di riga, stabilisci la sicurezza a livello di cella (filtraggio delle celle).

La seguente schermata della console Lake Formation mostra un filtro dati che esegue il filtraggio a livello di cella. Per le interrogazioni sulla `orders` tabella, limita l'accesso alla `customer_name` colonna e i risultati della query restituiscono solo le righe in cui la colonna contiene «pharma».

`product_type`

Create data filter



Data filter name

Enter a name that describes this data access filter.

restrict-pharma

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or under-scores (_), and be less than 256 characters.

Target database

Select the database that contains the target table.

Choose databases



Load more

sales



054881201579

Target table

Select the table for which the data filter will be created.

Choose tables



Load more

orders



054881201579

Column-level access

Choose whether this filter should have column-level restrictions.

- ☐ Access to all columns
Filter won't have any column restrictions.
- ☐ Include columns
Filter will only allow access to specific columns.
- ☒ Exclude columns
Filter will allow access to all but specific columns.

Select columns

Choose one or more columns



customer_name



string

Nota l'uso di virgolette singole per racchiudere la stringa letterale, 'pharma'

Puoi utilizzare la console Lake Formation per creare questo filtro di dati oppure puoi fornire il seguente oggetto di richiesta all'operazione `CreateDataCellsFilter` API.

```
{
  "Name": "restrict-pharma",
  "DatabaseName": "sales",
  "TableName": "orders",
  "TableCatalogId": "111122223333",
  "RowFilter": {"FilterExpression": "product_type='pharma'"},
  "ColumnWildcard": {
    "ExcludedColumnNames": ["customer_name"]
  }
}
```

Puoi creare tutti i filtri di dati di cui hai bisogno per una tabella. A tal fine, è necessaria l'`SELECT` autorizzazione con l'opzione di concessione su una tabella. Per impostazione predefinita, gli amministratori di Data Lake dispongono dell'autorizzazione per creare filtri di dati su tutte le tabelle di quell'account. In genere si utilizza solo un sottoinsieme dei possibili filtri di dati quando si concedono le autorizzazioni sulla tabella a un responsabile. Ad esempio, è possibile creare un secondo filtro di dati per la `orders` tabella che è un `row-security-only` filtro di dati. Facendo riferimento alla schermata precedente, è possibile scegliere l'opzione Accesso a tutte le colonne e includere un'espressione di filtro di riga di `product_type<>'pharma'`. Il nome di questo filtro di dati potrebbe essere `no-pharma`. Limita l'accesso a tutte le righe la cui `product_type` colonna è impostata su «pharma».

L'oggetto di richiesta per l'operazione `CreateDataCellsFilter` API per questo filtro di dati è il seguente.

```
{
  "Name": "no-pharma",
  "DatabaseName": "sales",
  "TableName": "orders",
  "TableCatalogId": "111122223333",
  "RowFilter": {"FilterExpression": "product_type<>'pharma'"},
  "ColumnNames": ["customer_id", "customer_name", "order_num",
    "product_id", "purchase_date", "product_type",
    "product_manufacturer", "quantity", "price"]
}
```

È quindi possibile concedere la licenza `SELECT` sulla `orders` tabella con il filtro `restrict-pharma` dati a un utente amministrativo e `SELECT` sulla `orders` tabella con il filtro `no-pharma` dati a utenti non amministrativi. Agli utenti del settore sanitario, potresti `SELECT` concedere alla `orders` tabella l'accesso completo a tutte le righe e le colonne (nessun filtro per i dati) o magari aggiungere un altro filtro di dati che limiti l'accesso alle informazioni sui prezzi.

È possibile includere o escludere colonne annidate quando si specifica la sicurezza a livello di colonna e di riga all'interno di un filtro dati. Nell'esempio seguente, l'accesso al `product.offer` campo viene specificato utilizzando nomi di colonna qualificati (racchiusi tra virgolette doppie). Questo è importante per i campi annidati per evitare che si verifichino errori quando i nomi delle colonne contengono caratteri speciali e per mantenere la compatibilità con le versioni precedenti delle definizioni di sicurezza a livello di colonna.

```
{
  "Name": "example_dcf",
  "DatabaseName": "example_db",
  "TableName": "example_table",
  "TableCatalogId": "111122223333",
  "RowFilter": { "FilterExpression": "customer.customerName <> 'John'" },
  "ColumnNames": ["customer", "\"product\".\"offer\""]
}
```

Consulta anche

- [Gestione dei filtri di dati](#)

Supporto PartiQL nelle espressioni di filtro di riga

È possibile creare espressioni di filtro di riga utilizzando un sottoinsieme di tipi di dati, operatori e aggregazioni PartiQL. Lake Formation non consente alcuna funzione PartiQL definita dall'utente o standard nell'espressione del filtro. È possibile utilizzare gli operatori di confronto per confrontare le colonne con costanti (ad esempio, `views >= 10000`), ma non è possibile confrontare le colonne con altre colonne.

Un'espressione di filtro Row può essere un'espressione semplice o un'espressione composta. La lunghezza totale dell'espressione deve essere inferiore a 2048 caratteri.

Espressione semplice

Un'espressione semplice avrà il seguente formato: `<column name > <comparison operator ><value >`

- Nome della colonna

Può essere una colonna di dati di primo livello, una colonna di partizione o una colonna nidificata presente nello schema della tabella e deve appartenere alle colonne [Tipi di dati supportati](#) elencate di seguito.

- Operatore di confronto

Gli operatori supportati sono i seguenti: `=`, `>`, `<`, `>=`, `<=`, `<>`, `!=`, `BETWEEN`, `IN`, `LIKE`, `NOT`, `IS [NOT] NULL`

- Tutti i confronti tra stringhe e le corrispondenze di LIKE modelli fanno distinzione tra maiuscole e minuscole. Non è possibile utilizzare l'operatore `IS [NOT] NULL` sulle colonne di partizione.
- Valore della colonna

Il valore della colonna deve corrispondere al tipo di dati del nome della colonna.

Espressione composta

Un'espressione composta avrà il formato: `( <simple expression >) <AND/OR >( <simple expression >)`. Le espressioni composte possono essere ulteriormente combinate utilizzando operatori logici `AND/OR`.

Tipi di dati supportati

I filtri di riga che fanno riferimento a una AWS Glue Data Catalog tabella che contiene un tipo di dati non supportato genereranno un errore. Di seguito sono riportati i tipi di dati supportati per le colonne e le costanti della tabella, che sono mappati ai Amazon Redshift tipi di dati:

- `STRING`, `CHAR`, `VARCHAR`
- `INT`, `LONG`, `BIGINT`, `FLOAT`, `DECIMAL`, `DOUBLE`
- `BOOLEAN`
- `STRUCT`

Per ulteriori informazioni sui tipi di dati in Amazon Redshift, consulta Tipi di [dati nella Amazon Redshift Database](#) Developer Guide.

Espressioni di filtro di riga

Example

Di seguito sono riportati alcuni esempi di espressioni di filtro di riga valide per una tabella con colonne: `country` (String), `id` (Long), `year` (partition column of type Integer), `month` (partition column of type Integer)

- `year > 2010 and country != 'US'`
- `(year > 2010 and country = 'US') or (month < 8 and id > 23)`
- `(country between 'Z' and 'U') and (year = 2018)`
- `(country like '%ited%') and (year > 2000)`

Example

Di seguito sono riportati alcuni esempi validi di espressioni di filtro di riga per una tabella con colonne annidate: `year > 2010 and customer.customerId <> 1`

Non è necessario fare riferimento ai campi nidificati nelle colonne di partizione quando si definiscono espressioni nidificate a livello di riga.

Le costanti di stringa devono essere racchiuse tra virgolette singole.

Parole chiave riservate

Se l'espressione del filtro di riga contiene parole chiave PartiQL, riceverai un errore di analisi poiché i nomi delle colonne potrebbero essere in conflitto con le parole chiave. Quando ciò accade, evita i nomi delle colonne usando le virgolette doppie. Alcuni esempi di parole chiave riservate sono «first», «last», «asc», «missing». Vedi la specifica PartiQL per un elenco di parole chiave riservate.

Riferimento PartiQL

Per ulteriori informazioni su PartiQL, vedere. <https://partiql.org/>

Autorizzazioni necessarie per eseguire interrogazioni su tabelle con filtraggio a livello di cella

Le seguenti autorizzazioni AWS Identity and Access Management (IAM) sono necessarie per eseguire query su tabelle con filtraggio a livello di cella.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:StartQueryPlanning",
        "lakeformation:GetQueryState",
        "lakeformation:GetWorkUnits",
        "lakeformation:GetWorkUnitResults"
      ],
      "Resource": "*"
    }
  ]
}
```

Per ulteriori informazioni sui permessi di Lake Formation, vedere [Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM](#).

Gestione dei filtri di dati

Per implementare la sicurezza a livello di colonna, riga e cella, puoi creare e gestire filtri di dati. Ogni filtro di dati appartiene a una tabella Data Catalog. È possibile creare più filtri di dati per una tabella e quindi utilizzarne uno o più per concedere le autorizzazioni sulla tabella. È inoltre possibile definire e applicare filtri di dati su colonne nidificate con struct tipi di dati che consentono agli utenti di accedere solo alle sottostrutture delle colonne nidificate.

È necessaria SELECT l'autorizzazione con l'opzione di concessione per creare o visualizzare un filtro dati. Per consentire ai responsabili del tuo account di visualizzare e utilizzare un filtro dati, puoi concedere l'DESCRIBE autorizzazione su di esso.

 Note

Lake Formation non supporta la concessione `Describe` dell'autorizzazione su un filtro dati, che viene condiviso da un altro account.

Puoi gestire i filtri dei dati utilizzando la AWS Lake Formation console, l'API o AWS Command Line Interface (AWS CLI).

Per informazioni sui filtri di dati, consulta [Filtri di dati in Lake Formation](#)

Creazione di un filtro dati

È possibile creare uno o più filtri di dati per ogni tabella del Data Catalog.

Per creare un filtro dati per una tabella Data Catalog (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, proprietario della tabella di destinazione o principale che dispone dell'autorizzazione Lake Formation sulla tabella di destinazione.

2. Nel riquadro di navigazione, in Catalogo dati, scegli Filtri dati.
3. Nella pagina Filtri dati, scegli Crea nuovo filtro.
4. Nella finestra di dialogo Crea filtro dati, inserisci le seguenti informazioni:

- Nome del filtro dati
- Database di destinazione: specifica il database che contiene la tabella.
- Tabella di destinazione
- Accesso a livello di colonna: lascia questa impostazione su Accesso a tutte le colonne per specificare solo il filtraggio delle righe. Scegliete Includi colonne o Escludi colonne per specificare il filtraggio di colonne o celle, quindi specificate le colonne da includere o escludere.

Colonne nidificate: se applichi il filtro a una tabella che contiene colonne nidificate, puoi specificare in modo esplicito le sottostrutture delle colonne struct nidificate all'interno di un filtro dati.

Quando concedi l'autorizzazione SELECT a un principale su questo filer, il principale che esegue la seguente query vedrà solo i dati relativi e non. `customer.customerName`
`customer.customerId`

```
SELECT "customer" FROM "example_db"."example_table";
```

Column-level access

Choose whether this filter should have column-level restrictions.

Column-level access

Choose whether this filter should have column-level restrictions.

- ☐ Access to all columns
- Filter won't have any column restrictions.
- ☒ Include columns
- Filter will only allow access to specific columns.
- ☐ Exclude columns
- Filter will allow access to all but specific columns.

Included columns (4/11)

Choose the columns for column-level access

Find column

< 1 >

	Name	Type
☐	customer	struct
☐	customerId	string
☒	customerName	string
☒	customerapplication	struct
☐	appId	string
☒	product	struct
☐	offer	struct
☐	listingId	string
☐	prodId	string
☐	type	string
☒	purchaseId	string

Row-level access

Choose whether this filter should have row-level restrictions.

- ☐ Access to all rows
- ☒ Filter rows

Row filter expression

Enter the rest of the following query statement `SELECT * FROM nested-table WHERE...`
Please see the documentation for examples of filter expressions.

customer.customerName <> 'John'

Quando concedi le autorizzazioni alla `customer` colonna, il principale riceve l'accesso alla colonna e ai campi annidati sotto la colonna (`customerName`, `customerID`).

- **Espressione di filtro di riga:** immettere un'espressione di filtro per specificare il filtraggio di righe o celle. Per i tipi di dati e gli operatori supportati, vedere [Supporto PartiQL nelle espressioni di filtro di riga](#). Scegli Accesso a tutte le righe per concedere l'accesso a tutti.

È possibile includere strutture di colonne parziali provenienti da colonne nidificate in un'espressione di filtro di riga per filtrare le righe che contengono un valore specifico.

Quando a un'entità vengono concesse le autorizzazioni per una tabella con un'espressione `Select * from example_nesttable where customer.customerName <> 'John'` di filtro di riga e l'accesso a livello di colonna è impostato su Accesso a tutte le colonne, i risultati della query mostrano solo le righe il cui risultato è `true`. `customerName <> 'John'`

La schermata seguente mostra un filtro dati che implementa il filtraggio delle celle. Nelle interrogazioni sulla `orders` tabella, nega l'accesso alla `customer_name` colonna e mostra solo le righe che contengono «pharma» nella colonna. `product_type`

Create data filter



Data filter name

Enter a name that describes this data access filter.

restrict-pharma

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or under-scores (_), and be less than 256 characters.

Target database

Select the database that contains the target table.

Choose databases



Load more

sales



054881201579

Target table

Select the table for which the data filter will be created.

Choose tables



Load more

orders



054881201579

Column-level access

Choose whether this filter should have column-level restrictions.

- ☐ Access to all columns
Filter won't have any column restrictions.
- ☐ Include columns
Filter will only allow access to specific columns.
- ☒ Exclude columns
Filter will allow access to all but specific columns.

Select columns

Choose one or more columns



customer_name
string



5. Scegli Create Filter (Crea filtro).

Per creare un filtro dati con politiche di filtro cellulare su un campo annidato

Questa sezione utilizza lo schema di esempio seguente per mostrare come creare un filtro di celle di dati:

```
[
  { name: "customer", type: "struct<customerId:string,customerName:string>" },
  { name: "customerApplication", type: "struct<appId:string>" },
  { name: "product", type:
"struct<offer:struct<prodId:string,listingId:string>,type:string>" },
  { name: "purchaseId", type: "string" },
]
```

1. Nella pagina Crea un filtro dati, inserisci un nome per il filtro dati.
2. Successivamente, utilizza il menu a discesa per scegliere il nome del database e il nome della tabella.
3. Nella sezione Accesso a livello di colonna, scegli Colonne incluse e seleziona una colonna nidificata (). `customer.customerName`
4. Nella sezione Accesso a livello di riga, scegli l'opzione Accesso a tutte le righe.
5. Scegli Create Filter (Crea filtro).

Quando concedi SELECT l'autorizzazione per questo filtro, il principale ottiene l'accesso a tutte le righe della `customerName` colonna.

6. Successivamente, definisci un altro filtro dati per lo stesso database/tabella.
7. Nella sezione Accesso a livello di colonna, scegli Colonne incluse e seleziona un'altra colonna nidificata (). `customer.customerId`
8. Nella sezione Accesso a livello di riga, scegliete Filtra righe e immettete un'espressione di filtro di riga (). `customer.customerId <> 5`
9. Scegli Create Filter (Crea filtro).

Quando concedete SELECT l'autorizzazione per questo filtro, il principale riceve l'accesso a tutte le righe e ai `customerId` campi ad eccezione della cella in cui il valore è 5 nella `customerId` colonna. `customerName`

Concessione delle autorizzazioni per il filtro dei dati

Puoi concedere le autorizzazioni `SELECT`, `DESCRIBE` e `DROP` Lake Formation sui filtri di dati ai responsabili.

All'inizio, solo tu puoi visualizzare i filtri di dati che crei per una tabella. Per consentire a un altro principale di visualizzare un filtro di dati e concedere le autorizzazioni di Data Catalog con il filtro dati, devi:

- `SELECT` Concedi una tabella al principale con l'opzione di concessione e applica il filtro dati alla concessione.
- Concedi l'`DROP` autorizzazione `DESCRIBE` o l'autorizzazione sul filtro dati al responsabile.

È possibile concedere l'`SELECT` autorizzazione a un AWS account esterno. Un amministratore del data lake di quell'account può quindi concedere tale autorizzazione ad altri responsabili dell'account. Quando si concede a un account esterno, è necessario includere l'opzione di concessione in modo che l'amministratore dell'account esterno possa trasferire ulteriormente l'autorizzazione agli altri utenti dell'account. his/her Quando concedi la concessione a un titolare del tuo account, la concessione con l'opzione di concessione è facoltativa.

Puoi concedere e revocare le autorizzazioni sui filtri di dati utilizzando la AWS Lake Formation console, l'API o (). AWS Command Line Interface AWS CLI

Console

1. Accedi a Console di gestione AWS e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel pannello di navigazione, sotto Autorizzazioni, scegli Autorizzazioni Data lake.
3. Nella pagina Autorizzazioni, nella sezione Autorizzazioni dati, scegli Concedi.
4. Nella pagina Concedi le autorizzazioni per i dati, scegli i principali a cui concedere le autorizzazioni.
5. Nella sezione LF-Tags o risorse del catalogo, scegliete Risorse del catalogo dati denominate. Scegliete quindi il database, la tabella e il filtro dati per i quali desiderate concedere le autorizzazioni.

LF-Tags or catalog resources

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named data catalog resources
Manager permissions for specific databases or tables, in addition to fine-grained data access.

Databases

Select one or more databases.

Choose databases ▼

cloudtrail 106567286946 ✕

Load more

Tables - optional

Select one or more tables.

Choose tables ▼

cloudtrail_logs_awslogs 106567286946 ✕

Load more

Data filters - optional

Select one or more data filters.

Choose data filters ▼

cloudtrail_lakeformation_filter 106567286946 ✕

Load more

Create new

[Manage data filters](#) ↗

6. Nella sezione Autorizzazioni del filtro dati, scegli le autorizzazioni che desideri concedere ai principali selezionati.

Data filter permissions

Data filter permissions

Choose specific access permissions to grant.

☐ Select ☒ Describe ☒ Drop

Grantable permissions

Choose the permission that may be granted to others.

☐ Select ☐ Describe ☐ Drop

AWS CLI

- Inserisci un comando. `grant-permissions` Specificate `DataCellsFilter` per l'argomento `resource` e specificate `DESCRIBE` o `DROP` per l'argomento `Permissions` e, facoltativamente, per l'argomento `PermissionsWithGrantOption`.

L'esempio seguente concede l'opzione `DESCRIBE` di concessione all'utente `datalake_user1` sul filtro `restrict-pharma`, che appartiene alla `orders` tabella del `sales` database dell' AWS account `1111-2222-3333`.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

Di seguito è riportato il contenuto del file. `grant-params.json`

```
{
  "Principal": {"DataLakePrincipalIdentifier":
    "arn:aws:iam::111122223333:user/datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["DESCRIBE"],
  "PermissionsWithGrantOption": ["DESCRIBE"]
}
```

Concessione delle autorizzazioni relative ai dati fornite dai filtri di dati

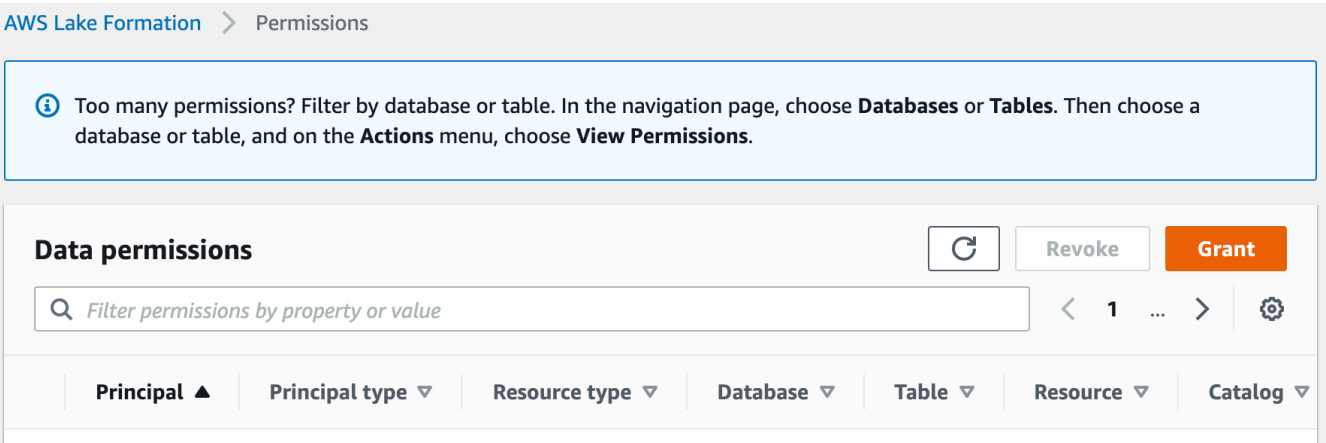
I filtri di dati rappresentano un sottoinsieme di dati all'interno di una tabella. Per fornire l'accesso ai dati ai principali, è necessario concedere `SELECT` le autorizzazioni a tali responsabili. Con questa autorizzazione i responsabili possono:

- Visualizza il nome effettivo della tabella nell'elenco delle tabelle condivise con il loro account.
- Crea filtri di dati sulla tabella condivisa e concedi le autorizzazioni ai relativi utenti su tali filtri di dati.

Console

Per concedere le autorizzazioni SELECT

1. Vai alla pagina Autorizzazioni nella console di Lake Formation, quindi scegli Concedi.



2. Seleziona i principali a cui desideri fornire l'accesso e seleziona Named data catalog resources.

LF-Tags or catalog resources

☐ Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ Named data catalog resources
Manager permissions for specific databases or tables, in addition to fine-grained data access.

Databases

Select one or more databases.

Choose databases ▼

cloudtrail 106567286946 ✕

Load more

Tables - optional

Select one or more tables.

Choose tables ▼

cloudtrail_logs_awslogs 106567286946 ✕

Load more

Data filters - optional

Select one or more data filters.

Choose data filters ▼

cloudtrail_lakeformation_filter 106567286946 ✕

Load more

Create new

[Manage data filters](#)

3. Per fornire l'accesso ai dati rappresentati dal filtro, scegli **Seleziona** in **Autorizzazioni del filtro dati**.

Data filter permissions

Data filter permissions

Choose specific access permissions to grant.

☒ Select
 ☐ Describe
 ☐ Drop

Grantable permissions

Choose the permission that may be granted to others.

☐ Select
 ☐ Describe
 ☐ Drop

 Select permissions on data filters will grant access to the table 'cloudtrail_logs_awslogs'.

CLI

Inserisci `grant-permissions` un comando. `DataCellsFilterSpecify` l'argomento risorsa e specificate `SELECT` l'argomento Autorizzazioni.

L'esempio seguente concede `SELECT` l'opzione `grant` all'utente `datalake_user1` sul filtro `datirestrict-pharma`, che appartiene alla `orders` tabella del `sales` database in. Account AWS 1111-2222-3333

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

Di seguito sono riportati i contenuti del file `grant-params.json`.

```
{
  "Principal": {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/datalake_user1"
  },
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
}
```

```
"Permissions": ["SELECT"]
}
```

Visualizzazione dei filtri di dati

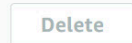
Puoi utilizzare la console Lake Formation o AWS CLI l'API Lake Formation per visualizzare i filtri dei dati.

Per visualizzare i filtri di dati, devi essere un amministratore di Data Lake o disporre delle autorizzazioni richieste sui filtri di dati.

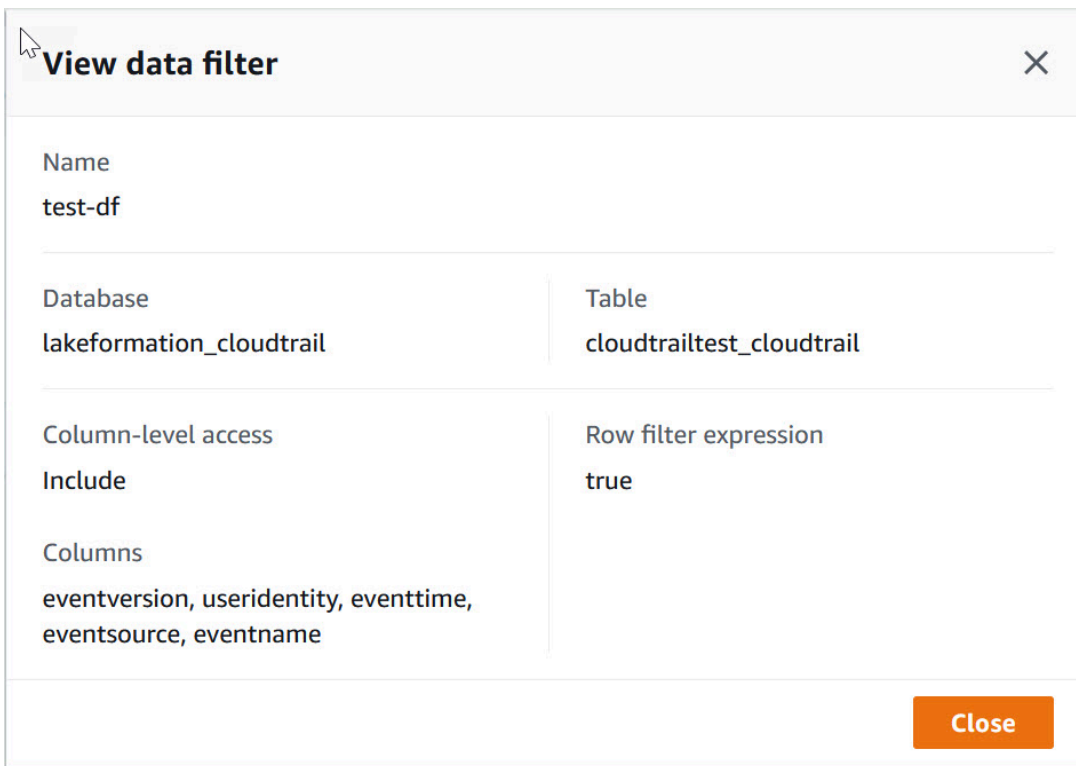
Console

1. Accedi a Console di gestione AWS e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel pannello di navigazione, in Catalogo dati, scegli Filtri dati.

La pagina mostra i filtri di dati a cui hai accesso.

Data filters (1)							
 Find filter				< 1 > 			
	Filter name	Table	Database	Table catalog ID			
	test-df	cloudtrailtest_cloudtrail	lakeformation_cloudtrail				

3. Per visualizzare i dettagli del filtro dati, scegli il filtro dati, quindi scegli Visualizza. Viene visualizzata una nuova finestra con informazioni dettagliate sul filtro dati.



The image shows a 'View data filter' dialog box with a close button (X) in the top right corner. It contains the following fields:

- Name:** test-df
- Database:** lakeformation_cloudtrail
- Table:** cloudtrailtest_cloudtrail
- Column-level access:** Include
- Row filter expression:** true
- Columns:** eventversion, useridentity, eventtime, eventsource, eventname

A 'Close' button is located at the bottom right of the dialog box.

AWS CLI

Immettete un `list-data-cells-filter` comando e specificate una risorsa della tabella.

L'esempio seguente elenca i filtri di dati per la `cloudtrailtest_cloudtrail` tabella.

```
aws lakeformation list-data-cells-filter --table '{ "CatalogId":"123456789012",
"DatabaseName":"lakeformation_cloudtrail", "Name":"cloudtrailtest_cloudtrail"}
```

API/SDK

Utilizzate l'`ListDataCellsFilterAPI` e specificate una risorsa per la tabella.

L'esempio seguente usa Python per elencare i primi 20 filtri di dati per la `myTable` tabella.

```
response = client.list_data_cells_filter(
    Table = {
        'CatalogId': '111122223333',
        'DatabaseName': 'mydb',
        'Name': 'myTable'
    },
    MaxResults=20
```

)

Elenco delle autorizzazioni per il filtro dei dati

Puoi utilizzare la console Lake Formation per visualizzare le autorizzazioni concesse sui filtri di dati.

Per visualizzare le autorizzazioni su un filtro dati, devi essere un amministratore di Data Lake o disporre delle autorizzazioni richieste sul filtro dati.

Console

1. Accedi a Console di gestione AWS e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nel pannello di navigazione, sotto Autorizzazioni, scegli Autorizzazioni dati.
3. Nella pagina Autorizzazioni dati, fai clic o tocca nel campo di ricerca e nel menu Proprietà scegli Tipo di risorsa.
4. Nel menu Tipo di risorsa, scegli Tipo di risorsa: filtro delle celle di dati.

Sono elencati i filtri di dati per i quali disponi delle autorizzazioni. Potrebbe essere necessario scorrere orizzontalmente per visualizzare le colonne Autorizzazioni e Concedibili.

Data Permissions (58)

Revoke

Grant

7 matches

Resource type: Data cell filter

Clear filter

	Principal	Resource type	Database	Table	Resource	Catalog	Permissions
☐	datalake_admin	Data cell filter	sales	orders	no-pharma	111122223333	Describe, Drop, Select
☐	datalake_admin	Data cell filter	sales	orders	restrict-pharma	111122223333	Describe, Drop, Select
☐	datalake_user1	Data cell filter	sales	orders	restrict-pharma	111122223333	Describe
☐	datalake_user2	Data cell filter	sales	orders	restrict-pharma	111122223333	Select

AWS CLI

- Immettete un comando. list-permissions Specificate DataCellsFilter per l'resourceargomento e specificate DESCRIBE o DROP per l'Permissionsargomento e, facoltativamente, per l'PermissionsWithGrantOptionargomento.

L'esempio seguente elenca DESCRIBE le autorizzazioni con l'opzione grant sul filtro dati. restrict-pharma I risultati sono limitati alle autorizzazioni concesse per il

principale `datalake_user1` e la `orders` tabella del `sales` database nell' AWS account 1111-2222-3333.

```
aws lakeformation list-permissions --cli-input-json file://list-params.json
```

Di seguito sono riportati i contenuti del file `grant-params.json`

```
{
  "Principal": {"DataLakePrincipalIdentifier":
    "arn:aws:iam::111122223333:user/datalake_user1"},
  "Resource": {
    "DataCellsFilter": {
      "TableCatalogId": "111122223333",
      "DatabaseName": "sales",
      "TableName": "orders",
      "Name": "restrict-pharma"
    }
  },
  "Permissions": ["DESCRIBE"],
  "PermissionsWithGrantOption": ["DESCRIBE"]
}
```

Visualizzazione delle autorizzazioni per database e tabelle in Lake Formation

È possibile visualizzare le autorizzazioni di Lake Formation concesse su un database o una tabella di Data Catalog. Puoi farlo utilizzando la console Lake Formation, l'API o il AWS Command Line Interface (AWS CLI).

Utilizzando la console, è possibile visualizzare le autorizzazioni a partire dalle pagine Database o Tabelle o dalla pagina Autorizzazioni dati.

Note

Se non sei un amministratore di database o il proprietario della risorsa, puoi visualizzare le autorizzazioni che altri principali hanno sulla risorsa solo se disponi dell'autorizzazione Lake Formation sulla risorsa con l'opzione di concessione.

Oltre alle autorizzazioni Lake Formation richieste, sono necessarie le autorizzazioni AWS Identity and Access Management (IAM) `glue:GetDatabases`, `glue:GetDatabase`, `glue:GetTables`, `glue:GetTable`, e `lakeformation:ListPermissions`.

Per visualizzare le autorizzazioni su un database (console, a partire dalla pagina Database)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore del database o come utente che dispone di un'autorizzazione Lake Formation sul database con l'opzione di concessione.

2. Nel riquadro di navigazione, scegliere Databases (Database).
3. Scegli un database e nel menu Azioni scegli Visualizza autorizzazioni.

Note

Se scegli un link alla risorsa del database, Lake Formation visualizza le autorizzazioni sul link alla risorsa, non sul database di destinazione del link alla risorsa.

La pagina delle autorizzazioni dei dati elenca tutte le autorizzazioni di Lake Formation per il database. Il nome del database e l'ID del catalogo (ID AWS account) del proprietario del database vengono visualizzati come etichette sotto la casella di ricerca. I riquadri indicano che è stato applicato un filtro alle autorizzazioni di elenco solo per quel database. Puoi regolare il filtro chiudendo un riquadro o scegliendo Cancella filtro.

Data permissions (1)							
Choose a database or table for which to review, grant or revoke user permissions.							
Find by properties < 1 > ⚙							
Database: logs X		Catalog ID: 11122223333 X		Clear filter			
Principal	Principal type	Resource type	Resource	Owner account ID	Permissions	Grantable	
☐ Administrator	IAM user	Database	logs	11122223333	Alter, Create table, Drop	Alter, Create table, Drop	

Per visualizzare le autorizzazioni su un database (console, a partire dalla pagina Autorizzazioni dati)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore del database o come utente che dispone di un'autorizzazione Lake Formation sul database con l'opzione di concessione.

2. Nel riquadro di navigazione, scegli Autorizzazioni dati.
3. Posiziona il cursore nella casella di ricerca nella parte superiore della pagina e nel menu Proprietà visualizzato, scegli Database.
4. Nel menu Database visualizzato, scegli un database.

 Note

Se scegli un link alla risorsa del database, Lake Formation visualizza le autorizzazioni sul link alla risorsa, non sul database di destinazione del link alla risorsa.

La pagina delle autorizzazioni dei dati elenca tutte le autorizzazioni di Lake Formation per il database. Il nome del database viene visualizzato come riquadro sotto la casella di ricerca. Il riquadro indica che è stato applicato un filtro per elencare le autorizzazioni solo per quel database. Puoi rimuovere il filtro chiudendo il riquadro o scegliendo Cancella filtro.

Per visualizzare le autorizzazioni su una tabella (console, a partire dalla pagina Tabelle)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore della tabella o come utente con un'autorizzazione Lake Formation sulla tabella con l'opzione di concessione.

2. Nel pannello di navigazione, seleziona Tabelle.
3. Scegli una tabella e nel menu Azioni scegli Visualizza autorizzazioni.

 Note

Se scegli un link alla risorsa alla tabella, Lake Formation visualizza le autorizzazioni sul link alla risorsa, non sulla tabella di destinazione del link alla risorsa.

La pagina Autorizzazioni dati elenca tutte le autorizzazioni di Lake Formation per la tabella. Il nome della tabella, il nome del database che contiene la tabella e l'ID del catalogo (ID AWS account) del proprietario della tabella vengono visualizzati come etichette sotto la casella di ricerca. Le etichette indicano che è stato applicato un filtro alle autorizzazioni di elenco solo per quella tabella. Puoi regolare il filtro chiudendo un'etichetta o scegliendo Cancella filtro.

Principal	Principal type	Resource type	Resource	Owner account ID	Permissions	Grantable
Administrator	IAM user	Table	alexa-logs	111122223333	Super	Super

Per visualizzare le autorizzazioni su una tabella (console, a partire dalla pagina Autorizzazioni dati)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake, creatore della tabella o come utente con un'autorizzazione Lake Formation sulla tabella con l'opzione di concessione.

2. Nel riquadro di navigazione, scegli Autorizzazioni dati.
3. Posiziona il cursore nella casella di ricerca nella parte superiore della pagina e nel menu Proprietà visualizzato, scegli Database.
4. Nel menu Database visualizzato, scegli un database.

Important

Se desideri visualizzare le autorizzazioni su una tabella condivisa con il tuo AWS account da un account esterno, devi scegliere il database nell'account esterno che contiene la tabella, non un collegamento di risorsa al database.

La pagina delle autorizzazioni dei dati elenca tutte le autorizzazioni di Lake Formation per il database.

5. Posiziona nuovamente il cursore nella casella di ricerca e nel menu Proprietà visualizzato, scegli Tabella.
6. Nel menu Tabelle visualizzato, scegli una tabella.

La pagina Autorizzazioni dati elenca tutte le autorizzazioni di Lake Formation per la tabella. Il nome della tabella e il nome del database che contiene la tabella vengono visualizzati come riquadri sotto la casella di ricerca. I riquadri indicano che è stato applicato un filtro alle autorizzazioni di elenco solo per quella tabella. Puoi regolare il filtro chiudendo un riquadro o scegliendo Cancella filtro.

Per visualizzare le autorizzazioni su una tabella ()AWS CLI

- Immettere un `list-permissions` comando.

L'esempio seguente elenca le autorizzazioni su una tabella condivisa da un account esterno. La `CatalogId` proprietà è l' AWS ID dell'account esterno e il nome del database si riferisce al database dell'account esterno che contiene la tabella.

```
aws lakeformation list-permissions --resource-type TABLE --resource '{ "Table":  
  {"DatabaseName":"logs", "Name":"alexa-logs", "CatalogId":"123456789012"} }'
```

Revoca dell'autorizzazione utilizzando la console Lake Formation

Puoi utilizzare la console per revocare tutti i tipi di autorizzazioni di Lake Formation: autorizzazioni Data Catalog, autorizzazioni per policy tag, autorizzazioni per filtri dati e autorizzazioni relative alla posizione.

Per revocare le autorizzazioni di Lake Formation su una risorsa (console)

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake o come utente a cui sono state concesse le autorizzazioni con l'opzione di concessione sulla risorsa.

2. Nel pannello di navigazione, in Autorizzazioni, scegli Autorizzazioni Data lake, Tag e autorizzazioni LF o Posizioni dati.
3. Seleziona l'autorizzazione o la posizione, quindi scegli Revoca.
4. Nella finestra di dialogo che si apre, scegli Revoca.

Condivisione dei dati tra account in Lake Formation

Le funzionalità cross-account di Lake Formation consentono agli utenti di condividere in modo sicuro i data lake distribuiti tra più AWS organizzazioni o direttamente con i responsabili IAM in un altro account AWS, fornendo un accesso granulare ai metadati del Data Catalog e ai dati sottostanti. Le grandi aziende in genere utilizzano più Account AWS account e molti di questi account potrebbero aver bisogno di accedere a un data lake gestito da un singolo account. Account AWS Gli utenti e i job di AWS Glue estrazione, trasformazione e caricamento (ETL) possono eseguire query e unire tabelle su più account e sfruttare comunque le protezioni dei dati a livello di tabella e colonna di Lake Formation.

Quando concedi le autorizzazioni di Lake Formation su una risorsa Data Catalog a un account esterno o direttamente a un responsabile IAM in un altro account, Lake Formation utilizza il servizio AWS Resource Access Manager (AWS RAM) per condividere la risorsa. Se l'account assegnatario appartiene alla stessa organizzazione dell'account concedente, la risorsa condivisa è immediatamente disponibile per l'assegnatario. Se l'account del beneficiario non appartiene alla stessa organizzazione, AWS RAM invia un invito all'account del beneficiario per accettare o rifiutare la concessione di risorse. Quindi, per rendere disponibile la risorsa condivisa, l'amministratore del data lake nell'account del beneficiario deve utilizzare la console o accettare l' AWS RAM invito. AWS CLI

Lake Formation supporta la condivisione delle risorse del Data Catalog con account esterni in modalità di accesso ibrido. La modalità di accesso ibrido offre la flessibilità necessaria per abilitare selettivamente le autorizzazioni di Lake Formation per database e tabelle del tuo. AWS Glue Data Catalog

Con la modalità di accesso Hybrid, ora disponi di un percorso incrementale che ti consente di impostare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti.

Per ulteriori informazioni, consulta [Modalità di accesso ibrida](#).

Condivisione diretta tra account

I responsabili autorizzati possono condividere le risorse in modo esplicito con un responsabile IAM in un account esterno. Questa funzionalità è utile quando il proprietario di un account desidera avere il controllo su chi nell'account esterno può accedere alle risorse. Le autorizzazioni ricevute dal preside IAM saranno costituite da un'unione di concessioni dirette e concessioni a livello di account, che verranno trasferite a cascata ai principali. Solo il destinatario della concessione di autorizzazione può

visualizzare le concessioni dirette tra account. Il principale che riceve la condivisione di risorse non può condividere la risorsa con altri destinatari.

Metodi per condividere le risorse del Data Catalog

Con un'unica operazione di concessione di Lake Formation, puoi concedere autorizzazioni tra account sulle seguenti risorse del Data Catalog.

- Un database
- Una tabella singola (con filtro opzionale per le colonne)
- Alcune tabelle selezionate
- Tutte le tabelle di un database (utilizzando il carattere jolly Tutte le tabelle)

Esistono due opzioni per condividere database e tabelle con un altro account Account AWS o con i principali IAM di un altro account.

- Controllo degli accessi basato su tag Lake Formation (LF-TBAC) (consigliato)

Il controllo degli accessi basato su tag Lake Formation è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. Puoi utilizzare il controllo degli accessi basato su tag per condividere le risorse del Data Catalog (database, tabelle e colonne) con responsabili IAM esterni, Account AWS Organizzazioni e unità organizzative (OUs). In Lake Formation, questi attributi sono chiamati LF-tag. Per ulteriori informazioni, consulta [Gestione di un data lake utilizzando il controllo degli accessi basato su tag Lake Formation](#).

Note

Il metodo LF-TBAC per la concessione delle autorizzazioni di Data Catalog viene utilizzato per le concessioni tra account. AWS Resource Access Manager Lake Formation ora supporta la concessione di autorizzazioni tra account a Organizzazioni e unità organizzative utilizzando il metodo LF-TBAC. Per abilitare questa funzionalità, è necessario aggiornare le impostazioni della versione dell'account Cross alla versione 3 o successiva. Per ulteriori informazioni, consulta [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#).

- Risorse denominate Lake Formation

La condivisione dei dati tra account di Lake Formation utilizzando il metodo delle risorse denominate consente di concedere le autorizzazioni di Lake Formation con un'opzione di concessione su tabelle e database di Data Catalog a dirigenti Account AWS, organizzazioni o unità organizzative IAM esterni. L'operazione di concessione condivide automaticamente tali risorse.

Note

Puoi anche consentire al AWS Glue crawler di accedere a un data store in un account diverso utilizzando le credenziali di Lake Formation. Per ulteriori informazioni, consulta la sezione Scansione [tra account nella Guida per gli sviluppatori](#). AWS Glue

I servizi integrati come Athena e Amazon Redshift Spectrum richiedono collegamenti alle risorse per poter includere risorse condivise nelle query. Per ulteriori informazioni sui link alle risorse, consulta.

[Come funzionano i link alle risorse in Lake Formation](#)

Per considerazioni e limitazioni, vedere [Buone pratiche e considerazioni sulla condivisione dei dati tra account](#).

Argomenti

- [Prerequisiti](#)
- [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#)
- [Condivisione delle tabelle e dei database del Data Catalog tra i nostri principali IAM provenienti Account AWS da account esterni](#)
- [Concessione delle autorizzazioni su un database o una tabella condivisa con il tuo account](#)
- [Concessione delle autorizzazioni per i collegamenti alle risorse](#)
- [Accesso ai dati sottostanti di una tabella condivisa](#)
- [Registrazione su più account CloudTrail](#)
- [Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation](#)
- [Visualizzazione di tutte le sovvenzioni tra account utilizzando l'operazione API GetResourceShares](#)

Argomenti correlati

- [Panoramica delle autorizzazioni di Lake Formation](#)

- [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#)
- [Creazione di collegamenti alle risorse](#)
- [Risoluzione dei problemi di accesso tra account](#)

Prerequisiti

Prima che l' AWS account possa condividere le risorse di Data Catalog (cataloghi, database e tabelle) con un altro account o i responsabili di un altro account e prima di poter accedere alle risorse condivise con il proprio account, devono essere soddisfatti i seguenti prerequisiti.

Requisiti generali per la condivisione dei dati tra account

- Per condividere i database e le tabelle di Data Catalog in modalità di accesso ibrido e condividere oggetti nei cataloghi federati, è necessario aggiornare le impostazioni della versione dell'account Cross alla versione 4.
- Prima di concedere autorizzazioni per più account su una risorsa Data Catalog, devi revocare tutte le autorizzazioni di Lake Formation dal `IAMAllowedPrincipals` gruppo per la risorsa. Se il principale chiamante dispone di autorizzazioni multiaccount per accedere a una risorsa e l'`IAMAllowedPrincipals` autorizzazione esiste sulla risorsa, Lake Formation lancia `AccessDeniedException`.

Questo requisito è applicabile solo quando si registra la posizione dei dati sottostante in modalità Lake Formation. Se si registra la posizione dei dati in modalità ibrida, le autorizzazioni di `IAMAllowedPrincipals` gruppo possono esistere sul database o sulla tabella condivisi.

- Per i database che contengono tabelle che si intende condividere, è necessario impedire che alle nuove tabelle venga assegnato di default `Super aIAMAllowedPrincipals`. Sulla console Lake Formation, modifica il database e disattiva Usa solo il controllo di accesso IAM per le nuove tabelle in questo database o inserisci il seguente AWS CLI comando, sostituendolo `database` con il nome del database. Se la posizione dei dati sottostante è registrata in modalità di accesso ibrida, non è necessario modificare questa impostazione predefinita. In modalità di accesso ibrido, Lake Formation consente di applicare selettivamente le autorizzazioni di Lake Formation e le politiche di autorizzazione IAM per Amazon S3 e sulla stessa risorsa. AWS Glue

```
aws glue update-database --name database --database-input  
'{"Name": "database", "CreateTableDefaultPermissions": []}'
```


- Per concedere le autorizzazioni su più account, il concedente deve disporre delle autorizzazioni richieste (IAM) sul servizio. AWS Identity and Access Management AWS Glue AWS RAM La policy AWS AWSLakeFormationCrossAccountManager gestita concede le autorizzazioni richieste.

Gli amministratori di Data Lake negli account che ricevono condivisioni di risorse utilizzando AWS RAM devono disporre della seguente politica aggiuntiva. Consente all'amministratore di accettare gli inviti alla condivisione AWS RAM delle risorse. Consente inoltre all'amministratore di abilitare la condivisione delle risorse con le organizzazioni.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ram:AcceptResourceShareInvitation",
        "ram:RejectResourceShareInvitation",
        "ec2:DescribeAvailabilityZones",
        "ram:EnableSharingWithAwsOrganization"
      ],
      "Resource": "*"
    }
  ]
}
```

- Se si desidera condividere le risorse di Data Catalog con le AWS Organizations nostre unità organizzative, è necessario abilitare la condivisione con le organizzazioni AWS RAM.

Per informazioni su come abilitare la condivisione con le organizzazioni, consulta [Abilitare la condivisione con AWS le organizzazioni](#) nella Guida per l'AWS RAM utente.

È necessario disporre dell'`ram:EnableSharingWithAwsOrganization` autorizzazione per abilitare la condivisione con le organizzazioni.

- Per condividere le risorse direttamente con un responsabile IAM in un altro account, devi aggiornare le impostazioni della versione dell'account Cross alla versione 3. Questa impostazione è disponibile nella pagina delle impostazioni del catalogo dati. Se utilizzi la versione 1, consulta

le istruzioni per aggiornare l'impostazione [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#).

- Non è possibile condividere le risorse del Data Catalog crittografate con una chiave gestita dal AWS Glue servizio con un altro account. È possibile condividere solo le risorse del Data Catalog crittografate con la chiave di crittografia del cliente e l'account che riceve la condivisione delle risorse deve disporre delle autorizzazioni sulla chiave di crittografia del Data Catalog per decrittografare gli oggetti.

Condivisione dei dati tra account utilizzando i requisiti LF-TBAC

- Per condividere le risorse di Data Catalog con AWS Organizations le unità organizzative (OUs), devi aggiornare le impostazioni della versione dell'account Cross alla versione 3 o successiva.
- Per condividere le risorse di Data Catalog con la versione 3 delle impostazioni della versione dell'account Cross, il concedente deve disporre delle autorizzazioni IAM definite `AWSLakeFormationCrossAccountManager` nella politica AWS gestita del proprio account.
- Se utilizzi la versione 1 o la versione 2 delle impostazioni della versione dell'account Cross, devi disporre di una politica delle risorse di Data Catalog (`glue:PutResourcePolicy`) che abiliti LF-TBAC. Per ulteriori informazioni, consulta [Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation](#).
- Se attualmente utilizzi una politica delle risorse di AWS Glue Data Catalog per condividere risorse e desideri concedere autorizzazioni tra account utilizzando la versione 3 delle impostazioni della versione Cross account, devi aggiungere l'**glue:ShareResource** autorizzazione nelle Impostazioni del Data Catalog utilizzando l'operazione `glue:PutResourcePolicy` API, come mostrato nella sezione. [Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation](#) Questa politica non è richiesta se il tuo account non ha concesso concessioni tra account utilizzando la politica delle risorse di AWS Glue Data Catalog (`glue:PutResourcePolicy` autorizzazione all'uso delle versioni 1 e 2) per concedere l'accesso a più account.

```
{
  "Effect": "Allow",
  "Action": [
    "glue:ShareResource"
  ],
  "Principal": {"Service": [
    "ram.amazonaws.com"
  ]},
}
```

```
"Resource": [  
  "arn:aws:glue:<region>:<account-id>:table/*/*",  
  "arn:aws:glue:<region>:<account-id>:database/*",  
  "arn:aws:glue:<region>:<account-id>:catalog"  
]  
}
```

- Se il tuo account ha effettuato condivisioni tra account utilizzando la politica delle risorse di AWS Glue Data Catalog e attualmente utilizzi il metodo Named Resource o LF-TBAC con impostazioni Cross account versione 3 per condividere le risorse, che utilizza AWS RAM per condividere le risorse, devi impostare l'EnableHybridargomento su quando richiami l'operazione API. 'true' glue:PutResourcePolicy Per ulteriori informazioni, consulta [Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation](#).

Configurazione richiesta in ogni account che accede alla risorsa condivisa

- Se condividi risorse con Account AWS, almeno un utente dell'account consumer deve essere un amministratore del data lake per visualizzare le risorse condivise. Per informazioni su come creare un amministratore di data lake, consulta [Crea un amministratore del data lake](#).

L'amministratore del data lake può concedere le autorizzazioni di Lake Formation sulle risorse condivise ad altri responsabili dell'account. Gli altri responsabili non possono accedere alle risorse condivise finché l'amministratore del data lake non concede loro le autorizzazioni sulle risorse.

- I servizi integrati come Athena e Redshift Spectrum richiedono collegamenti alle risorse per poter includere risorse condivise nelle query. I responsabili devono creare un collegamento di risorsa nel proprio catalogo dati a una risorsa condivisa da un'altra risorsa. Account AWS Per ulteriori informazioni sui collegamenti alle risorse, consulta [Come funzionano i link alle risorse in Lake Formation](#).
- Quando una risorsa viene condivisa direttamente con un principale IAM, per interrogare la tabella utilizzando Athena, il principale deve creare un collegamento alla risorsa. Per creare un collegamento a una risorsa, il principale necessita dell'CREATE_DATABASEautorizzazione CREATE_TABLE o del Lake Formation e dell'autorizzazione glue:CreateTable o glue>CreateDatabase IAM.

Se l'account produttore condivide una tabella diversa dello stesso database con lo stesso principale o con un altro principale, tale principale può interrogare immediatamente la tabella.

 Note

Per l'amministratore del data lake e per i responsabili a cui l'amministratore del data lake ha concesso le autorizzazioni, le risorse condivise vengono visualizzate nel Data Catalog come se fossero risorse locali (di proprietà). I job di estrazione, trasformazione e caricamento (ETL) possono accedere ai dati sottostanti delle risorse condivise.

Per le risorse condivise, le pagine Tabelle e database della console Lake Formation visualizzano l'ID dell'account del proprietario.

Quando si accede ai dati sottostanti di una risorsa condivisa, gli eventi di CloudTrail registro vengono generati sia nell'account del destinatario della risorsa condivisa che nell'account del proprietario della risorsa. Gli CloudTrail eventi possono contenere l'ARN del principale che ha avuto accesso ai dati, ma solo se l'account del destinatario sceglie di includere l'ARN principale nei registri. Per ulteriori informazioni, consulta [Registrazione su più account CloudTrail](#).

Aggiornamento delle impostazioni della versione di condivisione dei dati tra account

Di tanto in tanto, AWS Lake Formation aggiorna le impostazioni di condivisione dei dati tra account per distinguere le modifiche apportate all' AWS RAM utilizzo e per supportare gli aggiornamenti apportati alla funzionalità di condivisione dei dati tra account. Quando Lake Formation esegue questa operazione, crea una nuova versione delle impostazioni della versione dell'account Cross.

Principali differenze tra le impostazioni delle versioni per più account

Per ulteriori informazioni su come funziona la condivisione dei dati tra account in diverse impostazioni di versione di Cross account, consulta le seguenti sezioni.

 Note

Per condividere i dati con un altro account, il concedente deve avere `AWSLakeFormationCrossAccountManager` gestito le autorizzazioni relative alla policy IAM. Questo è un prerequisito per tutte le versioni.

L'aggiornamento delle impostazioni della versione dell'account Cross non influisce sulle autorizzazioni del destinatario sulle risorse condivise. Ciò è applicabile quando si esegue l'aggiornamento dalla versione 1 alla versione 2, dalla versione 2 alla versione 3 e dalla

versione 1 alla versione 3. Per l'aggiornamento delle versioni, consulta le considerazioni elencate di seguito.

Versione 1

Metodo di risorsa denominato: associa ogni autorizzazione concessa tra account Lake Formation a una condivisione di AWS RAM risorse. L'utente (concedente, ruolo o responsabile) non richiede autorizzazioni aggiuntive.

Metodo LF-TBAC: le concessioni di autorizzazioni di Lake Formation tra account non vengono utilizzate per condividere dati. AWS RAM L'glue:PutResourcePolicyutente deve avere l'autorizzazione.

Vantaggi dell'aggiornamento delle versioni: versione iniziale, non applicabile.

Considerazioni sull'aggiornamento delle versioni: Versione iniziale - non applicabile

Versione 2

Metodo della risorsa denominata: ottimizza il numero di condivisioni di AWS RAM risorse mappando più concessioni di autorizzazioni tra account con un'unica condivisione di risorse. AWS RAM L'utente non richiede autorizzazioni aggiuntive.

Metodo LF-TBAC: le concessioni di autorizzazioni di Lake Formation tra account non vengono utilizzate per condividere dati. AWS RAM L'glue:PutResourcePolicyutente deve avere l'autorizzazione.

Vantaggi dell'aggiornamento delle versioni: configurazione scalabile tra più account grazie all'utilizzo ottimale della capacità. AWS RAM

Considerazioni sull'aggiornamento delle versioni: gli utenti che desiderano concedere le autorizzazioni Lake Formation su più account devono disporre delle autorizzazioni nella politica gestita. AWSLakeFormationCrossAccountManager AWS In caso contrario, è necessario disporre ram:AssociateResourceShare delle ram:DisassociateResourceShare autorizzazioni necessarie per condividere correttamente le risorse con un altro account.

Versione 3

Metodo della risorsa denominata: ottimizza il numero di condivisioni di AWS RAM risorse mappando più concessioni di autorizzazioni tra account con un'unica condivisione di risorse. AWS RAM L'utente non richiede autorizzazioni aggiuntive.

Metodo LF-TBAC: Lake Formation utilizza AWS RAM per le sovvenzioni tra account. L'utente deve aggiungere la dichiarazione glue: all'autorizzazione. ShareResource glue:PutResourcePolicy Il destinatario deve accettare gli inviti alla condivisione delle risorse da AWS RAM.

Vantaggi dell'aggiornamento delle versioni: supporta le seguenti funzionalità:

- Consente di condividere le risorse in modo esplicito con un principale IAM in un account esterno.

Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per le risorse di Data Catalog](#).

- Abilita le condivisioni tra account utilizzando il metodo LF-TBAC per Organizzazioni o unità organizzative (). OUs
- Elimina il sovraccarico derivante dal mantenimento di politiche aggiuntive per le sovvenzioni tra account. AWS Glue

Considerazioni sull'aggiornamento delle versioni: quando si utilizza il metodo LF-TBAC per condividere risorse, se il concedente utilizza una versione precedente alla versione 3 e il destinatario utilizza la versione 3 o successiva, il concedente riceve il seguente messaggio di errore: «Richiesta di concessione tra account non valida. L'account consumer dispone della versione opt-in per più account: v3. Effettua l'aggiornamento CrossAccountVersion DataLakeSetting alla versione minima v3 (Servizio: AmazonDataCatalog; Codice di stato: 400; Codice di errore: InvalidInputException)». Tuttavia, se il concedente utilizza la versione 3 e il destinatario utilizza la versione 1 o la versione 2, le sovvenzioni tra account diversi che utilizzano i tag LF vanno a buon fine.

Le sovvenzioni tra account effettuate utilizzando il metodo della risorsa denominata sono compatibili tra diverse versioni. Anche se l'account concedente utilizza una versione precedente (versione 1 o 2) e l'account del destinatario utilizza una versione più recente (versione 3 o successiva), la funzionalità di accesso tra account funziona senza problemi di compatibilità o errori.

Per condividere le risorse direttamente con i responsabili IAM di un altro account, solo il concedente deve utilizzare la versione 3.

Le sovvenzioni tra account effettuate utilizzando il metodo LF-TBAC richiedono agli utenti di disporre di una politica delle risorse nell'account. AWS Glue Data Catalog Quando si esegue l'aggiornamento alla versione 3, LF-TBAC concede gli utilizzi. AWS RAM Per consentire

l'esecuzione delle sovvenzioni AWS RAM basate su più account, è necessario aggiungere la `glue:ShareResource` dichiarazione alle politiche esistenti in materia di risorse del Data Catalog, come mostrato nella sezione. [Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation](#)

Versione 4

Il concedente necessita della versione 4 o successiva per condividere le risorse di Data Catalog in modalità di accesso ibrido o condividere oggetti in un catalogo federato.

Ottimizza la condivisione delle risorse AWS RAM

Le nuove versioni (versione 2 e successive) delle sovvenzioni tra account utilizzano in modo ottimale la AWS RAM capacità per massimizzare l'utilizzo tra account. Quando condividi una risorsa con un responsabile IAM Account AWS o esterno, Lake Formation può creare una nuova condivisione di risorse o associare la risorsa a una condivisione esistente. Associandosi alle azioni esistenti, Lake Formation riduce il numero di inviti alla condivisione delle risorse che un consumatore deve accettare.

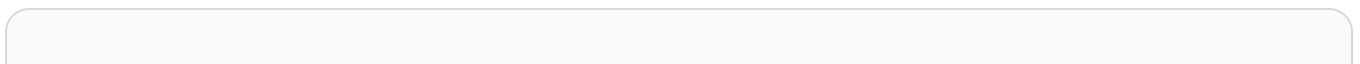
Abilita AWS RAM le condivisioni tramite TBAC o condividi le risorse direttamente con i principali

Per condividere le risorse direttamente con i responsabili IAM in un altro account o per abilitare le condivisioni TBAC tra account su Organizations o unità organizzative, devi aggiornare le impostazioni della versione Cross account alla versione 3. Per ulteriori informazioni sui limiti delle AWS RAM risorse, consulta. [Buone pratiche e considerazioni sulla condivisione dei dati tra account](#)

Autorizzazioni necessarie per l'aggiornamento delle impostazioni delle versioni tra account

Se un concedente di autorizzazioni su più account ha `AWSLakeFormationCrossAccountManager` gestito le autorizzazioni relative alle policy IAM, non è richiesta alcuna configurazione aggiuntiva delle autorizzazioni per il ruolo o il responsabile del concedente delle autorizzazioni su più account. Tuttavia, se il concedente che concede più account non utilizza la policy gestita, affinché la nuova versione della concessione tra più account abbia successo, al ruolo o al responsabile del concedente devono essere concesse le seguenti autorizzazioni IAM.

JSON



```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
        "ram:AssociateResourceShare",
        "ram:DisassociateResourceShare",
        "ram:GetResourceShares"
      ],
      "Resource": "*",
      "Condition": {
        "StringLike": {
          "ram:ResourceShareName": "LakeFormation*"
        }
      }
    }
  ]
}
```

Per abilitare la nuova versione

Segui questi passaggi per aggiornare le impostazioni della versione dell'account Cross tramite la AWS Lake Formation console o il AWS CLI.

Console

1. Scegli la versione 2, la versione 3 o la versione 4 nelle impostazioni della versione dell'account Cross nella pagina delle impostazioni del catalogo dati. Se selezioni la versione 1, Lake Formation utilizzerà la modalità di condivisione delle risorse predefinita.

AWS Lake Formation > Data catalog settings

Data catalog settings

Default permissions for newly created databases and tables

These settings maintain existing AWS Glue Data Catalog behavior. You can still set individual permissions on databases and tables, which will take effect when you revoke the Super permission from IAMAllowedPrincipals. See [Changing Default Settings for Your Data Lake](#).

- ☐ Use only IAM access control for new databases
- ☐ Use only IAM access control for new tables in new databases

Default permissions for AWS CloudTrail

These settings specify the information being shown in AWS CloudTrail.

Resource owners

Enter resource owners you wish to share your CloudTrail access details with.

 *Enter an AWS account ID*

Enter one or more AWS account IDs. Press Enter after each ID.

Cross account version settings

Version 1

Version 2

Version 3

Version 3

cross account permissions. See

Cancel

Save

2. Scegli Save (Salva).

AWS Command Line Interface (AWS CLI)

Utilizzate il `put-data-lake-settings` AWS CLI comando per impostare il parametro. `CROSS_ACCOUNT_VERSION` I valori accettati sono 1, 2, 3 e 4.

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
  file://settings
{
```

```
"DataLakeAdmins": [
  {
    "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/test"
  }
],
"CreateDatabaseDefaultPermissions": [],
"CreateTableDefaultPermissions": [],
"Parameters": {
  "CROSS_ACCOUNT_VERSION": "3"
}
}
```

Important

Dopo aver scelto la versione 2 o la versione 3, tutte le nuove sovvenzioni per risorse denominate passeranno attraverso la nuova modalità di concessione tra account. Per utilizzare in modo ottimale la AWS RAM capacità delle condivisioni esistenti su più account, ti consigliamo di revocare le sovvenzioni concesse con la versione precedente e di riassegnarle nella nuova modalità.

Condivisione delle tabelle e dei database del Data Catalog tra i nostri principali IAM provenienti Account AWS da account esterni

Questa sezione include istruzioni su come concedere autorizzazioni su più account sulle risorse di Data Catalog a un AWS account esterno, un responsabile IAM, un' AWS organizzazione o un'unità organizzativa. L'operazione di concessione condivide automaticamente tali risorse.

Argomenti

- [Condivisione dei dati tramite controllo degli accessi basato su tag](#)
- [Condivisione dei dati tra account utilizzando il metodo della risorsa denominato](#)

Condivisione dei dati tramite controllo degli accessi basato su tag

AWS Lake Formation il controllo degli accessi basato su tag (LF-TBAC) è una strategia di autorizzazione che definisce le autorizzazioni in base agli attributi. I passaggi seguenti spiegano come concedere le autorizzazioni per più account utilizzando i tag LF.

Configurazione obbligatoria sull'account producer/grantor

1. Aggiungi tag LF.

1. Accedi alla console di Lake Formation come amministratore di data lake o creatore di tag LF.
2. Nella barra di navigazione a sinistra, scegli Autorizzazioni e LF-Tag e autorizzazioni.
3. Scegli Aggiungi tag LF.

Per istruzioni dettagliate sulla creazione di tag LF, consulta. [Creazione di tag LF](#)

2. Concedi le autorizzazioni Descrivi and/or Associate le coppie chiave-valore LF-Tag ai principali IAM del tuo account o degli account esterni.

La concessione delle autorizzazioni sulle coppie chiave-valore LF-Tag consente ai principali di visualizzare i tag LF e di assegnarli alle risorse del Data Catalog (database, tabelle e colonne).

3. Successivamente, l'amministratore del data lake o un responsabile IAM con autorizzazione Associate può assegnare il tag LF a database, tabelle o colonne. Per ulteriori informazioni, consulta [Assegnazione di tag LF alle risorse del Data Catalog](#).
4. Successivamente, concedi l'autorizzazione ai dati agli account esterni utilizzando le espressioni LF-Tag. Ciò consente al concessionario o al destinatario delle autorizzazioni di accedere alle risorse del Data Catalog contrassegnate con le stesse chiavi e gli stessi valori.

1. Nel riquadro di navigazione, scegli Autorizzazioni e Autorizzazioni dati.
2. Scegli Concessione.
3. Nella pagina Concedi autorizzazioni, per Principal, scegli Account esterni e inserisci l' Account AWS ID del beneficiario o il ruolo IAM del principale o l'Amazon Resource Name (ARN) per il principale (ARN principale) se effettui una concessione diretta tra account a un principale esterno. Devi premere Invio dopo aver inserito l'ID dell'account.

Principals

Choose the principals to grant permissions.

☐ **IAM users and roles**
Users or roles from this AWS account.

☐ **IAM Identity Center**
- *new*
Users and groups configured in IAM Identity Center.

☐ **SAML users and groups**
SAML users and group or QuickSight ARNs.

☒ **External accounts**
AWS account, AWS organization or IAM principal outside of this account

AWS account, AWS organization, or IAM principal outside of this account

Enter one or more AWS account IDs, AWS organization IDs, or IAM principal ARNs. Press Enter after each ID or ARN.

Granting data permissions to organizations is not supported when granting permissions by using LF-Tags.

LF-Tags or catalog resources

Choose a method to grant permissions.

☒ **Resources matched by LF-Tags (recommended)**
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☐ **Named Data Catalog resources**
Manage permissions for specific databases or tables, in addition to fine-grained data access.

☒ **LF-Tag key-value pairs**
☐ **Saved LF-Tag expressions - new**

Key

Department ▼

Values

Choose LF-Tag values ▼

Sales ✕

Marketing ✕

Add LF-Tag key-value pair

4. Per i tag LF o le risorse del catalogo, scegliete Risorse abbinate ai tag LF (consigliato).
 - a. Scegliete l'opzione Coppie chiave-valore del tag LF o Espressioni di tag LF salvate.
 - b. Se scegliete le coppie chiave-valore del tag LF, inserite la chiave e i valori del tag LF associato alla risorsa Data Catalog condivisa con l'account beneficiario.

Al beneficiario vengono concesse le autorizzazioni sulle risorse del Data Catalog a cui è stato assegnato un tag LF corrispondente nell'espressione LF-Tag. Se l'espressione LF-Tag specifica più valori per chiave di tag, uno qualsiasi dei valori del tag può corrispondere.

5. Scegliete le autorizzazioni a livello di database o di tabella da concedere alle risorse che corrispondono all'espressione LF-Tag.

 Important

Poiché l'amministratore del data lake deve concedere le autorizzazioni sulle risorse condivise ai responsabili dell'account beneficiario, devi sempre concedere le autorizzazioni su più account con l'opzione di concessione.

Per ulteriori informazioni, consulta [Concessione delle autorizzazioni LF-Tag tramite la console](#).

 Note

I mandanti che ricevono sovvenzioni dirette su più account non avranno l'opzione Autorizzazioni concesse.

Configurazione obbligatoria sull'account receiving/grantee

1. Accedi alla console di Lake Formation come amministratore del data lake dell'account consumatore.
2. Successivamente, ricevi la condivisione delle risorse nell'account consumatore.
 1. Apri la AWS RAM console.
 2. Nel riquadro di navigazione, in Condivisi con me, scegli Condivisioni di risorse.
 3. Seleziona le condivisioni di risorse, scegli Accetta condivisione di risorse.
3. Quando condividi una risorsa con un altro account, la risorsa appartiene ancora all'account del produttore e non è visibile nella console Athena. Per rendere visibile la risorsa nella console Athena, devi creare un link alla risorsa che punti alla risorsa condivisa. Per istruzioni sulla creazione di un collegamento a una risorsa, consulta e [Creazione di un collegamento di risorsa a una tabella condivisa del Catalogo dati](#) [Creazione di un collegamento di risorsa a un database](#) [Data Catalog condiviso](#)
 1. Scegli Database o Tabelle nel Catalogo dati.
 2. Nella Databases/Tables pagina, scegli Crea, Link alle risorse.
 3. Immettete le seguenti informazioni per un collegamento alla risorsa del database:

- Nome del link alla risorsa: un nome univoco per il link alla risorsa.
 - Catalogo di destinazione: il catalogo in cui stai creando il link alla risorsa.
 - Regione del database condivisa: la regione del database condivisa con te se stai creando il link alla risorsa in un'altra regione.
 - Database condiviso: scegli il database condiviso.
 - ID di catalogo del database condiviso: immettere l'ID del catalogo per il database condiviso.
4. Scegli Create (Crea). È possibile visualizzare il collegamento alla risorsa appena creato nell'elenco dei database.

Allo stesso modo, è possibile creare un collegamento di risorsa a una tabella condivisa.

4. Ora concedi l'autorizzazione Descrivi sul link della risorsa ai principali IAM con cui condividi la risorsa.
1. Nella pagina Database/Tabelle, seleziona il link alla risorsa e nel menu Azioni scegli Concedi.
 2. Nella sezione Concedi autorizzazioni, seleziona Utenti e ruoli IAM.
 3. Scegli il ruolo IAM a cui desideri concedere l'accesso al link della risorsa.
 4. Nella sezione Autorizzazioni del collegamento alle risorse, seleziona Descrivi.
 5. Scegli Concessione.
5. Successivamente, concedi le autorizzazioni chiave-valore di LF-Tag ai principali dell'account consumatore.

Dovresti essere in grado di trovare i tag LF condivisi con te nell'account utente sulla console di Lake Formation, sotto Autorizzazioni, LF-tags e permessi. Puoi associare i tag condivisi dal concedente alle risorse condivise dall'account concedente che includono: database, tabelle e colonne. È possibile concedere ulteriormente le autorizzazioni sulle risorse ad altri responsabili.

LF-Tags | LF-Tag expressions - new | **LF-Tag permissions** | LF-Tag creators

LF-Tag permissions (8)  [View](#) [Revoke](#) [Grant permissions](#)

View and manage the permissions granted on LF-Tags. [Learn more](#)

	Principal ▲	Principal type ▼	Resource name ▼	Resource type ▼	LF-Tag permissions ▼	LF-Tag value permissions ▼
☐		AWS account	Department=Marketing, Sales	LF-Tag key value pair	-	Describe
☐		AWS account	Department=Marketing, Sales	LF-Tag key value pair	-	Associate
☐	arn:aws:iam::()/Admin	IAM role	Department	LF-Tag	Alter, Drop	-
☐	arn:aws:iam::()/Admin	IAM role	Department	LF-Tag	-	Describe
☐	arn:aws:iam::()/Admin	IAM role	Department	LF-Tag	-	Associate
☐	arn:aws:iam::()/Admin	IAM role	Department	LF-Tag	-	Grant with LF-Tag expression
☐	arn:aws:iam::()/Admin	IAM role	LF-Test	LF-Tag	-	Describe
☐	arn:aws:iam::()/Admin	IAM role	LF-Test	LF-Tag	-	Associate

1. Nel pannello di navigazione, in Autorizzazioni, Autorizzazioni dati, scegli Concedi.
2. Nella pagina Concedi autorizzazioni, scegli Utenti e ruoli IAM.
3. Quindi, scegli gli utenti e i ruoli IAM nel tuo account a cui concedere l'accesso ai database/ tabelle condivisi.
4. Successivamente, per i tag LF o le risorse del catalogo, scegli Risorse abbinate ai tag LF.
5. Quindi, scegli la chiave e i valori del tag LF che è condiviso con te.
6. Quindi, scegli le autorizzazioni per database e tabelle che desideri concedere agli utenti e ai ruoli IAM. Puoi anche scegliere le autorizzazioni Grantable che consentono agli utenti e ai ruoli IAM di concedere autorizzazioni ad altri utenti/ruoli.
7. Scegli Concessione.
8. Puoi visualizzare le autorizzazioni concesse in Autorizzazioni dati sulla console Lake Formation.

Condivisione dei dati tra account utilizzando il metodo della risorsa denominato

È possibile concedere le autorizzazioni direttamente ai responsabili di un altro AWS account o a utenti esterni o. Account AWS AWS Organizations Concedere i permessi di Lake Formation a Organizzazioni o unità organizzative equivale a concedere il permesso Account AWS a tutti i membri di quell'organizzazione o unità organizzativa.

Quando concedi autorizzazioni ad account o organizzazioni esterne, devi includere l'opzione Autorizzazioni concedibili. Solo l'amministratore del data lake dell'account esterno può accedere alle

risorse condivise finché l'amministratore non concede le autorizzazioni sulle risorse condivise ad altri responsabili dell'account esterno.

Note

L'opzione delle autorizzazioni concedibili non è supportata quando si concedono le autorizzazioni direttamente ai principali IAM da account esterni.

Segui le istruzioni per concedere le autorizzazioni [Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#) per più account utilizzando il metodo di risorsa denominato.

Concessione delle autorizzazioni su un database o una tabella condivisa con il tuo account

Dopo che una risorsa Data Catalog appartenente a un altro AWS account è stata condivisa con il tuo AWS account, in qualità di amministratore del data lake, puoi concedere le autorizzazioni sulla risorsa condivisa ad altri responsabili del tuo account. Tuttavia, non puoi concedere le autorizzazioni sulla risorsa ad altri AWS account o organizzazioni.

Puoi utilizzare la AWS Lake Formation console, l'API o AWS Command Line Interface (AWS CLI) per concedere le autorizzazioni.

Per concedere le autorizzazioni su un database condiviso (denominato metodo di risorsa, console)

- Segui le istruzioni in [Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#). Nell'elenco dei database sotto i tag LF o le risorse del catalogo, assicuratevi di selezionare il database nell'account esterno, non un collegamento alla risorsa per il database.

Se non vedi il database nell'elenco dei database, assicurati di aver accettato l'invito AWS Resource Access Manager (AWS RAM) alla condivisione delle risorse per il database. Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

Inoltre, per le ALTER autorizzazioni CREATE_TABLE e, segui le istruzioni riportate in [Concessione delle autorizzazioni per la localizzazione dei dati \(stesso account\)](#) e assicurati di inserire l'ID dell'account proprietario nel campo Ubicazione dell'account registrato.

Per concedere le autorizzazioni su una tabella condivisa (denominata metodo di risorsa, console)

- Segui le istruzioni in [Concessione delle autorizzazioni per le tabelle utilizzando il metodo di risorsa denominato](#). Nell'elenco dei database sotto i tag LF o le risorse del catalogo, assicuratevi di selezionare il database nell'account esterno, non un collegamento alla risorsa per il database.

Se non vedi la tabella nell'elenco delle tabelle, assicurati di aver accettato l'invito alla condivisione delle AWS RAM risorse per la tabella. Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

Inoltre, per l'ALTERautorizzazione, segui le istruzioni riportate e assicurati di inserire l'ID dell'account proprietario nel campo Ubicazione dell'account registrato. [Concessione delle autorizzazioni per la localizzazione dei dati \(stesso account\)](#)

Per concedere autorizzazioni su risorse condivise (metodo LF-TBAC, console)

- Segui le istruzioni in [Concessione delle autorizzazioni per Data Catalog](#) . Nella sezione LF-Tag o risorse del catalogo, concedi l'espressione esatta del tag LF che l'account esterno ha concesso al tuo account o un sottoinsieme di quell'espressione.

Ad esempio, se un account esterno ha concesso l'espressione LF-Tag `module=customers AND environment=production` al vostro account con l'opzione di concessione, in qualità di amministratore del data lake, potete concedere la stessa espressione `module=customers` o `environment=production` a un responsabile del vostro account. È possibile concedere solo le stesse autorizzazioni o un sottoinsieme delle autorizzazioni di Lake Formation (ad esempio, `SELECTALTER`, e così via) concesse alle risorse tramite l'espressione LF-Tag.

Per concedere le autorizzazioni su una tabella condivisa (metodo denominato resource,) AWS CLI

- Utilizzare un comando simile al seguente: In questo esempio:
 - L'ID del tuo AWS account è 1111-2222-3333.
 - L'account proprietario della tabella e che l'ha concessa al tuo account è 1234-5678-9012.
 - L'SELECTautorizzazione sulla tabella condivisa viene concessa all'utente. `pageviews dataLake_user1` Quell'utente è il principale del tuo account.
 - La `pageviews` tabella si trova nel `analytics` database, di proprietà dell'account 1234-5678-9012.

```
aws lakeformation grant-permissions --principal  
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1  
--permissions "SELECT" --resource '{ "Table": {"CatalogId":"123456789012",  
"DatabaseName":"analytics", "Name":"pageviews"}}'
```

Si noti che l'account proprietario deve essere specificato nella proprietà dell'argomento.
CatalogId resource

Concessione delle autorizzazioni per i collegamenti alle risorse

Segui questi passaggi per concedere AWS Lake Formation le autorizzazioni su uno o più link di risorse a un responsabile del tuo account. AWS

Dopo aver creato un link a una risorsa, solo tu puoi visualizzarlo e accedervi. (Ciò presuppone che il controllo di accesso Use only IAM per le nuove tabelle in questo database non sia abilitato per il database.) Per consentire agli altri responsabili del tuo account di accedere al link alla risorsa, concedi almeno l'DESCRIBE autorizzazione.

Important

La concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sulla tabella o sul database di destinazione (collegato). È necessario concedere le autorizzazioni sulla destinazione separatamente.

Puoi concedere le autorizzazioni utilizzando la console Lake Formation, l'API o AWS Command Line Interface (AWS CLI).

console

Per concedere le autorizzazioni per i collegamenti alle risorse utilizzando la console Lake Formation

1. Esegui una delle seguenti operazioni:

- Per i collegamenti alle risorse del database, segui la procedura riportata in [Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#). Per effettuare le seguenti operazioni:
 1. Seleziona il collegamento alla risorsa dall'elenco dei database in Data Catalog, Databases.
 2. Scegli Concedi per aprire la pagina Concedi le autorizzazioni.
 3. Specificate i principali per concedere le autorizzazioni.
 4. I campi Cataloghi e Database vengono compilati.
 - Per i collegamenti alle risorse delle tabelle, procedi nel [Concessione delle autorizzazioni per le tabelle utilizzando il metodo di risorsa denominato](#) seguente modo:
 1. Seleziona il collegamento alla risorsa dall'elenco delle tabelle in Data Catalog, Tabelle.
 2. Apri la pagina Concedi le autorizzazioni.
 3. Specificare i principali.
 4. I campi Cataloghi, Database, Tabelle vengono compilati.
 5. Specificare i principali.
2. In Autorizzazioni, seleziona le autorizzazioni da concedere. Facoltativamente, seleziona autorizzazioni concedibili.

LF-Tags or catalog resources

Choose a method to grant permissions.

☐ **Resources matched by LF-Tags (recommended)**
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.

☒ **Named Data Catalog resources**
Manage permissions for specific databases or tables, in addition to fine-grained data access.

Catalogs

Choose catalogs

Default catalog

Databases

Select one or more databases.

Choose databases

resourcelink

Tables - optional

Select one or more tables.

Choose tables

Views - optional

Select one or more views.

Choose views

Data filters - optional

Select one or more data filters.

Choose data filters

Load more

Create new

[Manage data filters](#)

Resource link permissions

Resource link permissions

Choose specific access permissions to grant.

☐ Describe ☐ Drop

Grantable permissions

Choose the permission that can be granted to others.

☐ Describe ☐ Drop

Cancel

Grant

3. Scegli Concessione.

AWS CLI

Per concedere le autorizzazioni per i collegamenti alle risorse utilizzando AWS CLI

- Esegui il `grant-permissions` comando, specificando un link alla risorsa come risorsa.

Example

Questo esempio concede `DESCRIBE` all'utente `datalake_user1` sulla tabella il collegamento alla risorsa nel database `incidents-link issues` nell' AWS account `1111-2222-3333`.

```
aws lakeformation grant-permissions --principal
DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
--permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"issues",
"Name":"incidents-link"}}'
```

 Consulta anche:

- [Creazione di collegamenti alle risorse](#)
- [Riferimento alle autorizzazioni di Lake Formation](#)

Accesso ai dati sottostanti di una tabella condivisa

Supponiamo che l' AWS account A condivida una tabella del catalogo dati con l'account B, ad esempio concedendo l'opzione di concessione `SELECT` sulla tabella all'account B. Affinché un responsabile dell'account B sia in grado di leggere i dati sottostanti della tabella condivisa, devono essere soddisfatte le seguenti condizioni:

- L'amministratore del data lake dell'account B deve accettare la condivisione. (Questo non è necessario se gli account A e B fanno parte della stessa organizzazione o se la concessione è stata concessa con il metodo di controllo degli accessi basato su tag Lake Formation.)
- L'amministratore del data lake deve concedere nuovamente al principale l'`SELECT` autorizzazione Lake Formation concessa dall'account A sulla tabella condivisa.
- Il principale deve disporre delle seguenti autorizzazioni IAM sulla tabella, sul database che la contiene e sull'account A Data Catalog.

 Note

Nella seguente politica IAM:

- Sostituisci *<account-id-A>* con l' AWS ID dell'account A.
- Sostituisci *<region>* con una regione valida.
- Sostituisci *<database>* con il nome del database nell'account A che contiene la tabella condivisa.
- Sostituisci *<table>* con il nome della tabella condivisa.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "glue:GetDatabase",
        "glue:GetDatabases"
      ],
      "Resource": [
        "arn:aws:glue:us-east-1:111122223333:table/<database>/<table>",
        "arn:aws:glue:us-east-1:111122223333:database/<database>",
        "arn:aws:glue:us-east-1:111122223333:catalog"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

```
    ]  
  }  
  ]  
}
```

 Consulta anche:

- [Accettazione di un invito alla condivisione di risorse da AWS RAM](#)

Registrazione su più account CloudTrail

Lake Formation fornisce un audit trail centralizzato di tutti gli accessi tra account ai dati nel tuo data lake. Quando un AWS account destinatario accede ai dati in una tabella condivisa, Lake Formation copia l' CloudTrail evento nei log dell' CloudTrail account proprietario. Gli eventi copiati includono interrogazioni sui dati da parte di servizi integrati come Amazon Redshift Spectrum Amazon Athena e accessi ai dati tramite processi. AWS Glue

CloudTrail gli eventi per le operazioni tra account sulle risorse di Data Catalog vengono copiati in modo analogo.

In qualità di proprietario delle risorse, se abiliti la registrazione a livello di oggetto in Amazon S3, puoi eseguire query che uniscono gli eventi S3 agli eventi di Lake CloudTrail Formation per determinare CloudTrail gli account che hanno avuto accesso ai tuoi bucket S3.

Argomenti

- [Inclusione delle identità principali nei log di più account CloudTrail](#)
- [Interrogazione dei CloudTrail log per l'accesso a più account Amazon S3](#)

Inclusione delle identità principali nei log di più account CloudTrail

Per impostazione predefinita, gli CloudTrail eventi tra account aggiunti ai log del destinatario della risorsa condivisa e copiati nei log del proprietario della risorsa contengono solo l'ID AWS principale dell'account esterno, non il nome Amazon Resource Name (ARN) leggibile dall'uomo del principale (ARN principale). Quando condividi risorse all'interno di confini affidabili, ad esempio all'interno della stessa organizzazione o team, puoi scegliere di includere l'ARN principale negli CloudTrail eventi. Gli

account dei proprietari delle risorse possono quindi tenere traccia dei principali account dei destinatari che accedono alle risorse di proprietà.

Important

In qualità di destinatario di risorse condivise, per visualizzare l'ARN principale negli eventi nei propri CloudTrail registri, è necessario scegliere di condividere l'ARN principale con l'account del proprietario.

Se l'accesso ai dati avviene tramite un collegamento alle risorse, nell'account del destinatario della risorsa condivisa vengono registrati due eventi: uno per l'accesso al collegamento alle risorse e uno per l'accesso alla risorsa di destinazione. L'evento per l'accesso al link di risorsa include l'ARN principale. L'evento per l'accesso alle risorse di destinazione non include l'ARN principale senza l'opt-in. L'evento di accesso al collegamento alle risorse non viene copiato nell'account del proprietario.

Di seguito è riportato un estratto di un CloudTrail evento predefinito tra più account (senza opt-in). L'account che esegue l'accesso ai dati è 1111-2222-3333. Questo è il registro visualizzato sia nell'account chiamante che nell'account del proprietario della risorsa. Lake Formation compila i log in entrambi gli account nel caso di più account.

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  },
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  ...
  ...
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"
  },
  ...
}
```


In qualità di consumatore di risorse condivise, quando scegli di includere l'ARN principale, l'estratto diventa il seguente. Il `lakeFormationPrincipal` campo rappresenta il ruolo finale o l'utente che esegue la query tramite Amazon Athena, Amazon Redshift Spectrum o job. AWS Glue

```
{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  },
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  ...
  ...
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"
  },
  ...
}
```

Per attivare l'inclusione del principale ARNs nei log di più account CloudTrail

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come Administrator utente o come utente con la policy Administrator Access IAM.

2. Nel pannello di navigazione scegli Impostazioni.
3. Nella pagina delle impostazioni del catalogo dati, nella AWS CloudTrail sezione Autorizzazioni predefinite per, per Proprietari delle risorse, inserisci uno o più account IDs del proprietario della AWS risorsa.

Premi Invio dopo ogni ID account.

4. Scegli Save (Salva).

Ora CloudTrail gli eventi tra account archiviati nei log sia per il destinatario della risorsa condivisa che per il proprietario della risorsa contengono l'ARN principale.

Interrogazione dei CloudTrail log per l'accesso a più account Amazon S3

In qualità di proprietario di risorse condivise, puoi interrogare CloudTrail i log di S3 per determinare gli account che hanno avuto accesso ai tuoi bucket Amazon S3 (a condizione che tu abbia abilitato la registrazione a livello di oggetto in Amazon S3). Questo vale solo per le sedi S3 che hai registrato con Lake Formation. Se i consumatori di risorse condivise scelgono di includere il principal ARNs nei CloudTrail log di Lake Formation, puoi determinare i ruoli o gli utenti che hanno avuto accesso ai bucket.

Quando esegui query con Amazon Athena, puoi unire gli eventi Lake Formation e CloudTrail gli eventi S3 nella CloudTrail proprietà del nome della sessione. Le query possono anche filtrare gli eventi di Lake Formation su `eventName="GetDataAccess"` gli eventi S3 su `eventName="GetObject"` o `eventName="Put Object"`

Di seguito è riportato un estratto di un CloudTrail evento cross-account di Lake Formation in cui è stato effettuato l'accesso ai dati in una posizione S3 registrata.

```
{
  "eventSource": "lakeformation.amazonaws.com",
  "eventName": "GetDataAccess",
  .....
  .....
  "additionalEventData": {
    "requesterService": "GLUE_JOB",
    "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",
    "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-B8JSAjo5QA"
  }
}
```

Il valore `lakeFormationRoleSessionName` chiave, `AWSLF-00-GL-111122223333-B8JSAjo5QA`, può essere unito al nome della sessione nella `principalId` chiave dell'evento S3. CloudTrail Di seguito è riportato un estratto dell'evento S3. CloudTrail Mostra la posizione del nome della sessione.

```
{
  "eventSource": "s3.amazonaws.com",
  "eventName": "GetObject"
  .....
  .....
```

```

"principalId": "AROAQS0X5XXUR7D6RMYLR:AWSLF-00-GL-111122223333-B8JSAjo5QA",
"arn": "arn:aws:sets::111122223333:assumed-role/Deformationally/AWSLF-00-GL-111122223333-B8JSAjo5QA",
"session Context": {
  "session Issuer": {
    "type": "Role",
    "principalId": "AROAQS0X5XXUR7D6RMYLR",
    "arn": "arn:aws:iam::111122223333:role/aws-service-role/lakeformation.amazonaws.com/Deformationally",
    "accountId": "111122223333",
    "user Name": "Deformationally"
  },
  .....
  .....
}

```

Il nome della sessione è formattato come segue:

```
AWSLF-<version-number>-<query-engine-code>-<account-id>-<suffix>
```

version-number

La versione di questo formato, attualmente 00. Se il formato del nome della sessione cambia, sarà la versione successiva 01.

query-engine-code

Indica l'entità che ha avuto accesso ai dati. I valori correnti sono:

GL	AWS Glue Lavoro ETL
AT	Athena
RE	Amazon Redshift Spectrum

account-id

L'ID AWS dell'account che ha richiesto le credenziali a Lake Formation.

suffix

Una stringa generata casualmente.

Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation

È possibile concedere l'accesso a più account alle risorse del Data Catalog e ai dati sottostanti utilizzando o. AWS Glue AWS Lake Formation

InAWS Glue, concedi autorizzazioni per più account creando o aggiornando una politica delle risorse di Data Catalog. In Lake Formation, concedi autorizzazioni per più account utilizzando il modello di autorizzazioni Lake Formation GRANT/REVOKE e il funzionamento dell'Grant PermissionsAPI.

Tip

Ti consigliamo di affidarti esclusivamente alle autorizzazioni di Lake Formation per proteggere il tuo data lake.

Puoi visualizzare le sovvenzioni di Lake Formation su più account utilizzando la console Lake Formation o la console AWS Resource Access Manager (AWS RAM). Tuttavia, queste pagine della console non mostrano le autorizzazioni per più account concesse dalla politica delle risorse di AWS Glue Data Catalog. Allo stesso modo, puoi visualizzare le concessioni tra account nella politica delle risorse di Data Catalog utilizzando la pagina Impostazioni della AWS Glue console, ma quella pagina non mostra le autorizzazioni per più account concesse utilizzando Lake Formation.

Per assicurarti di non perdere nessuna sovvenzione durante la visualizzazione e la gestione delle autorizzazioni tra più account, Lake Formation AWS Glue richiede che tu esegua le seguenti azioni per indicare che sei a conoscenza e che stai autorizzando le sovvenzioni tra account sia da Lake Formation che. AWS Glue

Quando si concedono autorizzazioni su più account utilizzando la politica delle risorse di Data Catalog AWS Glue

Se il tuo account (account concedente o account produttore) non ha concesso concessioni tra account diversi che vengono utilizzate AWS RAM per condividere le risorse, puoi salvare una politica sulle risorse di Data Catalog come di consueto in. AWS Glue Tuttavia, se sono già state concesse sovvenzioni che prevedono la condivisione AWS RAM delle risorse, è necessario effettuare una delle seguenti operazioni per garantire che la politica di salvataggio delle risorse abbia esito positivo:

- Quando salvi la politica delle risorse nella pagina Impostazioni della AWS Glue console, la console emette un avviso che indica che le autorizzazioni nella politica si aggiungeranno a tutte le

autorizzazioni concesse utilizzando la console Lake Formation. È necessario scegliere Procedi per salvare la politica.

- Quando si salva la politica delle risorse utilizzando l'operazione `glue:PutResourcePolicy` API, è necessario impostare il `EnableHybrid` campo su 'TRUE' (type = string).

Per aggiornare una politica delle risorse esistente, utilizza l'operazione `glue:GetResourcePolicy` API per recuperare prima la politica corrente, quindi modificala se necessario prima di chiamare `glue:PutResourcePolicy`.

Note

Quando crei politiche AWS Glue sulle risorse per l'accesso tra account diversi, concedi solo le autorizzazioni minime richieste per il tuo caso d'uso specifico.

Per ulteriori informazioni, consulta [PutResourcePolicy Action \(Python: put_resource_policy\)](#) nella Developer Guide.AWS Glue

Quando si concedono autorizzazioni per più account utilizzando il metodo delle risorse denominate Lake Formation

Se nel tuo account (account produttore) non è presente alcuna politica sulle risorse di Data Catalog, Lake Formation garantisce che tu proceda come al solito. Tuttavia, se esiste una politica in materia di risorse di Data Catalog, devi aggiungere la seguente dichiarazione per consentire che le concessioni tra account abbiano esito positivo se vengono effettuate con il metodo delle risorse denominato.
<region>Sostituiscila con un nome di regione valido e **<account-id>** con l'ID AWS del tuo account (ID dell'account produttore).

```
{
  "Effect": "Allow",
  "Action": [
    "glue:ShareResource"
  ],
  "Principal": {"Service": [
    "ram.amazonaws.com"
  ]},
  "Resource": [
    "arn:aws:glue:<region>:<account-id>:table/*/*",
    "arn:aws:glue:<region>:<account-id>:database/*",
```

```
"arn:aws:glue:<region>:<account-id>:catalog"  
]  
}
```

Senza questa dichiarazione aggiuntiva, la sovvenzione di Lake Formation ha esito positivo, ma viene bloccata e l'account del destinatario non può accedere alla risorsa concessa. AWS RAM

Important

Quando si utilizza il metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC) per concedere concessioni tra account, è necessario disporre di una politica delle risorse del Data Catalog con almeno le autorizzazioni specificate in [Prerequisiti](#)

Consulta anche:

- [Controllo dell'accesso ai metadati](#) (per una discussione sul metodo Named Resource rispetto al metodo di controllo degli accessi basato su tag Lake Formation (LF-TBAC)).
- [Visualizzazione di tabelle e database condivisi del Data Catalog](#)
- [Utilizzo delle impostazioni del catalogo dati sulla console](#) nella Guida per gli sviluppatori AWS Glue
- [Concessione dell'accesso a più account](#) nella Guida per gli AWS Glue sviluppatori (ad esempio, le politiche relative alle risorse di Data Catalog)

Visualizzazione di tutte le sovvenzioni tra account utilizzando l'operazione API GetResourceShares

Se la tua azienda concede autorizzazioni per più account utilizzando sia una politica AWS Glue Data Catalog delle risorse che le sovvenzioni Lake Formation, l'unico modo per visualizzare tutte le sovvenzioni su più account in un unico posto è utilizzare l'operazione API. `glue:GetResourceShares`

Quando concedi le autorizzazioni di Lake Formation su più account utilizzando il metodo di risorsa denominato, AWS Resource Access Manager (AWS RAM) crea una politica delle risorse AWS Identity and Access Management (IAM) e la memorizza nel tuo AWS account. La politica concede le autorizzazioni necessarie per accedere alla risorsa. AWS RAM crea una politica delle risorse

separata per ogni concessione tra account. È possibile visualizzare tutte queste politiche utilizzando l'operazione `glue:GetResourceShares` API.

Note

Questa operazione restituisce anche la politica delle risorse di Data Catalog. Tuttavia, se hai abilitato la crittografia dei metadati nelle impostazioni del Catalogo dati e non disponi dell'autorizzazione sulla AWS KMS chiave, l'operazione non restituirà la politica delle risorse del Catalogo dati.

Per visualizzare tutte le sovvenzioni concesse tra account

- Immettere il seguente comando AWS CLI .

```
aws glue get-resource-policies
```

Di seguito è riportato un esempio di politica delle risorse che AWS RAM crea e memorizza quando si concedono autorizzazioni su una tabella `t` nel database `db1` all' AWS account `1111-2222-3333`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetTableVersion",
        "glue:GetTableVersions",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "glue:SearchTables"
      ],
      "Principal": { "AWS": [
        "111122223333"
      ] },
    }
  ]
}
```

```
"Resource": [  
  "arn:aws:glue:us-east-1:111122223333:table/db1/t"  
]  
}  
]  
}
```

 Consulta anche:

- [GetResourceShares Azione \(Python: `get\_resource\_policies`\)](#) nella Guida per gli sviluppatori AWS Glue

Accesso e visualizzazione di tabelle e database di Data Catalog condivisi

Per l'amministratore del data lake e per i responsabili a cui sono state concesse le autorizzazioni, le risorse condivise con il tuo AWS account vengono visualizzate nel Data Catalog come se fossero risorse del tuo account. La console mostra l'account proprietario della risorsa.

Puoi visualizzare le risorse condivise con il tuo account utilizzando la console Lake Formation. Puoi anche utilizzare la console AWS Resource Access Manager (AWS RAM) per visualizzare sia le risorse condivise con il tuo account sia le risorse che hai condiviso con altri AWS account utilizzando il metodo della risorsa denominata.

Important

Quando qualcuno utilizza il metodo della risorsa denominata per concedere autorizzazioni multiaccount su una risorsa Data Catalog al tuo account o alla tua AWS organizzazione, Lake Formation utilizza il servizio AWS Resource Access Manager (AWS RAM) per condividere la risorsa. Se il tuo account appartiene alla stessa AWS organizzazione dell'account che concede l'autorizzazione, la risorsa condivisa è immediatamente disponibile.

Tuttavia, se il tuo account non fa parte della stessa organizzazione, AWS RAM invia un invito all'account per accettare o rifiutare la condivisione delle risorse. Quindi, per rendere disponibile la risorsa condivisa, l'amministratore del data lake del tuo account deve utilizzare la AWS RAM console o la CLI per accettare l'invito.

La console Lake Formation visualizza un avviso se c'è un invito alla condivisione di AWS RAM risorse in attesa di essere accettato. Solo gli utenti autorizzati a visualizzare AWS RAM gli inviti ricevono l'avviso.

 Consulta anche:

- [Condivisione di tabelle e database del Data Catalog tra più account AWS](#)
- [Condivisione dei dati tra account in Lake Formation](#)
- [Accesso ai dati sottostanti di una tabella condivisa](#)
- [Controllo dell'accesso ai metadati](#)(per informazioni sul metodo delle risorse denominato rispetto al metodo LF-TBAC per la condivisione delle risorse.)

Argomenti

- [Accettazione di un invito alla condivisione di risorse da AWS RAM](#)
- [Visualizzazione di tabelle e database condivisi del Data Catalog](#)

Accettazione di un invito alla condivisione di risorse da AWS RAM

Se una risorsa del Catalogo dati è condivisa con il tuo AWS account e il tuo account non fa parte della stessa AWS organizzazione dell'account di condivisione, non potrai accedere alla risorsa condivisa finché non accetti un invito alla condivisione delle risorse da AWS Resource Access Manager (AWS RAM). In qualità di amministratore del data lake, devi prima AWS RAM richiedere gli inviti in sospeso e poi accettare l'invito.

Puoi utilizzare la AWS RAM console, l'API o AWS Command Line Interface (AWS CLI) per visualizzare e accettare gli inviti.

Per visualizzare e accettare un invito alla condivisione di risorse da AWS RAM (console)

1. Assicurati di disporre delle autorizzazioni AWS Identity and Access Management (IAM) necessarie per visualizzare e accettare gli inviti alla condivisione delle risorse.

Per informazioni sulle politiche IAM consigliate per gli amministratori di data lake, consulta [the section called “Autorizzazioni di amministratore di Data Lake”](#)

2. Segui le istruzioni riportate nella sezione [Accettazione e rifiuto degli inviti nella Guida](#) per l'AWS RAM utente.

Per visualizzare e accettare un invito alla condivisione di risorse da AWS RAM (AWS CLI)

1. Assicurati di disporre delle autorizzazioni AWS Identity and Access Management (IAM) necessarie per visualizzare e accettare gli inviti alla condivisione delle risorse.

Per informazioni sulle politiche IAM consigliate per gli amministratori di data lake, consulta [the section called "Autorizzazioni di amministratore di Data Lake"](#)

2. Immetti il comando seguente per visualizzare gli inviti alla condivisione delle risorse in sospeso.

```
aws ram get-resource-share-invitations
```

L'output visualizzato dovrebbe essere simile al seguente.

```
{
  "resourceShareInvitations": [
    {
      "resourceShareInvitationArn": "arn:aws:ram:us-east-1:111122223333:resource-share-invitation/a93aa60a-1bd9-46e8-96db-a4e72eec1d9f",
      "resourceShareName": "111122223333-123456789012-uswuU",
      "resourceShareArn": "arn:aws:ram:us-east-1:111122223333:resource-share/2a4ab5fb-d859-4751-84f7-8760b35fc1fe",
      "senderAccountId": "111122223333",
      "receiverAccountId": "123456789012",
      "invitationTimestamp": 1589576601.79,
      "status": "PENDING"
    }
  ]
}
```

Annota lo stato di. PENDING

3. Copia il valore della resourceShareInvitationArn chiave negli appunti.
4. Incollate il valore nel comando seguente *<invitation-arn>*, sostituitelo e immettete il comando.

```
aws ram accept-resource-share-invitation --resource-share-invitation-arn <invitation-arn>
```

L'output visualizzato dovrebbe essere simile al seguente.

```
{
  "resourceShareInvitations": [
    {
      "resourceShareInvitationArn": "arn:aws:ram:us-east-1:111122223333:resource-share-invitation/a93aa60a-1bd9-46e8-96db-a4e72eec1d9f",
      "resourceShareName": "111122223333-123456789012-uswuU",
      "resourceShareArn": "arn:aws:ram:us-east-1:111122223333:resource-share/2a4ab5fb-d859-4751-84f7-8760b35fc1fe",
      "senderAccountId": "111122223333",
      "receiverAccountId": "123456789012",
      "invitationTimestamp": 1589576601.79,
      "status": "ACCEPTED"
    }
  ]
}
```

Annotate lo stato diACCEPTED.

Visualizzazione di tabelle e database condivisi del Data Catalog

Puoi visualizzare le risorse condivise con il tuo account utilizzando la console di Lake Formation o la AWS CLI. Puoi anche utilizzare la console AWS Resource Access Manager (AWS RAM) o la CLI per visualizzare sia le risorse condivise con il tuo account sia le risorse che hai condiviso con altri AWS account.

Per visualizzare le risorse condivise utilizzando la console Lake Formation

1. Apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.

Accedi come amministratore del data lake o utente a cui sono state concesse le autorizzazioni su una tabella condivisa.

2. Per visualizzare le risorse condivise con il tuo AWS account, esegui una delle seguenti operazioni:

- Per visualizzare le tabelle condivise con il tuo account, nel riquadro di navigazione, scegli Tabelle.
- Per visualizzare i database condivisi con il tuo account, nel riquadro di navigazione, scegli Database.

La console mostra un elenco di database o tabelle presenti nel tuo account e condivisi con il tuo account. Per le risorse condivise con il tuo account, la console mostra l'ID dell' AWS account del proprietario nella colonna ID account proprietario (la terza colonna nella schermata seguente).

	Name	Database	Owner account ...	Shared resource	Shared resource owner
☐	adviews	analytics	111122223333	-	-
☐	pageviews	analytics	111122223333	-	-
☐	blackholes	hubble	123456789012	-	-
☐	celestial-events	hubble	123456789012	-	-
☐	suns	hubble	123456789012	-	-

3. Per visualizzare le risorse condivise con altri AWS account o organizzazioni, nel riquadro di navigazione, scegli Autorizzazioni dati.

Le risorse che hai condiviso sono elencate nella pagina Autorizzazioni dati con il numero di account esterno mostrato nella colonna Principale, come mostrato nell'immagine seguente.

	Principal	Principal type	Resource type	Resource	Owner account ID	Permissions
☐	datalake_admin	IAM user	Table	clickthroughs	123456789012	Super, Alter, Delete, Drop, Insert
☐	datalake_admin	IAM user	Column	analytics.click throughs.*	123456789012	Select
☐	111122223333	AWS account	Table	clickthroughs	123456789012	Insert
☐	111122223333	AWS account	Column	analytics.click throughs.*	123456789012	Select

Per visualizzare le risorse condivise utilizzando la AWS RAM console

1. Assicurati di disporre delle autorizzazioni AWS Identity and Access Management (IAM) necessarie per visualizzare le risorse condivise utilizzando AWS RAM.

È necessario disporre almeno dell'autorizzazione `ram:ListResources`. Questa autorizzazione è inclusa nella policy gestita da AWS, `AWSLakeFormationCrossAccountManager`.

2. [Accedi a Console di gestione AWS e apri la AWS RAM console da casa. https://console.aws.amazon.com/ram/](https://console.aws.amazon.com/ram/)
3. Esegui una di queste operazioni:
 - Per visualizzare le risorse condivise, nel riquadro di navigazione, in Condivise da me, scegli Risorse condivise.
 - Per visualizzare le risorse condivise con te, nel riquadro di navigazione, in Condivise con me, scegli Risorse condivise.

Creazione di collegamenti alle risorse

I link alle risorse sono oggetti del Data Catalog che sono collegamenti a database e tabelle di metadati, in genere a database e tabelle condivisi di altri account. AWS Aiutano a consentire l'accesso tra account diversi ai dati nel data lake in tutte le regioni. AWS

Note

Lake Formation supporta l'interrogazione delle tabelle del Data Catalog tra AWS le regioni. È possibile accedere ai database e alle tabelle del Data Catalog da qualsiasi AWS regione creando collegamenti di risorse in quelle aree che puntano a database e tabelle condivisi in diverse regioni.

Argomenti

- [Come funzionano i link alle risorse in Lake Formation](#)
- [Creazione di un collegamento di risorsa a una tabella condivisa del Catalogo dati](#)
- [Creazione di un collegamento di risorsa a un database Data Catalog condiviso](#)
- [Gestione dei link alle risorse in AWS Glue APIs](#)

Come funzionano i link alle risorse in Lake Formation

Un collegamento a una risorsa è un oggetto del catalogo dati che è un collegamento a un database o a una tabella locale o condivisa. Dopo aver creato un link di risorsa a un database o a una tabella, è possibile utilizzare il nome del link di risorsa ovunque si utilizzi il nome del database o della tabella. Oltre alle tabelle di tua proprietà o alle tabelle condivise con te, i link alle risorse delle tabelle vengono restituiti **glue:GetTables()** e vengono visualizzati come voci nella pagina Tabelle della console Lake Formation. I collegamenti alle risorse ai database agiscono in modo simile.

La creazione di un collegamento di risorsa a un database o a una tabella consente di effettuare le seguenti operazioni:

- Assegna un nome diverso a un database o a una tabella nel tuo Data Catalog. Ciò è particolarmente utile se AWS account diversi condividono database o tabelle con lo stesso nome o se più database dell'account hanno tabelle con lo stesso nome.
- Accedi ai database e alle tabelle del Data Catalog da qualsiasi AWS regione creando collegamenti alle risorse in quelle aree che puntano al database e alle tabelle in un'altra regione. Puoi eseguire query in qualsiasi regione con questi link alle risorse utilizzando Athena, Amazon EMR ed AWS Glue eseguire job ETL Spark, senza copiare i dati di origine né i metadati in Glue Data Catalog.
- Utilizza AWS servizi integrati come Amazon Athena Amazon Redshift Spectrum per eseguire query che accedono a database o tabelle condivisi. Alcuni servizi integrati non possono accedere direttamente a database o tabelle tra account. Tuttavia, possono accedere ai link delle risorse presenti nel tuo account ai database e alle tabelle di altri account.

Note

Non è necessario creare un collegamento a una risorsa per fare riferimento a un database o a una tabella condivisa negli script di AWS Glue estrazione, trasformazione e caricamento (ETL). Tuttavia, per evitare ambiguità quando più AWS account condividono un database o una tabella con lo stesso nome, è possibile creare e utilizzare un collegamento alla risorsa o specificare l'ID del catalogo quando si richiamano le operazioni ETL.

L'esempio seguente mostra la pagina Tabelle della console di Lake Formation, che elenca due collegamenti a risorse. I nomi dei link alle risorse sono sempre visualizzati in corsivo. Ogni collegamento alla risorsa viene visualizzato insieme al nome e al proprietario della risorsa condivisa collegata. In questo esempio, un amministratore di Data Lake nell' AWS account 1111-2222-3333

ha condiviso le `incidents` tabelle `inventory` and con l'account 1234-5678-9012. Un utente di quell'account ha quindi creato collegamenti alle risorse a tali tabelle condivise.

Tables (30)

Actions ▼

Create table using a crawler

Create table

Find table by properties

< 1 >

	Name ▼	Database ▼	Owner account ... ▼	Shared resource ▼	Shared resource owner
☐	inventory-link	retail	123456789012	inventory	111122223333
☐	incidents-link	issues-local	123456789012	incidents	111122223333
☐	site-logs	logs	123456789012	-	-
☐	alexa-logs	logs	123456789012	-	-

Di seguito sono riportate le note e le restrizioni relative ai collegamenti alle risorse:

- I link alle risorse sono necessari per consentire a servizi integrati come Athena e Redshift Spectrum di interrogare i dati sottostanti delle tabelle condivise. Le query in questi servizi integrati sono costruite sulla base dei nomi dei link alle risorse.
- Supponendo che l'impostazione Usa solo il controllo di accesso IAM per le nuove tabelle in questo database sia disattivata per il database che lo contiene, solo il principale che ha creato un collegamento alla risorsa può visualizzarlo e accedervi. Per consentire agli altri responsabili del tuo account di accedere a un link a una risorsa, concedi l'`DESCRIBE` autorizzazione. Per consentire ad altri di eliminare un link a una risorsa, concedi l'`DROP` autorizzazione. Gli amministratori di Data Lake possono accedere a tutti i link alle risorse presenti nell'account. Per eliminare un collegamento a una risorsa creato da un altro principale, l'amministratore del data lake deve prima concedersi l'`DROP` autorizzazione sul collegamento alla risorsa. Per ulteriori informazioni, consulta [Riferimento alle autorizzazioni di Lake Formation](#).

Important

La concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sul database o sulla tabella di destinazione (collegati). È necessario concedere le autorizzazioni sulla destinazione separatamente.

- Per creare un collegamento a una risorsa, è necessaria l'`CREATE_DATABASE` autorizzazione `CREATE_TABLE` o la Lake Formation, nonché l'autorizzazione `glue:CreateTable` or `glue:CreateDatabase` AWS Identity and Access Management (IAM).

- Puoi creare collegamenti a risorse locali (di proprietà) del Data Catalog, nonché a risorse condivise con il tuo AWS account.
- Quando crei un link a una risorsa, non viene eseguito alcun controllo per verificare se la risorsa condivisa di destinazione esiste o se disponi di autorizzazioni per più account sulla risorsa. Ciò consente di creare il collegamento alla risorsa e la risorsa condivisa in qualsiasi ordine.
- Se si elimina un collegamento a una risorsa, la risorsa condivisa collegata non viene eliminata. Se si elimina una risorsa condivisa, i link delle risorse a tale risorsa non vengono eliminati.
- È possibile creare catene di collegamenti di risorse. Tuttavia, non è utile farlo, perché APIs segue solo il primo collegamento alla risorsa.

 Consulta anche:

- [Concessione delle autorizzazioni per le risorse di Data Catalog](#)

Creazione di un collegamento di risorsa a una tabella condivisa del Catalogo dati

Puoi creare un link di risorsa a una tabella condivisa in qualsiasi AWS regione utilizzando la AWS Lake Formation console, l'API o AWS Command Line Interface ().AWS CLI

Per creare un collegamento di risorsa a una tabella condivisa (console)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come preside che dispone dell'CREATE_TABLEautorizzazione Lake Formation sul database per contenere il link alla risorsa.
2. Nel riquadro di navigazione, scegli Tabelle in Data Catalog, quindi scegli Crea, Link alle risorse.
3. Nella pagina Crea collegamento alla risorsa, fornisci le seguenti informazioni:

Nome del link alla risorsa

Immettete un nome che rispetti le stesse regole del nome della tabella. Il nome può essere lo stesso della tabella condivisa di destinazione.

Database

Il database nel Data Catalog locale che deve contenere il collegamento alla risorsa.

Regione proprietaria della tabella condivisa

Se stai creando il link alla risorsa in un'altra regione, seleziona la regione della tabella condivisa di destinazione.

Tabella condivisa

Seleziona una tabella condivisa dall'elenco o inserisci un nome di tabella locale (di proprietà) o condivisa.

L'elenco contiene tutte le tabelle condivise con il tuo account. Annota il database e l'ID dell'account del proprietario elencati in ogni tabella. Se non vedi una tabella che sai è stata condivisa con il tuo account, controlla quanto segue:

- Se non sei un amministratore del data lake, verifica che l'amministratore del data lake ti abbia concesso le autorizzazioni Lake Formation sulla tabella.
- Se sei un amministratore del data lake e il tuo account non fa parte della stessa AWS organizzazione dell'account concedente, assicurati di aver accettato l'invito alla condivisione delle risorse AWS Resource Access Manager (AWS RAM) per la tabella. Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

Database della tabella condivisa

Se hai selezionato una tabella condivisa dall'elenco, questo campo viene popolato con il database della tabella condivisa nell'account esterno. Altrimenti, inserisci un database locale (per un collegamento di risorse a una tabella locale) o il database della tabella condivisa nell'account esterno.

Proprietario della tabella condivisa

Se hai selezionato una tabella condivisa dall'elenco, questo campo viene compilato con l'ID dell'account proprietario della tabella condivisa. Altrimenti, inserisci l'ID AWS del tuo account (per il collegamento di una risorsa a una tabella locale) o l'ID dell' AWS account che ha condiviso la tabella.

4. Scegli Crea per creare il link alla risorsa.

È quindi possibile visualizzare il nome del collegamento alla risorsa nella colonna Nome della pagina Tabelle.

5. (Facoltativo) Concedi l'DESCRIBE autorizzazione Lake Formation sul link della risorsa ai responsabili che devono essere in grado di visualizzare il collegamento e accedere alla tabella di destinazione.

Tuttavia, la concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sul database o sulla tabella di destinazione (collegati). È necessario concedere le autorizzazioni sul database di destinazione separatamente affinché il collegamento tabella/risorsa sia visibile in Athena.

Per creare un collegamento di risorsa a una tabella condivisa nella stessa regione ()AWS CLI

1. Utilizzare un comando simile al seguente:

```
aws glue create-table --database-name myissues --table-input
'{"Name": "my_customers", "TargetTable":
{"CatalogId": "111122223333", "DatabaseName": "issues", "Name": "customers"}}'
```

Questo comando crea un link di risorsa denominato `my_customers` alla tabella condivisa `customers`, che si trova `issues` nel database dell' AWS account 1111-2222-3333. Il collegamento alla risorsa è memorizzato nel database locale. `myissues`

2. (Facoltativo) Concedi l'DESCRIBE autorizzazione Lake Formation sul link della risorsa ai responsabili che devono essere in grado di visualizzare il collegamento e accedere alla tabella di destinazione.

Tuttavia, la concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sulla tabella di destinazione (collegata). È necessario concedere le autorizzazioni sul database di destinazione separatamente affinché il collegamento tabella/risorsa sia visibile in Athena.

Per creare un collegamento di risorsa a una tabella condivisa in una regione diversa ()AWS CLI

1. Utilizzare un comando simile al seguente:

```
aws glue create-table --region eu-west-1 --cli-input-json '{
  "CatalogId": "111122223333",
  "DatabaseName": "ireland_db",
  "TableInput": {
    "Name": "rl_useast1salestb_ireland",
```

```
"TargetTable": {  
  "CatalogId": "444455556666",  
  "DatabaseName": "useast1_salesdb",  
  "Region": "us-east-1",  
  "Name": "useast1_salestb"  
}  
}  
'
```

Questo comando crea un collegamento di risorse denominato `rl_useast1salestb_ireland` nella regione Europa (Irlanda) alla `useast1_salestb` tabella condivisa, che si trova nel `useast1_salesdb` database dell'account 444455556666 AWS nella regione Stati Uniti orientali (Virginia settentrionale). Il collegamento alla risorsa è memorizzato nel database locale. `ireland_db`

2. Concedi l'`DESCRIBE` autorizzazione a Lake Formation ai responsabili che devono essere in grado di visualizzare il link e accedere alla destinazione del link tramite il link.

Tuttavia, la concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sulla tabella di destinazione (collegata). È necessario concedere le autorizzazioni sulla tabella di destinazione separatamente affinché il collegamento tabella/risorsa sia visibile in Athena.

 Consulta anche:

- [Come funzionano i link alle risorse in Lake Formation](#)
- [DESCRIBE](#)

Creazione di un collegamento di risorsa a un database Data Catalog condiviso

È possibile creare un collegamento di risorsa a un database condiviso utilizzando la AWS Lake Formation console, l'API o AWS Command Line Interface (AWS CLI).

Per creare un collegamento di risorsa a un database condiviso (console)

1. Apri la AWS Lake Formation console all'indirizzo <https://console.aws.amazon.com/lakeformation/>. Accedi come amministratore del data lake o come creatore di database.

Un creatore di database è un responsabile a cui è stata concessa l'`CREATE_DATABASE` autorizzazione Lake Formation.

2. Nel riquadro di navigazione, scegli Database, quindi scegli Crea, Link alle risorse.
3. Nella pagina Crea collegamento a una risorsa, fornisci le seguenti informazioni:

Nome del link alla risorsa

Immettete un nome che rispetti le stesse regole del nome del database. Il nome può essere lo stesso del database condiviso di destinazione.

Catalogo di destinazione

Seleziona il catalogo di destinazione per il collegamento alle risorse del database.

Regione del proprietario del database condiviso

Se stai creando il link alla risorsa in un'altra regione, seleziona la regione del database condiviso di destinazione.

Database condiviso

Scegli un database dall'elenco o inserisci un nome di database locale (di proprietà) o condiviso.

L'elenco contiene tutti i database condivisi con il tuo account. Annota l'ID dell'account del proprietario elencato in ogni database. Se non vedi un database che sai è stato condiviso con il tuo account, controlla quanto segue:

- Se non sei un amministratore del data lake, verifica che l'amministratore del data lake ti abbia concesso le autorizzazioni Lake Formation sul database.
- Se sei un amministratore del data lake e il tuo account non fa parte della stessa AWS organizzazione dell'account concedente, assicurati di aver accettato l'invito AWS Resource Access Manager (AWS RAM) alla condivisione delle risorse per il database. Per ulteriori informazioni, consulta [Accettazione di un invito alla condivisione di risorse da AWS RAM](#).

Proprietario del database condiviso

Se hai selezionato un database condiviso dall'elenco, questo campo viene compilato con l'ID dell'account proprietario del database condiviso. Altrimenti, inserisci l'ID AWS del tuo account (per un collegamento di risorsa a un database locale) o l'ID dell' AWS account che ha condiviso il database.

ID del catalogo del database condiviso

Immettere l'ID del catalogo per il database condiviso. Quando si crea un collegamento di risorsa a un database condiviso da un altro AWS account, è necessario specificare questo ID di catalogo per identificare il catalogo dati dell'account che contiene il database di origine.

Quando selezioni un database condiviso dal menu a discesa, il sistema inserisce automaticamente l'ID del catalogo dell'account che possiede e ha condiviso quel database con te.

Create resource link [Info](#)

Database resource link details

Create a database resource link in the Data Catalog.

Resource link name

If you plan to access the database from Amazon Athena, then the name should be under 256 characters and contain only lowercase letters (a-z), numbers (0-9), and underscore (_). For more information, see [Athena names](#).

Destination catalog

Select the destination catalog for the database resource link.

Shared database's region

Select the region of the shared database.

Shared database

Enter or choose a shared database.

Shared database's catalog ID

Enter the catalog ID of the shared database.

[Cancel](#)[Create](#)

4. Scegli Crea per creare il link alla risorsa.

È quindi possibile visualizzare il nome del collegamento alla risorsa nella colonna Nome della pagina Database.

5. (Facoltativo) Concedi a Lake Formation l'DESCRIBE autorizzazione per il collegamento alla risorsa ai responsabili della regione Europa (Irlanda) che devono essere in grado di visualizzare il collegamento e accedere al database di destinazione.

Tuttavia, la concessione di autorizzazioni per un collegamento a una risorsa non concede autorizzazioni per il database o la tabella di destinazione (collegati). È necessario concedere le autorizzazioni sul database di destinazione separatamente affinché il collegamento tabella/risorsa sia visibile in Athena.

Per creare un collegamento di risorsa a un database condiviso nella stessa regione ()AWS CLI

1. Utilizzare un comando simile al seguente:

```
aws glue create-database --database-input '{"Name":"myissues","TargetDatabase":
{"CatalogId":"111122223333","DatabaseName":"issues"}}'
```

Questo comando crea un collegamento a una risorsa denominato `myissues` al database condiviso `issues`, che si trova nell' AWS account 1111-2222-3333.

2. (Facoltativo) Concedi l'`DESCRIBE` autorizzazione a Lake Formation ai responsabili del link alla risorsa che devono essere in grado di visualizzare il collegamento e accedere al database o alla tabella di destinazione.

Tuttavia, la concessione delle autorizzazioni su un collegamento a una risorsa non concede le autorizzazioni sul database o sulla tabella di destinazione (collegati). È necessario concedere le autorizzazioni sul database di destinazione separatamente affinché il collegamento tabella/risorsa sia visibile in Athena.

Per creare un collegamento di risorsa a un database condiviso in un'altra regione ()AWS CLI

1. Utilizzare un comando simile al seguente:

```
aws glue create-database --region eu-west-1 --cli-input-json '{
  "CatalogId": "111122223333",
  "DatabaseInput": {
    "Name": "rl_useast1shared_irelanddb",
    "TargetDatabase": {
      "CatalogId": "444455556666",
      "DatabaseName": "useast1shared_db",
      "Region": "us-east-1"
    }
  }
}'
```

Questo comando crea un collegamento di risorse denominato `rl_useast1shared_irelanddb` nell' AWS account 111122223333 nella regione Europa (Irlanda) al database condiviso `useast1shared_db`, che si trova nell' AWS account 444455556666 nella regione Stati Uniti orientali (Virginia settentrionale).

2. Concedi il DESCRIBE permesso di Lake Formation ai dirigenti della regione Europa (Irlanda) che devono essere in grado di visualizzare il link e accedere alla destinazione del link tramite il link.

 Consulta anche:

- [Come funzionano i link alle risorse in Lake Formation](#)
- [DESCRIBE](#)

Gestione dei link alle risorse in AWS Glue APIs

Le tabelle seguenti spiegano come il AWS Glue Data Catalog APIs gestisce i collegamenti alle risorse di database e tabelle. Per tutte le operazioni Get * API, vengono restituiti solo i database e le tabelle per i quali il chiamante dispone delle autorizzazioni. Inoltre, quando si accede a un database o a una tabella di destinazione tramite un collegamento a una risorsa, è necessario disporre delle autorizzazioni sia AWS Identity and Access Management (IAM) che di Lake Formation sia sulla destinazione che sul collegamento alla risorsa. L'autorizzazione di Lake Formation richiesta per i collegamenti alle risorse è DESCRIBE. Per ulteriori informazioni, consulta [DESCRIBE](#).

Operazioni dell'API del database

Operazione API	Gestione dei collegamenti alle risorse
CreateDatabase	Se il database è un collegamento di risorse, crea il collegamento alla risorsa al database di destinazione designato.
UpdateDatabase	Se il database designato è un collegamento a una risorsa, segue il collegamento e aggiorna il database di destinazione. Se è necessario modificare il collegamento alla risorsa per collegarsi a un database diverso, è necessario eliminarlo e crearne uno nuovo.
DeleteDatabase	Elimina il link alla risorsa. Non elimina il database collegato (di destinazione).
GetDatabase	Se il chiamante dispone delle autorizzazioni sulla destinazione, segue il link per restituire le proprietà della destinazione. Altrimenti, restituisce le proprietà del link.

Operazione API	Gestione dei collegamenti alle risorse
GetDatabases	Restituisce un elenco di database, inclusi i collegamenti alle risorse. Per ogni collegamento di risorsa nel set di risultati, l'operazione segue il collegamento per ottenere le proprietà della destinazione del collegamento. Devi specificare <code>ResourceShareType = ALL</code> per vedere i database condivisi con il tuo account.

Tabella delle operazioni API

Operazione API	Gestione dei link alle risorse
CreateTable	Se il database è un collegamento a una risorsa, segue il collegamento al database e crea una tabella nel database di destinazione. Se la tabella è un collegamento di risorse, l'operazione crea il collegamento alle risorse nel database designato. La creazione di un collegamento alle risorse della tabella tramite un collegamento alle risorse del database non è supportata.
UpdateTable	Se la tabella o il database designato è un collegamento a una risorsa, aggiorna la tabella di destinazione. Se sia la tabella che il database sono collegamenti a risorse, l'operazione ha esito negativo.
DeleteTable	Se il database designato è un collegamento a una risorsa, segue il collegamento ed elimina la tabella o il collegamento alle risorse della tabella nel database di destinazione. Se la tabella è un collegamento a una risorsa, l'operazione elimina il collegamento alle risorse della tabella nel database designato. L'eliminazione di un collegamento alle risorse della tabella non elimina la tabella di destinazione.
BatchDeleteTable	Come DeleteTable .
GetTable	Se il database designato è un collegamento di risorse, segue il collegamento al database e restituisce la tabella o il collegamento alle risorse della tabella dal database di destinazione. Altrimenti,

Operazione API	Gestione dei link alle risorse
	se la tabella è un collegamento a una risorsa, l'operazione segue il collegamento e restituisce le proprietà della tabella di destinazione.
GetTables	Se il database designato è un collegamento a una risorsa, segue il collegamento al database e restituisce le tabelle e i collegamenti alle risorse della tabella dal database di destinazione. Se il database di destinazione è un database condiviso di un altro AWS account, l'operazione restituisce solo le tabelle condivise in quel database. Non segue i collegamenti alle risorse della tabella nel database di destinazione. Altrimenti, se il database designato è un database locale (di proprietà), l'operazione restituisce tutte le tabelle del database locale e segue ogni collegamento alle risorse della tabella per restituire le proprietà della tabella di destinazione.
SearchTables	Restituisce tabelle e collegamenti alle risorse delle tabelle. Non segue i link per restituire le proprietà della tabella di destinazione. Devi specificare <code>ResourceShareType = ALL</code> per vedere le tabelle condivise con il tuo account.
GetTableVersion	Come GetTable.
GetTableVersions	Come GetTable.
DeleteTableVersion	Come DeleteTable .
BatchDeleteTableVersion	Come DeleteTable .

Operazioni dell'API di partizione

Operazione API	Gestione dei link alle risorse
CreatePartition	Se il database designato è un collegamento di risorse, segue il collegamento al database e crea una partizione nella tabella designata nel database di destinazione. Se la tabella è un collegamento di risorse, l'operazione segue il collegamento alle risorse e crea la partizione nella tabella di destinazione. La

Operazione API	Gestione dei link alle risorse
	creazione di una partizione tramite un collegamento alle risorse della tabella e un collegamento alle risorse del database non è supportata.
BatchCreatePartitions	Come CreatePartition .
UpdatePartition	Se il database designato è un collegamento a una risorsa, segue il collegamento al database e aggiorna la partizione nella tabella designata nel database di destinazione. Se la tabella è un collegamento a una risorsa, l'operazione segue il collegamento alle risorse e aggiorna la partizione nella tabella di destinazione. L'aggiornamento di una partizione tramite un collegamento alle risorse della tabella e un collegamento alle risorse del database non è supportato.
DeletePartition	Se il database designato è un collegamento di risorse, segue il collegamento al database ed elimina la partizione nella tabella designata nel database di destinazione. Se la tabella è un collegamento di risorse, l'operazione segue il collegamento alla risorsa ed elimina la partizione nella tabella di destinazione. L'eliminazione di una partizione tramite un collegamento alle risorse della tabella e un collegamento alle risorse del database non è supportata.
BatchDeletePartitions	Come DeletePartition .
GetPartition	Se il database designato è un collegamento a una risorsa, segue il collegamento al database e restituisce le informazioni sulla partizione e dalla tabella designata. Altrimenti, se la tabella è un collegamento a una risorsa, l'operazione segue il collegamento e restituisce informazioni sulla partizione. Se sia la tabella che il database sono collegamenti a risorse, restituisce un set di risultati vuoto.

Operazione API	Gestione dei link alle risorse
GetPartitions	Se il database designato è un collegamento di risorse, segue il collegamento al database e restituisce le informazioni sulla partizione per tutte le partizioni nella tabella designata. Altrimenti, se la tabella è un collegamento di risorse, l'operazione segue il collegamento e restituisce informazioni sulla partizione. Se sia la tabella che il database sono collegamenti a risorse, restituisce un set di risultati vuoto.
BatchGetPartition	Come GetPartition .

Funzioni definite dall'utente, operazioni API.

Operazione API	Gestione dei link alle risorse
(Tutte le operazioni API)	Se il database è un collegamento a una risorsa, segue il collegamento alla risorsa ed esegue l'operazione sul database di destinazione.

 Consulta anche:

- [Come funzionano i link alle risorse in Lake Formation](#)

Accesso alle tabelle in tutte le regioni

Lake Formation supporta l'interrogazione delle tabelle del Data Catalog tra AWS le regioni. Puoi accedere ai dati in una regione da altre regioni utilizzando Amazon Athena, Amazon EMR ed AWS Glue ETL [creando collegamenti a risorse](#) in altre regioni che puntano ai database e alle tabelle di origine. Con l'accesso alle tabelle tra regioni, puoi accedere ai dati tra le regioni senza copiare i dati o i metadati sottostanti nel catalogo dati.

Ad esempio, puoi condividere un database o una tabella in un account produttore con un account consumatore nella Regione A. Dopo aver accettato l'invito alla condivisione delle risorse nella Regione A, l'amministratore del data lake dell'account consumer può creare collegamenti alle risorse

condivise nella Regione B. L'amministratore dell'account consumer può concedere le autorizzazioni sulla risorsa condivisa ai responsabili IAM di quell'account nella Regione A e può concedere le autorizzazioni per il collegamento alle risorse nella Regione B. Utilizzando il link alla risorsa, i principali nell'account consumer può interrogare i dati condivisi dalla regione B.

Puoi anche ospitare l'origine dati Amazon S3 nella Regione A in un account produttore e registrare la posizione dei dati in un account centrale nella Regione B. Puoi creare risorse Data Catalog nell'account centrale, impostare le autorizzazioni di Lake Formation e condividere dati con i consumatori nel tuo account o con account esterni nella Regione B. La funzionalità interregionale consente agli utenti di accedere a queste tabelle del Catalogo dati dalla Regione C utilizzando collegamenti alle risorse.

Utilizzando questa funzionalità, è possibile interrogare i database federati in Apache Hive Metastores tra regioni e anche unire tabelle nella regione locale con tabelle in un'altra regione durante l'esecuzione di query.

Lake Formation supporta le seguenti funzionalità con l'accesso alle tabelle tra regioni:

- Controllo degli accessi basato su tag LF
- Autorizzazioni di controllo degli accessi granulari
- Operazioni di scrittura sul database o sulla tabella condivisa con le autorizzazioni appropriate
- Condivisione dei dati tra account a livello di account e direttamente con i principali IAM

Gli utenti non amministrativi con `Create_Database` e `Create_Table` autorizzazioni possono creare collegamenti di risorse tra regioni.

Note

Puoi creare collegamenti a risorse interregionali in qualsiasi regione e accedere ai dati senza applicare le autorizzazioni di Lake Formation. Per i dati di origine in Amazon S3 che non sono registrati con Lake Formation, l'accesso è determinato dalle politiche di autorizzazione IAM per Amazon S3 e dalle azioni. AWS Glue

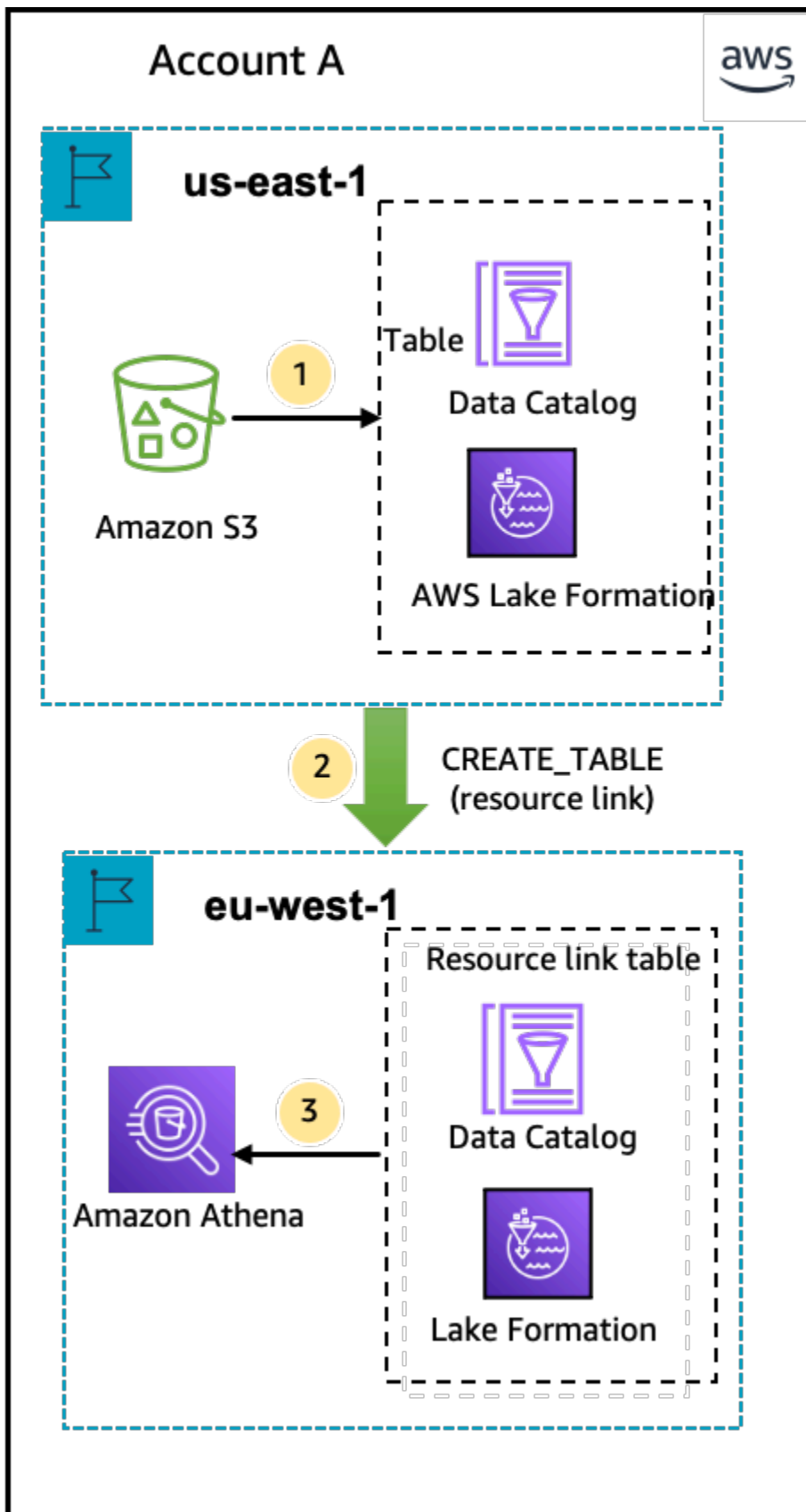
Per le limitazioni, consulta [Limitazioni di accesso ai dati tra regioni](#).

Flussi di lavoro

I seguenti diagrammi mostrano i flussi di lavoro per l'accesso ai dati tra AWS regioni dallo stesso AWS account e da un account esterno.

Flusso di lavoro per accedere alle tabelle condivise all'interno dello stesso account AWS

Nel diagramma seguente, i dati vengono condivisi con un utente dello stesso AWS account nella regione Stati Uniti orientali (Virginia settentrionale) e l'utente richiede i dati condivisi dalla regione Europa (Irlanda).



L'amministratore del data lake esegue le seguenti attività (passaggi 1-2):

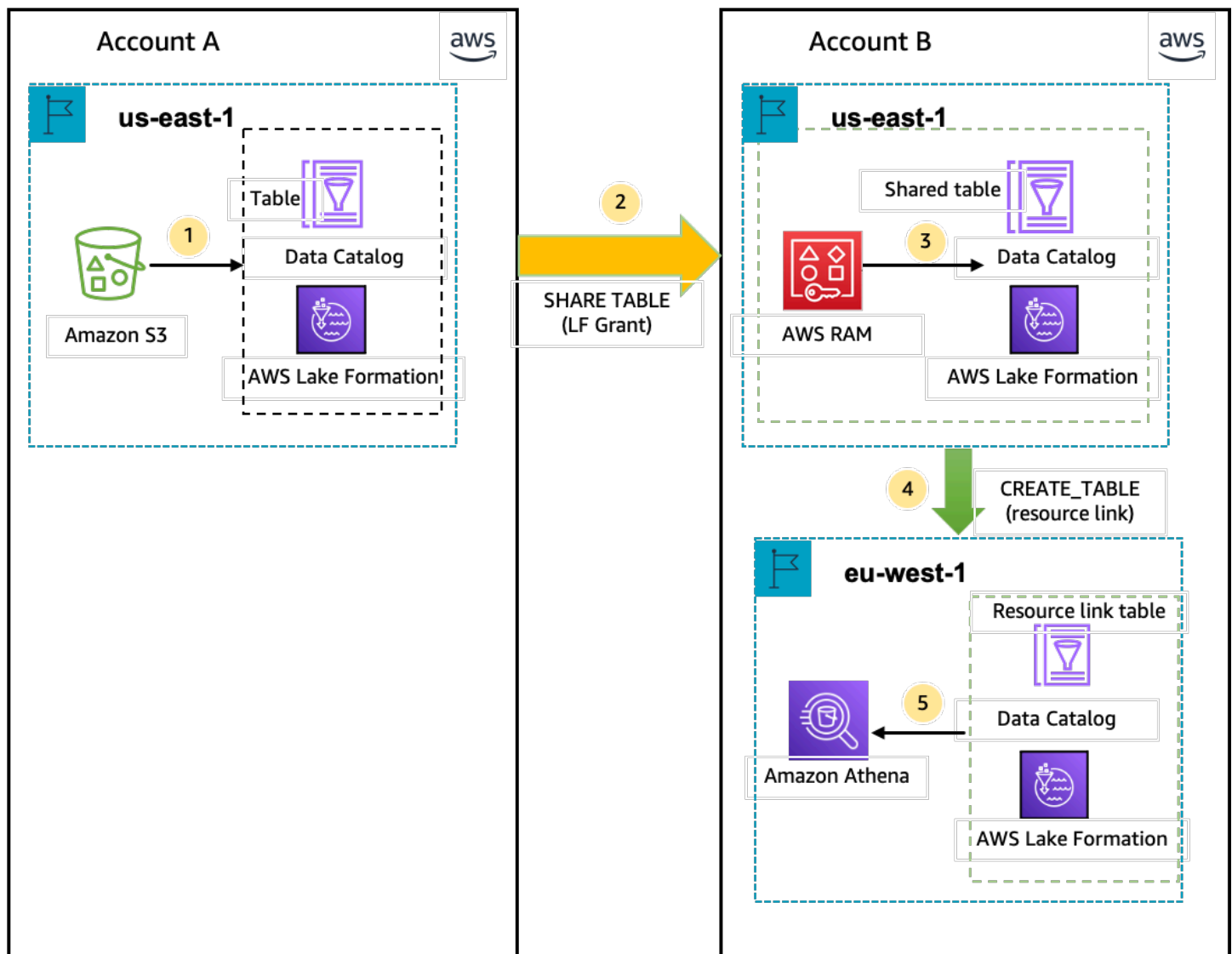
1. Un amministratore del data lake configura un AWS account con i database e le tabelle Data Catalog e registra una posizione dati Amazon S3 con Lake Formation nella regione Stati Uniti orientali (Virginia settentrionale).

Concede Select l'autorizzazione per una risorsa del Data Catalog (tabella dei prodotti nel diagramma) a un principale (utente) nello stesso account.

2. Crea un collegamento alla risorsa nella regione Europa (Irlanda) che punta alla tabella di origine nella regione Stati Uniti orientali (Virginia settentrionale). Concede l'DESCRIBE autorizzazione per il collegamento alla risorsa dalla regione Europa (Irlanda) al principale.
3. L'utente esegue una query sulla tabella dalla regione Europa (Irlanda) utilizzando Athena.

Flusso di lavoro per accedere alle tabelle condivise con un account esterno AWS

Nel diagramma seguente, l'account produttore (Account A) ospita il bucket Amazon S3, registra la posizione dei dati e condivide una tabella del catalogo dati con un account consumer (Account B) nella regione Stati Uniti orientali (Virginia settentrionale) e un utente dell'account consumatore (Account B) esegue una query sulla tabella dalla regione Europa (Irlanda).



1. Un amministratore del data lake configura un AWS account (account produttore) con le risorse Data Catalog e una data location Amazon S3 registrata presso Lake Formation nella regione Stati Uniti orientali (Virginia settentrionale).
2. L'amministratore del data lake dell'account produttore condivide una tabella Data Catalog con un account consumatore.
3. L'amministratore del data lake dell'account consumer accetta l'invito alla condivisione dei dati nella regione Stati Uniti orientali (Virginia settentrionale) e concede l'Select autorizzazione per la tabella condivisa a un principale della stessa regione.
4. L'amministratore del data lake dell'account consumer crea un link di risorsa nella regione Europa (Irlanda) che punta alla tabella condivisa di destinazione nella regione Stati Uniti orientali (Virginia

settentrionale) e concede DESCRIBE all'utente l'autorizzazione per il collegamento alla risorsa dalla regione Europa (Irlanda).

5. L'utente interroga i dati dalla regione Europa (Irlanda) utilizzando Athena.

Configurazione dell'accesso alle tabelle tra regioni

Per accedere ai dati da una regione diversa, devi prima configurare i database e le tabelle del catalogo dati nella regione in cui registri la tua posizione dati su Amazon S3. Puoi condividere i database e le tabelle del Data Catalog con i principali del tuo account o di un altro account. Quindi, devi creare amministratori di data lake che possano creare collegamenti a risorse che puntino alla posizione dei dati condivisi di destinazione nelle regioni in cui gli utenti interrogano i dati.

Per interrogare i dati condivisi all'interno dello stesso account da una regione diversa

In questa sezione, la regione della tabella condivisa di destinazione viene denominata Regione A e gli utenti eseguono query dalla Regione B.

1. Configurazione dell'account nella Regione A (dove si creano e si condividono i dati)

Un amministratore del data lake deve completare le seguenti azioni:

- a. Registra una posizione dati Amazon S3.

Per ulteriori informazioni, consulta [Aggiungere una posizione Amazon S3 al tuo data lake](#).

- b. Crea database e tabelle nell'account. Questa operazione può essere eseguita anche da un utente non amministrativo che dispone delle autorizzazioni per creare database e tabelle.
- c. Concedi le autorizzazioni relative ai dati su una tabella ai principali con. Grantable permissions

Per ulteriori informazioni, consultare [Concessione delle autorizzazioni per le risorse di Data Catalog](#).

2. Configurazione dell'account nella Regione B (dove si accede ai dati)

Un amministratore del data lake deve completare le seguenti azioni:

- a. Crea un link di risorsa nella Regione B che punti alla tabella condivisa di destinazione nella Regione A. Specificare la regione proprietaria della tabella condivisa nella schermata Crea tabella.

Create table

Table details

Create a table in the AWS Glue Data Catalog.

☐ Table
Create a table in my account.

☒ Resource link
Create a resource link to a shared table.

Resource link name

Enter resource link name

Name may contain letters (A-Z), numbers (0-9), hyphens (-), or underscores (_), and must be less than 256 characters long.

Database

Resource link will be contained in this database.

Enter or choose a database

Shared table owner region

Select the region where the table is shared

US West (N. California)

Shared table

Enter or choose a shared table.

Enter or choose a shared table.

Shared table's database

Enter the database containing the shared table.

Enter the database that contains the shared table

Shared table's owner ID

Enter the AWS account ID of the shared table owner.

Enter an AWS account ID

Cancel

Create

Per istruzioni sulla creazione di collegamenti di risorse a database e tabelle, vedere [Creazione di collegamenti alle risorse](#).

- b. Concedi Describe l'autorizzazione ai responsabili IAM sul link alla risorsa nella Regione B.

Per ulteriori informazioni sulla concessione delle autorizzazioni sui link alle risorse, consulta [Concessione delle autorizzazioni per i collegamenti alle risorse](#)

I responsabili IAM nella regione B possono interrogare la tabella di destinazione tramite il collegamento utilizzando Athena.

Per accedere ai dati di più account da una regione diversa

1. Configurazione dell'account produttore/concedente

Un amministratore del data lake deve completare le seguenti azioni:

- a. Configura l'account produttore/concedente nella Regione A.
- b. Registra una posizione dati Amazon S3 nella regione A.
- c. Crea database e tabelle. Questa operazione può essere eseguita da un utente non amministrativo che dispone delle autorizzazioni necessarie per creare tabelle.
- d. Concedi le autorizzazioni relative ai dati all'account consumatore/beneficiario su una tabella nella Regione A con `Grantable permissions`

Per ulteriori informazioni, consulta [Condivisione delle tabelle e dei database del Data Catalog tra i nostri principali IAM provenienti Account AWS da account esterni](#).

2. Configurazione dell'account consumatore/beneficiario

Un amministratore del data lake deve completare le seguenti azioni:

- a. Accetta l'invito alla condivisione delle risorse dalla AWS RAM Regione A.
- b. Crea un link alla risorsa nella Regione B che punti alla tabella condivisa. La regione B è dove gli utenti vorranno interrogare la tabella.
- c. Concedi le autorizzazioni relative ai dati sulla tabella condivisa ai responsabili IAM nella regione A.

Note

È necessario concedere le autorizzazioni alla tabella condivisa nella stessa regione in cui la tabella è stata condivisa.

- d. Concedi le autorizzazioni ai responsabili sul link alla risorsa nella Regione B.

I responsabili dell'account consumatore nella regione B interrogano quindi la tabella condivisa dalla regione B utilizzando Athena.

Sicurezza in AWS Lake Formation

La sicurezza del cloud AWS è la massima priorità. In qualità di AWS cliente, puoi beneficiare di un data center e di un'architettura di rete progettati per soddisfare i requisiti delle organizzazioni più sensibili alla sicurezza.

La sicurezza è una responsabilità condivisa tra AWS te e te. Il [modello di responsabilità condivisa](#) descrive questo come sicurezza del cloud e sicurezza nel cloud:

- **Sicurezza del cloud:** AWS è responsabile della protezione dell'infrastruttura che gestisce AWS i servizi nel AWS cloud. AWS ti fornisce 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 ulteriori informazioni sui programmi di conformità applicabili AWS Lake Formation, consulta [AWS Services in Scope by Compliance Program](#).
- **Sicurezza nel cloud:** la tua responsabilità è determinata dal AWS servizio che utilizzi. L'utente è anche responsabile di altri fattori, tra cui la riservatezza dei dati, i requisiti della propria azienda e le leggi e normative vigenti.

Questa documentazione ti aiuta a capire come applicare il modello di responsabilità condivisa quando usi Lake Formation. I seguenti argomenti mostrano come configurare Lake Formation per soddisfare i tuoi obiettivi di sicurezza e conformità. Imparerai anche a utilizzare altri AWS servizi che ti aiutano a monitorare e proteggere le tue risorse di Lake Formation.

Argomenti

- [Protezione dei dati in Lake Formation](#)
- [Sicurezza dell'infrastruttura in AWS Lake Formation](#)
- [Prevenzione del confused deputy tra servizi](#)
- [Registrazione degli eventi di sicurezza AWS Lake Formation](#)

Protezione dei dati in Lake Formation

Il modello di [responsabilità AWS condivisa modello](#) di di si applica alla protezione dei dati in AWS Lake Formation. Come descritto in questo modello, AWS è responsabile della protezione dell'infrastruttura globale che gestisce tutto il Cloud AWS. L'utente è responsabile del controllo dei contenuti ospitati su questa infrastruttura. L'utente è inoltre responsabile della configurazione della

protezione e delle attività di gestione per i Servizi AWS utilizzati. Per ulteriori informazioni sulla privacy dei dati, vedi le [Domande frequenti sulla privacy dei dati](#). Per informazioni sulla protezione dei dati in Europa, consulta il post del blog relativo al [Modello di responsabilità condivisa AWS e GDPR](#) nel Blog sulla sicurezza AWS .

Ai fini della protezione dei dati, consigliamo di proteggere Account AWS le credenziali e configurare i singoli utenti con AWS IAM Identity Center o AWS Identity and Access Management (IAM). In tal modo, a ogni utente verranno assegnate solo le autorizzazioni necessarie per svolgere i suoi compiti. Ti suggeriamo, inoltre, di proteggere i dati nei seguenti modi:

- Utilizza l'autenticazione a più fattori (MFA) con ogni account.
- Usa SSL/TLS per comunicare con le risorse. AWS È richiesto TLS 1.2 ed è consigliato TLS 1.3.
- Configura l'API e la registrazione delle attività degli utenti con. AWS CloudTrail Per informazioni sull'utilizzo dei CloudTrail percorsi per acquisire AWS le attività, consulta [Lavorare con i CloudTrail percorsi](#) nella Guida per l'AWS CloudTrail utente.
- Utilizza soluzioni di AWS crittografia, insieme a tutti i controlli di sicurezza predefiniti all'interno Servizi AWS.
- Utilizza i servizi di sicurezza gestiti avanzati, come Amazon Macie, che aiutano a individuare e proteggere i dati sensibili archiviati in Amazon S3.
- Se hai bisogno di moduli crittografici convalidati FIPS 140-3 per accedere AWS tramite un'interfaccia a riga di comando o un'API, usa un endpoint FIPS. Per ulteriori informazioni sugli endpoint FIPS disponibili, consulta il [Federal Information Processing Standard \(FIPS\) 140-3](#).

Ti consigliamo di non inserire mai informazioni riservate o sensibili, ad esempio gli indirizzi e-mail dei clienti, nei tag o nei campi di testo in formato libero, ad esempio nel campo Nome. Ciò include quando lavori con Lake Formation o altro Servizi AWS utilizzando la console, l'API o AWS SDKs. AWS CLI I dati inseriti nei tag o nei campi di testo in formato libero utilizzati per i nomi possono essere utilizzati per la fatturazione o i log di diagnostica. Quando fornisci un URL a un server esterno, ti suggeriamo vivamente di non includere informazioni sulle credenziali nell'URL per convalidare la tua richiesta al server.

Crittografia dei dati inattivi

AWS Lake Formation supporta la crittografia dei dati nelle seguenti aree:

- Dati nel tuo data lake Amazon Simple Storage Service (Amazon S3).

Lake Formation supporta la crittografia dei dati con [AWS Key Management Service](#) (AWS KMS). I dati vengono in genere scritti nel data lake tramite AWS Glue lavori di estrazione, trasformazione e caricamento (ETL). Per informazioni su come crittografare i dati scritti da AWS Glue jobs, [consulta Encrypting Data Written by Crawlers, Jobs and Development Endpoints](#) nella Developer Guide. AWS Glue

- Il AWS Glue Data Catalog, dove Lake Formation archivia le tabelle di metadati che descrivono i dati nel data lake.

Per ulteriori informazioni, [consulta Encrypting Your Data Catalog](#) nella AWS Glue Developer Guide.

Per aggiungere una posizione Amazon S3 come storage nel tuo data lake, devi registrare la posizione con AWS Lake Formation. Puoi quindi utilizzare le autorizzazioni di Lake Formation per un controllo granulare degli accessi agli AWS Glue Data Catalog oggetti che puntano a questa posizione e ai dati sottostanti nella posizione.

Lake Formation supporta la registrazione di una posizione Amazon S3 che contiene dati crittografati. Per ulteriori informazioni, consulta [Registrazione di una posizione Amazon S3 crittografata](#).

Sicurezza dell'infrastruttura in AWS Lake Formation

In quanto servizio gestito, AWS Lake Formation è protetto dalle procedure di sicurezza della rete AWS globale descritte nel white paper [Amazon Web Services: Overview of Security Processes](#).

Utilizzi chiamate API AWS pubblicate per accedere a Lake Formation attraverso la rete. I client devono supportare Transport Layer Security (TLS) 1.0 o versioni successive. È consigliabile TLS 1.2 o versioni successive. I client devono, inoltre, supportare le suite di cifratura con PFS (Perfect Forward Secrecy), 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. In alternativa, è possibile utilizzare [AWS Security Token Service](#) (AWS STS) per generare le credenziali di sicurezza temporanee per sottoscrivere le richieste.

Prevenzione del confused deputy tra servizi

Il problema confused deputy è un problema di sicurezza in cui un'entità che non dispone dell'autorizzazione per eseguire un'azione può costringere un'entità maggiormente privilegiata a eseguire l'azione. Nel AWS, l'impersonificazione tra servizi può portare al confuso problema del vice. La rappresentazione tra servizi può verificarsi quando un servizio (il servizio chiamante) effettua una chiamata a un altro servizio (il servizio chiamato). Il servizio chiamante può essere manipolato per utilizzare le proprie autorizzazioni e agire sulle risorse di un altro cliente, a cui normalmente non avrebbe accesso. Per evitare che ciò accada, AWS mette a disposizione strumenti che consentono di proteggere i dati relativi a tutti i servizi con responsabili del servizio a cui è stato concesso l'accesso alle risorse del vostro account.

Ti consigliamo di utilizzare le chiavi di contesto delle condizioni globali [aws:SourceArn](#) e [aws:SourceAccount](#) nelle policy delle risorse per limitare le autorizzazioni con cui AWS Lake Formation fornisce un altro servizio alla risorsa. Se si utilizzano entrambe le chiavi di contesto delle condizioni globali, il valore `aws:SourceAccount` e l'account nel valore `aws:SourceArn` devono utilizzare lo stesso ID account nella stessa istruzione di policy.

Attualmente, Lake Formation supporta `aws:SourceArn` solo il seguente formato:

```
arn:aws:lakeformation:aws-region:account-id:*
```

L'esempio seguente mostra come utilizzare le chiavi di contesto `aws:SourceArn` e `aws:SourceAccount` global condition in Lake Formation per prevenire il confuso problema del vice.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ConfusedDeputyPreventionExamplePolicy",
      "Effect": "Allow",
      "Principal": {
        "Service": "lakeformation.amazonaws.com"
      },
      "Action": [
        "sts:AssumeRole"
      ],
    },
  ],
}
```



```
"Condition": {
  "StringEquals": {
    "aws:SourceAccount": "account-id"
  },
  "ArnEquals": {
    "aws:SourceArn": "arn:aws:lakeformation:aws-region:account-id:"
  }
}
]
```

Registrazione degli eventi di sicurezza AWS Lake Formation

AWS Lake Formation è integrato con AWS CloudTrail, un servizio che fornisce una registrazione delle azioni intraprese da un utente, ruolo o AWS servizio in Lake Formation. CloudTrail acquisisce tutte le chiamate API per Lake Formation come eventi. Le chiamate acquisite includono chiamate dalla console di Lake Formation AWS Command Line Interface, e chiamate in codice alle operazioni dell'API Lake Formation.

Per ulteriori informazioni sulla registrazione degli eventi in Lake Formation, vedere [Registrazione delle chiamate API AWS Lake Formation utilizzando AWS CloudTrail](#).

Note

GetTableObjectsUpdateTableObjects, e GetWorkUnitResults sono operazioni sul piano dati ad alto volume. Le chiamate verso queste APIs destinazioni non sono attualmente registrate. CloudTrail Per ulteriori informazioni sulle operazioni del piano dati in CloudTrail, vedere [Registrazione degli eventi relativi ai dati per i percorsi nella Guida per l'AWS CloudTrail utente](#).

Le modifiche apportate a Lake Formation a supporto di CloudTrail eventi aggiuntivi saranno documentate su [Cronologia dei documenti per AWS Lake Formation](#).

Integrazione di servizi di terze parti con Lake Formation

L'integrazione con AWS Lake Formation consente ai servizi di terze parti di accedere in modo sicuro ai dati nei loro data lake basati su Amazon S3. Puoi utilizzare Lake Formation come motore di autorizzazione per gestire o applicare le autorizzazioni al tuo data lake con AWS servizi integrati come AWS Glue ETL, Amazon Athena, Amazon EMR e Redshift Spectrum. Lake Formation offre due opzioni per l'integrazione dei servizi:

1. Le impostazioni di integrazione delle applicazioni Lake Formation: Lake Formation può fornire credenziali temporanee ridotte sotto forma di token AWS STS alle sedi Amazon S3 registrate in base alle autorizzazioni effettive, in modo che le applicazioni autorizzate possano accedere ai dati per conto degli utenti.
2. Applicazione centralizzata: le operazioni di [interrogazione delle API](#) di Lake Formation recuperano i dati da Amazon S3 e filtrano i risultati in base a autorizzazioni effettive. Il motore o l'applicazione che si integra con l'operazione dell'API di interrogazione può dipendere da Lake Formation per valutare le autorizzazioni dell'identità chiamante e filtrare in modo sicuro i dati in base a tali autorizzazioni. I motori di query di terze parti vedono e funzionano solo su dati filtrati.

Il vending di credenziali di Lake Formation non si integra con le query sql spark. La vendita di credenziali funziona solo con le query che vengono eseguite attraverso la libreria ETL. AWS Glue

Argomenti

- [Utilizzo dell'integrazione delle applicazioni Lake Formation](#)

Utilizzo dell'integrazione delle applicazioni Lake Formation

Lake Formation consente ai servizi di terze parti di integrarsi con Lake Formation e ottenere l'accesso temporaneo ai dati di Amazon S3 per conto dei propri utenti tramite l'utilizzo [GetTemporaryGlueTableCredentials](#) e [GetTemporaryGluePartitionCredentials](#) le operazioni. Ciò consente ai servizi di terze parti di utilizzare la stessa funzionalità di vendita di autorizzazioni e credenziali utilizzata dagli altri servizi di AWS analisi. Questa sezione descrive come utilizzare queste operazioni API per integrare un motore di query di terze parti con. Lake Formation

Queste operazioni API sono disabilitate per impostazione predefinita. Esistono due opzioni per autorizzare Lake Formation a integrare le applicazioni:

- Configura i tag di sessione IAM che vengono convalidati ogni volta che vengono richiamate le operazioni dell'API di integrazione delle applicazioni

Per ulteriori informazioni, consulta [Abilitazione delle autorizzazioni per un motore di query di terze parti per chiamare le operazioni dell'API di integrazione delle applicazioni](#).

- Abilita l'opzione che consente ai motori esterni di accedere ai dati nelle ubicazioni Amazon S3 con accesso completo alla tabella

Questa opzione consente ai motori di query e alle applicazioni di ottenere credenziali senza tag di sessione IAM se l'utente ha accesso completo alla tabella. Fornisce vantaggi in termini di prestazioni ai motori di query e alle applicazioni e semplifica l'accesso ai dati. Amazon EMR su Amazon EC2 è in grado di sfruttare questa impostazione.

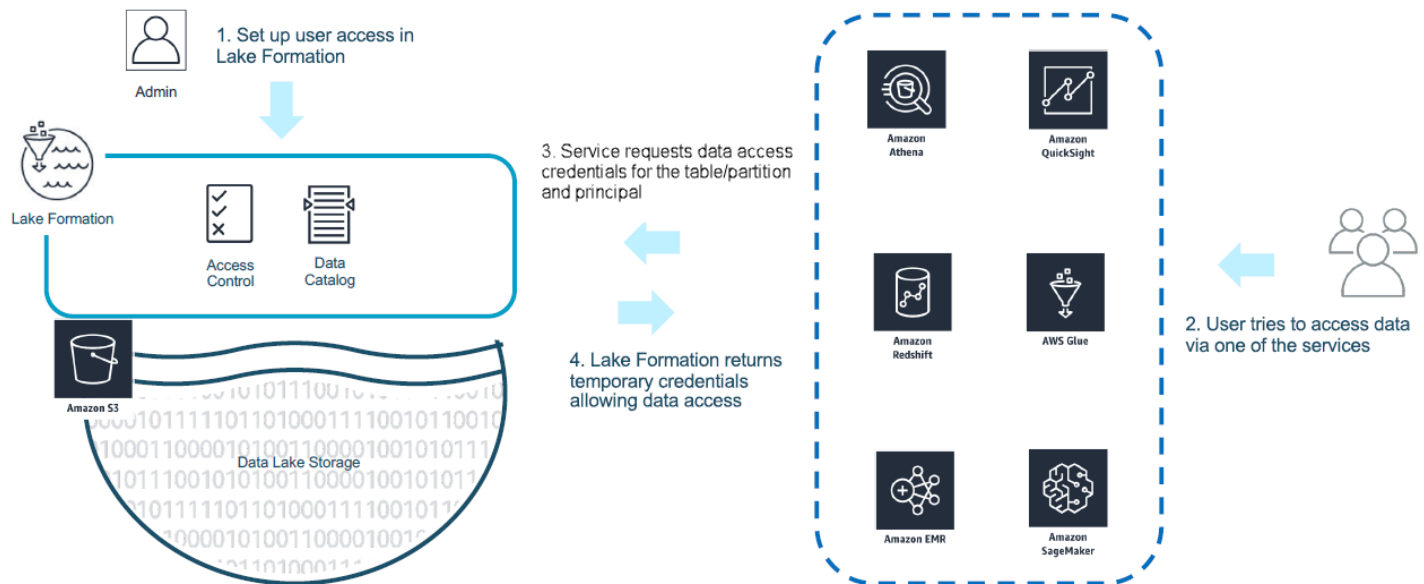
Per ulteriori informazioni, consulta [Integrazione delle applicazioni per l'accesso completo alla tabella](#).

Argomenti

- [Come funziona l'integrazione delle applicazioni Lake Formation](#)
- [Ruoli e responsabilità nell'integrazione delle applicazioni Lake Formation](#)
- [Lake Formation flusso di lavoro per le operazioni API di integrazione delle applicazioni](#)
- [Registrazione di un motore di query di terze parti](#)
- [Abilitazione delle autorizzazioni per un motore di query di terze parti per chiamare le operazioni dell'API di integrazione delle applicazioni](#)
- [Integrazione delle applicazioni per l'accesso completo alla tabella](#)

Come funziona l'integrazione delle applicazioni Lake Formation

Questa sezione descrive come utilizzare le operazioni API di integrazione delle applicazioni per integrare un'applicazione di terze parti (motore di query) con Lake Formation.



1. L'Lake Formation amministratore svolge le seguenti attività:

- Registra una sede Amazon S3 con Lake Formation fornendo un ruolo IAM (utilizzato per le credenziali di vendita) con le autorizzazioni appropriate per accedere ai dati all'interno della posizione Amazon S3
- Registra un'applicazione di terze parti per poter chiamare le operazioni API di vendita delle credenziali di Lake Formation. Per informazioni, consultare [the section called "Registrazione di un motore di query di terze parti"](#).
- Concede agli utenti le autorizzazioni per accedere a database e tabelle

Ad esempio, se si desidera pubblicare un set di dati sulle sessioni utente che include alcune colonne contenenti informazioni di identificazione personale (PII), per limitare l'accesso, si assegna a queste colonne un tag [LF-TBAC](#) denominato «classificazione» con il valore «sensibile». Successivamente, si definisce un'autorizzazione che consenta a un analista aziendale di accedere ai dati delle sessioni utente, ma si escludono le colonne contrassegnate con classificazione = sensibile.

2. Un principale (utente) invia una richiesta a un servizio integrato.
3. L'applicazione integrata invia la richiesta a Lake Formation chiedendo le informazioni sulla tabella e le credenziali per accedere alla tabella.
4. Se il principale richiedente è autorizzato ad accedere alla tabella, Lake Formation restituisce le credenziali all'applicazione integrata, che consente l'accesso ai dati.

 Note

Lake Formation non accede ai dati sottostanti quando vende le credenziali.

5. Il servizio integrato legge i dati da Amazon S3, filtra le colonne in base alle politiche ricevute e restituisce i risultati al principale.

 Important

Lake Formation Le operazioni API di vendita delle credenziali abilitano un modello di applicazione distribuita con un modello esplicito di negazione dell'errore (fail-close). Questo introduce un modello di sicurezza a tre parti tra clienti, servizi di terze parti e Lake Formation. I servizi integrati sono affidabili per applicare correttamente le Lake Formation autorizzazioni (applicazione distribuita).

Il servizio integrato è responsabile del filtraggio dei dati letti da Amazon S3 in base alle politiche restituite Lake Formation prima che i dati filtrati vengano restituiti all'utente. I servizi integrati seguono un modello di chiusura fallimentare, il che significa che devono fallire la query se non sono in grado di applicare le autorizzazioni richieste. Lake Formation

Ruoli e responsabilità nell'integrazione delle applicazioni Lake Formation

Di seguito sono riportati i ruoli e le responsabilità associate per abilitare l'integrazione di applicazioni di terze parti con AWS Lake Formation.

Ruolo	Responsabilità
Il cliente	• Abilita l'impostazione di integrazione dell'applicazione Lake Formation (vedi the section called “Registrazione di un motore di query di terze parti”). • Registra esplicitamente le terze parti approvate con Lake Formation (vedi the section called “Registrazione di un motore di query di terze parti”). • Testa e convalida soluzioni di terze parti con autorizzazioni Lake Formation.

Ruolo	Responsabilità
La terza parte	• Monitora e verifica l'utilizzo da parte di terzi delle operazioni dell'API di vendita delle credenziali di Lake Formation. • Documenta pubblicamente la funzionalità supportata per ogni revisione del software e fornisce istruzioni per attivarla correttamente. • Pubblicizza in modo accurato le funzionalità supportate quando richiama le operazioni API di vendita delle credenziali di Lake Formation (secondo la documentazione). • Archivia e gestisce in modo sicuro le credenziali vendute per evitare fughe di credenziali e aumento dei privilegi. • Applica le autorizzazioni in base alle funzionalità supportate e restituisce agli utenti solo dati filtrati • Non riesce a eseguire la query quando non è possibile applicare correttamente le autorizzazioni richieste
AWS Lake Formation	• Deriva e restituisce correttamente le autorizzazioni effettive per un determinato principale. • Convalida le funzionalità supportate da terze parti in base al funzionamento dell'API. call-by-call • Restituisce credenziali IAM limitate solo quando le funzionalità pubblicizzate del motore corrispondono a quelle definite nelle risorse del catalogo, altrimenti restituisce un errore.

Lake Formation flusso di lavoro per le operazioni API di integrazione delle applicazioni

Di seguito è riportato il flusso di lavoro per le operazioni dell'API di integrazione delle applicazioni:

1. Un utente invia una query o una richiesta di dati utilizzando un motore di query integrato di terze parti. Il motore di query assume un ruolo IAM che rappresenta l'utente o un gruppo di utenti e recupera credenziali affidabili da utilizzare quando si chiamano le operazioni dell'API di integrazione delle applicazioni.

2. Il motore di query chiama `eGetUnfilteredTableMetadata`, se si tratta di una tabella partizionata, chiama `GetUnfilteredPartitionsMetadata` per recuperare i metadati e le informazioni sulle policy dal Data Catalog.
3. Lake Formation esegue l'autorizzazione per la richiesta. Se l'utente non dispone delle autorizzazioni appropriate sulla tabella, `AccessDeniedException` viene generato.
4. Come parte della richiesta, il motore di query invia il filtro che supporta. È possibile inviare due flag all'interno di un array: `COLUMN_PERMISSIONS` e `CELL_FILTER_PERMISSION`. Se il motore di query non supporta nessuna di queste funzionalità e nella tabella esiste una policy per la funzionalità, viene generato un messaggio e la query ha esito negativo. `PermissionTypeMismatchException` Questo serve per evitare la perdita di dati.
5. La risposta restituita contiene quanto segue:
 - L'intero schema della tabella in modo che i motori di query possano utilizzarlo per analizzare i dati dall'archiviazione.
 - Un elenco di colonne autorizzate a cui l'utente ha accesso. Se l'elenco delle colonne autorizzate è vuoto, indica che l'utente dispone `DESCRIBE` delle autorizzazioni, ma non delle `haSELECT`, e la query ha esito negativo.
 - Una bandiera `IsRegisteredWithLakeFormation`, che indica se Lake Formation può fornire credenziali ai dati di queste risorse. Se restituisce false, le credenziali dei clienti devono essere utilizzate per accedere ad Amazon S3.
 - Un elenco degli `CellFilters` eventuali elementi da applicare alle righe di dati. Questo elenco contiene colonne e un'espressione per valutare ogni riga. Questo campo deve essere compilato solo se `CELL_FILTER_PERMISSION` viene inviato come parte della richiesta ed è presente un filtro di dati sulla tabella per l'utente chiamante.
6. Dopo aver recuperato i metadati, il motore di query chiama `GetTemporaryGlueTableCredentials` o `GetTemporaryGluePartitionCredentials` per ottenere AWS le credenziali per recuperare i dati dalla posizione Amazon S3.
7. Il motore di query legge gli oggetti pertinenti da Amazon S3, filtra i dati in base alle politiche ricevute nella fase 2 e restituisce i risultati all'utente.

Le operazioni dell'API di integrazione delle applicazioni Lake Formation contengono contenuti aggiuntivi per la configurazione dell'integrazione con motori di query di terze parti. Puoi vedere i dettagli dell'operazione nella sezione Operazioni dell'[API di vendita delle credenziali](#).

Registrazione di un motore di query di terze parti

Prima che un motore di query di terze parti possa utilizzare le operazioni dell'API di integrazione delle applicazioni, devi abilitare esplicitamente le autorizzazioni affinché il motore di query chiami le operazioni API per tuo conto. Questa operazione viene eseguita in pochi passaggi:

1. È necessario specificare AWS gli account e i tag di sessione IAM che richiedono l'autorizzazione a chiamare le operazioni dell'API di integrazione delle applicazioni tramite la AWS Lake Formation console AWS CLI o l'API/SDK.
2. Quando il motore di query di terze parti assume il ruolo di esecuzione nel tuo account, il motore di query deve allegare un tag di sessione registrato presso Lake Formation che rappresenta il motore di terze parti. Lake Formation utilizza questo tag per verificare che la richiesta provenga da un motore approvato. Per ulteriori informazioni sui tag di sessione, consulta Tag di [sessione](#) nella Guida per l'utente IAM.
3. Quando si configura un ruolo di esecuzione di un motore di query di terze parti, è necessario disporre del seguente set minimo di autorizzazioni nella policy IAM:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": {
    "Effect": "Allow",
    "Action": [
      "lakeformation:GetDataAccess",
      "glue:GetTable",
      "glue:GetTables",
      "glue:GetDatabase",
      "glue:GetDatabases",
      "glue>CreateDatabase",
      "glue:GetUserDefinedFunction",
      "glue:GetUserDefinedFunctions",
      "glue:GetPartition",
      "glue:GetPartitions"
    ],
    "Resource": "*"
  }
}
```

4. Imposta una policy di affidabilità dei ruoli sul ruolo di esecuzione del motore di query per avere un controllo di accesso preciso su quale coppia chiave-valore del tag di sessione può essere

associata a questo ruolo. Nell'esempio seguente, a questo ruolo è consentito associare solo la chiave del tag di sessione "LakeFormationAuthorizedCaller" e il valore "engine1" del tag di sessione e non è consentita nessun'altra coppia di valori chiave del tag di sessione.

```
{
  "Sid": "AllowPassSessionTags",
  "Effect": "Allow",
  "Principal": {
    "AWS": "arn:aws:iam::111122223333:role/query-execution-role"
  },
  "Action": "sts:TagSession",
  "Condition": {
    "StringLike": {
      "aws:RequestTag/LakeFormationAuthorizedCaller": "engine1"
    }
  }
}
```

Quando si LakeFormationAuthorizedCaller chiama l'operazione STS: AssumeRole API per recuperare le credenziali da utilizzare dal motore di query, il tag di sessione deve essere incluso nella AssumeRole richiesta. La credenziale temporanea restituita può essere utilizzata per effettuare richieste API di integrazione Lake Formation delle applicazioni.

Lake FormationLe operazioni dell'API di integrazione delle applicazioni richiedono che il principale chiamante sia un ruolo IAM. Il ruolo IAM deve includere un tag di sessione con un valore predeterminato che è stato registrato conLake Formation. Questo tag consente di Lake Formation verificare che il ruolo utilizzato per chiamare le operazioni dell'API di integrazione delle applicazioni sia autorizzato a farlo.

Abilitazione delle autorizzazioni per un motore di query di terze parti per chiamare le operazioni dell'API di integrazione delle applicazioni

Segui questi passaggi per consentire a un motore di query di terze parti di richiamare le operazioni dell'API di integrazione delle applicazioni tramite la AWS Lake Formation console, l' AWS CLI API/ SDK.

Console

Per registrare il tuo account per il filtraggio dei dati esterni:

1. Accedi a Console di gestione AWS, e apri la console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nella barra di navigazione a sinistra, espandi Amministrazione, quindi scegli Impostazioni di integrazione delle applicazioni.
3. Nella pagina delle impostazioni di integrazione delle applicazioni, scegli l'opzione Consenti ai motori esterni di filtrare i dati nelle località Amazon S3 registrate con. Lake Formation
4. Inserisci i tag di sessione che hai creato per il motore di terze parti. Per informazioni sui tag di sessione, consultate [Passare i tag di sessione in AWS STS](#) nella Guida AWS Identity and Access Management per l'utente.
5. Inserisci l'account IDs per gli utenti che possono utilizzare il motore di terze parti per accedere a informazioni sui metadati non filtrate e alle credenziali di accesso ai dati delle risorse dell'account corrente.

Puoi anche utilizzare il campo ID AWS account per configurare l'accesso tra account.

Application integration settings [Learn more](#)

Application integration settings

Use the options below to control which third-party engines are allowed to read and filter data in Amazon S3 locations registered with Lake Formation.

☒ Allow external engines to filter data in Amazon S3 locations registered with Lake Formation

Check this box to allow third-party engines to access data in Amazon S3 locations that are registered with Lake Formation.

Session tag values

Enter one or more strings that match the LakeFormationAuthorizedCaller session tag defined for third-party engines.

Clear all

engine 1 X

engine 2 X

session 1 X

Enter one or several string values separated by comma.

AWS account IDs

Enter the external AWS account IDs from where third-party engines are allowed to access locations registered with Lake Formation.

Clear all

111111111111 X
Account

222222222222 X
Account

Enter one or more AWS account IDs. Press enter after each ID.

☐ Allow external engines to access data in Amazon S3 locations with full table access.

When you enable this option, Lake Formation will return credentials to the integrated application directly without IAM session tag validation.

Cancel

Save

CLI

Utilizzate il comando `put-data-lake-settings` CLI per impostare i seguenti parametri.

Ci sono tre campi da configurare quando si utilizza questo AWS CLI comando:

- `allow-external-data-filtering` — (booleano) Indica che un motore di terze parti può accedere a informazioni sui metadati non filtrate e alle credenziali di accesso ai dati delle risorse dell'account corrente.
- `external-data-filtering-allow-list` — (array) Un elenco di account IDs che possono accedere alle informazioni sui metadati non filtrate e alle credenziali di accesso ai dati delle risorse dell'account corrente quando si utilizza un motore di terze parti. Quando

AllowExternalDataFiltering è impostata su true, la ExternalDataFilteringAllowList proprietà deve includere almeno un ID di account. Non è consentito un elenco vuoto.

- **authorized-sessions-tag-value-list**— (array) Un elenco di valori (stringhe) dei tag di sessione autorizzati. Se una credenziale di ruolo IAM è stata associata a una coppia chiave-valore autorizzata, se il tag di sessione è incluso nell'elenco, alla sessione viene concesso l'accesso alle informazioni sui metadati non filtrate e alle credenziali di accesso ai dati sulle risorse dell'account configurato. La chiave del tag di sessione autorizzata è definita come `*LakeFormationAuthorizedCaller*`
- **AllowFullTableExternalDataAccess**- (boolean) Indica se consentire a un motore di query di terze parti di ottenere le credenziali di accesso ai dati senza tag di sessione quando un chiamante dispone delle autorizzazioni complete di accesso ai dati.

Esempio:

```
aws lakeformation put-data-lake-settings --cli-input-json file://
datalakesettings.json

{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111111111111:user/lakeAdmin"
      }
    ],
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": [],
    "TrustedResourceOwners": [],
    "AllowExternalDataFiltering": true,
    "ExternalDataFilteringAllowList": [
      {"DataLakePrincipalIdentifier": "111111111111"}
    ],
    "AuthorizedSessionTagValueList": ["engine1"],
    "AllowFullTableExternalDataAccess": false
  }
}
```

API/SDK

Utilizzate l'operazione `PutDataLakeSetting` API per impostare i seguenti parametri.

Esistono tre campi da configurare quando si utilizza questa operazione API:

- `AllowExternalDataFiltering`— (Boolean) Indica se un motore di terze parti può accedere a informazioni sui metadati non filtrate e alle credenziali di accesso ai dati delle risorse dell'account corrente.
- `ExternalDataFilteringAllowList`— (array) Un elenco di account IDs che possono accedere alle informazioni sui metadati non filtrate e alle credenziali di accesso ai dati delle risorse dell'account corrente utilizzando un motore di terze parti.
- `AuthorizedSectionsTagValueList`— (array) Un elenco di valori di tag autorizzati (stringhe). Se una credenziale di ruolo IAM è stata associata a un tag autorizzato, alla sessione viene concesso l'accesso alle informazioni sui metadati non filtrate e alle credenziali di accesso ai dati sulle risorse dell'account configurato. La chiave del tag di sessione autorizzata è definita come. `*LakeFormationAuthorizedCaller*`
- `AllowFullTableExternalDataAccess`- (boolean) Indica se consentire a un motore di query di terze parti di ottenere le credenziali di accesso ai dati senza tag di sessione quando un chiamante dispone delle autorizzazioni complete di accesso ai dati.

Esempio:

```
//Enable session tag on existing data lake settings
public void sessionTagSetUpForExternalFiltering(AWSLakeFormationClient
lakeformation) {
    GetDataLakeSettingsResult getDataLakeSettingsResult =
    lfClient.getDataLakeSettings(new GetDataLakeSettingsRequest());
    DataLakeSettings dataLakeSettings =
    getDataLakeSettingsResult.getDataLakeSettings();

    //set account level flag to allow external filtering
    dataLakeSettings.setAllowExternalDataFiltering(true);

    //set account that are allowed to call credential vending or Glue
    GetFilteredMetadata API
    List<DataLakePrincipal> allowlist = new ArrayList<>();
    allowlist.add(new
    DataLakePrincipal().withDataLakePrincipalIdentifier("111111111111"));
}
```

```
dataLakeSettings.setWhitelistedForExternalDataFiltering(allowlist);

//set registered session tag values
List<String> registeredTagValues = new ArrayList<>();
registeredTagValues.add("engine1");
dataLakeSettings.setAuthorizedSessionTagValueList(registeredTagValues);

lakeformation.putDataLakeSettings(new
PutDataLakeSettingsRequest().withDataLakeSettings(dataLakeSettings));
}
```

Integrazione delle applicazioni per l'accesso completo alla tabella

Segui questi passaggi per consentire ai motori di query di terze parti di accedere ai dati senza la convalida dei tag di sessione IAM:

Console

1. Accedi alla console Lake Formation all'indirizzo <https://console.aws.amazon.com/lakeformation/>.
2. Nella barra di navigazione a sinistra, espandi Amministrazione e scegli Impostazioni di integrazione delle applicazioni.
3. Nella pagina delle impostazioni di integrazione delle applicazioni, scegli l'opzione Consenti ai motori esterni di accedere ai dati nelle posizioni Amazon S3 con accesso completo alla tabella.

Quando abiliti questa opzione, Lake Formation restituisce le credenziali all'applicazione di interrogazione direttamente senza la convalida del tag di sessione IAM.

Application integration settings [Learn more](#)

Application integration settings

Use the options below to control which third-party engines are allowed to read and filter data in Amazon S3 locations registered with Lake Formation.

☐ Allow external engines to filter data in Amazon S3 locations registered with Lake Formation

Check this box to allow third-party engines to access data in Amazon S3 locations that are registered with Lake Formation.

Session tag values

Enter one or more strings that match the LakeFormationAuthorizedCaller session tag defined for third-party engines.

[Clear all](#)
 engine 1 ✕

 engine 2 ✕

 session 1 ✕

Enter one or several string values separated by comma.

AWS account IDs

Enter the external AWS account IDs from where third-party engines are allowed to access locations registered with Lake Formation.

[Clear all](#)
 111111111111 ✕
Account

 222222222222 ✕
Account

Enter one or more AWS account IDs. Press enter after each ID.

☒ Allow external engines to access data in Amazon S3 locations with full table access.

When you enable this option, Lake Formation will return credentials to the integrated application directly without IAM session tag validation.

[Cancel](#)
[Save](#)

AWS CLI

Utilizzate il comando `put-data-lake-settings` CLI per impostare il `AllowFullTableExternalDataAccess` parametro.

```
aws lakeformation put-data-lake-settings --cli-input-json file://put-data-lake-
settings.json --region ap-northeast-1
{
  "DataLakeSettings": {
    "DataLakeAdmins": [
      {
        "DataLakePrincipalIdentifier": "arn:aws:iam::111111111111:user/
lakeAdmin"
      }
    ]
  }
}
```

```
    ],  
    "AllowFullTableExternalDataAccess": true  
  }  
}
```


Collaborazione con altri AWS servizi

AWS servizi come Amazon Athena AWS Glue, Amazon Redshift Spectrum e Amazon EMR possono essere AWS Lake Formation utilizzati per accedere in modo sicuro ai dati nelle sedi Amazon S3 registrate con Lake Formation. Con Lake Formation, puoi definire e gestire le autorizzazioni di controllo degli accessi granulari (FGAC) per le tue tabelle in. AWS Glue Data Catalog Ciascuno di questi AWS servizi è un chiamante affidabile di Lake Formation e Lake Formation fornisce l'accesso ai dati archiviati in Amazon S3 tramite credenziali temporanee. Per ulteriori informazioni, consulta [Come funziona l'integrazione delle applicazioni Lake Formation](#).

Per usufruire di queste funzionalità, Lake Formation richiede innanzitutto la registrazione della posizione Amazon S3 e l'assegnazione delle autorizzazioni appropriate al principale IAM per accedere alla tabella, al database e alla posizione Amazon S3. Per ulteriori informazioni, consultare [Gestione delle autorizzazioni di Lake Formation](#).

Le seguenti tabelle elencano i tipi di autorizzazioni Lake Formation supportate da Amazon Athena AWS Glue, Amazon EMR e Amazon Redshift Spectrum per accedere ai dati AWS Glue da tabelle standard e tabelle transazionali ([Apache Iceberg](#), [Apache Hudi](#) e [Linux foundation Delta Lake](#)) con dati archiviati in Amazon S3 e metadati di tabella nel Data Catalog.

AWS servizi e tipi AWS Glue di autorizzazione supportati per tabelle e viste standard

AWS servizio	autorizzazioni a livello di tabella	Autorizzazioni a livello di colonna	Autorizzazioni a livello di riga e cella
Athena SQL	Accesso in lettura/s crittura	Accesso in lettura	Accesso alla lettura
Athena Spark	Non supportato	Non supportato	Non supportato
Redshift Spectrum su un cluster fornito o Amazon Redshift serverless	Accesso in lettura/s crittura	Accesso in lettura	Accesso alla lettura
Apache Spark su Amazon EMR () EC2	Accesso in lettura/s crittura	Accesso in lettura	Accesso alla lettura

AWS servizio	autorizzazioni a livello di tabella	Autorizzazioni a livello di colonna	Autorizzazioni a livello di riga e cella
Apache Hive su Amazon EMR () EC2	Accesso in lettura/s crittura	Accesso in lettura	Non supportato
Apache Spark su EMR Serverless	Accesso in lettura/s crittura	Accesso in lettura	Accesso alla lettura
Apache Hive su EMR Serverless	Non supportato	Non supportato	Non supportato
Amazon EMR su EKS	Non supportato	Non supportato	Non supportato
AWS Glue ETL	Accesso in lettura/s crittura	AWS Glue 5.0 o versioni successive supportano l'accesso in lettura.	AWS Glue 5.0 o versioni successive supportano l'accesso in lettura.

Considerazioni e limitazioni

- Athena Spark non supporta l'interrogazione delle tabelle del Data Catalog con autorizzazioni Lake Formation.
- Gli utenti basati su SAML di Athena possono leggere le fonti di dati protette utilizzando le autorizzazioni di Lake Formation abilitando la federazione basata su SAML 2.0. Gli utenti SAML possono inserire dati nelle tabelle Parquet.
- Apache Spark su EMR Serverless non supporta l'interrogazione delle viste del catalogo dati.
- Apache Hive su EMR Serverless non supporta l'interrogazione di tabelle con autorizzazioni Lake Formation.
- AWS Glue 5.0 o versioni successive supportano controlli di accesso dettagliati sulle tabelle Iceberg e Hive nel Data Catalog che sono supportate da S3. Questa funzionalità consente di configurare i controlli di accesso a livello di tabella, riga, colonna e cella per le query di lettura all'interno dei job di Apache Spark. AWS Glue

Per ulteriori informazioni, consulta [Versioni di AWS Glue](#).

AWS servizi e tipi di autorizzazione supportati per i formati di tabelle transazionali

AWS servizio	Iceberg	Hudi	Delta Lake (nativo)	Delta Lake (tabelle con collegamenti simbolici)
Athena SQL	Supporta la lettura di tabelle con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura richiedono l'accesso completo alla tabella.	Supporta operazioni di lettura e creazione su tabelle con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.	Athena (versione 3 del motore) supporta la lettura di tabelle native di Delta Lake con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.	Athena (versione 3 del motore) supporta la lettura di tabelle Symlink Delta Lake con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.
Redshift Spectrum su un cluster fornito	Supporta la lettura di tabelle con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.	Supporta la lettura di tabelle con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.	Non sono supportate	Supporta la lettura delle tabelle Delta Lake tramite symlink manifest con autorizzazioni a livello di tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.
Apache Spark su Amazon EMR () EC2	Supporta la lettura di tabelle con autorizzazioni a livello di	Supporta la lettura di tabelle con autorizzazioni a livello di	Supporta la lettura di tabelle con autorizzazioni a livello di	Supporta la lettura di tabelle con autorizzazioni a livello di

AWS servizio	Iceberg	Hudi	Delta Lake (nativo)	Delta Lake (tabelle con collegamenti simbolici)
	tabella, colonna, riga e cella. Le operazioni di scrittura richiedono o l'accesso completo alla tabella.	tabella, colonna, riga e cella. Le operazioni di scrittura richiedono o l'accesso completo alla tabella.	tabella, colonna, riga e cella. Le operazioni di scrittura non sono supportate.	tabella, colonna, riga e cella. Le operazioni di scrittura richiedono o l'accesso completo alla tabella.
AWS Glue ETL	AWS Glue 5.0 o versioni successive supportano la lettura di tabelle con autorizzazioni a livello di tabella, colonna, riga e cella.	Supporta tabelle con autorizzazioni read/write a livello di tabella.	Supporta tabelle con autorizzazioni read/write a livello di tabella.	Supporta tabelle con autorizzazioni read/write a livello di tabella.

Argomenti

- [Utilizzo AWS Lake Formation con Amazon Athena](#)
- [Utilizzo AWS Lake Formation con Amazon Redshift Spectrum](#)
- [AWS Lake Formation Utilizzo con AWS Glue](#)
- [Utilizzo AWS Lake Formation con Amazon EMR](#)
- [Utilizzo AWS Lake Formation con Quick Suite](#)
- [Utilizzo AWS Lake Formation con Lake AWS CloudTrail](#)

Utilizzo AWS Lake Formation con Amazon Athena

[Amazon Athena](#) è un servizio di query senza server che ti aiuta ad analizzare dati strutturati, semistrutturati e non strutturati archiviati in Amazon S3. È possibile utilizzare Athena SQL per interrogare i dati dai formati di dati CSV, JSON, Parquet e Avro. [Athena SQL supporta anche formati di tabella come Apache Hive, ApacheHudi e Apache Iceberg](#). Athena si integra con l' AWS Glue Data Catalog archiviazione dei metadati dei tuoi set di dati in Amazon S3. Athena può utilizzare Lake Formation per definire e mantenere le politiche di controllo degli accessi su tali set di dati.

Ecco alcuni casi d'uso comuni in cui è possibile utilizzare Lake Formation con Athena.

- Utilizza le autorizzazioni di Lake Formation per accedere alle risorse del Data Catalog (database e tabelle) da Athena. È possibile utilizzare il metodo della risorsa denominata o LF-Tags per definire le autorizzazioni su database e tabelle. Per ulteriori informazioni, consulta:
 - [Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#)
 - [Controllo degli accessi basato su tag Lake Formation](#)

Note

Le autorizzazioni di Lake Formation si applicano solo quando si utilizza Athena SQL per interrogare i dati di origine da Amazon S3 e i metadati nel Data Catalog. Athena Spark non supporta l'interrogazione delle tabelle del Data Catalog con autorizzazioni Lake Formation. Le autorizzazioni di Lake Formation supportano operazioni di lettura e scrittura su database e tabelle.

Note

Non puoi applicare filtri di dati quando usi LF-Tags per gestire le autorizzazioni sulle risorse di Data Catalog.

- Controlla i risultati delle query utilizzando [Filtri di dati in Lake Formation](#) per proteggere le tabelle nei tuoi data lake Amazon S3 concedendo autorizzazioni a livello di colonna, riga e cella. Consulta le [limitazioni relative alla proiezione delle partizioni](#) nella Guida per l'utente di Amazon Athena.
- Applica un controllo granulare degli accessi sui dati disponibili per l'utente Athena basato su SAML durante l'esecuzione di query federate.

I driver JDBC e ODBC Athena supportano la configurazione dell'accesso federato all'origine dati utilizzando Identity Provider (IdP) basato su SAML. Usa Quick Suite integrato con Lake Formation con il tuo ruolo IAM esistente o con gli utenti o i gruppi SAML per visualizzare i risultati delle query Athena.

 Note

Le autorizzazioni di Lake Formation per utenti e gruppi SAML verranno applicate solo quando invii query ad Athena utilizzando il driver JDBC o ODBC.

Per ulteriori informazioni, consulta [Utilizzo dei driver Lake Formation e Athena JDBC e ODBC per l'accesso federato](#) ad Athena.

 Note

Attualmente, l'autorizzazione dell'accesso alle identità SAML in Lake Formation non è supportata nelle seguenti regioni:

- Medio Oriente (Bahrein) - me-south-1
- Asia Pacifico (Hong Kong) - ap-east-1
- Africa (Città del Capo) - af-south-1
- Cina (Ningxia) - cn-nordovest-1
- Asia Pacifico (Osaka) - ap-northeast-3

- [Condivisione dei dati tra account in Lake Formation](#) Da utilizzare per interrogare le tabelle in un altro account.

 Note

Per ulteriori informazioni sulle limitazioni relative all'utilizzo delle autorizzazioni di Lake Formation per Views, consulta [Considerazioni e limitazioni](#).

Support per i formati di tabelle transazionali

L'applicazione delle autorizzazioni Lake Formation consente di proteggere i dati transazionali nei data lake basati su Amazon S3. La tabella seguente elenca i formati di tabelle transazionali supportati nelle autorizzazioni Athena e Lake Formation. Lake Formation applica queste autorizzazioni quando gli utenti di Athena eseguono le loro query.

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Athena
Apache Hudi	Un formato utilizzato per semplificare l'elaborazione incrementale dei dati e lo sviluppo di pipeline di dati. Athena supporta operazioni di creazione e lettura utilizzando i formati di tabella Apache Hudi su set di dati Amazon S3 per i tipi di tabella Hudi Copy on Write (CoW) e Merge On Read (MoR). Athena non supporta le operazioni di scrittura sulle tabelle Hudi. Usa Athena per interrogare i set di dati Hudi.	Utilizzalo Filtraggio dei dati e sicurezza a livello di cella in Lake Formation per proteggere e la tabella Hudi utilizzando le autorizzazioni a livello di tabella, colonna, riga e cella.
Apache Iceberg	Un formato di tabella aperta che gestisce grandi raccolte di file sotto forma di tabelle e supporta le moderne operazioni analitiche dei data lake come l'inserimento, l'aggiornamento, l'eliminazione e le query sui viaggi nel tempo a livello di record.	Sono supportate le autorizzazioni a livello di tabella, colonna, riga e cella. Attualmente, Lake Formation non supporta la gestione delle autorizzazioni su operazioni di scrittura come VACUUM UPDATE e OPTIMIZE sulle tabelle in Open Table Formats. MERGE

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Athena
	Per ulteriori informazioni sul supporto di Athena per le tabelle Iceberg, vedere Utilizzo delle tabelle Iceberg.	

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Athena
Linux Foundation Delta Lake	Delta Lake è un progetto open source che aiuta a implementare moderne architetture di data lake comunemente costruite su Amazon S3 o Hadoop Distributed File System (HDFS). Athena supporta le tabelle Delta Lake create utilizzando una definizione di tabella manifest basata su symlink a AWS Glue Data Catalog partendo da una tabella Delta Lake. Per ulteriori informazioni, consulta Scansionare le tabelle Delta Lake utilizzando i crawler. AWS Glue Athena (versione 3 del motore) supporta la lettura delle tabelle native di Delta Lake. Per ulteriori informazioni, consulta Introduzione al supporto tabellare nativo di Delta Lake con i AWS Glue crawler.	Le autorizzazioni a livello di tabella, colonna, riga e cella sono supportate per le tabelle symlink e le tabelle native di Delta Lake.

Risorse aggiuntive

Post di blog, video e workshop

- [Interroga un set di dati Apache Hudi in un data lake Amazon S3 con Amazon Athena](#)
- [Crea un data lake Apache Iceberg utilizzando Amazon Athena, Amazon EMR e AWS Glue](#)
- [Inserisci, aggiorna, elimina su Amazon S3 con Athena e Apache Iceberg](#)
- Workshop Lake Formation sul [controllo degli accessi basato su LF-tag](#) sull'interrogazione di un data lake.

Utilizzo AWS Lake Formation con Amazon Redshift Spectrum

[Amazon Redshift](#) Spectrum ti consente di interrogare e recuperare dati nei data lake Amazon S3 senza caricare dati nei nodi del cluster Amazon Redshift.

Redshift Spectrum supporta due modi per registrare un catalogo di AWS Glue dati esterno abilitato con Lake Formation.

- Utilizzo di un ruolo IAM collegato al cluster con autorizzazione al Data Catalog

Per creare un ruolo IAM, segui i passaggi descritti nella procedura riportata di seguito.

[Controllo dell'accesso a AWS Glue Data Catalog](#)

- Utilizzando un'identità IAM federata configurata per gestire l'accesso a risorse esterne AWS Glue Data Catalog

Redshift Spectrum supporta l'interrogazione delle tabelle di Lake Formation utilizzando identità IAM federate. Le identità IAM possono essere un utente IAM o un ruolo IAM. Per ulteriori informazioni sulla federazione delle identità IAM in Redshift Spectrum, consulta [Usare un'identità federata per gestire l'accesso di Amazon Redshift alle risorse locali e le tabelle esterne di Redshift](#) Spectrum.

Grazie all'integrazione di Lake Formation con Redshift Spectrum, puoi definire le autorizzazioni di controllo degli accessi a livello di riga, colonna e cella sulle tabelle dopo la registrazione dei dati con Lake Formation.

Per ulteriori informazioni, consulta [Usare Redshift Spectrum](#) con. AWS Lake Formation

Redshift Spectrum supporta letture o SELECT interrogazioni sulle tabelle di schemi esterne gestite da Lake Formation.

Per ulteriori informazioni, consulta [Creazione di schemi esterni per Redshift Spectrum](#).

Support per tipi di tabelle transazionali

Questa tabella elenca i formati di tabelle transazionali supportati in Redshift Spectrum e le autorizzazioni Lake Formation applicabili.

Formati di tabella supportati

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Redshift Spectrum
Apache Hudi	Un formato utilizzato per semplificare l'elaborazione incrementale dei dati e lo sviluppo di pipeline di dati. Redshift Spectrum supporta le operazioni di inserimento, eliminazione e storno di scrittura utilizzando il formato di tabella Apache Hudi Copy on Write (CoW) su Amazon S3. Per ulteriori informazioni, consulta Creazione di tabelle esterne per i dati gestiti in Apache Hudi.	Filtraggio dei dati e sicurezza a livello di cella in Lake Formation Utilizzatelo per proteggere le tabelle Hudi utilizzando le autorizzazioni a livello di tabella, colonna, riga e cella.
Apache Iceberg	Un formato di tabella aperta che gestisce grandi raccolte di file sotto forma di tabelle e supporta le moderne operazioni analitiche di data	Redshift Spectrum supporta le tabelle Apache Iceberg per le interrogazioni.

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Redshift Spectrum
	lake come l'inserimento, l'aggiornamento, l'eliminazione e le query sui viaggi nel tempo a livello di record. Per ulteriori informazioni, consulta Usare le tabelle Apache Iceberg con Amazon Redshift.	
Linux Foundation Delta Lake	Delta Lake è un progetto open source che aiuta a implementare moderne architetture di data lake comunemente costruite su Amazon S3 o Hadoop Distributed File System (HDFS). Redshift Spectrum supporta l'interrogazione delle tabelle Delta Lake. Per ulteriori informazioni, consulta Creazione di tabelle esterne per i dati gestiti in Delta Lake.	Sono supportate le autorizzazioni a livello di tabella, colonna, riga e cella.

Risorse aggiuntive

Post di blog e workshop

- [Centralizza la governance per il tuo data lake utilizzando, al AWS Lake Formation contempo, un'architettura di dati moderna con Amazon Redshift Spectrum](#)
- [Usa Redshift Spectrum per interrogare le tabelle Apache HUDI Copy On Write \(CoW\) nel data lake Amazon S3](#)

AWS Lake Formation Utilizzo con AWS Glue

I data engineer e i DevOps professionisti utilizzano AWS Glue Extract, Transform and Load (ETL) con Apache Spark per eseguire trasformazioni sui propri set di dati in Amazon S3 e caricare i dati trasformati in data lake e data warehouse per analisi, apprendimento automatico e sviluppo di applicazioni. Poiché diversi team accedono allo stesso set di dati in Amazon S3, è fondamentale concedere e limitare le autorizzazioni in base ai rispettivi ruoli.

AWS Lake Formation è basato su e AWS Glue i servizi interagiscono nei seguenti modi:

- Lake Formation e AWS Glue condividono lo stesso Data Catalog.
- Le seguenti funzionalità della console di Lake Formation richiamano la AWS Glue console:
 - Lavori — Per ulteriori informazioni, consulta [Aggiungere lavori](#) nella Guida per gli AWS Glue sviluppatori.
 - Crawler — Per ulteriori informazioni, consulta [Cataloging Tables with a Crawler nella Developer Guide](#).AWS Glue
- I flussi di lavoro generati quando si utilizza un blueprint di Lake Formation sono AWS Glue flussi di lavoro. Puoi visualizzare e gestire questi flussi di lavoro sia nella console di Lake Formation che nella AWS Glue console.
- Le trasformazioni di machine learning sono fornite con Lake Formation e si basano su operazioni AWS Glue API. Puoi creare e gestire le trasformazioni dell'apprendimento automatico sulla AWS Glue console. Per ulteriori informazioni, consulta [Machine Learning Transforms](#) nella AWS Glue Developer Guide.

È possibile utilizzare il controllo granulare degli accessi di Lake Formation per gestire le risorse esistenti nel catalogo dati e le posizioni di dati di Amazon S3.

Note

AWS Glue 5.0 o versioni successive supportano controlli di accesso dettagliati sulle tabelle Iceberg e Hive supportate da S3. Questa funzionalità consente di configurare i controlli di accesso a livello di tabella, riga, colonna e cella per le query di lettura all'interno dei job di Apache Spark. AWS Glue

Support per tipi di tabelle transazionali

L'applicazione delle autorizzazioni Lake Formation consente di proteggere i dati transazionali nei data lake basati su Amazon S3. La tabella seguente elenca i formati di tabelle transazionali supportati AWS Glue e le autorizzazioni di Lake Formation. Lake Formation applica queste autorizzazioni per AWS Glue le operazioni.

Formati di tabella supportati

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in AWS Glue
Apache Hudi	Un formato di tabella aperta utilizzato per semplificare l'elaborazione incrementale dei dati e lo sviluppo di pipeline di dati. Per esempi, vedete Using the Hudi framework in AWS Glue	Le autorizzazioni a livello di tabella sono disponibili per le tabelle Hudi. Per ulteriori informazioni, consulta Limitazioni.
Apache Iceberg	Un formato di tabella aperta che gestisce grandi raccolte di file sotto forma di tabelle. Per esempi, vedete Using the Iceberg framework in AWS Glue.	AWS Glue la versione 5.0 e successive consentono di configurare i controlli di accesso a livello di tabella, riga, colonna e cella per le query di lettura all'interno dei job for Apache Spark AWS Glue for Iceberg tables. Per ulteriori informazioni, consulta Limitazioni.
Linux Foundation Delta Lake	Delta Lake è un progetto open source che aiuta a implementare moderne architetture di data lake comunemente costruite su Amazon S3	Le autorizzazioni a livello di tabella sono disponibili per le tabelle Delta Lake. Per ulteriori informazioni, consulta Limitazioni.

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in AWS Glue
	o Hadoop Distributed File System (HDFS). Per esempi, consulta Usare il framework Delta Lake in. AWS Glue	

Risorse aggiuntive

Post e repository del blog

- [Usa il AWS Glue connettore per leggere e scrivere tabelle Apache Iceberg con transazioni ACID ed eseguire viaggi nel tempo](#)
- [Scrittura su tabelle Apache Hudi utilizzando un connettore personalizzato AWS Glue](#)
- AWS repository del [modello Cloudformation e dell'esempio di codice pyspark](#) per analizzare i dati di streaming utilizzando Apache Hudi e AWS Glue Amazon S3.

Utilizzo AWS Lake Formation con Amazon EMR

Amazon EMR è una piattaforma cluster AWS gestita flessibile su cui è possibile eseguire qualsiasi codice personalizzato su framework di big data supportati come Hadoop Map-Reduce, Spark, Hive, Presto, ecc. Le organizzazioni utilizzano Amazon EMR anche per eseguire applicazioni di elaborazione dati in batch e in streaming su un cluster altamente distribuito. Utilizzando Apache Spark su Amazon EMR, puoi eseguire trasformazioni di dati e codice personalizzato su database e tabelle le cui autorizzazioni sono gestite da Lake Formation.

Esistono tre opzioni per la distribuzione di Amazon EMR:

- EMR attivo EC2
- EMR Serverless
- Amazon EMR su EKS

Per ulteriori informazioni, consulta [Integrazione di Amazon EMR con Lake Formation](#) o Utilizzo di [EMR Serverless](#) con per un controllo granulare degli accessi AWS Lake Formation

Support per i formati di tabelle transazionali

Le versioni 6.15.0 e successive di Amazon EMR includono il supporto per le autorizzazioni di controllo degli accessi a livello di tabella, riga, colonna e cella di Lake Formation sui formati di tabella [Apache Hudi](#), [Apache Iceberg](#) e [Delta Lake](#) quando leggi e scrivi dati con Spark SQL.

Per le limitazioni, consulta [Considerazioni per Amazon EMR with Lake Formation](#).

Formati di tabella supportati

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Amazon EMR
Apache Hudi	Un formato di tabella aperta utilizzato per semplificare l'elaborazione incrementale dei dati e lo sviluppo di pipeline di dati. Per un elenco delle operazioni supportate, consulta Apache Hudi e Lake Formation.	Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con Apache Hudi.
Apache Iceberg	Un formato di tabella aperta che gestisce grandi raccolte di file sotto forma di tabelle. Per un elenco delle operazioni supportate, consulta Apache Iceberg e Lake Formation.	Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con Apache Iceberg.
Linux Foundation Delta Lake	Delta Lake è un progetto open source che aiuta a implementare moderne architetture di data lake comunemente costruite su Amazon S3	Amazon EMR supporta il controllo degli accessi a livello di tabella, riga, colonna e cella con le tabelle Delta Lake.

Formato della tabella	Descrizione e operazioni consentite	Autorizzazioni Lake Formation supportate in Amazon EMR
	o Hadoop Distributed File System (HDFS). Per un elenco delle operazioni supportate, consulta Delta Lake and Lake Formation.	

Risorse aggiuntive

Guida per l'utente, post di blog e workshop

- [Integrazione con Amazon EMR tramite Runtime Roles](#)
- [Inizia rapidamente a usare Apache Hudi, Apache Iceberg e Delta Lake con Amazon EMR su EKS](#)
- [Utilizzo di Delta Lake OSS con EMR Serverless](#)

Utilizzo AWS Lake Formation con Quick Suite

Quick Suite supporta l'esplorazione di set di dati gestiti dalle autorizzazioni Lake Formation in Amazon S3 utilizzando Athena.

Sia gli utenti delle edizioni Standard che Enterprise di Quick Suite si integrano con Lake Formation, ma in modo leggermente diverso.

- Edizione Enterprise: concedi autorizzazioni granulari di controllo degli accessi (FGAC) a singoli utenti e gruppi di Quick Suite per accedere a database e tabelle.
- Edizione standard: concedi le autorizzazioni ai ruoli IAM per accedere a database e tabelle.

Note

Per impostazione predefinita, Quick Suite utilizza un ruolo denominato `aws-quicksight-service-role-v0`. Puoi anche definire ruoli personalizzati con le autorizzazioni richieste che consentono a Quick Suite di accedere ad Athena.

Per ulteriori informazioni, consulta [Autorizzazione](#) delle connessioni tramite AWS Lake Formation

Risorse aggiuntive

Post del blog

- [Abilita autorizzazioni dettagliate per gli autori di Quick Suite in AWS Lake Formation](#)
- [Analizza i tuoi dati in modo sicuro con Quick Suite AWS Lake Formation](#)

Utilizzo AWS Lake Formation con Lake AWS CloudTrail

AWS CloudTrail Lake supporta l'esplorazione degli archivi di dati di eventi utilizzando Amazon Athena autorizzazioni granulari in. AWS Lake Formation

Note

CloudTrail Lake può essere solo interrogato. Amazon Athena

Per registrare il tuo archivio dati di eventi CloudTrail Lake con Lake Formation, consulta [Federare un data store di eventi](#).

Registrazione delle chiamate API AWS Lake Formation utilizzando AWS CloudTrail

AWS Lake Formation è integrato con AWS CloudTrail, un servizio che fornisce una registrazione delle azioni intraprese da un utente, ruolo o AWS servizio in Lake Formation. CloudTrail acquisisce tutte le chiamate API Lake Formation come eventi. Le chiamate acquisite includono le chiamate dalla console di Lake Formation e AWS Command Line Interface le chiamate in codice alle azioni dell'API Lake Formation. Se crei un trail, puoi abilitare la distribuzione continua di CloudTrail eventi a un bucket Amazon S3, inclusi gli eventi per Lake Formation. Se non configuri un percorso, puoi comunque visualizzare gli eventi più recenti nella CloudTrail console nella cronologia degli eventi. Utilizzando le informazioni raccolte da CloudTrail, è possibile determinare la richiesta che è stata effettuata a Lake Formation, l'indirizzo IP da cui è stata effettuata la richiesta, chi ha effettuato la richiesta, quando è stata effettuata e dettagli aggiuntivi.

Per ulteriori informazioni CloudTrail, consulta la [Guida AWS CloudTrail per l'utente](#).

Informazioni su Lake Formation in CloudTrail

CloudTrail è abilitato per impostazione predefinita quando si crea un nuovo AWS account. Quando si verifica un'attività in Lake Formation, tale attività viene registrata come CloudTrail evento insieme ad altri eventi AWS di servizio nella cronologia degli eventi. Un evento rappresenta una singola richiesta da un'origine e include informazioni sull'operazione, la data e l'ora dell'operazione e i parametri della richiesta. Inoltre, ogni evento o voce di registro contiene informazioni su chi ha generato la richiesta. Le informazioni di identità consentono di determinare quanto segue:

- Se la richiesta è stata effettuata con credenziali utente root o AWS Identity and Access Management (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 AWS servizio.

Per ulteriori informazioni, consulta [Elemento CloudTrail userIdentity](#).

Puoi visualizzare, cercare e scaricare gli eventi recenti relativi al tuo AWS account. Per ulteriori informazioni, consulta [Visualizzazione degli eventi con la cronologia degli CloudTrail eventi](#).

Per una registrazione continua degli eventi nel tuo AWS account, inclusi gli eventi per Lake Formation, crea un percorso. Un trail consente di CloudTrail inviare file di log a un bucket Amazon S3. Per impostazione predefinita, quando crei un trail nella console, il trail si applica a tutte le regioni AWS. Il trail registra gli eventi di tutte le regioni della AWS partizione e consegna i file di log al bucket Amazon S3 specificato. Inoltre, puoi configurare altri AWS servizi, ad esempio analizzare ulteriormente i Amazon Athena dati sugli eventi raccolti nei log e agire in base a essi. CloudTrail CloudTrail può anche inviare file di registro ad Amazon CloudWatch Logs and CloudWatch Events.

Per ulteriori informazioni, consulta gli argomenti seguenti:

- [Panoramica della creazione di un percorso](#)
- [CloudTrail servizi e integrazioni supportati](#)
- [Configurazione delle notifiche Amazon SNS per CloudTrail](#)
- [Ricezione di file di CloudTrail registro da più regioni](#) e [ricezione di file di CloudTrail registro da più account](#)

Comprendere gli eventi di Lake Formation

Tutte le azioni dell'API Lake Formation vengono registrate CloudTrail e sono documentate nella AWS Lake Formation Developer Guide. Ad esempio, le chiamate alle PutDataLakeSettings RevokePermissions azioni e generano voci nei file di CloudTrail registro. GrantPermissions

L'esempio seguente mostra un CloudTrail evento per l'GrantPermissionsazione. La voce include l'utente che ha concesso l'autorizzazione (datalake_admin), il principale a cui è stata concessa l'autorizzazione (datalake_user1) e l'autorizzazione concessa (CREATE_TABLE). La voce mostra anche che la concessione non è riuscita perché il database di destinazione non è stato specificato nell'resourceargomento.

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "IAMUser",
    "principalId": "AIDAZKE67KM3P775X74U2",
    "arn": "arn:aws:iam::111122223333:user/datalake_admin",
    "accountId": "111122223333",
    "accessKeyId": "...",
    "userName": "datalake_admin"
  },
```

```

    "eventTime": "2021-02-06T00:43:21Z",
    "eventSource": "lakeformation.amazonaws.com",
    "eventName": "GrantPermissions",
    "awsRegion": "us-east-1",
    "sourceIPAddress": "72.21.198.65",
    "userAgent": "aws-cli/1.19.0 Python/3.6.12
Linux/4.9.230-0.1.ac.223.84.332.metal1.x86_64 botocore/1.20.0",
    "errorCode": "InvalidInputException",
    "errorMessage": "Resource must have one of the have either the catalog, table or
database field populated.",
    "requestParameters": {
      "principal": {
        "dataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
      },
      "resource": {},
      "permissions": [
        "CREATE_TABLE"
      ]
    },
    "responseElements": null,
    "requestID": "b85e863f-e75d-4fc0-9ff0-97f943f706e7",
    "eventID": "8d2ccefc0-55f3-42d3-9ede-3a6faedaa5c1",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
    "recipientAccountId": "111122223333"
  }
}

```

L'esempio successivo mostra una voce di CloudTrail registro per l'GetDataAccessazione. I principali non chiamano direttamente questa API. Piuttosto, GetDataAccess viene registrato ogni volta che un AWS servizio principale o integrato richiede credenziali temporanee per accedere ai dati in una posizione di data lake registrata con Lake Formation.

```

{
  "eventVersion": "1.05",
  "userIdentity": {
    "type": "AWSAccount",
    "principalId": "AROAQGFTBBBG0BWV2EMZA:GlueJobRunnerSession",
    "accountId": "111122223333"
  },
  "eventSource": "lakeformation.amazonaws.com",

```

```
    "eventName": "GetDataAccess",  
    ...  
    ...  
    "additionalEventData": {  
        "requesterService": "GLUE_JOB",  
        "lakeFormationPrincipal": "arn:aws:iam::111122223333:role/ETL-Glue-Role",  
        "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"  
    },  
    ...  
}
```

Monitoraggio dell'accesso al catalogo federato di Amazon Redshift

I cataloghi federati di Amazon Redshift nei log di GetDataAccess Lake Formation AWS Glue Data Catalog non generano eventi. CloudTrail Per tracciare l'accesso ai dati per le tabelle Amazon Redshift nel Data Catalog, monitora GetTable ed BatchGetTable events, che acquisiscono i modelli di accesso ai metadati per le tabelle federate di Redshift, tra cui il nome della tabella, il nome del database e le informazioni sull'ID del catalogo.

Vedi anche

- [Registrazione su più account CloudTrail](#)

Le migliori pratiche, considerazioni e limitazioni di Lake Formation

Usa questa sezione per trovare rapidamente le migliori pratiche, le considerazioni e le limitazioni all'interno di AWS Lake Formation.

Consulta le [quote di servizio](#) per il numero massimo di risorse o operazioni di servizio per la tua Account AWS.

Argomenti

- [Buone pratiche e considerazioni sulla condivisione dei dati tra account](#)
- [Limitazioni dei ruoli legati ai servizi](#)
- [Limitazioni di accesso ai dati tra regioni](#)
- [Visualizzazioni, considerazioni e limitazioni di Data Catalog](#)
- [Limitazioni del filtraggio dei dati](#)
- [Considerazioni e limitazioni della modalità di accesso ibrido](#)
- [Limitazioni per l'inserimento dei dati del data warehouse di Amazon Redshift nel AWS Glue Data Catalog](#)
- [Limitazioni dell'integrazione del catalogo S3 Tables](#)
- [I metadati di Hive archiviano: considerazioni e limitazioni sulla condivisione dei dati](#)
- [Limitazioni della condivisione dei dati di Amazon Redshift](#)
- [Limitazioni dell'integrazione di IAM Identity Center](#)
- [Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation](#)
- [Considerazioni, limitazioni e aree supportate sul controllo degli accessi basato sugli attributi](#)

Buone pratiche e considerazioni sulla condivisione dei dati tra account

Le funzionalità cross-account di Lake Formation consentono agli utenti di condividere in modo sicuro i data lake distribuiti tra più AWS organizzazioni o direttamente con i responsabili IAM in un altro account Account AWS, fornendo un accesso granulare ai metadati del Data Catalog e ai dati sottostanti.

Prendi in considerazione le seguenti best practice per l'utilizzo della condivisione di dati tra account di Lake Formation:

- Non c'è limite al numero di concessioni di autorizzazioni di Lake Formation che puoi concedere ai mandanti per tuo conto. AWS Tuttavia, Lake Formation utilizza la capacità AWS Resource Access Manager (AWS RAM) per le sovvenzioni tra account che il tuo account può effettuare con il metodo delle risorse denominato. Per massimizzare la AWS RAM capacità, segui queste migliori pratiche per il metodo delle risorse denominate:
- Utilizza la nuova modalità di concessione tra account (versione 3 e successive nelle impostazioni della versione Cross account) per condividere una risorsa con un esterno. Account AWS Per ulteriori informazioni, consulta [Aggiornamento delle impostazioni della versione di condivisione dei dati tra account](#).
- Suddividi AWS gli account in organizzazioni e concedi le autorizzazioni alle organizzazioni o alle unità organizzative. Una concessione a un'organizzazione o a un'unità organizzativa conta come un'unica concessione.

La concessione a organizzazioni o unità organizzative elimina inoltre la necessità di accettare un AWS Resource Access Manager (AWS RAM) invito alla condivisione di risorse per la sovvenzione. Per ulteriori informazioni, consulta [Accesso e visualizzazione di tabelle e database di Data Catalog condivisi](#).

- Invece di concedere le autorizzazioni su molte singole tabelle di un database, utilizzate la speciale jolly Tutte le tabelle per concedere le autorizzazioni su tutte le tabelle del database. La concessione su tutte le tabelle viene considerata un'unica concessione. Per ulteriori informazioni, consulta [Concessione delle autorizzazioni per le risorse di Data Catalog](#).

Note

Per ulteriori informazioni sulla richiesta di un limite più elevato per il numero di condivisioni di risorse in AWS RAM, vedere le [quote AWS di servizio](#) nel. Riferimenti generali di AWS

- È necessario creare un link di risorsa a un database condiviso affinché tale database appaia negli Amazon Athena editor di query di Amazon Redshift Spectrum. Analogamente, per poter eseguire query su tabelle condivise utilizzando Athena e Redshift Spectrum, è necessario creare collegamenti di risorse alle tabelle. I collegamenti alle risorse vengono quindi visualizzati nell'elenco delle tabelle degli editor di query.

Invece di creare collegamenti alle risorse per molte singole tabelle a scopo di interrogazione, è possibile utilizzare il carattere jolly Tutte le tabelle per concedere le autorizzazioni su tutte le tabelle di un database. Quindi, quando crei un link a una risorsa per quel database e lo selezioni nell'editor di query, avrai accesso a tutte le tabelle del database per la tua query. Per ulteriori informazioni, consulta [Creazione di collegamenti alle risorse](#).

- Quando condividi risorse direttamente con i responsabili di un altro account, il responsabile IAM dell'account destinatario potrebbe non avere l'autorizzazione a creare collegamenti alle risorse per poter interrogare le tabelle condivise utilizzando Athena e Amazon Redshift Spectrum. Invece di creare un link a una risorsa per ogni tabella condivisa, l'amministratore del data lake può creare un database segnaposto e concedere CREATE_TABLE l'autorizzazione al gruppo. ALLIAMPrincipal. Quindi, tutti i responsabili IAM presenti nell'account del destinatario possono creare collegamenti alle risorse nel database placeholder e iniziare a interrogare le tabelle condivise.

Vedi il comando CLI di esempio per concedere le autorizzazioni a in. ALLIAMPrincipal
[Concessione delle autorizzazioni al database utilizzando il metodo della risorsa denominata](#)

- Quando le autorizzazioni per più account vengono concesse direttamente a un principale, solo il destinatario della concessione può visualizzare tali autorizzazioni. L'amministratore del data lake nell' AWS account del destinatario non può visualizzare queste concessioni dirette.
- Athena e Redshift Spectrum supportano il controllo degli accessi a livello di colonna, ma solo per l'inclusione, non per l'esclusione. Il controllo degli accessi a livello di colonna non è supportato nei job ETL. AWS Glue
- Quando una risorsa viene condivisa con il tuo AWS account, puoi concedere le autorizzazioni sulla risorsa solo agli utenti del tuo account. Non puoi concedere le autorizzazioni sulla risorsa ad altri AWS account, alle organizzazioni (nemmeno alla tua organizzazione) o al IAMAllowedPrincipal gruppo.
- Non puoi concedere DROP o Super su un database a un account esterno.
- Revoca le autorizzazioni tra account prima di eliminare un database o una tabella. Altrimenti, è necessario eliminare le condivisioni di risorse orfane in. AWS Resource Access Manager

Consulta anche

- [Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation](#)

- [CREATE_TABLE](#) nella sezione [Riferimento alle autorizzazioni di Lake Formation](#) per ulteriori regole e limitazioni di accesso tra account.

Limitazioni dei ruoli legati ai servizi

Un ruolo collegato ai servizi è un tipo speciale di ruolo IAM a cui è collegato direttamente. AWS Lake Formation Questo ruolo dispone di autorizzazioni predefinite che consentono a Lake Formation di eseguire azioni per tuo conto su tutti AWS i servizi.

Le seguenti limitazioni si applicano quando si utilizza un ruolo collegato ai servizi (SLR) per registrare le posizioni dei dati con Lake Formation.

- Non è possibile modificare le politiche relative ai ruoli collegati ai servizi una volta create.
- Un ruolo collegato al servizio non supporta la condivisione di risorse di catalogo crittografate tra account. Le risorse crittografate richiedono autorizzazioni AWS KMS chiave specifiche. I ruoli collegati ai servizi dispongono di autorizzazioni predefinite che non includono la possibilità di utilizzare risorse di catalogo crittografate su più account.
- Quando si registrano più sedi Amazon S3, l'utilizzo del ruolo collegato al servizio può far sì che si superino rapidamente i limiti delle policy IAM. Ciò accade perché, con i ruoli collegati ai servizi, AWS scrive la policy per te e la incrementa come un unico grande blocco che include tutte le tue registrazioni. Puoi scrivere politiche gestite dai clienti in modo più efficiente, distribuire le autorizzazioni su più politiche o utilizzare ruoli diversi per regioni diverse.
- Amazon EMR on non EC2 può accedere ai dati, registri posizioni dati con ruoli collegati ai servizi.
- Le operazioni relative ai ruoli collegati ai servizi aggirano le politiche di controllo del servizio. AWS
- Quando registri le posizioni dei dati con un ruolo collegato al servizio, aggiorna le policy IAM con una certa coerenza. Per ulteriori informazioni, consulta la documentazione sulla [risoluzione dei problemi di IAM nella Guida per l'utente IAM](#).
- Non puoi configurare `SET_CONTEXT = TRUE` le impostazioni del data lake Lake Formation quando usi ruoli collegati ai servizi e stai utilizzando IAM Identity Center. Il motivo è che i ruoli collegati ai servizi hanno politiche di fiducia immutabili che sono incompatibili con la propagazione dell'identità affidabile necessaria per il controllo con i principali di IAM Identity Center. `SetContext`

Limitazioni di accesso ai dati tra regioni

Lake Formation supporta l'interrogazione su tutte Regioni AWS le tabelle del Data Catalog. Puoi accedere ai dati in una regione da altre regioni utilizzando Amazon Athena Amazon EMR ed AWS Glue ETL creando collegamenti di risorse in altre regioni che puntano ai database e alle tabelle di origine. Con l'accesso alle tabelle tra regioni, puoi accedere ai dati di tutte le regioni senza copiare i dati o i metadati sottostanti nel Data Catalog.

Le seguenti limitazioni si applicano all'accesso alle tabelle tra aree geografiche.

- Lake Formation non supporta l'interrogazione di tabelle Data Catalog da un'altra regione utilizzando Amazon Redshift Spectrum.
- Nella console Lake Formation, le viste del database e della tabella non mostrano i nomi dei database/tabelle Region di origine.
- Per visualizzare l'elenco delle tabelle in un database condiviso di un'altra regione, devi prima creare un collegamento di risorsa al database condiviso, quindi selezionare il collegamento alla risorsa e scegliere Visualizza tabelle.
- Lake Formation non supporta le chiamate ai link di risorse interregionali effettuate dagli utenti SAML.
- La funzionalità interregionale di Lake Formation non comporta costi aggiuntivi per i trasferimenti di dati.

Visualizzazioni, considerazioni e limitazioni di Data Catalog

Le seguenti considerazioni e limitazioni si applicano alle visualizzazioni del catalogo dati.

- Non è possibile creare una vista del Data Catalog dalla console Lake Formation. È possibile creare viste utilizzando AWS CLI o SDK.
- È possibile creare viste del catalogo dati da 10 tabelle. È un limite rigido. Le tabelle di riferimento sottostanti per una vista possono appartenere allo stesso database o a database diversi all'interno dello stesso AWS account.
- Per ulteriori considerazioni e limitazioni specifiche sulla creazione di viste del catalogo dati utilizzando Redshift, consulta [la sezione Considerazioni e limitazioni delle viste del catalogo dati](#) nella Amazon Redshift Database Developer Guide. Per Athena, consulta la sezione [Considerazioni e limitazioni delle visualizzazioni del catalogo dati](#) nella Guida per l'utente di Amazon Athena.

- Puoi creare viste del Data Catalog su tabelle registrate con Lake Formation sia in modalità di accesso ibrido che in modalità Lake Formation.

Quando si utilizzano le viste di Data Catalog con la modalità di accesso ibrida Lake Formation, si consiglia di assicurarsi che i principali che utilizzano la visualizzazione abbiano attivato le autorizzazioni Lake Formation per le tabelle di base a cui si fa riferimento nella vista senza concedere l'accesso. Ciò garantisce che le tabelle di base non vengano rivelate ai consumatori tramite le autorizzazioni IAM. AWS Glue

- Non ci sono restrizioni sulla versione di condivisione tra account per la condivisione delle visualizzazioni.
- Le visualizzazioni vengono modificate in modo analogo alle tabelle di Data Catalog, quando si utilizza l'ALTER VIEWistruzione per un dialetto di visualizzazione già creato. Non è possibile tornare a una vista precedente perché la versione della vista cambia con le modifiche dei dati sottostanti. È possibile eliminare una versione di visualizzazione e per impostazione predefinita verrà utilizzata la prossima versione più recente disponibile. Quando modifichi la versione di visualizzazione, assicurati che i dati siano sincronizzati con lo schema della versione di visualizzazione selezionata.
- Non viene introdotto alcun nuovo Data Catalog APIs . Gli esistenti CreateTable UpdateTable DeleteTable e GetTable APIs vengono aggiornati.
- Amazon Redshift crea sempre viste con colonne varchar da tabelle con stringhe. È necessario trasmettere colonne di stringhe a varchar con una lunghezza esplicita quando si aggiungono dialetti da altri motori.
- La concessione delle autorizzazioni per il data lake all'All tablesinterno di un database comporterà che il beneficiario disponga delle autorizzazioni su tutte le tabelle e le viste all'interno del database.
- Non puoi creare viste:
 - Questo fa riferimento ad altre viste.
 - Quando la tabella di riferimento è un collegamento a una risorsa.
 - Quando la tabella di riferimento si trova in un altro account.
 - Da metastori Hive esterni.
- I ruoli di definizione tra account non sono supportati per le viste Redshift Spectrum Dialect.
- I collegamenti alle risorse per il dialetto Athena nell'editor di query Athena non sono supportati. Per utilizzare i ruoli di definizione tra account per il dialetto Athena, aggiungi l'account che ospita le tabelle di base come origine dati in Athena.

Limitazioni del filtraggio dei dati

Quando concedi le autorizzazioni di Lake Formation su una tabella Data Catalog, puoi includere specifiche di filtraggio dei dati per limitare l'accesso a determinati dati nei risultati delle query e nei motori integrati con Lake Formation. Lake Formation utilizza il filtraggio dei dati per ottenere la sicurezza a livello di colonna, la sicurezza a livello di riga e la sicurezza a livello di cella. Puoi definire e applicare filtri di dati sulle colonne nidificate se i dati di origine contengono strutture nidificate.

Note e restrizioni per il filtraggio a livello di colonna

Esistono tre modi per specificare il filtraggio delle colonne:

- Utilizzando filtri di dati
- Utilizzando un semplice filtraggio a colonne o un filtraggio a colonne annidate.
- Utilizzando. TAGs

Il semplice filtraggio delle colonne specifica solo un elenco di colonne da includere o escludere. Sia la console di Lake Formation che l'API AWS CLI supportano il semplice filtraggio delle colonne. Per vedere un esempio, consulta [Grant with Simple Column Filtering](#).

Le seguenti note e restrizioni si applicano al filtraggio delle colonne:

- AWS Glue 5.0 o versioni successive supportano il controllo granulare degli accessi tramite Lake Formation solo per le tabelle Apache Hive e Apache Iceberg.
- Per concedere SELECT con l'opzione grant e il filtraggio delle colonne, è necessario utilizzare un elenco di inclusione, non un elenco di esclusione. Senza l'opzione di concessione, è possibile utilizzare elenchi di inclusione o esclusione.
- Per concedere SELECT su una tabella con filtro a colonne, devi aver ottenuto l'autorizzazione SELECT sulla tabella con l'opzione di concessione e senza alcuna restrizione di riga. Devi avere accesso a tutte le righe.
- Se SELECT concedi l'opzione di concessione e il filtraggio delle colonne a un principale del tuo account, tale principale deve specificare il filtraggio delle colonne per le stesse colonne o per un sottoinsieme delle colonne concesse quando concede a un altro principale. Se concedi l'opzione SELECT di concessione e il filtraggio delle colonne a un account esterno, l'amministratore del data lake dell'account esterno può concedere SELECT tutte le colonne a un altro responsabile del proprio account. Tuttavia, anche se SELECT su tutte le colonne, tale principale avrà visibilità solo sulle colonne concesse all'account esterno.

- Non è possibile applicare il filtraggio delle colonne sulle chiavi di partizione.
- A un principale con l'`SELECT` autorizzazione su un sottoinsieme di colonne di una tabella non può essere concessa l'`INSERT` autorizzazione `ALTER`, `DROP` `DELETE`, o per quella tabella. Per un'entità principale con l'`INSERT` autorizzazione `ALTER`, `DROP` `DELETE`, o su una tabella, se si concede l'`SELECT` autorizzazione con il filtro delle colonne, l'autorizzazione non ha alcun effetto.

Le seguenti note e restrizioni si applicano al filtraggio delle colonne annidate:

- È possibile includere o escludere cinque livelli di campi nidificati in un filtro dati.

Example

`Col1.Col1_1.Col1_1_1.Col1_1_1_1.Col1_1_1_1_1`

- Non è possibile applicare il filtraggio delle colonne ai campi annidati all'interno delle colonne delle partizioni.
- Se lo schema della tabella contiene un nome di colonna di primo livello («cliente»).» address») che presenta lo stesso schema di rappresentazione di un campo nidificato all'interno di un filtro dati (una colonna nidificata con un nome di colonna di primo livello `customer` e un nome di campo nidificato `address` viene specificata come `"customer"."address"` in un filtro dati), non è possibile specificare in modo esplicito l'accesso alla colonna di primo livello o al campo nidificato perché entrambi sono rappresentati utilizzando lo stesso modello negli elenchi. inclusion/exclusion Questo è ambiguo e Lake Formation non può risolvere se stai specificando la colonna di primo livello o il campo annidato.
- Se una colonna di primo livello o un campo nidificato contiene una doppia virgoletta all'interno del nome, devi includere una seconda virgoletta quando specifichi l'accesso a un campo nidificato all'interno dell'elenco di inclusione ed esclusione di un filtro di celle di dati.

Example

Esempio di nome di colonna nidificata con virgolette doppie: `a.b.double"quote`

Example

Esempio di rappresentazione di una colonna annidata all'interno di un filtro dati:

`"a"."b"."double""quote"`

Limitazioni del filtraggio a livello di cella

Tieni presente le seguenti note e restrizioni per il filtraggio a livello di riga e di cella.

- La sicurezza a livello di cella non è supportata nelle colonne annidate, nelle viste e nei link alle risorse.
- Tutte le espressioni supportate nelle colonne di primo livello sono supportate anche nelle colonne nidificate. Tuttavia, non è necessario fare riferimento ai campi nidificati nelle colonne di partizione quando si definiscono espressioni nidificate a livello di riga.
- La sicurezza a livello di cella è disponibile in tutte le regioni quando si utilizza la versione 3 del motore Athena o Amazon Redshift Spectrum. Per altri servizi, la sicurezza a livello di cella è disponibile solo nelle regioni menzionate in [Regioni supportate](#).
- Le istruzioni `SELECT INTO` non sono supportate.
- I `array` tipi di `map` dati e non sono supportati nelle espressioni di filtro di riga. Il tipo di `struct` dati è supportato.
- Non esiste un limite al numero di filtri di dati che possono essere definiti su una tabella, ma esiste un limite di 100 filtri di dati per un singolo principale su una tabella.
- Per applicare un filtro dati con un'espressione di filtro di riga, è necessario disporre `SELECT` dell'opzione `grant` su tutte le colonne della tabella. Questa restrizione non si applica agli amministratori degli account esterni quando la concessione è stata concessa all'account esterno.
- Se un principale è membro di un gruppo e sia al principale che al gruppo vengono concesse le autorizzazioni su un sottoinsieme di righe, i permessi di riga effettivi del principale sono l'unione dei permessi del principale e dei permessi del gruppo.
- I seguenti nomi di colonna sono limitati in una tabella per il filtraggio a livello di riga e di cella:
 - `ctid`
 - `oid`
 - `xmin`
 - `cmin`
 - `xmax`
 - `cmax`
 - `tableoid`
 - `inserire xid`
 - `elimina xid`

- importoide
- ID univoco per gatti rossi
- Se si applica l'espressione di filtro composta da tutte le righe su una tabella contemporaneamente ad altre espressioni di filtro con predicati, l'espressione composta da tutte le righe prevarrà su tutte le altre espressioni di filtro.
- Quando le autorizzazioni su un sottoinsieme di righe vengono concesse a un AWS account esterno e l'amministratore del data lake dell'account esterno concede tali autorizzazioni a un principale in quell'account, il predicato di filtro effettivo del principale è l'intersezione tra il predicato dell'account e qualsiasi predicato concesso direttamente al principale.

Ad esempio, se l'account dispone delle autorizzazioni di riga con il predicato `dept= 'hr'` e al principale è stata concessa separatamente l'autorizzazione `percountry= 'us'`, l'account principale ha accesso solo alle righe con `dept= 'hr'` e `ecountry= 'us'`.

Per ulteriori informazioni sul filtraggio a livello di cella, vedere. [Filtraggio dei dati e sicurezza a livello di cella in Lake Formation](#)

Per considerazioni e limitazioni relative all'esecuzione di query su tabelle utilizzando Amazon Redshift Spectrum con policy di sicurezza a livello di riga, consulta [Considerazioni e limitazioni sull'uso delle politiche RLS nella Amazon Redshift Database](#) Developer Guide.

Considerazioni e limitazioni della modalità di accesso ibrido

La modalità di accesso ibrido offre la flessibilità necessaria per abilitare selettivamente le autorizzazioni di Lake Formation per database e tabelle del tuo. AWS Glue Data Catalog. Con la modalità di accesso ibrida, ora disponi di un percorso incrementale che ti consente di impostare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti.

Le seguenti considerazioni e limitazioni si applicano alla modalità di accesso ibrida.

Limitazioni

- Aggiorna la registrazione delle sedi Amazon S3: non puoi modificare i parametri di una sede registrata con Lake Formation utilizzando un ruolo collegato al servizio.

- Opzione di attivazione quando si utilizzano i tag LF — Quando è possibile concedere i permessi di Lake Formation utilizzando i tag LF, è possibile attivare i principali per applicare i permessi di Lake Formation in un passaggio consecutivo scegliendo database e tabelle a cui sono allegati i tag LF.
- Accesso in modalità di accesso ibrido: l'accesso alla modalità di accesso ibrida in Lake Formation è limitato agli utenti con autorizzazioni di amministratore del data lake o di amministratore di sola lettura.
- Principi di opt-in: attualmente, solo un ruolo di amministratore del data lake può assegnare i principali alle risorse.
- Attivazione di tutte le tabelle di un database: nelle concessioni tra account, quando si concedono autorizzazioni e in tutte le tabelle di un database, è necessario attivare anche il database affinché le autorizzazioni funzionino.

Considerazioni

- Aggiornamento della posizione Amazon S3 registrata con Lake Formation alla modalità di accesso ibrida: non consigliamo di convertire una posizione dati Amazon S3 già registrata con Lake Formation in modalità di accesso ibrida, sebbene sia possibile farlo.
- Comportamenti delle API quando una posizione dei dati viene registrata in modalità di accesso ibrida
 - CreateTable — La località è considerata registrata presso Lake Formation indipendentemente dal flag della modalità di accesso ibrida e dallo stato di attivazione. Pertanto, l'utente richiede l'autorizzazione alla localizzazione dei dati per creare una tabella.
 - CreatePartition/BatchCreatePartitions/UpdatePartitions (quando la posizione della partizione viene aggiornata in modo che punti alla posizione registrata con hybrid): la posizione Amazon S3 è considerata registrata presso Lake Formation indipendentemente dal flag della modalità di accesso ibrida e dallo stato di attivazione. Pertanto, l'utente richiede l'autorizzazione alla localizzazione dei dati per creare o aggiornare un database.
 - CreateDatabase/UpdateDatabase (quando la posizione del database viene aggiornata in modo che punti alla posizione registrata in modalità di accesso ibrida): la posizione viene considerata registrata con Lake Formation indipendentemente dal flag della modalità di accesso ibrida e dallo stato di attivazione. Pertanto, l'utente richiede l'autorizzazione alla localizzazione dei dati per creare o aggiornare un database.
 - UpdateTable (quando una posizione della tabella viene aggiornata in modo che punti alla posizione registrata in modalità di accesso ibrida): la posizione viene considerata registrata con Lake Formation indipendentemente dal flag della modalità di accesso ibrida e dallo stato di

attivazione. Pertanto, l'utente richiede l'autorizzazione alla localizzazione dei dati per aggiornare la tabella. Se la posizione della tabella non è aggiornata o punta a una posizione non registrata con Lake Formation, l'utente non richiede l'autorizzazione alla posizione dei dati per aggiornare la tabella.

Limitazioni per l'inserimento dei dati del data warehouse di Amazon Redshift nel AWS Glue Data Catalog

Puoi catalogare e gestire l'accesso ai dati analitici nei data warehouse di Amazon Redshift utilizzando il AWS Glue Data Catalog. Si applicano le limitazioni seguenti:

- La condivisione tra account non è supportata a livello di catalogo federato. Tuttavia, puoi condividere singoli database e tabelle dall'interno di un catalogo federato tra s. Account AWS
- È necessario disporre delle impostazioni della versione 4 di Cross account per condividere database o tabelle nel catalogo federato tra Account AWS s.
- Il Data Catalog supporta la creazione solo di cataloghi di primo livello.
- È possibile aggiornare solo la descrizione dei cataloghi in Redshift Managed Storage (RMS).
- L'impostazione delle autorizzazioni per i cataloghi federati, nonché per i database e le tabelle del catalogo federato da raggruppare non è supportata. `IAMAllowedPrincipals`
- Le operazioni DDL (Data Definition Language) sul catalogo da motori come Athena, Amazon EMR Spark o altri, inclusa l'impostazione delle configurazioni del catalogo, non sono supportate.
- L'esecuzione di operazioni DDL su tabelle RMS utilizzando Athena non è supportata.
- La creazione di viste materializzate non è supportata, sia che avvenga tramite Athena, Apache Spark, AWS Glue Data Catalog Amazon Redshift consumer o Amazon Redshift.
- Athena non supporta un'esperienza multicatalogo. Può connettersi solo a un singolo catalogo specifico alla volta. Athena non può accedere o eseguire query su più cataloghi contemporaneamente.
- Le operazioni di etichettatura e ramificazione sulle tabelle Iceberg tramite Athena e Amazon Redshift non sono supportate.
- Time Travel sulle tabelle RMS non è supportato.
- I cataloghi a più livelli con tabelle Data Lake non sono supportati. Tutti i dati archiviati in Amazon S3 per l'uso con le tabelle dei data lake devono risiedere nell'impostazione predefinita AWS Glue Data Catalog e non possono essere organizzati in cataloghi a più livelli.

- In Amazon Redshift, le condivisioni di dati non vengono aggiunte allo spazio dei nomi registrato. Cluster e namespace sono sinonimi, una volta pubblicato un cluster in, non è possibile aggiungere nuovi dati. AWS Glue Data Catalog
- Amazon EMR on EC2 non supporta l'unione tra tabelle RMS e tabelle Amazon S3. Solo EMR Serverless supporta questa funzionalità.
- Gli schemi e le tabelle esterni non sono supportati.
- Le tabelle RMS sono accessibili solo dall'endpoint di estensione nel catalogo REST di AWS Glue Iceberg.
- Le tabelle Hive non sono accessibili dai motori di terze parti collegati al catalogo REST di Iceberg. AWS Glue
- Sarà supportato il livello di isolamento read_commit sulle tabelle RMS tramite Spark.
- I nomi dei database Redshift vengono trattati senza distinzione tra maiuscole e minuscole in AWS Glue Data Catalog, limitati a 128 caratteri e possono essere alfanumerici con trattini (-) e caratteri di sottolineatura (_).
- I nomi dei cataloghi non fanno distinzione tra maiuscole e minuscole, sono limitati a 50 caratteri e possono essere alfanumerici con trattini (-) e caratteri di sottolineatura (_).
- Amazon Redshift non supporta l'utilizzo dei comandi GRANT e REVOKE in stile SQL di Lake Formation per gestire le autorizzazioni di accesso alle tabelle pubblicate su. AWS Glue Data Catalog
- Le policy di sicurezza a livello di riga e di mascheramento dinamico dei dati collegate al cluster Amazon Redshift del produttore (fonte) non verranno applicate. Invece, le autorizzazioni di accesso definite in Lake Formation verranno applicate ai dati condivisi.
- L'esecuzione di operazioni DDL (Data Definition Language) e Data Manipulation Language (DML) sui collegamenti alle tabelle non è supportata.
- Se le parole chiave riservate non vengono evase correttamente, si verificheranno errori o errori.
- La crittografia dei dati in scenari multicatalogo non è supportata.

Limitazioni dell'integrazione del catalogo S3 Tables

Amazon S3 Tables si integra con AWS Glue Data Catalog (Data Catalog) e registra il catalogo come posizione dati di Lake Formation. Puoi configurare questa registrazione dalla console di Lake Formation o utilizzando il servizio APIs.

 Note

Se disponi di una policy basata sulle risorse IAM o S3 Tables che limita gli utenti IAM e i ruoli IAM in base ai tag principali, devi associare gli stessi tag principali al ruolo IAM utilizzato da Lake Formation per accedere ai tuoi dati S3 (ad esempio LakeFormationDataAccessRole) e concedere a questo ruolo le autorizzazioni necessarie. Ciò è necessario affinché la politica di controllo degli accessi basata su tag funzioni correttamente con l'integrazione analitica di S3 Tables.

Le seguenti limitazioni si applicano all'integrazione del catalogo S3 Tables con Data Catalog e Lake Formation:

- AWS Glue e Lake Formation non supportano i nomi di colonna composti da maiuscole e minuscole e convertono tutti i nomi delle colonne in minuscolo. È necessario verificare che i nomi delle colonne della tabella siano univoci quando vengono convertiti in lettere minuscole. Usa `customer_id` invece di `customerId`. L'uso di nomi di colonna composti da maiuscole e minuscole era supportato solo durante la versione di anteprima.
- L' `CreateCatalog` API non può creare bucket di tabelle in Amazon S3.
- L' `SearchTables` API non può cercare nelle tabelle S3.

I metadati di Hive archiviano: considerazioni e limitazioni sulla condivisione dei dati

Con la federazione dei AWS Glue Data Catalog metadati (Data Catalog federation), puoi connettere il Data Catalog a metastore esterni che archiviano i metadati per i tuoi dati Amazon S3 e gestire in modo sicuro le autorizzazioni di accesso ai dati utilizzando AWS Lake Formation.

Le seguenti considerazioni e limitazioni si applicano ai database federati creati dai database Hive:

Considerazioni

- AWS SAM supporto delle applicazioni: sei responsabile della disponibilità delle risorse applicative che AWS SAM distribuisce (Amazon API Gateway e della funzione Lambda). Assicurati che la connessione tra il metastore AWS Glue Data Catalog e Hive funzioni quando gli utenti eseguono le query.

- Requisiti della versione di Hive metastore: puoi creare database federati solo utilizzando Apache Hive versione 3 e successive.
- Requisiti del database mappato: ogni database Hive deve essere mappato su un nuovo database in Lake Formation.
- Supporto federativo a livello di database: puoi connetterti a Hive metastore solo a livello di database.
- Autorizzazioni su database federati: le autorizzazioni applicate a un database federato o alle tabelle in un database federato persistono anche quando viene eliminato un database o una tabella di origine. Quando il database o la tabella di origine vengono ricreati, non è necessario concedere nuovamente le autorizzazioni. Quando una tabella federata con autorizzazioni Lake Formation viene eliminata all'origine, le autorizzazioni di Lake Formation sono ancora visibili e puoi revocarle se necessario.

Se un utente elimina un database federato, tutte le autorizzazioni corrispondenti vengono perse. La ricreazione dello stesso database con lo stesso nome non ripristinerà le autorizzazioni di Lake Formation. Gli utenti dovranno configurare nuovamente le nuove autorizzazioni.

- IAMAllowedAutorizzazioni di gruppo principali sui database federati: in base a `DataLakeSettings`, Lake Formation potrebbe impostare le autorizzazioni per tutti i database e le tabelle a un gruppo virtuale denominato `IAMAllowedPrincipal`. `IAMAllowedPrincipal` si riferisce a tutti i responsabili IAM che hanno accesso alle risorse del Data Catalog tramite le politiche principali e le politiche delle risorse IAM. AWS Glue Se queste autorizzazioni esistono su un database o una tabella, a tutti i principali viene concesso l'accesso al database o alla tabella.

Tuttavia, Lake Formation non consente `IAMAllowedPrincipal` le autorizzazioni sulle tabelle nei database federati. Quando crei database federati, assicurati di passare il `CreateTableDefaultPermissions` parametro come elenco vuoto.

Per ulteriori informazioni, consulta [Modifica delle impostazioni predefinite per il data lake](#).

- Unire tabelle nelle query: puoi unire le tabelle metastore di Hive alle tabelle native di Data Catalog per eseguire le query.

Limitazioni

- Limitazione alla sincronizzazione dei metadati tra il metastore Hive AWS Glue Data Catalog e il metastore Hive: dopo aver stabilito la connessione al metastore Hive, è necessario creare un database federato per sincronizzare i metadati nel metastore Hive con il. AWS Glue Data Catalog

Le tabelle del database federato vengono sincronizzate in fase di esecuzione quando gli utenti eseguono le query.

- Limitazione alla creazione di nuove tabelle in un database federato: non sarà possibile creare nuove tabelle in database federati.
- Limitazione delle autorizzazioni dei dati: il supporto per le autorizzazioni sulle viste delle tabelle dei metastore di Hive non è disponibile.

Limitazioni della condivisione dei dati di Amazon Redshift

AWS Lake Formation consente di gestire in modo sicuro i dati in un datashare di Amazon Redshift. Amazon Redshift è un servizio di data warehouse completamente gestito su scala di petabyte nel cloud. AWS Utilizzando la funzionalità di condivisione dei dati, Amazon Redshift ti aiuta a condividere i dati tra di loro. Account AWS Per ulteriori informazioni sulla condivisione dei dati di Amazon Redshift, consulta [Panoramica sulla condivisione dei dati in Amazon Redshift](#).

Le seguenti note e restrizioni si applicano ai database federati creati da condivisioni di dati Amazon Redshift:

- Requisiti del database mappato: ogni datashare Amazon Redshift deve essere mappato su un nuovo database in Lake Formation. Ciò è necessario per mantenere nomi di tabella univoci quando la rappresentazione degli oggetti datashare è appiattita nel database Data Catalog.
- Limitazione alla creazione di nuove tabelle in un database federato: non sarà possibile creare nuove tabelle in database federati.
- Autorizzazioni sui database federati: le autorizzazioni applicate a un database federato o alle tabelle in un database federato persistono anche quando viene eliminato una tabella di origine o un database. Quando il database o la tabella di origine vengono ricreati, non è necessario concedere nuovamente le autorizzazioni. Quando una tabella federata con autorizzazioni Lake Formation viene eliminata all'origine, le autorizzazioni di Lake Formation saranno ancora visibili e potrai revocarle se necessario.

Se un utente elimina un database federato, tutte le autorizzazioni corrispondenti vengono perse. La ricreazione dello stesso database con lo stesso nome non ripristinerà le autorizzazioni di Lake Formation. Gli utenti dovranno configurare nuovamente le nuove autorizzazioni.

- IAMAllowedAutorizzazioni di gruppo principali sui database federati: in base aDataLakeSettings, Lake Formation potrebbe impostare le autorizzazioni per tutti i database e le tabelle a un gruppo virtuale denominato. IAMAllowedPrincipal IAMAllowedPrincipalSi

riferisce a tutti i responsabili IAM che hanno accesso alle risorse del Data Catalog tramite le politiche principali e le politiche delle risorse IAM. AWS Glue Se queste autorizzazioni esistono su un database o una tabella, a tutti i principali viene concesso l'accesso al database o alla tabella.

Tuttavia, Lake Formation non consente `IAMAllowedPrincipal` le autorizzazioni sulle tabelle nei database federati. Quando crei database federati, assicurati di passare il `CreateTableDefaultPermissions` parametro come elenco vuoto.

Per ulteriori informazioni, consulta [Modifica delle impostazioni predefinite per il data lake](#).

- Filtraggio dei dati: in Lake Formation, puoi concedere autorizzazioni su una tabella in un database federato con filtri a livello di colonna e di riga. Tuttavia, non è possibile combinare il filtraggio a livello di colonna e a livello di riga per limitare l'accesso con granularità a livello di cella alle tabelle all'interno di database federati.
- Identificatore di distinzione tra maiuscole e minuscole: gli oggetti datashare di Amazon Redshift gestiti da Lake Formation supporteranno i nomi delle tabelle e delle colonne solo in lettere minuscole. Non attivare l'identificatore di distinzione tra maiuscole e minuscole per database, tabelle e colonne nelle condivisioni di dati Amazon Redshift, se verranno condivise e gestite utilizzando Lake Formation.
- Supporto per le query: puoi eseguire query sulle condivisioni di dati Amazon Redshift gestite da Lake Formation con Amazon Redshift. Athena non supporta l'interrogazione di datashare Amazon Redshift gestite da Lake Formation.

Per ulteriori informazioni sulle limitazioni quando si lavora con le condivisioni di dati in Amazon Redshift, [consulta Limitazioni per la condivisione dei dati](#) nella Amazon Redshift Database Developer Guide.

Limitazioni dell'integrazione di IAM Identity Center

Con AWS IAM Identity Center, puoi connetterti ai provider di identità (IdPs) e gestire centralmente l'accesso per utenti e gruppi attraverso i servizi di AWS analisi. Puoi configurarlo AWS Lake Formation come applicazione abilitata in IAM Identity Center e gli amministratori del data lake possono concedere autorizzazioni granulari a utenti e gruppi autorizzati sulle risorse. AWS Glue Data Catalog

Le seguenti limitazioni si applicano all'integrazione di Lake Formation con IAM Identity Center:

- Non puoi assegnare utenti e gruppi di IAM Identity Center come amministratori di data lake o amministratori di sola lettura in Lake Formation.

Gli utenti e i gruppi di IAM Identity Center possono interrogare le risorse crittografate del Data Catalog se utilizzi un ruolo IAM che AWS Glue può assumere per tuo conto per la crittografia e la decrittografia del Data Catalog. AWS le chiavi gestite non supportano la propagazione affidabile delle identità.

- Gli utenti e i gruppi di IAM Identity Center possono richiamare solo le operazioni API elencate nella `AWSIAMIdentityCenterAllowListForIdentityContext` policy fornita da IAM Identity Center.
- Lake Formation consente ai ruoli IAM di account esterni di agire come portatori per conto degli utenti e dei gruppi di IAM Identity Center per l'accesso alle risorse del Data Catalog, ma le autorizzazioni possono essere concesse solo sulle risorse del Data Catalog all'interno dell'account proprietario. Se provi a concedere le autorizzazioni agli utenti e ai gruppi di IAM Identity Center sulle risorse Data Catalog in un account esterno, Lake Formation genera il seguente errore: «Le sovvenzioni tra account non sono supportate per il principale».
- Quando si utilizza Lake Formation con IAM Identity Center, la configurazione di assegnazione dell'applicazione è impostata come `false` impostazione predefinita. Se modifichi questa configurazione direttamente tramite l'[API IAM Identity Center](#), devi quindi gestire tutte le assegnazioni delle applicazioni manualmente utilizzando l'API. Lake Formation non sincronizza o gestisce automaticamente le modifiche alle assegnazioni effettuate al di fuori dei suoi flussi di lavoro standard, il che può influire sui modelli di accesso e sui flussi di autorizzazione all'interno dell'ambiente di data lake.

Buone pratiche e considerazioni per il controllo degli accessi basato su tag Lake Formation

È possibile creare, gestire e assegnare LF-tags per controllare l'accesso ai database, alle tabelle e alle colonne del Data Catalog.

Considerate le seguenti best practice quando utilizzate il controllo degli accessi basato su tag Lake Formation:

- Tutti i tag LF devono essere predefiniti prima di poter essere assegnati alle risorse del Data Catalog o concessi ai responsabili.

L'amministratore del data lake può delegare le attività di gestione dei tag creando creatori di tag LF con le autorizzazioni IAM richieste. Gli ingegneri e gli analisti dei dati decidono le caratteristiche e le relazioni dei tag LF. I creatori di LF-tag creano quindi e mantengono i tag LF in Lake Formation.

- È possibile assegnare più LF-tag alle risorse del Data Catalog. È possibile assegnare un solo valore per una particolare chiave a una particolare risorsa.

Ad esempio, è possibile assegnare `module=Orders, region=Westdivision=Consumer`, e così via a un database, una tabella o una colonna. Non puoi `module=Orders, Customers` assegnare.

- Non è possibile assegnare tag LF alle risorse quando si crea la risorsa. È possibile aggiungere tag LF solo alle risorse esistenti.
- È possibile concedere espressioni LF-Tag, non solo singoli tag LF, a un principale.

Un'espressione LF-Tag ha un aspetto simile alla seguente (in pseudo-codice).

```
module=sales AND division=(consumer OR commercial)
```

Un principale a cui è concessa questa espressione LF-Tag può accedere solo alle risorse del Data Catalog (database, tabelle e colonne) assegnate e a una delle due opzioni. `module=sales division=consumer division=commercial` Se desideri che il principale sia in grado di accedere a risorse che dispongono `module=sales` o meno `division=commercial` di includerle entrambe nella stessa concessione. Fai due sovvenzioni, una per `module=sales` e una per `division=commercial`.

L'espressione più semplice del tag LF è costituita da un solo tag LF, ad esempio. `module=sales`

- Un principale a cui sono concesse le autorizzazioni per un tag LF con più valori può accedere alle risorse del Data Catalog con uno di questi valori. Ad esempio, se a un utente viene concesso un tag LF con `key= module` e `values=orders, customers`, l'utente ha accesso alle risorse assegnate o. `module=orders module=customers`
- È necessario disporre dell'`Grant with LF-Tag expressions` autorizzazione per concedere le autorizzazioni relative ai dati sulle risorse del Catalogo dati utilizzando il metodo LF-TBAC. L'amministratore del data lake e il creatore di LF-Tag ricevono implicitamente questa autorizzazione. Un principale che dispone dell'`Grant with LFTag expressions` autorizzazione può concedere le autorizzazioni relative ai dati sulle risorse utilizzando:
 - il metodo di risorsa denominato
 - il metodo LF-TBAC, ma utilizzando solo la stessa espressione LF-TAG

Ad esempio, supponiamo che l'amministratore del data lake conceda la seguente concessione (in pseudo-codice).

```
GRANT (SELECT ON TABLES) ON TAGS module=customers, region=west,south TO user1 WITH  
GRANT OPTION
```

In questo caso, `user1` può concedere `SELECT` tabelle ad altri principali utilizzando il metodo `LF-TBAC`, ma solo con l'espressione `LF-TAG` completa. `module=customers, region=west,south`

- Se a un principale vengono concesse le autorizzazioni su una risorsa sia con il metodo `LF-TBAC` che con il metodo della risorsa denominata, i permessi che il principale ha sulla risorsa sono l'unione dei permessi concessi con entrambi i metodi.
- Lake Formation supporta la concessione `DESCRIBE` e `ASSOCIATE` l'eliminazione dei tag `LF` tra gli account e la concessione di autorizzazioni sulle risorse del Data Catalog tra gli account utilizzando il metodo `LF-TBAC`. In entrambi i casi, il principale è l'`ID` dell'account. AWS

Note

Lake Formation supporta sovvenzioni tra account a organizzazioni e unità organizzative utilizzando il metodo `LF-TBAC`. Per utilizzare questa funzionalità, è necessario aggiornare le impostazioni della versione di Cross account alla versione 3.

Per ulteriori informazioni, consulta [Condivisione dei dati tra account in Lake Formation](#).

- Le risorse del Data Catalog create in un account possono essere etichettate solo utilizzando i tag `LF` creati nello stesso account. I tag `LF` creati in un account non possono essere associati a risorse condivise di un altro account.
- L'utilizzo del controllo degli accessi basato su tag Lake Formation (`LF-TBAC`) per concedere l'accesso tra account alle risorse del Data Catalog richiede aggiunte alla politica delle risorse del Data Catalog per il tuo account. AWS Per ulteriori informazioni, consulta [Prerequisiti](#).
- Le chiavi `LF-Tag` e i valori `LF-Tag` non possono superare i 50 caratteri di lunghezza.
- Il numero massimo di tag `LF` che possono essere assegnati a una risorsa del Data Catalog è 50.
- I seguenti limiti sono limiti flessibili:
 - Il numero massimo di tag `LF` che è possibile creare è 1000.

- Il numero massimo di valori che possono essere definiti per un LF-tag è 1000.
- Le chiavi e i valori dei tag vengono convertiti in lettere minuscole quando vengono memorizzati.
- È possibile assegnare un solo valore per un tag LF a una particolare risorsa.
- Se vengono concessi più LF-tag a un principale con un'unica concessione, il principale può accedere solo alle risorse del Data Catalog che contengono tutti i tag LF.
- Se la valutazione di un'espressione LF-Tag comporta l'accesso solo a un sottoinsieme di colonne della tabella, ma l'autorizzazione Lake Formation concessa in caso di corrispondenza è una delle autorizzazioni che richiedevano l'accesso completo alla colonna, vale a dire, `Alter` `Drop` `InsertDelete`, allora nessuna di queste autorizzazioni viene concessa. Invece, viene concesso solo `Describe`. Se l'autorizzazione concessa è `All (Super)`, allora solo `Select` e `Describe` vengono concesse.
- I wildcard non vengono utilizzati con i tag LF. Per assegnare un tag LF a tutte le colonne di una tabella, si assegna il tag LF alla tabella e tutte le colonne della tabella ereditano il tag LF. Per assegnare un tag LF a tutte le tabelle di un database, si assegna il tag LF al database e tutte le tabelle del database ereditano quel tag LF.
- È possibile creare fino a 1000 espressioni di tag LF in un account.
- È possibile utilizzare fino a 50 espressioni di tag LF per concedere autorizzazioni a un principale sulle risorse del Data Catalog.
- Quando si concedono o si revocano le autorizzazioni su un'espressione di tag LF in linea, la dimensione dell'espressione del tag LF non può superare i 900 byte. Per concedere i permessi su espressioni di tag LF più grandi, utilizzate le espressioni di tag LF salvate. Per ulteriori informazioni, consulta [Creazione di espressioni con tag LF](#).
- Per aggiungere LF-Tag ai cataloghi federati Redshift esistenti creati prima della versione di disponibilità generale del supporto LF-Tag per i cataloghi federati, è necessario contattare il team di supporto per ricevere assistenza. AWS

Considerazioni, limitazioni e aree supportate sul controllo degli accessi basato sugli attributi

Le seguenti considerazioni e limitazioni si applicano al controllo degli accessi basato sugli attributi (ABAC).

- ABAC non supporta la concessione dell'accesso tramite le politiche LF-tag.
- Le autorizzazioni concesse non sono disponibili con ABAC.

- ABAC non supporta la concessione di autorizzazioni agli utenti di IAM Identity Center.
- Quando si utilizzano le sovvenzioni ABAC su una tabella in Lake Formation, Lake Formation non concede DESCRIBE autorizzazioni al database o al catalogo principale. Ciò differisce dagli scenari non ABAC, in cui Lake Formation fornisce DESCRIBE autorizzazioni implicite alle risorse principali.
- Tutti i principali con la chiave AmazonDataZoneProject tag vengono sempre considerati come se fossero stati inseriti in Lake Formation per tutte le risorse del Data Catalog.
- ABAC supporta solo gli attributi di stringa.

Risoluzione dei problemi relativi a Lake Formation

Se riscontri problemi quando lavori con AWS Lake Formation, consulta gli argomenti di questa sezione.

Argomenti

- [Risoluzione dei problemi generali](#)
- [Risoluzione dei problemi di accesso tra account](#)
- [Risoluzione dei problemi relativi a blueprint e flussi di lavoro](#)
- [Problemi noti per AWS Lake Formation](#)
- [Messaggio di errore aggiornato](#)

Risoluzione dei problemi generali

Usa le informazioni qui per aiutarti a diagnosticare e risolvere vari problemi di Lake Formation.

Errore: autorizzazioni Lake Formation insufficienti su <Amazon S3 location>

È stato effettuato un tentativo di creare o modificare una risorsa Data Catalog senza autorizzazioni di localizzazione dei dati sulla posizione Amazon S3 indicata dalla risorsa.

Se un database o una tabella Data Catalog punta a una posizione Amazon S3, quando concedi le autorizzazioni a Lake Formation CREATE_TABLE oppure ALTER devi anche concedere l'AUTHORIZAZIONE_DATA_LOCATION_ACCESS autorizzazione sulla posizione. Se concedi queste autorizzazioni ad account esterni o a organizzazioni, devi includere l'opzione di concessione.

Dopo aver concesso queste autorizzazioni a un account esterno, l'amministratore del data lake di quell'account deve quindi concedere le autorizzazioni ai principali (utenti o ruoli) dell'account. Quando concedi l'AUTHORIZAZIONE_DATA_LOCATION_ACCESS autorizzazione ricevuta da un altro account, devi specificare l>ID di catalogo (ID account) dell'AWS account proprietario. L'account del proprietario è l'account che ha registrato la sede.

Per ulteriori informazioni, consulta [Controllo sottostante dell'accesso ai dati](#) e [Concessione delle autorizzazioni per la localizzazione dei dati](#).

Errore: «Autorizzazioni per la chiave di crittografia insufficienti per l'API Glue»

È stato fatto un tentativo di concedere autorizzazioni a Lake Formation senza autorizzazioni AWS Identity and Access Management (IAM) sulla chiave di AWS KMS crittografia per un Data Catalog crittografato.

La mia query Amazon Athena o quella di Amazon Redshift che utilizza i manifesti non riesce

Lake Formation non supporta le query che utilizzano i manifesti.

Errore: «Autorizzazioni di Lake Formation insufficienti: è richiesta la creazione di un tag sul catalogo»

L'utente/ruolo deve essere un amministratore del data lake.

Errore durante l'eliminazione di amministratori di data lake non validi

È necessario eliminare contemporaneamente tutti gli amministratori di data lake non validi (ruoli IAM eliminati definiti come amministratori di data lake). Se si tenta di eliminare separatamente gli amministratori di data lake non validi, Lake Formation genera un errore principale non valido.

Risoluzione dei problemi di accesso tra account

Usa le informazioni qui per aiutarti a diagnosticare e risolvere i problemi di accesso su più account.

Argomenti

- [Ho concesso un'autorizzazione Lake Formation per più account ma il destinatario non può vedere la risorsa](#)
- [I responsabili dell'account del destinatario possono vedere la risorsa Data Catalog ma non possono accedere ai dati sottostanti](#)
- [Errore: «Associazione non riuscita perché il chiamante non era autorizzato» quando si accetta un AWS RAM invito alla condivisione delle risorse](#)
- [Errore: «Non autorizzato a concedere le autorizzazioni per la risorsa»](#)

- [Errore: «Accesso negato per recuperare le informazioni AWS sull'organizzazione»](#)
- [Errore: «Organizzazione <organization-ID>non trovata»](#)
- [Errore: «Autorizzazioni Lake Formation insufficienti: combinazione non valida»](#)
- [ConcurrentModificationException sulle richieste di concessione/revoca ad account esterni](#)
- [Errore durante l'utilizzo di Amazon EMR per accedere ai dati condivisi tramite più account](#)

Ho concesso un'autorizzazione Lake Formation per più account ma il destinatario non può vedere la risorsa

- L'utente dell'account del destinatario è un amministratore di Data Lake? Solo gli amministratori del data lake possono vedere la risorsa al momento della condivisione.
- Stai condividendo con un account esterno alla tua organizzazione utilizzando il metodo della risorsa denominata? In tal caso, l'amministratore del data lake dell'account destinatario deve accettare un invito alla condivisione delle risorse in AWS Resource Access Manager (AWS RAM).

Per ulteriori informazioni, consulta [the section called “Accettazione di un invito alla condivisione di AWS RAM risorse”](#).

- Stai utilizzando politiche relative alle risorse a livello di account (Data Catalog) in AWS Glue? In caso affermativo, se si utilizza il metodo delle risorse denominate, è necessario includere una dichiarazione speciale nella politica che autorizzi AWS RAM a condividere le politiche per conto dell'utente.

Per ulteriori informazioni, consulta [the section called “Gestione delle autorizzazioni tra account utilizzando entrambi AWS Glue e Lake Formation”](#).

- Disponi delle autorizzazioni AWS Identity and Access Management (IAM) necessarie per concedere l'accesso a più account?

Per ulteriori informazioni, consulta [the section called “Prerequisiti”](#).

- La risorsa a cui hai concesso le autorizzazioni non deve avere alcuna autorizzazione Lake Formation concessa al IAMAllowedPrincipals gruppo.
- C'è una deny dichiarazione sulla risorsa nella politica a livello di account?

I responsabili dell'account del destinatario possono vedere la risorsa Data Catalog ma non possono accedere ai dati sottostanti

I principali dell'account del destinatario devono disporre delle autorizzazioni richieste AWS Identity and Access Management (IAM). Per informazioni dettagliate, consultare [Accesso ai dati sottostanti di una tabella condivisa](#).

Errore: «Associazione non riuscita perché il chiamante non era autorizzato» quando si accetta un AWS RAM invito alla condivisione delle risorse

Dopo aver concesso l'accesso a una risorsa a un altro account, quando l'account ricevente tenta di accettare l'invito alla condivisione delle risorse, l'azione fallisce.

```
$ aws ram get-resource-share-associations --association-type PRINCIPAL --resource-
share-arns arn:aws:ram:aws-region:444444444444:resource-share/e1d1f4ba-xxxx-xxxx-xxxx-
xxxxxxxx5d8d
{
  "resourceShareAssociations": [
    {
      "resourceShareArn": "arn:aws:ram:aws-region:444444444444:resource-share/
e1d1f4ba-xxxx-xxxx-xxxx-xxxxxxxx5d8d
",
      "resourceShareName": "LakeFormation-MMCC0XQBH3Y",
      "associatedEntity": "5815803XXXXX",
      "associationType": "PRINCIPAL",
      "status": "FAILED",
      "statusMessage": "Association failed because the caller was not
authorized.",
      "creationTime": "2021-07-12T02:20:10.267000+00:00",
      "lastUpdatedTime": "2021-07-12T02:20:51.830000+00:00",
      "external": true
    }
  ]
}
```

L'errore si verifica perché `glue:PutResourcePolicy` viene richiamato da AWS Glue quando l'account ricevente accetta l'invito alla condivisione delle risorse. Per risolvere il problema, consenti l'`glue:PutResourcePolicy` intervento del ruolo assunto dall'account produttore/concedente.

Errore: «Non autorizzato a concedere le autorizzazioni per la risorsa»

È stato effettuato un tentativo di concedere autorizzazioni per più account su un database o una tabella di proprietà di un altro account. Quando un database o una tabella vengono condivisi con il tuo account, in qualità di amministratore del data lake, puoi concedere le autorizzazioni solo agli utenti del tuo account.

Errore: «Accesso negato per recuperare le informazioni AWS sull'organizzazione»

Il tuo account è un account di gestione di AWS Organizations e non disponi delle autorizzazioni necessarie per recuperare le informazioni sull'organizzazione, come le unità organizzative nell'account.

Per ulteriori informazioni, consulta [Required permissions for cross-account grants](#).

Errore: «Organizzazione <organization-ID>non trovata»

È stato effettuato un tentativo di condividere una risorsa con un'organizzazione, ma la condivisione con le organizzazioni non è abilitata. Abilita la condivisione delle risorse con le organizzazioni.

Per ulteriori informazioni, consulta [Enable Sharing with AWS Organizations](#) nella Guida AWS RAM per l'utente.

Errore: «Autorizzazioni Lake Formation insufficienti: combinazione non valida»

Un utente ha condiviso una risorsa Data Catalog mentre al IAMAllowedPrincipals gruppo venivano concesse le autorizzazioni di Lake Formation per la risorsa. L'utente deve revocare tutte le autorizzazioni di Lake Formation IAMAllowedPrincipals prima di condividere la risorsa.

ConcurrentModificationException sulle richieste di concessione/revoca ad account esterni

Quando gli utenti effettuano più richieste di sovvenzione simultanee. and/or revoke permission requests for a principal on LF-Tag policies, then Lake Formation throws ConcurrentModificationException. Users need to catch the exception and retry the failed the grant/

revoke Utilizza versioni batch delle operazioni [GrantPermissions/RevokePermissionsAPI](#) [BatchGrantPermissionse](#) [BatchRevokePermissions](#) allevia questo problema in una certa misura riducendo il numero di richieste di concessione/revoca simultanee.

Errore durante l'utilizzo di Amazon EMR per accedere ai dati condivisi tramite più account

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

Risoluzione dei problemi relativi a blueprint e flussi di lavoro

Utilizza le informazioni riportate qui per aiutarti a diagnosticare e risolvere i problemi relativi ai blueprint e al flusso di lavoro.

Argomenti

- [<role-ARN>Il mio progetto non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:»](#)
- [Il mio flusso di lavoro non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:<role-ARN>»](#)
- [Un crawler del mio flusso di lavoro non è riuscito con il messaggio «La risorsa non esiste o il richiedente non è autorizzato ad accedere alle autorizzazioni richieste»](#)
- [Un crawler del mio flusso di lavoro non è riuscito con il messaggio «Si è verificato un errore \(AccessDeniedException\) durante la chiamata dell' CreateTable operazione...»](#)

<role-ARN>Il mio progetto non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:»

È stato effettuato un tentativo di creare un blueprint da parte di un utente che non dispone di autorizzazioni sufficienti per assegnare il ruolo scelto.

Aggiorna la policy IAM dell'utente per poter passare il ruolo o chiedi all'utente di scegliere un ruolo diverso con le autorizzazioni passrole richieste.

Per ulteriori informazioni, consulta [the section called “Riferimento ai personaggi di Lake Formation e alle autorizzazioni IAM”](#).

Il mio flusso di lavoro non è riuscito con «User: <user-ARN>is not authorized to perform: iam: PassRole on resource:<role-ARN>»

Il ruolo che hai specificato per il flusso di lavoro non aveva una politica in linea che consentisse al ruolo di passare da solo.

Per ulteriori informazioni, consulta [the section called “\(Facoltativo\) Crea un ruolo IAM per i flussi di lavoro”](#).

Un crawler del mio flusso di lavoro non è riuscito con il messaggio «La risorsa non esiste o il richiedente non è autorizzato ad accedere alle autorizzazioni richieste»

Una possibile causa è che il ruolo passato non disponeva di autorizzazioni sufficienti per creare una tabella nel database di destinazione. Concedi al ruolo l'CREATE_TABLE autorizzazione per il database.

Un crawler del mio flusso di lavoro non è riuscito con il messaggio «Si è verificato un errore (AccessDeniedException) durante la chiamata dell' CreateTable operazione...»

Una possibile causa è che il ruolo del workflow non disponeva delle autorizzazioni per la localizzazione dei dati sulla posizione di archiviazione di destinazione. Concedi le autorizzazioni per la localizzazione dei dati al ruolo.

Per ulteriori informazioni, consulta [the section called “DATA_LOCATION_ACCESS”](#).

Problemi noti per AWS Lake Formation

Consulta questi problemi noti per AWS Lake Formation.

Argomenti

- [Limitazione al filtraggio dei metadati delle tabelle](#)
- [Problema con la ridenominazione di una colonna esclusa](#)

- [Problema con l'eliminazione delle colonne nelle tabelle CSV](#)
- [Le partizioni delle tabelle devono essere aggiunte in un percorso comune](#)
- [Problema con la creazione di un database durante la creazione del flusso di lavoro](#)
- [Problema con l'eliminazione e la successiva creazione di un utente](#)
- [Le operazioni dell'API Data Catalog non aggiornano il valore del parametro `IsRegisteredWithLakeFormation`](#)
- [Le operazioni di Lake Formation non supportano AWS Glue Schema Registry](#)

Limitazione al filtraggio dei metadati delle tabelle

AWS Lake Formation le autorizzazioni a livello di colonna possono essere utilizzate per limitare l'accesso a colonne specifiche di una tabella. Quando un utente recupera i metadati sulla tabella utilizzando la console o un'API come `glue:GetTable`, l'elenco delle colonne nell'oggetto tabella contiene solo i campi a cui ha accesso. È importante comprendere i limiti di questo filtraggio dei metadati.

Sebbene Lake Formation renda disponibili i metadati sulle autorizzazioni delle colonne ai servizi integrati, l'effettivo filtraggio delle colonne nelle risposte alle query è responsabilità del servizio integrato. I client di Lake Formation che supportano il filtraggio a livello di colonna, tra cui Amazon Athena, Amazon Redshift Spectrum e Amazon EMR, filtrano i dati in base alle autorizzazioni delle colonne registrate con Lake Formation. Gli utenti non saranno in grado di leggere dati a cui non dovrebbero avere accesso. Attualmente, AWS Glue ETL non supporta il filtraggio delle colonne.

Note

I cluster EMR non sono gestiti completamente da AWS. Pertanto, è responsabilità degli amministratori EMR proteggere adeguatamente i cluster per evitare l'accesso non autorizzato ai dati.

Alcune applicazioni o formati potrebbero memorizzare metadati aggiuntivi, inclusi nomi e tipi di colonne, nella mappa come proprietà della `Parameters` tabella. Queste proprietà vengono restituite non modificate e sono accessibili da qualsiasi utente con `SELECT` autorizzazione su qualsiasi colonna.

Ad esempio, [Avro SerDe](#) memorizza una rappresentazione JSON dello schema della tabella in una proprietà della tabella denominata `avro.schema.literal`, disponibile per tutti gli utenti con

accesso alla tabella. Si consiglia di evitare di memorizzare informazioni riservate nelle proprietà delle tabelle e di tenere presente che gli utenti possono apprendere lo schema completo delle tabelle in formato Avro. Questa limitazione è specifica per i metadati relativi a una tabella.

AWS Lake Formation rimuove qualsiasi proprietà della tabella a partire da `spark.sql.sources.schema` quando risponde a una richiesta `glue:GetTable` o a una richiesta simile se il chiamante non dispone `SELECT` delle autorizzazioni su tutte le colonne della tabella. Ciò impedisce agli utenti di accedere a metadati aggiuntivi sulle tabelle create con Apache Spark. Se eseguite su Amazon EMR, le applicazioni Apache Spark possono ancora leggere queste tabelle, ma alcune ottimizzazioni potrebbero non essere applicate e i nomi delle colonne con distinzione tra maiuscole e minuscole non sono supportati. Se l'utente ha accesso a tutte le colonne della tabella, Lake Formation restituisce la tabella non modificata con tutte le proprietà della tabella.

Problema con la ridenominazione di una colonna esclusa

Se si utilizzano le autorizzazioni a livello di colonna per escludere una colonna e quindi rinominarla, la colonna non viene più esclusa dalle query, ad esempio. `SELECT *`

Problema con l'eliminazione delle colonne nelle tabelle CSV

Se crei una tabella del catalogo dati in formato CSV e poi elimini una colonna dallo schema, le query potrebbero restituire dati errati e le autorizzazioni a livello di colonna potrebbero non essere rispettate.

Soluzione alternativa: create invece una nuova tabella.

Le partizioni delle tabelle devono essere aggiunte in un percorso comune

Lake Formation si aspetta che tutte le partizioni di una tabella si trovino in un percorso comune impostato nel campo della posizione della tabella. Quando si utilizza il crawler per aggiungere partizioni a un catalogo, ciò funziona perfettamente. Tuttavia, se si aggiungono partizioni manualmente e queste partizioni non si trovano nella posizione impostata nella tabella principale, l'accesso ai dati non funziona.

Problema con la creazione di un database durante la creazione del flusso di lavoro

Quando crei un flusso di lavoro da un blueprint utilizzando la console Lake Formation, puoi creare il database di destinazione se non esiste. Quando lo fai, l'utente che ha effettuato l'accesso ottiene

L'CREATE_TABLE autorizzazione sul database che viene creato. Tuttavia, il crawler generato dal flusso di lavoro assume il ruolo del flusso di lavoro quando tenta di creare una tabella. Questa operazione non riesce perché il ruolo non dispone dell'CREATE_TABLE autorizzazione per il database.

Soluzione alternativa: se si crea il database tramite la console durante la configurazione del flusso di lavoro, prima di eseguire il flusso di lavoro, è necessario concedere al ruolo associato al flusso di lavoro l'CREATE_TABLE autorizzazione sul database appena creato.

Problema con l'eliminazione e la successiva creazione di un utente

Lo scenario seguente genera permessi errati per Lake Formation restituiti da:
`lakeformation:ListPermissions`

1. Crea un utente e concedi le autorizzazioni di Lake Formation.
2. Eliminare l'utente.
3. Ricrea l'utente con lo stesso nome.

`ListPermissions` restituisce due voci, una per il vecchio utente e una per il nuovo utente. Se si tenta di revocare le autorizzazioni concesse al vecchio utente, le autorizzazioni vengono revocate al nuovo utente.

Le operazioni dell'API Data Catalog non aggiornano il valore del parametro **IsRegisteredWithLakeFormation**

È noto che le operazioni dell'API Data Catalog, ad esempio `GetTables` e `SearchTables` non, aggiornano il valore del `IsRegisteredWithLakeFormation` parametro e restituiscono il valore predefinito, che è falso. Si consiglia di utilizzare l'`GetTableAPI` per visualizzare il valore corretto del `IsRegisteredWithLakeFormation` parametro.

Le operazioni di Lake Formation non supportano AWS Glue Schema Registry

Le operazioni di Lake Formation non supportano le AWS Glue tabelle che contengono un `SchemaReference StorageDescriptor` da utilizzare nello [Schema Registry](#).

Messaggio di errore aggiornato

AWS Lake Formation ha aggiornato le eccezioni specifiche delle risorse al messaggio di `EntityNotFound` errore generale per le seguenti operazioni API per soddisfare gli obiettivi di sicurezza e conformità.

- `RevokePermissions`
- `GrantPermissions`
- `GetResourceLFTags`
- `GetTable`
- `GetDatabase`

AWS Lake Formation API

Note

Il [riferimento API](#) aggiornato per il AWS Lake Formation servizio è ora disponibile.

Indice

- [Autorizzazioni APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Impostazioni del data lake APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Integrazione con IAM Identity Center APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Modalità di accesso ibrida APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Vendita di credenziali APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Etichettatura APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Filtro dati APIs](#)
 - [Operazioni](#)
 - [Tipi di dati](#)
- [Tipi di dati comuni](#)
 - [ErrorDetail struttura](#)

- [Modelli di stringa](#)

Autorizzazioni APIs

La sezione Permissions API descrive le operazioni e i tipi di dati necessari per la concessione e la revoca delle autorizzazioni in AWS Lake Formation. Consulta la [Guida di riferimento dell'API Lake Formation](#) per tutte le operazioni e i tipi di dati delle AWS Lake Formation API.

Operazioni

- [GrantPermissions](#)
- [RevokePermissions](#)
- [BatchGrantPermissions](#)
- [BatchRevokePermissions](#)
- [GetEffectivePermissionsForPath](#)
- [ListPermissions](#)
- [GetDataLakePrincipal](#)

Tipi di dati

- [Risorsa](#)
- [DatabaseResource](#)
- [TableResource](#)
- [TableWithColumnsResource](#)
- [DataCellsFilterResource](#)
- [DataLocationResource](#)
- [DataLakePrincipal](#)
- [PrincipalPermissions](#)
- [PrincipalResourcePermissions](#)
- [DetailsMap](#)
- [ColumnWildcard](#)
- [BatchPermissionsRequestEntry](#)

- [BatchPermissionsFailureEntry](#)

Impostazioni del data lake APIs

Questa sezione contiene le operazioni dell'API Data Lake Settings e i tipi di dati per la gestione degli amministratori del data lake.

Operazioni

- [GetDataLakeSettings](#)
- [PutDataLakeSettings](#)

Tipi di dati

- [DataLakeSettings](#)

Integrazione con IAM Identity Center APIs

Questa sezione contiene le operazioni per creare e gestire l'integrazione di Lake Formation con IAM Identity Center.

Operazioni

- [CreateLakeFormationIdentityCenterConfiguration](#)
- [DeleteLakeFormationIdentityCenterConfiguration](#)
- [DescribeLakeFormationIdentityCenterConfiguration](#)
- [UpdateLakeFormationIdentityCenterConfiguration](#)

Tipi di dati

- [ExternalFilteringConfiguration](#)

Modalità di accesso ibrida APIs

La sezione API della modalità di accesso ibrida descrive le operazioni e i tipi di dati necessari per configurare la modalità di accesso ibrida in AWS Lake Formation. Consulta la [Guida di riferimento dell'API Lake Formation](#) per tutte le operazioni e i tipi di dati delle AWS Lake Formation API.

Operazioni

- [CreateLakeFormationOptIn](#)
- [DeleteLakeFormationOptIn](#)
- [ListLakeFormationOptIns](#)

Tipi di dati

- [Risorsa](#)
- [DatabaseResource](#)
- [TableResource](#)
- [Informazioni sulle risorse](#)
- [LakeFormationOptInsInfo](#)
- [DataLocationResource](#)

Vendita di credenziali APIs

La sezione Credential Vending API descrive le operazioni e i tipi di dati relativi all'utilizzo del AWS Lake Formation servizio per la vendita di credenziali e la registrazione e la gestione di una risorsa data lake.

Operazioni

- [RegisterResource](#)
- [DeregisterResource](#)
- [ListResources](#)
- [GetUnfilteredTableMetadata](#)
- [GetUnfilteredPartitionsMetadata](#)

- [GetTemporaryGluePartitionCredentials](#)
- [GetTemporaryGlueTableCredentials](#)
- [UpdateResource](#)

Tipi di dati

- [FilterCondition](#)
- [RowFilter](#)
- [ResourceInfo](#)

Etichettatura APIs

La sezione Tagging API descrive le operazioni e i tipi di dati relativi a una strategia di autorizzazione che definisce un modello di autorizzazioni sugli attributi o sui tag di coppia chiave-valore.

Operazioni

- [LFTagGet Expression](#)
- [Elenca LFTag le espressioni](#)
- [Elimina LFTag espressione](#)
- [Aggiorna LFTag espressione](#)
- [Crea LFTag espressione](#)
- [AddLFTagsToResource](#)
- [RemoveLFTagsFromResource](#)
- [GetResourceLFTags](#)
- [ElencoLFTags](#)
- [CreaLFTag](#)
- [Ottenere LFTag](#)
- [AggiornaLFTag](#)
- [EliminaLFTag](#)
- [SearchTablesByLFTags](#)

- [SearchDatabasesByLFTags](#)

Tipi di dati

- [LFTagKeyResource](#)
- [LFTagPolicyResource](#)
- [TaggedTable](#)
- [TaggedDatabase](#)
- [LFTag](#)
- [LFTagCoppia](#)
- [LFTagErrore](#)
- [ColonnaLFTag](#)

Filtro dati APIs

Il filtro dati APIs descrive come gestire i filtri delle celle di dati in AWS Lake Formation.

Operazioni

- [CreateDataCellsFilter](#)
- [DeleteDataCellsFilter](#)
- [ListDataCellsFilter](#)
- [GetDataCellsFilter](#)
- [UpdateDataCellsFilter](#)

Tipi di dati

- [DataCellsFilter](#)
- [RowFilter](#)

Tipi di dati comuni

Common Data Types descrive i vari tipi di dati comuni in AWS Lake Formation.

ErrorDetail struttura

Contiene dettagli su un errore.

Campi

- **ErrorCode:** stringa UTF-8, non inferiore a 1 o superiore a 255 byte di lunghezza, corrispondente a [Single-line string pattern](#).

Il codice associato a questo errore.

- **ErrorMessage:** stringa di descrizione, non superiore a 2048 byte di lunghezza, corrispondente a [URI address multi-line string pattern](#).

Messaggio che descrive l'errore.

Modelli di stringa

L'API usa le seguenti espressioni regolari per definire i contenuti validi per vari membri e parametri di stringa:

- Modello di stringa a una riga: "[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*"
- IndirizzoModello di stringa a più righe per indirizzo URI: "[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*"
- Schema di stringhe personalizzato #3 — "^w+\.w+\.w+\$»
- Schema di stringhe personalizzato #4 — "^w+\.w+\$»
- Schema di stringhe personalizzato #5 — "arn:aws:iam::[0-9]*:role/.*)"»
- Schema di stringhe personalizzato #6 — "arn:aws:iam::[0-9]*:user/.*)"»
- Schema di stringhe personalizzato #7 — "arn:aws:iam::[0-9]*:group/.*)"»
- Pattern di stringa personalizzato n. 8 — "arn:aws:iam::[0-9]*:saml-provider/.*)"»
- Pattern di stringa personalizzato n. 9 — "^([\p{L}\p{Z}\p{N}_.:/=+~@%]*)\$"
- Pattern di stringa personalizzato n. 10: "^([\p{L}\p{Z}\p{N}_.:~@%]*)\$"
- Pattern di stringa personalizzato n. 11: "[\p{L}\p{N}\p{P}]*"

Regioni supportate

Questa sezione contiene informazioni sul supporto Regioni AWS e sulle funzionalità di Lake Formation.

Disponibilità generale

Per i servizi Regioni AWS supportati da AWS Lake Formation, consulta [Elenco dei AWS servizi disponibili per regione](#).

Per un elenco degli endpoint del servizio Lake Formation per ogni regione e delle quote del servizio Lake Formation, vedi [AWS Lake Formation endpoint](#) e quote.

AWS GovCloud (US)

[Per una panoramica delle differenze tra AWS GovCloud \(US\) Region e standard Regioni AWS, vedi How differences for. AWS Lake FormationAWS GovCloud \(US\)](#)

Ottimizzazione delle transazioni e dello storage

Le tabelle gestite, il supporto delle transazioni e le funzionalità di ottimizzazione dello storage per Lake Formation sono disponibili nelle seguenti Regioni AWS versioni:

Nome della Regione	Parametro della regione	Endpoint
Stati Uniti orientali (Virginia settentrionale)	us-east-1	lakeformation.us-east-1.amazonaws.com
		lakeformation-fips.us-east-1.amazonaws.com
Stati Uniti orientali (Ohio)	us-east-2	lakeformation.us-east-2.amazonaws.com
		lakeformation-fips.us-east-2.amazonaws.com

Nome della Regione	Parametro della regione	Endpoint
Stati Uniti occidentali (Oregon)	us-west-2	lakeformation.us-west-2.amazonaws.com lakeformation-fips.us-west-2.amazonaws.com
Asia Pacifico (Mumbai)	ap-south-1	lakeformation.ap-south-1.amazonaws.com
Asia Pacific (Seoul)	ap-northeast-2	lakeformation.ap-northeast-2.amazonaws.com
Asia Pacifico (Singapore)	ap-southeast-1	lakeformation.ap-southeast-1.amazonaws.com
Asia Pacifico (Sydney)	ap-southeast-2	lakeformation.ap-southeast-2.amazonaws.com
Asia Pacifico (Tokyo)	ap-northeast-1	lakeformation.ap-northeast-1.amazonaws.com
Europa (Francoforte)	eu-central-1	lakeformation.eu-central-1.amazonaws.com
Europa (Irlanda)	eu-west-1	lakeformation.eu-west-1.amazonaws.com
Europa (Londra)	eu-west-2	lakeformation.eu-west-2.amazonaws.com
Europa (Stoccolma)	eu-north-1	lakeformation.eu-north-1.amazonaws.com

Nome della Regione	Parametro della regione	Endpoint
Canada (Centrale)	ca-central-1	lakeformation.ca-central-1.amazonaws.com
Sud America (San Paolo)	sa-east-1	lakeformation.sa-east-1.amazonaws.com

Cronologia dei documenti per AWS Lake Formation

La tabella seguente descrive le modifiche importanti alla documentazione per AWS Lake Formation

Modifica	Descrizione	Data
Documentazione aggiornata per il controllo degli accessi basato sugli attributi	Lake Formation consente di concedere autorizzazioni ai principali con attributi corrispondenti sulle risorse del Data Catalog. Per ulteriori informazioni, consulta Controllo degli accessi basato sugli attributi .	17 aprile 2025
Integrazione di Amazon S3 Tables con e AWS Lake FormationAWS Glue Data Catalog	Ora puoi integrare e catalogare le tabelle S3 come AWS Glue Data Catalog oggetti e registrare il catalogo come posizione dati di Lake Formation. Per ulteriori informazioni, consulta Creazione di un catalogo Amazon S3 Tables nel . AWS Glue Data Catalog	13 marzo 2025
AWSLakeFormationCrossAccountManager Politica aggiornata di Lake Formation.	Lake Formation ha migliorato la AWSLakeFormationCrossAccountManager politica sostituendo l'operatore delle StringLike condizioni con l'ArnLikeoperatore che consente a IAM di eseguire il controllo del formato ARN.	25 gennaio 2025

Modifica aggiornata della politica	Ha documentato la modifica alla AWSLakeFormationDataAdmin politica.	3 dicembre 2024
Aggiornamenti multicatalogo	AWS Glue Data Catalog ti consente di creare cataloghi federati e unificare i dati tra i data lake Amazon S3 e i data warehouse Amazon Redshift, nonché di integrare i dati dei tuoi database operativi come Amazon DynamoDB e fonti di dati di terze parti come Snowflake, MySQL e così via. Per ulteriori AWS Glue Data Catalog informazioni , consulta Bringing your data into the.	3 dicembre 2024
Documentazione aggiornata per le espressioni dei tag LF	È possibile salvare le espressioni LF-Tag e riutilizzarle per concedere autorizzazioni sulle risorse del Data Catalog. Per ulteriori informazioni, vedete Gestione delle espressioni dei tag LF.	7 novembre 2024
Documentazione aggiornata per le visualizzazioni del catalogo dati	Puoi creare viste in uso oltre AWS Glue APIs a AWS Glue Data Catalog DDLs utilizzare e Amazon Athena e Amazon Redshift. Per ulteriori informazioni, consulta Building Data Catalog views.	7 agosto 2024

[È stata aggiunta documentazione per la vendita di credenziali verificabili](#)

Lake Formation ti consente di includere il contesto dell'utente di IAM Identity Center negli CloudTrail eventi e quindi di tenere traccia degli utenti che accedono alle tue risorse. Per ulteriori informazioni, consulta [Inclusione del contesto utente di IAM Identity Center nei CloudTrail log](#).

14 luglio 2024

[Modifica aggiornata della politica](#)

[È stata documentata la modifica \(aggiunta di una dichiarazione IDs e rimosse le autorizzazioni ridondanti\) alle politiche and. AWSLakeFormationCrossAccountManager AWSLakeFormationDataAdmin](#)

14 marzo 2024

[Configurazione aggiornata di Lake Formation](#)

Sono stati aggiornati i passaggi nella AWS Lake Formation sezione [Configurazione](#).

7 febbraio 2024

[Modifica aggiornata della politica](#)

Sono state aggiunte nuove autorizzazioni alla politica in linea del ruolo collegato al servizio. Per ulteriori informazioni, consulta [Using service-linked roles for Lake Formation](#).

7 febbraio 2024

[Modifica aggiornata della politica](#)

Ha documentato la modifica alla [LakeFormationDataAccessServiceRolePolicy](#) politica.

2 febbraio 2024

[Limitazioni di Consolidated Lake Formation](#)

Creata una sezione unificata per i limiti e le considerazioni di Lake Formation. Per ulteriori informazioni, consulta le [limitazioni di Lake Formation](#).

15 dicembre 2023

[Documentazione aggiunta per la compattazione di Iceberg](#)

Per migliorare le prestazioni di lettura, servizi di AWS analisi come Athena e Amazon EMR e processi AWS Glue ETL, AWS Glue Data Catalog offre la compattazione gestita (un processo che compatta piccoli oggetti Amazon S3 in oggetti più grandi) per le tabelle Iceberg nel Data Catalog. Per ulteriori informazioni, consulta [Ottimizzazione delle tabelle Iceberg](#).

25 novembre 2023

[Documentazione aggiunta per l'integrazione con IAM Identity Center](#)

Le integrazioni di IAM Identity Center consentono a utenti e gruppi di accedere alle risorse del Data Catalog applicando le autorizzazioni di Lake Formation. Per ulteriori informazioni, consulta l'integrazione con [IAM Identity Center](#).

25 novembre 2023

[È stata aggiunta documentazione per le visualizzazioni del catalogo dati](#)

Puoi creare viste AWS Glue Data Catalog che fanno riferimento a un massimo di 10 tabelle utilizzando gli editor SQL per Amazon Athena o Amazon Redshift. Per ulteriori informazioni, consulta [Creazione](#) di viste.

25 novembre 2023

[Aggiornata la modifica alla politica](#)

Ha documentato la modifica alla [AWSLakeFormationCrossAccountManager](#) politica.

25 ottobre 2023

[Documentazione aggiunta per la modalità di accesso ibrido](#)

La modalità di accesso ibrido offre la flessibilità necessari a per abilitare selettivamente le autorizzazioni di Lake Formation per database e tabelle del tuo. AWS Glue Data Catalog Con la modalità di accesso ibrido, ora disponi di un percorso incrementale che ti consente di impostare le autorizzazioni di Lake Formation per un set specifico di utenti senza interrompere le politiche di autorizzazione di altri utenti o carichi di lavoro esistenti. [Per ulteriori informazioni, consulta Modalità di accesso ibrida.](#)

26 settembre 2023

[È stata aggiunta documentazione per la creazione di tabelle Apache Iceberg](#)

Ora puoi creare tabelle Apache Iceberg che utilizzano il formato di dati Apache Parquet AWS Glue Data Catalog con dati che risiedono in Amazon S3. [Per ulteriori informazioni, consulta Creazione di tabelle Iceberg.](#)

16 agosto 2023

[È stata aggiunta la documentazione per l'accesso ai dati tra regioni](#)

Lake Formation supporta l'interrogazione delle tabelle del Data Catalog tra AWS le regioni. Puoi accedere ai dati in una regione da altre regioni utilizzando Athena, Amazon EMR ed eseguire AWS Glue ETL creando collegamenti di risorse in altre regioni che puntano ai database e alle tabelle di origine. Puoi connettere il Data Catalog a metastore esterni che archiviano i metadati per i tuoi dati Amazon S3 e gestire in modo sicuro le autorizzazioni di accesso ai dati utilizzando. AWS Lake Formation [Per ulteriori informazioni, consulta Accesso alle tabelle tra regioni.](#)

30 giugno 2023

[Contenuti riorganizzati](#)

Capitoli riorganizzati della guida per adattarli al percorso utente di Lake Formation.

15 maggio 2023

[Documentazione aggiunta per la federazione HMS](#)

Puoi connettere il Data Catalog a metastore esterni che archiviano i metadati per i tuoi dati Amazon S3 e gestire in modo sicuro le autorizzazioni di accesso ai dati utilizzando. AWS Lake Formation Per ulteriori informazioni, consulta [Gestione delle autorizzazioni](#) sui set di dati che utilizzano metastore esterni.

15 aprile 2023

[Documentazione aggiunta per la condivisione dei dati di Amazon Redshift](#)

Ora puoi gestire in modo sicuro i dati in un datashare di Amazon Redshift utilizzando le autorizzazioni di Lake Formation. Lake Formation supporta la concessione di licenze di accesso ai tuoi dati tramite AWS Data Exchange. Per ulteriori informazioni, consulta [Condivisione dei dati in AWS Lake Formation](#).

30 novembre 2022

[Support per la condivisione dei dati tra account direttamente con i mandanti](#)

Sono state aggiunte informazioni sulla condivisione diretta dei dati con i responsabili IAM in un altro account. Per ulteriori informazioni, consulta [Condivisione dei dati tra account](#) in. AWS Lake Formation

10 novembre 2022

Support per la condivisione AWS RAM abilitata dei dati tramite TBAC	Sono state aggiunte informazioni sul metodo LF-TBAC per la concessione delle autorizzazioni di Data Catalog da utilizzare per le concessioni tra account. AWS Resource Access Manager	10 novembre 2022
È stata aggiunta una sezione sull'utilizzo di altri servizi	Sono state aggiunte informazioni su come AWS servizi come Athena, AWS Glue Redshift Spectrum e Amazon EMR possono utilizzare Lake Formation per accedere in modo sicuro ai dati nelle sedi Amazon S3 registrati e con Lake Formation. Per ulteriori informazioni, consulta Collaborazione con altri AWS servizi .	10 novembre 2022
???	Sono state aggiunte informazioni sulla risoluzione di un errore durante l'utilizzo di Amazon EMR per accedere ai dati di più account. Per ulteriori informazioni, consulta Errore durante l'utilizzo di Amazon EMR per accedere ai dati condivisi tramite più account .	7 novembre 2022

[Aggiornamenti alla condivisione di risorse tra account](#)

È stata aggiunta una descrizione del funzionamento [delle condivisioni di risorse tra account](#) in Lake Formation. Ha documentato la modifica alla [AWSLakeFormationCrossAccountManager](#) politica.

6 maggio 2022

[Nuovi tutorial](#)

Sono stati aggiunti nuovi tutorial per la creazione di tabelle gestite, la protezione dei data lake e la condivisione dei data lake. Per ulteriori dettagli, consulta la sezione Guida [introduttiva](#).

20 aprile 2022

[Pagina iniziale di New Lake Formation](#)

È stata aggiornata la landing page di [Lake Formation](#) per includere collegamenti a tutorial che forniscono step-by-step istruzioni su come creare un data lake, importare dati, condividere e proteggere i data lake utilizzando Lake Formation.

20 aprile 2022

[Support per la vendita di credenziali](#)

Sono state aggiunte informazioni sulla vendita di credenziali, che supporta Lake Formation per consentire ai servizi di terze parti di integrarsi con Lake Formation utilizzando le operazioni API di vendita delle credenziali. Per ulteriori informazioni, consulta [Come funziona il vending di credenziali in Lake Formation](#).

28 febbraio 2022

[Support per tabelle gestite e filtraggio avanzato dei dati](#)

Sono state aggiunte informazioni sulle tabelle governate, che supportano transazioni ACID, compattazione automatica dei dati e query con viaggi nel tempo. Sono state aggiunte informazioni sulla creazione di filtri di dati per supportare la sicurezza a livello di colonna, la sicurezza a livello di riga e la sicurezza a livello di cella. Per ulteriori informazioni, consulta [Governed Tables in Lake Formation](#) and [Data Filtering and Cell-Level Security in Lake Formation](#).

30 novembre 2021

[Support per endpoint di interfaccia VPC](#)

Sono state aggiunte informazioni sulla creazione di un endpoint di interfaccia cloud privato virtuale (VPC) per Lake Formation, in modo che la comunicazione tra il tuo VPC e Lake Formation avvenga in modo completo e sicuro all'interno della rete. AWS Per ulteriori informazioni, consulta [Using Lake Formation with VPC Endpoints](#).

11 ottobre 2021

[Supporto per le policy di endpoint VPC](#)

Sono state aggiunte informazioni sul supporto per le policy degli endpoint Virtual Private Cloud (VPC) in Lake Formation. Per ulteriori informazioni, consulta [Using Lake Formation with VPC Endpoints](#).

11 ottobre 2021

[Support per il controllo degli accessi basato su tag](#)

Il controllo degli accessi basato su tag Lake Formation offre un nuovo modo più scalabile per gestire l'accesso alle risorse del Data Catalog e ai dati sottostanti utilizzando i tag LF. Per ulteriori informazioni, consulta [Lake Formation Tag-Based Access Control](#).

7 maggio 2021

[Nuovo requisito di opt-in per il filtraggio dei dati su Amazon EMR.](#)

Sono state aggiunte informazioni sul requisito di attivazione per consentire ad Amazon EMR di filtrare i dati gestiti da Lake Formation. Per ulteriori informazioni, consulta [Consenti il filtraggio dei dati su Amazon EMR](#).

9 ottobre 2020

[Support per la concessione di autorizzazioni complete per più account sui database Data Catalog](#)

Sono state aggiunte informazioni sulla concessione delle autorizzazioni complete di Lake Formation sui database Data Catalog tra AWS gli account, tra cui. CREATE_TABLE Per ulteriori informazioni, consulta [Sharing Data Catalog Databases](#).

1 ottobre 2020

[Support per l'autenticazione Amazon Athena degli utenti tramite SAML.](#)

Sono state aggiunte informazioni sul supporto per gli utenti Athena che si connettono o tramite il driver JDBC o ODBC e si autenticano tramite provider di identità SAML come Okta e Microsoft Active Directory Federation Service (ADFS). Per ulteriori informazioni, consulta [AWS Service Integrations with Lake Formation.](#)

30 settembre 2020

[Support per l'accesso su più account con un catalogo dati crittografato](#)

Sono state aggiunte informazioni sulla concessione di autorizzazioni per più account quando il Data Catalog è crittografato. Per ulteriori informazioni, consulta [Prerequisiti per l'accesso su più account.](#)

30 luglio 2020

[Support per l'accesso al data lake da più account](#)

Sono state aggiunte informazioni sulla concessione AWS Lake Formation delle autorizzazioni sui database e sulle tabelle di Data Catalog ad AWS account e organizzazioni esterni e sull'accesso agli oggetti Data Catalog condivisi da account esterni. Per ulteriori informazioni, vedere [Cross-Account Access.](#)

7 luglio 2020

[Integrazione con Amazon Quick Suite](#)

Sono state aggiunte informazioni su come concedere le autorizzazioni di Lake Formation agli utenti di Quick Suite Enterprise Edition in modo che possano accedere ai set di dati che risiedono nelle sedi registrate di Amazon S3. Per ulteriori informazioni, consulta [Concessione](#) delle autorizzazioni del catalogo dati.

29 giugno 2020

[Aggiornamenti alla configurazione e ai capitoli introduttivi](#)

Sono stati riorganizzati e migliorati i capitoli Configurazione e Guida introduttiva. Sono state aggiornate le autorizzazioni consigliate AWS Identity and Access Management (IAM) per l'amministratore del data lake.

27 febbraio 2020

[Support per AWS Key Management Service](#)

Sono state aggiunte informazioni su come il supporto di Lake Formation per AWS Key Management Service (AWS KMS) semplifica la configurazione di servizi integrati per leggere e scrivere dati crittografati nelle sedi registrate di Amazon Simple Storage Service (Amazon S3). Sono state aggiunte informazioni su come registrare posizioni Amazon S3 crittografate con AWS KMS keys Per ulteriori informazioni, consulta [the section called “Aggiungere una posizione Amazon S3 al tuo data lake”](#).

27 febbraio 2020

[Aggiornamenti ai blueprint e alle politiche IAM degli amministratori del data lake](#)

Parametri di input chiariti per i blueprint di database incrementali. Sono state aggiornate le politiche IAM richieste per un amministratore di data lake.

20 dicembre 2019

[Riscrittura e aggiornamento del capitolo sulla sicurezza, revisioni del capitolo](#)

Sono stati migliorati i capitoli relativi alla sicurezza e all'aggiornamento.

29 ottobre 2019

[L'autorizzazione Super sostituisce tutte le autorizzazioni](#)

Sono stati aggiornati i capitoli Sicurezza e Aggiornamento in modo da riflettere la sostituzione dell'autorizzazione con All Super

10 ottobre 2019

[Aggiunte, correzioni e chiarimenti](#)

Sono state apportate aggiunte, correzioni e chiarimenti in base al feedback. È stato rivisto il capitolo sulla sicurezza. Sono stati aggiornati i capitoli Sicurezza e Aggiornamento in modo da rispecchiare la sostituzione del gruppo con. Everyone IAMAllowedPrincipals

11 settembre 2019

[Nuova guida](#)

Questa è la versione iniziale della Guida per gli sviluppatori di AWS Lake Formation .

8 agosto 2019

AWS Glossario

Per la AWS terminologia più recente, consultate il [AWS glossario](#) nella sezione Reference. Glossario AWS

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.