aws

Panduan Developer

AWS Lake Formation

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS Lake Formation: Panduan Developer

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Merek dagang dan tampilan dagang Amazon tidak boleh digunakan sehubungan dengan produk atau layanan apa pun yang bukan milik Amazon, dengan cara apa pun yang dapat menyebabkan kebingungan di antara pelanggan, atau dengan cara apa pun yang merendahkan atau mendiskreditkan Amazon. Semua merek dagang lain yang tidak dimiliki oleh Amazon merupakan hak milik masing-masing pemiliknya, yang mungkin atau mungkin tidak terafiliasi, terkait dengan, atau disponsori oleh Amazon.

Table of Contents

Apa itu AWS Lake Formation? 1
Fitur Lake Formation 2
Konsumsi dan manajemen data 2
Manajemen keamanan 3
Membawa data Anda ke dalam Katalog Data 4
Cara kerjanya 6
Alur kerja manajemen izin Lake Formation 6
Izin metadata 8
Manajemen akses penyimpanan 11
Cross-account berbagi data di Lake Formation 13
Komponen Lake Formation 14
Konsol Lake Formation 14
Lake Formation API dan Antarmuka Baris Perintah 14
AWS Layanan lainnya 14
Terminologi Lake Formation 15
Danau data 15
Akses data 15
Mode akses hibrid 15
Cetak biru 15
Alur kerja 16
Katalog Data 16
Data yang mendasari 16
Utama 16
Administrator danau data 17
AWS integrasi layanan dengan Lake Formation 17
Sumber daya Lake Formation tambahan 19
Blog 19
Pembicaraan teknologi dan webinar 20
Arsitektur modern 20
Sumber daya data mesh 20
Panduan praktik terbaik 20
Memulai dengan Lake Formation 20
Memulai 22
Selesaikan inisial AWS tugas konfigurasi 22

Mendaftar untuk Akun AWS ... 22
Penyiapan AWS Lake Formation ... 22
Menyiapkan sumber daya Formasi Danau menggunakan CloudFormation templat ... 23
Buat administrator danau data ... 24
Ubah model izin default atau gunakan mode akses hybrid ... 29
Tetapkan izin ke pengguna Lake Formation ... 30
Konfigurasikan lokasi Amazon S3 untuk danau data Anda ... 32
(Opsional) Pengaturan penyaringan data eksternal ... 32
(Opsional) Berikan akses ke kunci enkripsi Katalog Data ... 33
(Opsional) Membuat peran IAM untuk alur kerja ... 34
Memutakhirkan izin AWS Glue data ke model Lake Formation ... 35
Tentang izin default ... 36
Daftar izin yang ada ... 37
Siapkan izin Lake Formation ... 40
Berikan izin IAM kepada pengguna ... 40
Beralih ke model izin Lake Formation ... 41
Langkah 5: Amankan sumber daya Katalog Data baru ... 44
Langkah 6: Beri pengguna kebijakan IAM baru ... 45
Langkah 7: Bersihkan kebijakan IAM yang ada ... 46
Menyiapkan titik akhir Amazon VPC ()AWS PrivateLink ... 46
Pertimbangan untuk titik akhir VPC Lake Formation ... 47
Membuat titik akhir VPC antarmuka untuk Lake Formation ... 47
Membuat kebijakan titik akhir VPC untuk Lake Formation ... 47
Tutorial ... 49
Membuat danau data dari sebuah AWS CloudTrail sumber ... 50
Audiens yang dituju ... 51
Prasyarat ... 52
Langkah 1: Buat pengguna analis data ... 52
Langkah 2: Tambahkan izin untuk membaca AWS CloudTrail log ke peran alur kerja ... 53
Langkah 3: Buat bucket Amazon S3 untuk data lake ... 54
Langkah 4: Daftarkan jalur Amazon S3 ... 55
Langkah 5: Berikan izin lokasi data ... 55
Langkah 6: Buat database di Katalog Data ... 55
Langkah 7: Berikan izin data ... 56
Langkah 8: Gunakan cetak biru untuk membuat alur kerja ... 56
Langkah 9: Jalankan alur kerja ... 57

Langkah 10: Berikan SELECT pada tabel 58
Langkah 11: Kueri data lake Menggunakan Amazon Athena 59
Membuat data lake dari sumber JDBC 60
Audiens yang dituju 60
Prasyarat 61
Langkah 1: Buat pengguna analis data 62
Langkah 2: Buat koneksi di AWS Glue 63
Langkah 3: Buat bucket Amazon S3 untuk data lake 63
Langkah 4: Daftarkan jalur Amazon S3 64
Langkah 5: Berikan izin lokasi data 64
Langkah 6: Buat database di Katalog Data 64
Langkah 7: Berikan izin data 65
Langkah 8: Gunakan cetak biru untuk membuat alur kerja 65
Langkah 9: Jalankan alur kerja 67
Langkah 10: Berikan SELECT pada tabel 68
Langkah 11: Kueri data lake menggunakan Amazon Athena 68
Langkah 12: Kueri data di danau data menggunakan Amazon Redshift Spectrum 69
Langkah 13: Berikan atau cabut izin Lake Formation menggunakan Amazon Redshift Spectrum 73
Menyiapkan izin untuk format tabel terbuka di Lake Formation 73
Audiens yang dituju 74
Prasyarat 75
Langkah 1: Menyediakan sumber daya Anda 76
Langkah 2: Siapkan izin untuk tabel Iceberg 78
Langkah 3: Siapkan izin untuk tabel Hudi 84
Langkah 4: Siapkan izin untuk tabel Delta Lake 86
Langkah 5: Bersihkan AWS sumber daya 89
Mengelola data lake menggunakan kontrol akses berbasis tag 89
Audiens yang dituju 91
Prasyarat 92
Langkah 1: Menyediakan sumber daya Anda 92
Langkah 2: Daftarkan lokasi data Anda, buat ontologi LF-tag, dan berikan izin 93
Langkah 3: Buat database Lake Formation 97
Langkah 4: Berikan izin tabel 107
Langkah 5: Jalankan kueri di Amazon Athena untuk memverifikasi izin 109
Langkah 6: Bersihkan AWS sumber daya 110

Mengamankan data lake dengan kontrol akses tingkat baris 110
Audiens yang dituju 111
Prasyarat 112
Langkah 1: Menyediakan sumber daya Anda 113
Langkah 2: Kueri tanpa filter data 114
Langkah 3: Siapkan filter data dan berikan izin 116
Langkah 4: Kueri dengan filter data 118
Langkah 5: Bersihkan AWS sumber daya 119
Bagikan data Anda dengan aman menggunakan Lake Formation 119
Audiens yang dituju 120
Konfigurasikan pengaturan Lake Formation 122
Langkah 1: Menyediakan sumber daya Anda menggunakan AWS CloudFormation template 124
Langkah 2: Prasyarat berbagi lintas akun Lake Formation 126
Langkah 3: Terapkan berbagi lintas akun menggunakan metode kontrol akses berbasis tag 128
Langkah 4: Menerapkan metode sumber daya bernama 134
Langkah 5: Bersihkan AWS sumber daya 138
Berbagi sumber daya Katalog Data dengan eksternal Akun AWS menggunakan kontrol akses berbutir halus 138
Audiens yang dituju 139
Prasyarat 140
Langkah 1: Berikan akses halus ke akun lain 141
Langkah 2: Berikan akses halus ke pengguna di akun yang sama 143
Orientasi ke izin Lake Formation 144
Ikhtisar izin Lake Formation 145
Metode untuk kontrol akses berbutir halus 147
Kontrol akses metadata 150
Kontrol akses data yang mendasari 154
Referensi personas Lake Formation dan izin IAM 160
AWS Lake Formation persona 160
AWS kebijakan terkelola untuk Lake Formation 162
Personas menyarankan izin 170
Mengubah pengaturan default untuk data lake 181
Izin Formasi Danau Implisit 184
Referensi izin Lake Formation 186

Izin Lake Formation per jenis sumber daya 186
Lake Formation memberikan dan mencabut AWS CLI perintah 189
Izin Lake Formation 194
Mengintegrasikan Pusat Identitas IAM 208
Prasyarat untuk integrasi Pusat Identitas IAM dengan Formasi Danau 209
Menghubungkan Formasi Danau dengan Pusat Identitas IAM 213
Memperbarui integrasi Pusat Identitas IAM 217
Menghapus koneksi Lake Formation dengan IAM Identity Center 219
Memberikan izin kepada pengguna dan grup 220
Termasuk konteks pengguna Pusat Identitas IAM dalam log CloudTrail 223
Menambahkan lokasi Amazon S3 ke danau data Anda 225
Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi 226
Mendaftarkan lokasi Amazon S3 234
Mendaftarkan lokasi Amazon S3 terenkripsi 238
Mendaftarkan lokasi Amazon S3 di tempat lain AWS akun 243
Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening 245
Membatalkan pendaftaran lokasi Amazon S3 250
Mengakses lokasi Amazon S3 250
Mode akses hibrid 255
Kasus penggunaan mode akses hybrid umum 257
Cara kerja mode akses hybrid 258
Menyiapkan mode akses hybrid - skenario umum 260
Menghapus prinsip dan sumber daya dari mode akses hybrid 280
Melihat prinsip dan sumber daya dalam mode akses hybrid 281
Sumber daya tambahan 282
Membuat objek di AWS Glue Data Catalog 282
Membuat katalog 284
Membuat basis data 284
Membuat tabel 285
Membangun tampilan Katalog Data 293
Mengimpor data menggunakan alur kerja 341
Cetak biru dan alur kerja 342
Membuat alur kerja 343
Menjalankan alur kerja 347
Membawa data ke dalam Katalog Data 349
Federasi katalog 351

Fitur Federasi Katalog ... 351
Batasan ... 352
Federasi untuk Snowflake ... 352
Federasi ke Databricks ... 360
Membawa data Amazon Redshift ke dalam Katalog Data ... 368
Manfaat utama ... 371
Peran dan tanggung jawab ... 371
Prasyarat ... 372
Membuat katalog federasi Amazon Redshift ... 378
Melihat objek katalog ... 387
Memperbarui katalog federasi ... 389
Mengakses katalog federasi bersama ... 391
Menghapus katalog federasi ... 396
Menanyakan katalog federasi ... 397
Sumber daya tambahan ... 397
Federasi ke sumber data eksternal ... 397
Alur kerja ... 398
Prasyarat ... 399
Membuat katalog federasi ... 404
Melihat objek katalog ... 410
Menghapus katalog federasi ... 412
Menanyakan katalog federasi ... 413
Sumber daya tambahan ... 413
Integrasi Tabel Amazon S3 dengan Katalog Data dan Lake Formation ... 413
Cara kerja integrasi Katalog Data dan Lake Formation ... 415
Prasyarat ... 416
Mengaktifkan integrasi Tabel Amazon S3 ... 419
Membuat database dan tabel ... 424
Mendaftarkan keranjang meja Amazon S3 di akun lain AWS ... 427
Memberikan izin ... 430
Mengubah kontrol akses untuk integrasi Tabel S3 ... 433
Aktifkan Lake Formation dengan integrasi Tabel S3 ... 434
Perubahan dari Lake Formation ke IAM ... 435
Membuat katalog terkelola Amazon Redshift ... 439
Mengelola izin untuk data dalam data Amazon Redshift ... 443
Prasyarat ... 445

Menyiapkan izin untuk datashares Amazon Redshift ... 445
Menanyakan database federasi ... 449
Mengelola izin pada kumpulan data yang menggunakan metastor eksternal ... 450
Alur kerja ... 453
Prasyarat ... 453
Menghubungkan Katalog Data ke metastore Hive eksternal ... 456
Sumber daya tambahan ... 459
Mengelola izin Lake Formation ... 460
Memberikan izin lokasi data ... 460
Memberikan izin lokasi data (akun yang sama) ... 461
Memberikan izin lokasi data (akun eksternal) ... 464
Memberikan izin pada lokasi data yang dibagikan dengan akun Anda ... 467
Memberikan izin data ... 468
Izin IAM diperlukan untuk memberikan izin Lake Formation ... 470
Menggunakan metode sumber daya bernama ... 472
Tag-based kontrol akses ... 497
Memberikan izin data lake menggunakan metode LF-TBAC ... 558
Attribute-based kontrol akses ... 566
Skenario contoh izin ... 570
Pemfilteran data dan keamanan tingkat sel ... 572
Filter data ... 574
Dukungan PartiQL dalam ekspresi filter baris ... 578
Izin diperlukan untuk menanyakan tabel dengan pemfilteran tingkat sel ... 580
Mengelola filter data ... 581
Melihat Izin Database dan Tabel ... 596
Mencabut izin menggunakan konsol ... 600
Cross-account berbagi data ... 600
Prasyarat ... 603
Memperbarui pengaturan versi berbagi data lintas akun ... 607
Berbagi tabel dan database Katalog Data Akun AWS atau prinsip IAM dari akun eksternal ... 613
Memberikan izin pada database atau tabel yang dibagikan dengan akun Anda ... 619
Memberikan izin tautan sumber daya ... 621
Mengakses data dasar tabel bersama ... 624
Cross-account CloudTrail penebangan ... 626
Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau ... 630
Melihat semua hibah lintas akun menggunakan operasi API GetResourceShares ... 633

Mengakses dan melihat tabel dan database Katalog Data bersama 635
Menerima undangan berbagi AWS RAM sumber daya 636
Melihat tabel dan database Katalog Data bersama 638
Membuat tautan sumber daya 640
Cara kerja tautan sumber daya 640
Membuat tautan sumber daya ke tabel bersama 643
Membuat tautan sumber daya ke database bersama 646
Penanganan tautan sumber daya di AWS Glue API 650
Mengakses tabel di seluruh Wilayah 654
Alur kerja 655
Menyiapkan akses tabel lintas wilayah 659
Keamanan 662
Perlindungan Data 662
Enkripsi saat Istirahat 664
Keamanan Infrastruktur 664
Pencegahan "confused deputy" lintas layanan 665
Pencatatan peristiwa keamanan di AWS Lake Formation 666
Integrasi dengan Lake Formation 667
Menggunakan integrasi aplikasi Lake Formation 667
Cara kerja integrasi aplikasi Lake Formation 668
Peran dan tanggung jawab dalam integrasi aplikasi Lake Formation 670
Lake Formation alur kerja untuk operasi API integrasi aplikasi 671
Mendaftarkan mesin kueri pihak ketiga 673
Mengaktifkan izin untuk mesin kueri pihak ketiga untuk memanggil operasi API integrasi aplikasi 675
Integrasi aplikasi untuk akses tabel penuh 679
Bekerja dengan AWS layanan lain 681
Amazon Athena 684
Support untuk format tabel transaksional 686
Sumber daya tambahan 688
Amazon Redshift Spectrum 689
Support untuk tipe tabel transaksional 690
Sumber daya tambahan 691
AWS Glue 691
Support untuk tipe tabel transaksional 692
Sumber daya tambahan 694

Amazon EMR 694
Support untuk format tabel transaksional 695
Sumber daya tambahan 696
Quick 696
Sumber daya tambahan 697
AWS CloudTrail Danau 697
Pencatatan panggilan API AWS Lake Formation menggunakan AWS CloudTrail 698
Informasi Lake Formation di CloudTrail 698
Memahami peristiwa Lake Formation 699
Menelusuri akses data CloudTrail dengan Lake Formation 701
Praktik, pertimbangan, dan batasan terbaik Lake Formation 704
Praktik dan pertimbangan terbaik berbagi data lintas akun 704
Batasan peran terkait layanan 706
Keterbatasan akses data Lintas Wilayah 707
Katalog Data melihat pertimbangan dan batasan 708
Batasan penyaringan data 709
Catatan dan batasan untuk penyaringan tingkat kolom 709
Batasan penyaringan tingkat sel 711
Pertimbangan dan batasan mode akses hibrida 713
Keterbatasan untuk membawa data gudang data Amazon Redshift ke dalam AWS Glue Data Catalog 714
Keterbatasan integrasi katalog Tabel S3 716
Metadata sarang menyimpan pertimbangan dan batasan berbagi data 716
Batasan berbagi data Amazon Redshift 718
Keterbatasan integrasi Pusat Identitas IAM 719
Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation 720
Pertimbangan kontrol akses berbasis atribut, batasan, dan wilayah yang didukung 723
Memecahkan Masalah Lake Formation 725
Pemecahan masalah umum 725
Kesalahan: Izin Lake Formation tidak mencukupi <Amazon S3 location> 725
Kesalahan: “Izin kunci enkripsi tidak memadai untuk Glue API” 726
Kueri saya Amazon Athena atau Amazon Redshift yang menggunakan manifes gagal 726
Kesalahan: “Izin Lake Formation tidak mencukupi: Wajib membuat tag di katalog” 726
Kesalahan saat menghapus administrator danau data yang tidak valid 726
Kesalahan: “Penelepon sementara tidak berwenang untuk melakukan operasi ini pada sumber daya...” 726

Memecahkan masalah akses lintas akun 726
Saya memberikan izin Lake Formation lintas akun tetapi penerima tidak dapat melihat sumber dayanya 727
Prinsipal di akun penerima dapat melihat sumber daya Katalog Data tetapi tidak dapat mengakses data yang mendasarinya 728
Kesalahan: “Asosiasi gagal karena pemanggil tidak diotorisasi” saat menerima undangan berbagi AWS RAM sumber daya 728
Kesalahan: “Tidak berwenang untuk memberikan izin untuk sumber daya” 729
Kesalahan: “Akses ditolak untuk mengambil informasi AWS Organisasi” 729
Kesalahan: “Organisasi <organization-ID>tidak ditemukan” 729
Kesalahan: “Izin Lake Formation tidak mencukupi: Kombinasi ilegal” 729
ConcurrentModificationException atas grant/revoke permintaan ke akun eksternal 729
Kesalahan saat menggunakan Amazon EMR untuk mengakses data yang dibagikan melalui lintas akun 730
Memecahkan masalah cetak biru dan alur kerja 731
<role-ARN>Cetak biru saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: on resource:PassRole ” 731
<role-ARN>Alur kerja saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: PassRole on resource:” 732
Perayap dalam alur kerja saya gagal dengan “Sumber daya tidak ada atau pemohon tidak diizinkan untuk mengakses izin yang diminta” 732
Perayap di alur kerja saya gagal dengan “Terjadi kesalahan (AccessDeniedException) saat memanggil CreateTable operasi...” 732
Masalah yang diketahui untuk AWS Lake Formation 732
Batasan pada penyaringan metadata tabel 733
Masalah dengan mengganti nama kolom yang dikecualikan 734
Masalah dengan menghapus kolom dalam tabel CSV 734
Partisi tabel harus ditambahkan di bawah jalur umum 734
Masalah dengan membuat database selama pembuatan alur kerja 734
Masalah dengan menghapus dan kemudian membuat ulang pengguna 735
Operasi API Katalog Data tidak memperbarui nilai untuk `IsRegisteredWithLakeFormation` parameter 735
Operasi Lake Formation tidak mendukung AWS Glue Schema Registry 735
Pesan kesalahan yang diperbarui 735
Lake Formation API 737
Izin 738

— operasi — 738
— tipe data — 738
Pengaturan Data Lake 739
— operasi — 739
— tipe data — 739
Integrasi Pusat Identitas IAM 739
— operasi — 739
— tipe data — 739
Mode akses hibrid 740
— operasi — 740
— tipe data — 738
Penjual kredensi 740
— operasi — 740
— tipe data — 741
Penandaan 741
— operasi — 741
— tipe data — 742
Filter data APIs 742
— operasi — 742
— tipe data — 742
Tipe data umum 742
ErrorDetail 743
Pola string 743
Wilayah yang Didukung 744
Ketersediaan umum 744
AWS GovCloud (US) 744
Transaksi dan optimasi penyimpanan 744
Riwayat Dokumen 747
AWS Glosarium 763
..... dcclxiv

Apa itu AWS Lake Formation?

Selamat datang di Panduan AWS Lake Formation Pengembang.

AWS Lake Formation membantu Anda mengatur secara terpusat, mengamankan, dan berbagi data secara global untuk analitik dan pembelajaran mesin. Dengan Lake Formation, Anda dapat mengelola kontrol akses berbutir halus untuk data lake data Anda di Amazon Simple Storage Service (Amazon S3) dan metadatanya di. AWS Glue Data Catalog

Lake Formation menyediakan model izinnya sendiri yang menambah model izin IAM. Model izin Lake Formation memungkinkan akses halus ke data yang disimpan di danau data serta sumber data eksternal seperti gudang data Amazon Redshift, Amazon DynamoDB database, dan sumber data pihak ketiga melalui mekanisme hibah atau pencabutan sederhana, seperti sistem manajemen basis data relasional (RDBMS). Izin Lake Formation diberlakukan menggunakan kontrol granular di tingkat kolom, baris, dan sel di seluruh layanan AWS analitik dan pembelajaran mesin, termasuk Amazon Athena, Amazon Amazon Quick Redshift Spectrum, Amazon EMR, dan. AWS Glue

Dengan mode akses hibrida Lake Formation untuk AWS Glue Data Catalog (Katalog Data), Anda dapat mengamankan dan mengakses data yang dikatalogkan menggunakan izin Lake Formation dan kebijakan izin IAM untuk Amazon S3 dan tindakan. AWS Glue Dengan mode akses hybrid, administrator data dapat memasukkan izin Lake Formation secara selektif dan bertahap, dengan fokus pada satu kasus penggunaan data lake pada satu waktu.

Lake Formation juga memungkinkan Anda untuk berbagi data secara internal dan eksternal di beberapa AWS organisasi Akun AWS, atau langsung dengan prinsipal IAM di akun lain yang menyediakan akses halus ke metadata Katalog Data dan data yang mendasarinya.

Topik

- Fitur Lake Formation
- AWS Lake Formation Cara kerjanya:
- Komponen Lake Formation
- Terminologi Lake Formation
- AWS integrasi layanan dengan Lake Formation
- Sumber daya Lake Formation tambahan
- Memulai dengan Lake Formation

Fitur Lake Formation

Lake Formation membantu Anda memecah silo data dan menggabungkan berbagai jenis data terstruktur dan tidak terstruktur ke dalam repositori terpusat. Pertama, identifikasi penyimpanan data yang ada di Amazon S3 atau database relasional dan NoSQL, dan pindahkan data ke data lake Anda. Kemudian crawl, katalog, dan siapkan data untuk analitik. Selanjutnya, berikan pengguna Anda akses layanan mandiri yang aman ke data melalui pilihan layanan analitik mereka.

Anda dapat menggunakan konsol Lake Formation untuk membuat katalog federasi multi-level di Katalog Data, dan menyatukan data di seluruh danau data Amazon S3 dan gudang data Amazon Redshift. Anda juga dapat mengintegrasikan data dari database operasional Anda seperti Amazon DynamoDB, dan sumber data pihak ketiga seperti Google BigQuery, MySQL, dan lainnya. Katalog Data menyediakan repositori metadata terpusat yang membuat pengelolaan dan penemuan data di seluruh sistem yang berbeda menjadi lebih mudah.

Untuk informasi selengkapnya, lihat Membawa data Anda ke AWS Glue Data Catalog.

Topik

- Konsumsi dan manajemen data
- Manajemen keamanan
- Membawa data Anda ke dalam Katalog Data

Konsumsi dan manajemen data

Impor data dari database yang sudah ada AWS

Setelah Anda menentukan di mana basis data yang ada dan memberikan kredensil akses Anda, Lake Formation membaca data dan metadatanya (skema) untuk memahami isi sumber data. Kemudian mengimpor data ke danau data baru Anda dan mencatat metadata dalam katalog pusat. Dengan Lake Formation, Anda dapat mengimpor data dari MySQL, PostgreSQL, SQL Server, MariaDB, dan database Oracle yang berjalan di Amazon RDS atau dihosting di Amazon EC2. Pemuatan data massal dan tambahan didukung.

Impor data dari sumber eksternal lainnya

Anda dapat menggunakan Lake Formation untuk memindahkan data dari database lokal dengan menghubungkan dengan Java Database Connectivity (JDBC). Identifikasi sumber target Anda dan berikan kredensil akses di konsol, dan Lake Formation membaca dan memuat data Anda ke dalam

data lake. Untuk mengimpor data dari database selain yang tercantum di atas, Anda dapat membuat pekerjaan ETL khusus dengan. AWS Glue

Katalog dan beri label data Anda

Anda dapat menggunakan AWS Glue crawler untuk membaca data Anda di Amazon S3 dan mengekstrak skema database dan tabel serta menyimpan data tersebut dalam Katalog Data yang dapat dicari. Kemudian, gunakan Lake Formation Kontrol akses berbasis tag Lake Formation (TBAC) untuk mengelola izin pada database, tabel, dan kolom. Untuk informasi selengkapnya tentang menambahkan tabel ke Katalog Data, lihatMembuat objek di AWS Glue Data Catalog.

Manajemen keamanan

Tentukan dan kelola kontrol akses

Lake Formation menyediakan satu tempat untuk mengelola kontrol akses data di danau data Anda. Anda dapat menentukan kebijakan keamanan yang membatasi akses ke data di tingkat database, tabel, kolom, baris, dan sel. Kebijakan ini berlaku untuk pengguna dan peran IAM, dan untuk pengguna dan grup saat melakukan federasi melalui penyedia identitas eksternal. Anda dapat menggunakan kontrol berbutir halus untuk mengakses data yang diamankan oleh Lake Formation dalam Amazon Redshift Spectrum, Athena, ETL AWS Glue , dan Amazon EMR untuk Apache Spark. Setiap kali Anda membuat identitas IAM, pastikan untuk mengikuti praktik terbaik IAM. Untuk informasi selengkapnya, lihat Praktik terbaik keamanan di Panduan Pengguna IAM.

Mode akses hibrida

Mode akses hibrida Lake Formation memberikan fleksibilitas untuk mengaktifkan izin Lake Formation secara selektif untuk database dan tabel di Katalog Data Anda. Dengan mode akses hybrid, Anda sekarang memiliki jalur tambahan yang memungkinkan Anda mengatur izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu kebijakan izin pengguna atau beban kerja lain yang ada. Untuk informasi selengkapnya, lihat Mode akses hibrid.

Melaksanakan pencatatan audit

Lake Formation menyediakan log audit komprehensif CloudTrail untuk memantau akses dan menunjukkan kepatuhan terhadap kebijakan yang ditetapkan secara terpusat. Anda dapat mengaudit riwayat akses data di seluruh layanan analitik dan pembelajaran mesin yang membaca data di danau data Anda melalui Lake Formation. Ini memungkinkan Anda melihat pengguna atau peran mana yang telah mencoba mengakses data apa, dengan layanan mana, dan kapan. Anda dapat mengakses log audit dengan cara yang sama Anda mengakses CloudTrail log lain menggunakan

CloudTrail APIs dan konsol. Untuk informasi selengkapnya tentang CloudTrail log, lihatPencatatan panggilan API AWS Lake Formation menggunakan AWS CloudTrail.

Keamanan baris dan tingkat sel

Lake Formation menyediakan filter data yang memungkinkan Anda membatasi akses ke kombinasi kolom dan baris. Gunakan keamanan baris dan tingkat sel untuk melindungi data sensitif seperti Informasi Identifikasi Pribadi (PII). Untuk informasi selengkapnya tentang keamanan tingkat baris, lihat. Pemfilteran data dan keamanan tingkat sel di Lake Formation

Kontrol akses berbasis tanda

Gunakan kontrol akses berbasis atribut Lake Formation untuk mengelola ratusan atau bahkan ribuan izin data dengan membuat label khusus yang disebut LF-tag. Anda sekarang dapat menentukan LF-tag dan melampirkannya ke database, tabel, atau kolom. Kemudian, bagikan akses terkontrol di seluruh layanan analitik, pembelajaran mesin (ML), dan ekstrak, transformasi, dan muat (ETL) untuk konsumsi. LF-tag memastikan bahwa tata kelola data dapat diskalakan dengan mudah dengan mengganti definisi kebijakan dari ribuan sumber daya dengan beberapa tag logis. Lake Formation menyediakan pencarian berbasis teks melalui metadata ini, sehingga pengguna Anda dapat dengan cepat menemukan data yang perlu mereka analisis.

Kontrol Akses Berbasis Atribut

Gunakan kontrol akses berbasis atribut untuk memberikan akses pada objek Katalog Data. Attribute-based access control (ABAC) adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. AWS memanggil tag atribut ini. Anda dapat menggunakan ABAC untuk memberikan akses ke prinsipal dalam akun yang sama atau di akun lain pada sumber daya Katalog Data. Setiap prinsipal IAM dengan tag IAM yang cocok atau kunci tag sesi dan nilai mendapatkan akses ke sumber daya. Anda harus memiliki izin yang dapat diberikan pada sumber daya untuk membuat hibah ini.

Akses lintas akun

Kemampuan manajemen izin Lake Formation menyederhanakan pengamanan dan pengelolaan data lake terdistribusi di beberapa AWS akun melalui pendekatan terpusat, menyediakan kontrol akses berbutir halus ke Katalog Data dan lokasi Amazon S3. Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation.

Membawa data Anda ke dalam Katalog Data

Kemampuan federasi memungkinkan Anda membuat katalog federasi dan mengatur izin pada kumpulan data yang disimpan dalam sumber data yang berbeda seperti Amazon Redshift tanpa

memigrasikan data atau metadata ke Amazon S3 atau. AWS Glue Data Catalog Anda dapat menggunakan metode berikut untuk membawa data dan mengelola izin pada kumpulan data eksternal di Lake Formation:

Untuk informasi selengkapnya, lihat Membawa data Anda ke dalam AWS Glue Data Catalog.

- Membawa data di gudang data Amazon Redshift ke dalam AWS Glue Data Catalog — Daftarkan namespace Amazon Redshift yang ada atau cluster dengan Katalog Data, dan buat katalog federasi multi-level di Katalog Data.

  Anda dapat mengakses data Anda menggunakan mesin kueri apa pun yang kompatibel dengan spesifikasi OpenAPI katalog Apache Iceberg REST, seperti Amazon EMR Tanpa Server, dan Amazon Athena.

  Untuk informasi selengkapnya, lihat Membawa data Amazon Redshift ke dalam AWS Glue Data Catalog.
- Menggabungkan ke dalam Katalog Data dari sumber data eksternal — Hubungkan Katalog Data ke sumber data eksternal menggunakan AWS Glue koneksi, dan buat katalog federasi untuk mengelola izin akses secara terpusat pada kumpulan data menggunakan Lake Formation. Tidak diperlukan migrasi metadata ke dalam Katalog Data.

  Untuk informasi selengkapnya, lihat Menggabungkan ke sumber data eksternal di AWS Glue Data Catalog.
- Mengintegrasikan Bucket Tabel Amazon S3 dengan Katalog Data — Anda dapat mempublikasikan dan membuat katalog Tabel Amazon S3 sebagai objek Katalog Data dan mendaftarkan katalog sebagai lokasi data Lake Formation dari konsol Lake Formation atau menggunakan. AWS Glue APIs

  Untuk informasi selengkapnya, lihat Integrasi Tabel Amazon S3 dengan AWS Glue Data Catalog and AWS Lake Formation.
- Buat katalog untuk mengelola tabel Amazon Redshift di Katalog Data — Anda mungkin tidak memiliki klaster produsen Amazon Redshift atau datashare Amazon Redshift yang tersedia saat ini, tetapi ingin membuat dan mengelola tabel Amazon Redshift menggunakan Katalog Data. Anda dapat memulai dengan membuat katalog AWS Glue terkelola menggunakan `glue:CreateCatalog` API atau AWS Lake Formation konsol dengan menyetel jenis katalog sebagai `Managed` dan `Catalog source` sebagai Redshift.

Untuk informasi selengkapnya, lihat Membuat katalog terkelola Amazon Redshift di AWS Glue Data Catalog.

- Mengintegrasikan Lake Formation dengan berbagi data Amazon Redshift — Gunakan Lake Formation untuk mengelola database, tabel, kolom, dan izin akses tingkat baris secara terpusat dari datashares Amazon Redshift dan membatasi akses pengguna ke objek dalam datashare.
- Menghubungkan Katalog Data ke metastor eksternal — Sambungkan AWS Glue Data Catalog ke metastor eksternal untuk mengelola izin akses pada kumpulan data di Amazon S3 menggunakan Lake Formation. Tidak diperlukan migrasi metadata ke dalam Katalog Data.

  Untuk informasi selengkapnya, lihat Mengelola izin pada kumpulan data yang menggunakan metastor eksternal.
- Mengintegrasikan Lake Formation dengan AWS Data Exchange — Lake Formation mendukung lisensi akses ke data Anda melalui. AWS Data Exchange Jika Anda tertarik untuk melisensikan data Lake Formation Anda, lihat Apa yang ada AWS Data Exchange di Panduan AWS Data Exchange Pengguna.

AWS Lake Formation Cara kerjanya:

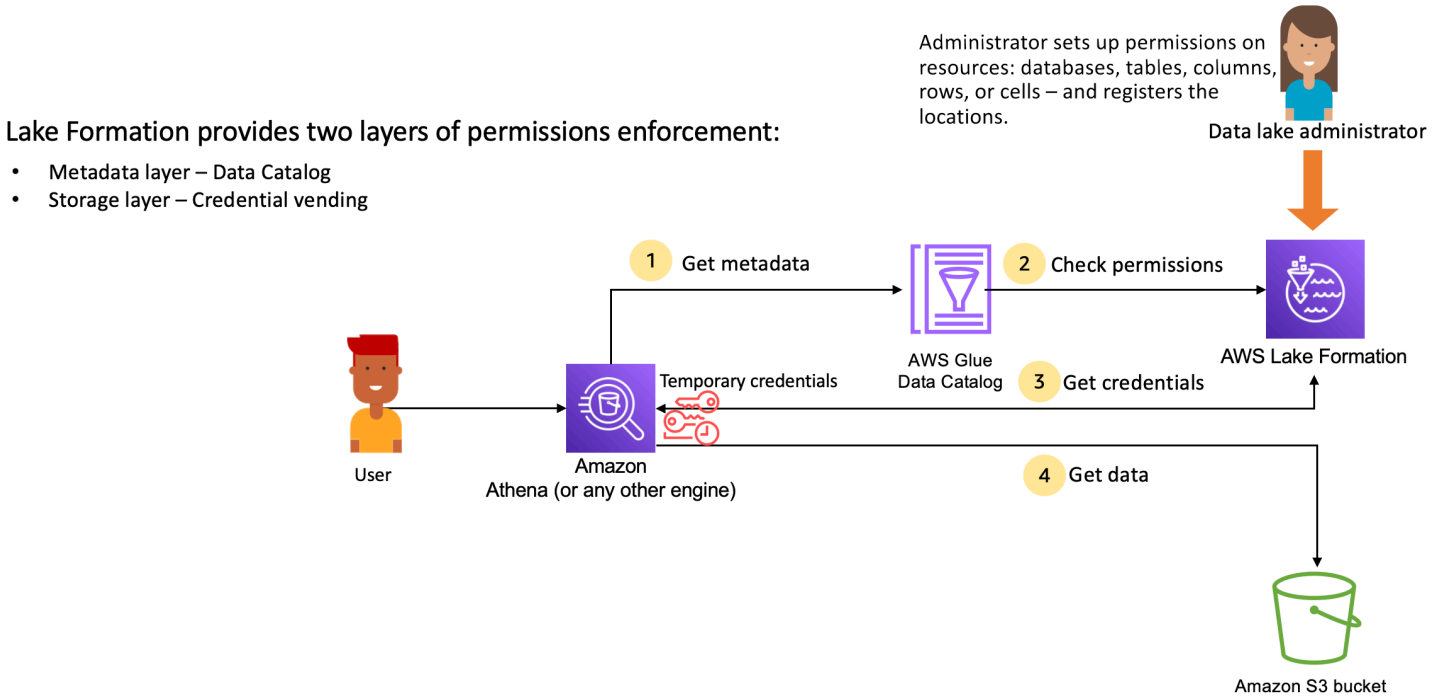
AWS Lake Formation menyediakan model izin sistem manajemen basis data relasional (RDBMS) untuk memberikan atau mencabut akses ke sumber daya Katalog Data seperti database, tabel, dan kolom dengan data dasar di Amazon S3. Izin Lake Formation yang mudah dikelola menggantikan kebijakan bucket Amazon S3 yang kompleks dan kebijakan IAM terkait.

Di Lake Formation, Anda dapat menerapkan izin pada dua level:

- Menegakkan izin tingkat metadata pada sumber daya Katalog Data seperti database dan tabel
- Mengelola izin akses penyimpanan pada data dasar yang disimpan di Amazon S3 atas nama mesin terintegrasi

Alur kerja manajemen izin Lake Formation

Lake Formation terintegrasi dengan mesin analitik untuk menanyakan penyimpanan data Amazon S3 dan objek metadata yang terdaftar di Lake Formation. Diagram berikut menggambarkan cara kerja manajemen izin di Lake Formation.



Izin Lake Formation mengelola langkah-langkah tingkat tinggi

Sebelum Lake Formation dapat memberikan kontrol akses untuk data di danau data Anda, administrator danau data atau pengguna dengan izin administratif akan menyiapkan kebijakan pengguna tabel Katalog Data individual untuk mengizinkan atau menolak akses ke tabel Katalog Data menggunakan izin Lake Formation.

Kemudian, administrator data lake atau pengguna yang didelegasikan oleh administrator memberikan izin Lake Formation kepada pengguna di database dan tabel Katalog Data, dan mendaftarkan lokasi Amazon S3 tabel dengan Lake Formation.

1. Dapatkan metadata — Prinsipal (pengguna) mengirimkan kueri atau skrip ETL ke mesin analitik terintegrasi seperti Amazon Athena, Amazon EMR, atau AWS Glue Amazon Redshift Spectrum. Mesin analitik terintegrasi mengidentifikasi tabel yang diminta dan mengirimkan permintaan metadata ke Katalog Data.
2. Periksa izin — Katalog Data memeriksa izin pengguna dengan Lake Formation, dan jika pengguna diberi wewenang untuk mengakses tabel, mengembalikan metadata yang diizinkan dilihat pengguna ke mesin.
3. Dapatkan kredensil — Katalog Data memungkinkan mesin mengetahui apakah tabel dikelola oleh Lake Formation atau tidak. Jika data yang mendasarinya terdaftar di Lake Formation, mesin analitik meminta Lake Formation untuk menyediakan akses data dengan memberikan akses sementara.

4. Dapatkan data — Jika pengguna berwenang untuk mengakses tabel, Lake Formation menyediakan akses sementara ke mesin analitik terintegrasi. Menggunakan akses sementara, mesin analitik mengambil data dari Amazon S3, dan melakukan pemfilteran yang diperlukan seperti pemfilteran kolom, baris, atau sel. Ketika mesin selesai menjalankan pekerjaan, ia mengembalikan hasilnya kembali ke pengguna. Proses ini disebut credential vending.

   Jika tabel tidak dikelola oleh Lake Formation, panggilan kedua dari mesin analitik dilakukan langsung ke Amazon S3. Kebijakan bucket Amazon S3 terkait dan kebijakan pengguna IAM dievaluasi untuk akses data.

   Setiap kali Anda menggunakan kebijakan IAM, pastikan bahwa Anda mengikuti praktik terbaik IAM. Untuk informasi selengkapnya tentang administrator, lihat Praktik terbaik keamanan di IAM dalam Panduan Pengguna IAM.

Topik

- Izin metadata
- Manajemen akses penyimpanan
- Cross-account berbagi data di Lake Formation

Izin metadata

Lake Formation menyediakan otorisasi dan kontrol akses untuk Katalog Data. Ketika peran IAM membuat panggilan API Katalog Data dari sistem apa pun, Katalog Data memverifikasi izin data pengguna dan hanya mengembalikan metadata yang pengguna memiliki izin untuk mengakses. Misalnya, jika peran IAM hanya memiliki akses ke satu tabel dalam database, dan layanan atau pengguna dengan asumsi peran melakukan `GetTables` operasi, respons hanya akan berisi satu tabel, terlepas dari jumlah tabel dalam database.

Pengaturan default - izin **`IAMAllowedPrincipal`** grup

AWS Lake Formation, secara default, menetapkan izin ke semua database dan tabel ke grup virtual bernama. `IAMAllowedPrincipal` Grup ini unik dan hanya terlihat di dalam Lake Formation. `IAMAllowedPrincipal`Grup ini mencakup semua kepala sekolah IAM yang memiliki akses ke sumber daya Katalog Data melalui kebijakan utama IAM dan kebijakan sumber daya. AWS Glue Jika izin ini ada pada database atau tabel, semua prinsipal akan diberikan akses ke database atau tabel.

Jika Anda ingin memberikan izin yang lebih terperinci pada database atau tabel, hapus `IAMAllowedPrincipal` izin dan, Lake Formation memberlakukan semua kebijakan lain yang terkait dengan database atau tabel tersebut. Misalnya, jika ada kebijakan yang memungkinkan Pengguna A mengakses Database A dengan `DESCRIBE` izin, dan `IAMAllowedPrincipal` ada dengan semua izin, Pengguna A akan terus melakukan semua tindakan lainnya, hingga `IAMAllowedPrincipal` izin dicabut.

Selain itu, secara default, `IAMAllowedPrincipal` grup memiliki izin pada semua database dan tabel baru saat dibuat. Ada dua konfigurasi yang mengontrol perilaku ini. Yang pertama adalah di akun dan Region-level yang memungkinkan ini untuk database yang baru dibuat, dan yang kedua adalah di tingkat database. Untuk mengubah pengaturan default, lihatUbah model izin default atau gunakan mode akses hybrid.

Memberikan izin

Administrator data lake dapat memberikan izin Katalog Data kepada prinsipal sehingga prinsipal dapat membuat dan mengelola database dan tabel, serta dapat mengakses data yang mendasarinya.

Izin basis data dan tingkat tabel

Saat Anda memberikan izin dalam Lake Formation, pemberi harus menentukan prinsipal untuk memberikan izin, sumber daya untuk memberikan izin, dan tindakan yang harus diakses oleh penerima hibah. Untuk sebagian besar sumber daya dalam Lake Formation, daftar utama dan sumber daya untuk memberikan izin serupa, tetapi tindakan yang dapat dilakukan penerima hibah berbeda berdasarkan jenis sumber daya. Misalnya, `SELECT` izin tersedia untuk tabel untuk membaca tabel, tetapi `SELECT` izin tidak diizinkan pada database. `CREATE_TABLE`Izin diizinkan pada database, tetapi tidak pada tabel.

Anda dapat memberikan AWS Lake Formation izin menggunakan dua metode:

- Metode sumber daya bernama - Memungkinkan Anda memilih nama database dan tabel sambil memberikan izin kepada pengguna.
- LF-Tag based access control (LF-TBAC) — Pengguna membuat LF-Tags, mengaitkannya dengan sumber daya Katalog Data, memberikan `Describe` izin LF-Tags, mengaitkan izin ke pengguna individu, dan menulis kebijakan izin LF yang digunakan LF-Tags untuk pengguna yang berbeda. LF-Tag-based Kebijakan tersebut berlaku untuk semua sumber daya Katalog Data yang terkait dengan LF-Tag nilai-nilai tersebut.

> **Note**
>
> LF-Tags unik untuk Lake Formation. Mereka hanya terlihat di Lake Formation dan tidak boleh bingung dengan tag AWS sumber daya.

LF-TBAC adalah fitur yang memungkinkan pengguna untuk mengelompokkan sumber daya ke dalam kategori yang ditentukan pengguna LF-Tags dan menerapkan izin pada grup sumber daya tersebut. Oleh karena itu, ini adalah cara terbaik untuk menskalakan izin di sejumlah besar sumber daya Katalog Data.

Untuk informasi selengkapnya, lihat Kontrol akses berbasis tag Lake Formation.

Saat Anda memberikan izin kepada kepala sekolah, Lake Formation mengevaluasi izin sebagai gabungan dari semua kebijakan untuk pengguna tersebut. Misalnya, jika Anda memiliki dua kebijakan pada tabel untuk prinsipal di mana satu kebijakan memberikan izin ke kolom col1, col2, dan col3 melalui metode sumber daya bernama, dan kebijakan lainnya memberikan izin ke tabel dan prinsipal yang sama ke col5, dan col6 melalui, izin efektif akan menjadi gabungan izin yang akan menjadi col1 LF-Tags, col2, col3, col5, dan col6. Ini juga termasuk filter data dan baris.

Izin lokasi data

Izin lokasi data memberi pengguna non-administratif kemampuan untuk membuat database dan tabel di lokasi Amazon S3 tertentu. Jika pengguna mencoba membuat database atau tabel di lokasi yang tidak memiliki izin untuk dibuat, tugas pembuatan gagal. Hal ini untuk mencegah pengguna membuat tabel di lokasi arbitrer dalam data lake dan memberikan kontrol atas di mana pengguna dapat membaca dan menulis data. Ada izin implisit saat membuat tabel di lokasi Amazon S3 dalam database tempat ia dibuat. Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

Buat izin tabel dan basis data

Pengguna non-administratif secara default tidak memiliki izin untuk membuat database atau tabel dalam database. Pembuatan database dikendalikan di tingkat akun menggunakan pengaturan Lake Formation sehingga hanya kepala sekolah yang berwenang yang dapat membuat database. Untuk informasi selengkapnya, lihat Membuat basis data. Untuk membuat tabel, prinsipal memerlukan `CREATE_TABLE` izin pada database tempat tabel sedang dibuat. Untuk informasi selengkapnya, lihat Membuat tabel.

Izin implisit dan eksplisit

Lake Formation memberikan izin implisit tergantung pada persona dan tindakan yang dilakukan persona. Misalnya, administrator data lake secara otomatis mendapatkan `DESCRIBE` izin ke semua sumber daya dalam Katalog Data, izin lokasi data ke semua lokasi, izin untuk membuat database dan tabel di semua lokasi, serta `Grant` izin pada sumber daya apa pun. `Revoke` Pembuat database secara otomatis mendapatkan semua izin database pada database yang mereka buat, dan pembuat tabel mendapatkan semua izin pada tabel yang mereka buat. Untuk informasi selengkapnya, lihat Izin Formasi Danau Implisit.

Izin yang dapat diberikan

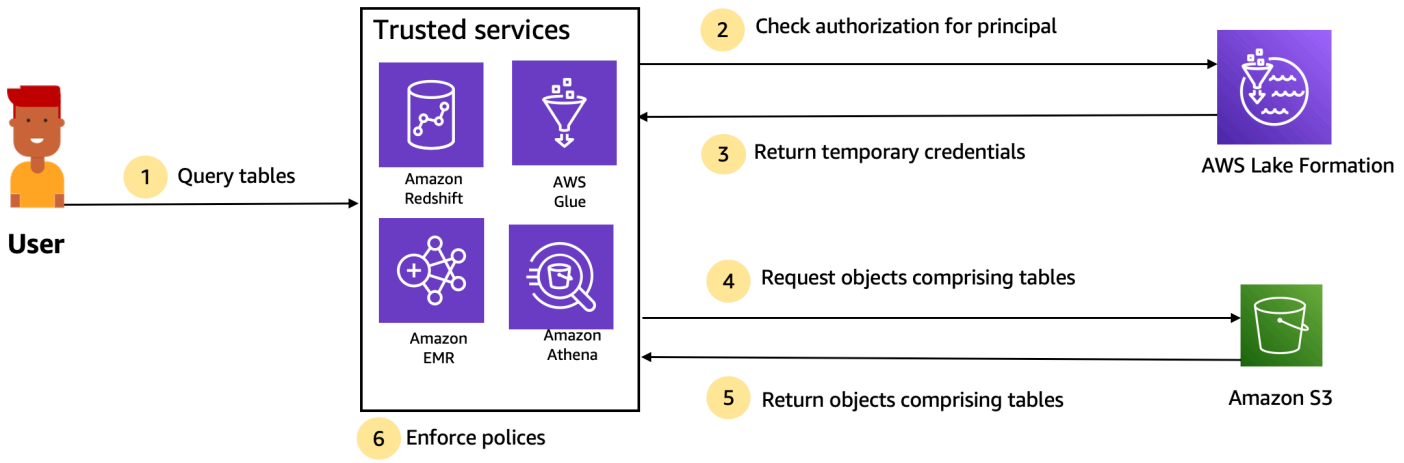
Administrator data lake memiliki kemampuan untuk mendelegasikan pengelolaan izin kepada pengguna non-administratif dengan memberikan izin yang dapat diberikan. Ketika prinsipal diberikan izin yang dapat diberikan pada sumber daya dan serangkaian izin, prinsipal tersebut memperoleh kemampuan untuk memberikan izin kepada prinsipal lain pada sumber daya tersebut.

Manajemen akses penyimpanan

Lake Formation menggunakan fungsionalitas penjual kredenal untuk menyediakan akses sementara ke data Amazon S3. Credential vending, atau token vending adalah pola umum yang memberikan kredensi sementara kepada pengguna, layanan, atau entitas lain untuk tujuan pemberian akses jangka pendek ke sumber daya.

Lake Formation memanfaatkan pola ini untuk menyediakan akses jangka pendek ke layanan AWS analitik seperti Athena untuk mengakses data atas nama kepala panggilan. Saat memberikan izin, pengguna tidak perlu memperbarui kebijakan bucket Amazon S3 atau kebijakan IAM mereka, dan mereka tidak memerlukan akses langsung ke Amazon S3.

Diagram berikut menunjukkan bagaimana Lake Formation menyediakan akses sementara ke lokasi terdaftar:



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

1. Seorang prinsipal (pengguna) memasukkan kueri atau permintaan data untuk tabel melalui layanan terintegrasi tepercaya seperti Athena, Amazon EMR, Redshift Spectrum, atau. AWS Glue
2. Layanan terintegrasi memeriksa otorisasi dari Lake Formation untuk tabel dan kolom yang diminta dan membuat penentuan otorisasi. Jika pengguna tidak berwenang, Lake Formation menolak akses ke data dan kueri gagal.
3. Setelah otorisasi berhasil dan otorisasi penyimpanan diaktifkan untuk tabel dan pengguna, layanan terintegrasi mengambil kredensi sementara dari Lake Formation untuk mengakses data.
4. Layanan terintegrasi menggunakan kredensil sementara dari Lake Formation untuk meminta objek dari Amazon S3.
5. Amazon S3 menyediakan objek Amazon S3 ke layanan terintegrasi. Objek Amazon S3 berisi semua data dari tabel.
6. Layanan terintegrasi melakukan penegakan kebijakan Lake Formation yang diperlukan, seperti tingkat kolom, penyaringan tingkat and/or sel tingkat baris. Layanan terintegrasi memproses kueri dan mengembalikan hasilnya kembali ke pengguna.

Aktifkan penegakan izin tingkat penyimpanan untuk tabel Katalog Data

Secara default, penegakan tingkat penyimpanan tidak diaktifkan untuk tabel dalam Katalog Data. Untuk mengaktifkan penegakan tingkat penyimpanan, Anda harus mendaftarkan lokasi Amazon S3 dari data sumber Anda dengan Lake Formation dan memberikan peran IAM. Storage-level izin akan diaktifkan untuk semua tabel dengan jalur lokasi tabel yang sama atau awalan lokasi Amazon S3.

Ketika layanan terintegrasi meminta akses ke lokasi data atas nama pengguna, layanan Lake Formation mengambil peran ini dan mengembalikan kredensialnya ke layanan yang diminta dengan

izin tercakup ke sumber daya sehingga akses data dapat dibuat. Peran IAM terdaftar harus memiliki semua akses yang diperlukan ke lokasi AWS KMS Amazon S3 termasuk kunci.

Untuk informasi selengkapnya, lihat Mendaftarkan lokasi Amazon S3.

Didukung AWS layanan

AWS layanan analitik seperti Athena, Redshift Spectrum, Amazon AWS Glue EMR,, Amazon Quick, dan Amazon SageMaker AI berintegrasi dengan AWS Lake Formation menggunakan operasi API penjual kredenal Lake Formation. Untuk melihat daftar lengkap AWS layanan yang terintegrasi dengan Lake Formation, dan tingkat granularitas dan format tabel yang mereka dukung, lihat. Bekerja dengan AWS layanan lain

Cross-account berbagi data di Lake Formation

Dengan Lake Formation, Anda dapat berbagi sumber daya Katalog Data (database dan tabel) dalam AWS akun dan di seluruh akun dalam pengaturan sederhana menggunakan metode sumber daya bernama atau LF-Tags. Anda dapat membagikan seluruh database atau memilih tabel dari database ke kepala IAM (peran dan pengguna IAM) di akun, ke akun lain AWS di tingkat akun, atau langsung ke prinsipal IAM di akun lain.

Anda juga dapat membagikan tabel Katalog Data dengan filter data untuk membatasi akses ke detail pada detail tingkat baris dan tingkat sel. Lake Formation menggunakan AWS Resource Access Manager (AWS RAM) untuk memfasilitasi pemberian izin antar akun. Saat sumber daya dibagi antara dua akun, AWS RAM kirim undangan ke akun penerima. Saat pengguna menerima undangan AWS RAM berbagi, AWS RAM berikan izin yang diperlukan ke Lake Formation agar sumber daya Katalog Data tersedia serta penegakan tingkat penyimpanan yang diaktifkan. Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation.

Ketika administrator data lake dari akun penerima menerima AWS RAM pembagian, sumber daya bersama tersedia di akun penerima. Administrator data lake memberikan izin Lake Formation lebih lanjut pada sumber daya bersama ke prinsipal IAM tambahan di akun penerima, jika administrator memiliki `GRANTABLE` izin pada sumber daya bersama.

Namun, kepala sekolah tidak dapat menanyakan sumber daya bersama menggunakan Athena atau Redshift Spectrum tanpa tautan sumber daya. Tautan sumber daya adalah entitas dalam Katalog Data dan mirip dengan Linux-Symlink konsep.

Administrator data lake dari akun penerima membuat tautan sumber daya pada sumber daya bersama. Administrator memberikan `Describe` izin pada tautan sumber daya dengan izin yang

diperlukan pada sumber daya bersama asli kepada pengguna tambahan. Pengguna di akun penerima kemudian dapat menggunakan tautan sumber daya untuk menanyakan sumber daya bersama menggunakan Athena dan Redshift Spectrum. Untuk informasi selengkapnya tentang tautan sumber daya, lihatMembuat tautan sumber daya.

Komponen Lake Formation

AWS Lake Formation bergantung pada interaksi beberapa komponen untuk membuat dan mengelola data lake Anda.

Konsol Lake Formation

Anda menggunakan konsol Lake Formation untuk menentukan dan mengelola data lake Anda serta memberikan serta mencabut izin Lake Formation. Anda dapat menggunakan cetak biru di konsol untuk menemukan, membersihkan, mengubah, dan menelan data. Anda juga dapat mengaktifkan atau menonaktifkan akses ke konsol untuk masing-masing pengguna Lake Formation.

Lake Formation API dan Antarmuka Baris Perintah

Lake Formation menyediakan operasi API melalui beberapa bahasa khusus SDKs dan AWS Command Line Interface ()AWS CLI. Lake Formation API bekerja sama dengan AWS Glue API. Lake Formation API berfokus terutama pada pengelolaan izin Lake Formation, sedangkan AWS Glue API menyediakan API katalog data dan infrastruktur terkelola untuk menentukan, menjadwalkan, dan menjalankan operasi ETL pada data Anda.

Untuk informasi tentang AWS Glue API, lihat Panduan AWS Glue Pengembang. Untuk informasi tentang penggunaan AWS CLI, lihat Referensi AWS CLI Perintah.

AWS Layanan lainnya

Lake Formation menggunakan layanan berikut:

- AWS Glueuntuk mengatur pekerjaan dan crawler untuk mengubah data menggunakan AWS Glue mengubah.
- IAM akan memberikan kebijakan izin kepada kepala sekolah Lake Formation. Model izin Lake Formation menambah model izin IAM untuk mengamankan data lake Anda.

Terminologi Lake Formation

Berikut ini adalah beberapa istilah penting yang akan Anda temui dalam panduan ini.

Danau data

Data lake adalah data persisten Anda yang disimpan di Amazon S3 dan dikelola oleh Lake Formation menggunakan Katalog Data. Danau data biasanya menyimpan hal-hal berikut:

- Data terstruktur dan tidak terstruktur
- Data mentah dan data yang diubah

Agar jalur Amazon S3 berada di dalam danau data, itu harus terdaftar di Lake Formation.

Akses data

Lake Formation menyediakan akses data yang aman dan terperinci melalui model izin hibah/pencabutan baru yang menambah kebijakan (IAM). AWS Identity and Access Management

Analis dan ilmuwan data dapat menggunakan portofolio lengkap layanan AWS analitik dan pembelajaran mesin, seperti Amazon Athena, untuk mengakses data. Kebijakan keamanan Lake Formation yang dikonfigurasi membantu memastikan bahwa pengguna hanya dapat mengakses data yang diizinkan untuk diakses.

Mode akses hibrid

Mode akses hibrid memungkinkan Anda mengamankan dan mengakses data yang dikatalogkan menggunakan izin Lake Formation dan izin IAM dan Amazon S3. Mode akses hibrid memungkinkan administrator data untuk memasukkan izin Lake Formation secara selektif dan bertahap, dengan fokus pada satu kasus penggunaan data lake pada satu waktu.

Cetak biru

Blueprint adalah template manajemen data yang memungkinkan Anda untuk dengan mudah menelan data ke dalam danau data. Lake Formation menyediakan beberapa cetak biru, masing-masing untuk jenis sumber yang telah ditentukan, seperti database relasional atau log. AWS CloudTrail Dari cetak biru, Anda dapat membuat alur kerja. Alur kerja terdiri dari AWS Glue crawler, pekerjaan, dan pemicu yang dihasilkan untuk mengatur pemuatan dan pembaruan data. Cetak biru mengambil sumber data, target data, dan jadwal sebagai input untuk mengonfigurasi alur kerja.

Alur kerja

Alur kerja adalah wadah untuk satu set yang terkait AWS Glue pekerjaan, crawler, dan pemicu. Anda membuat alur kerja di Lake Formation, dan dijalankan di AWS Glue layanan. Lake Formation dapat melacak status alur kerja sebagai entitas tunggal.

Saat Anda menentukan alur kerja, Anda memilih cetak biru yang menjadi dasarnya. Anda kemudian dapat menjalankan alur kerja sesuai permintaan atau sesuai jadwal.

Alur kerja yang Anda buat di Lake Formation terlihat di AWS Glue konsol sebagai grafik asiklik terarah (DAG). Menggunakan DAG, Anda dapat melacak kemajuan alur kerja dan melakukan pemecahan masalah.

Katalog Data

Katalog Data adalah penyimpanan metadata persisten Anda. Ini adalah layanan terkelola yang memungkinkan Anda menyimpan, membuat anotasi, dan berbagi metadata di AWS Cloud dengan cara yang sama seperti yang Anda lakukan di metastore Apache Hive. Ini menyediakan repositori seragam di mana sistem yang berbeda dapat menyimpan dan menemukan metadata untuk melacak data dalam silo data, dan kemudian menggunakan metadata itu untuk menanyakan dan mengubah data. Lake Formation menggunakan AWS Glue Katalog Data untuk menyimpan metadata tentang data lake, sumber data, transformasi, dan target.

Metadata tentang sumber data dan target adalah dalam bentuk database dan tabel. Tabel menyimpan informasi skema, informasi lokasi, dan banyak lagi. Database adalah kumpulan tabel. Lake Formation menyediakan hierarki izin untuk mengontrol akses ke database dan tabel di Katalog Data.

Setiap AWS akun memiliki satu Katalog Data per AWS Wilayah.

Data yang mendasari

Data yang mendasari mengacu pada data sumber atau data dalam danau data yang ditunjukkan oleh tabel Katalog Data.

Utama

Principal adalah pengguna AWS Identity and Access Management (IAM) atau peran atau pengguna Active Directory.

Administrator danau data

Administrator data lake adalah kepala sekolah yang dapat memberikan izin kepada kepala sekolah apa pun (termasuk diri sendiri) pada sumber daya Katalog Data atau lokasi data apa pun. Tentukan administrator data lake sebagai pengguna pertama Katalog Data. Pengguna ini kemudian dapat memberikan izin sumber daya yang lebih terperinci kepada prinsipal lain.

> **Note**
>
> Pengguna administratif IAM — pengguna dengan kebijakan `AdministratorAccess` AWS terkelola — bukan administrator data lake secara otomatis. Misalnya, mereka tidak dapat memberikan izin Lake Formation pada objek katalog kecuali mereka telah diberikan izin untuk melakukannya. Namun, mereka dapat menggunakan konsol Lake Formation atau API untuk menunjuk diri mereka sebagai administrator danau data.

Untuk informasi tentang kemampuan administrator danau data, lihatIzin Formasi Danau Implisit. Untuk informasi tentang menunjuk pengguna sebagai administrator data lake, lihatBuat administrator danau data.

AWS integrasi layanan dengan Lake Formation

Anda dapat menggunakan Lake Formation untuk mengelola database, tabel, dan izin akses tingkat kolom pada data yang disimpan di Amazon S3. Setelah data Anda terdaftar di Lake Formation, Anda dapat menggunakan layanan AWS analitis seperti AWS Glue, Amazon Athena, Amazon Redshift Spectrum, Amazon EMR untuk menanyakan data. AWS Layanan berikut terintegrasi dengan AWS Lake Formation dan menghormati izin Lake Formation.

AWS Layanan	Detail integrasi
AWS Glue	Topik referensi: Menggunakan AWS Lake Formation dengan AWS Glue AWS Gluedan Lake Formation berbagi Katalog Data yang sama. Untuk operasi konsol (seperti melihat daftar tabel) dan semua operasi API, AWS Glue pengguna hanya dapat mengakses database dan tabel tempat mereka memiliki izin Lake Formation.

AWS Layanan	Detail integrasi
Amazon Athena	Topik referensi: Menggunakan AWS Lake Formation dengan Amazon Athena Gunakan Lake Formation untuk mengizinkan atau menolak izin membaca data di Amazon S3. Saat Amazon Athena pengguna memilih AWS Glue katalog di editor kueri, mereka hanya dapat menanyakan database, tabel, dan kolom tempat mereka memiliki izin Lake Formation. Kueri menggunakan manifes tidak didukung. Saat ini, Lake Formation tidak mendukung pengelolaan izin pada operasi tulis seperti `VACUUMMERGE`, `UPDATE` dan `OPTIMIZE` pada tabel dalam Format Tabel Terbuka. Selain kepala sekolah yang mengautentikasi dengan Athena melalui (IAM), Lake Formation mendukung pengguna Athena yang terhubung AWS Identity and Access Management melalui driver JDBC atau ODBC dan mengautentikasi melalui SALL. Penyedia SALL yang didukung termasuk Okta dan Microsoft Active Directory Federation Service (AD FS).
Spektrum Pergeseran Merah Amazon	Topik referensi: Menggunakan AWS Lake Formation dengan Amazon Redshift Spectrum Saat pengguna Amazon Redshift membuat skema eksternal pada database di AWS Glue Data Catalog, mereka hanya dapat menanyakan tabel dan kolom dalam skema yang memiliki izin Lake Formation.
Edisi Perusahaan Cepat Amazon	Referensi: Menggunakan AWS Lake Formation dengan Cepat Saat pengguna Amazon Quick Enterprise Edition menanyakan kumpulan data di lokasi Amazon S3, pengguna harus memiliki izin Lake `SELECT` Formation pada data tersebut.

AWS Layanan	Detail integrasi
Amazon EMR	Referensi: Menggunakan AWS Lake Formation dengan Amazon EMR Anda dapat mengintegrasikan izin Lake Formation saat membuat klaster EMR Amazon dengan peran runtime. Peran runtime adalah peran IAM yang Anda kaitkan dengan pekerjaan atau kueri EMR Amazon, dan kemudian Amazon EMR menggunakan peran ini untuk mengakses sumber daya. AWS

Lake Formation juga berfungsi dengan AWS Key Management Service(AWS KMS) untuk memungkinkan Anda mengatur layanan terintegrasi ini dengan lebih mudah untuk mengenkripsi dan mendekripsi data di lokasi Amazon Simple Storage Service (Amazon S3).

Sumber daya Lake Formation tambahan

Untuk informasi lebih lanjut tentang AWS Lake Formation, kami menyarankan Anda untuk terus mempelajari lebih lanjut tentang konsep yang diperkenalkan dalam panduan ini dengan menggunakan sumber daya berikut:

Topik

- Blog
- Pembicaraan teknologi dan webinar
- Arsitektur modern
- Sumber daya data mesh
- Panduan praktik terbaik

Blog

- AWS Lake Formation 2022 tahun dalam ulasan
- Arsitektur data modern Multi-region yang sangat tangguh
- Berbagi lintas akun menggunakan LF-tag untuk mengarahkan prinsipal IAM
- Dasbor inventaris izin Lake Formation

- Mesh data yang digerakkan oleh peristiwa

Pembicaraan teknologi dan webinar

- Re:invent 2020 - Data lake: Mudah membangun, mengamankan, dan berbagi dengan AWS Lake Formation
- re:invent 2022 — Membangun dan mengoperasikan datalake di Amazon S3
- AWS Summit SF 2022 — Memahami dan mencapai arsitektur data modern
- AWS Summit ATL 2022 — Danau data modern dengan AWS Lake Formation, Amazon Redshift, dan AWS Glue
- AWS Summit ANZ 2022 — Danau data, rumah danau, dan data mesh: apa, mengapa, dan bagaimana?
- AWS Online Tech Talks — Menyederhanakan izin dan tata kelola di danau data Anda

Arsitektur modern

- Pola arsitektur modern

Sumber daya data mesh

- Membangun arsitektur data modern dan pola data mesh dalam skala menggunakan kontrol akses AWS Lake Formation berbasis tag
- Bagaimana JPMorgan Chase membangun arsitektur data mesh untuk mendorong nilai signifikan guna meningkatkan platform data perusahaan mereka
- Membangun data mesh di AWS

Panduan praktik terbaik

- AWS Lake Formation panduan praktik terbaik

Memulai dengan Lake Formation

Kami menyarankan Anda memulai dengan bagian berikut:

- AWS Lake Formation Cara kerjanya:— Pelajari tentang terminologi penting dan bagaimana berbagai komponen berinteraksi.
- Memulai dengan Formasi Danau— Dapatkan informasi tentang prasyarat, dan selesaikan tugas penyiapan penting.
- AWS Lake Formation tutorial— Ikuti step-by-step tutorial untuk mempelajari cara menggunakan Lake Formation.
- Keamanan di AWS Lake Formation— Pahami bagaimana Anda dapat membantu mengamankan akses ke data di Lake Formation.

Memulai dengan Formasi Danau

Jika Anda belum mendaftar AWS atau membutuhkan bantuan untuk memulai, pastikan untuk menyelesaikan tugas-tugas berikut.

Topik

- Selesaikan inisial AWS tugas konfigurasi
- Penyiapan AWS Lake Formation
- Memutakhirkan izin AWS Glue data ke model AWS Lake Formation
- AWS Lake Formation dan antarmuka titik akhir VPC ()AWS PrivateLink

Selesaikan inisial AWS tugas konfigurasi

Untuk menggunakan, AWS Lake Formation Anda harus terlebih dahulu menyelesaikan tugas-tugas berikut:

Topik

- Mendaftar untuk Akun AWS

Mendaftar untuk Akun AWS

Untuk memulainya AWS, Anda membutuhkan Akun AWS. Untuk informasi tentang membuat Akun AWS, lihat Mem ulai dengan an Akun AWS di Panduan AWS Account Management Referensi.

Penyiapan AWS Lake Formation

Bagian berikut memberikan informasi tentang pengaturan Formasi Danau untuk pertama kalinya. Tidak semua topik di bagian ini diperlukan untuk mulai menggunakan Formasi Danau. Anda dapat menggunakan petunjuk untuk mengatur model izin Lake Formation untuk mengelola AWS Glue Data Catalog objek dan lokasi data yang ada di Amazon Simple Storage Service (Amazon S3).

1. Buat administrator danau data
2. Ubah model izin default atau gunakan mode akses hybrid
3. the section called “Konfigurasikan lokasi Amazon S3 untuk danau data Anda”

4. the section called “Tetapkan izin ke pengguna Lake Formation”
5. the section called “Mengintegrasikan Pusat Identitas IAM”
6. the section called “(Opsional) Pengaturan penyaringan data eksternal”
7. the section called “(Opsional) Berikan akses ke kunci enkripsi Katalog Data”
8. (Opsional) Membuat peran IAM untuk alur kerja

Bagian ini menunjukkan cara mengatur sumber daya Formasi Danau dengan dua cara berbeda:

- Menggunakan AWS CloudFormation template
- Menggunakan konsol Formasi Danau

Untuk mengatur Formasi Danau menggunakan AWS konsol, bukaBuat administrator danau data.

Menyiapkan sumber daya Formasi Danau menggunakan CloudFormation templat

> **Note**
>
> Tum CloudFormation pukan melakukan langkah 1 hingga 6 di atas, kecuali langkah 2 dan 5. Lakukan Ubah model izin default atau gunakan mode akses hybrid dan the section called “Mengintegrasikan Pusat Identitas IAM” secara manual dari konsol Formasi Danau.

1. Masuk ke AWS CloudFormation konsol di https://console.aws.amazon.com/cloudformation sebagai administrator IAM di Wilayah AS Timur (Virginia Utara).
2. Pilih Launch Stack.
3. Pilih Berikutnya pada layar Buat tumpukan.
4. Masukkan Nama tumpukan.
5. Untuk DatalakeAdminName dan DatalakeAdminPassword, masukkan nama pengguna dan kata sandi Anda untuk pengguna admin danau data.
6. Untuk DatalakeUser1Name dan DatalakeUser1Password, masukkan nama pengguna dan kata sandi Anda untuk pengguna analis data lake.
7. Untuk DataLakeBucketName, masukkan nama bucket baru yang akan dibuat.
8. Pilih Berikutnya.

9. Pada halaman berikutnya, pilih `I acknowledge that CloudFormation might create IAM resources with custom names` dan pilih Ber ikutnya.
10. Tinjau detail di halaman terakhir dan pilih Saya mengakui bahwa AWS CloudFormation mungkin membuat sumber daya IAM.
11. Pilih Buat.

    Pembuatan tumpukan bisa memakan waktu hingga dua menit.

Pembersihan sumber daya

Jika Anda ingin membersihkan sumber daya tum CloudFormation pukan:

1. De-register bucket Amazon S3 yang dibuat dan didaftarkan oleh stack Anda sebagai lokasi danau data.
2. Hapus St CloudFormation ack. Ini akan menghapus semua sumber daya yang dibuat oleh tumpukan.

Buat administrator danau data

Administrator danau data awalnya adalah satu-satunya pengguna atau peran AWS Identity and Access Management (IAM) yang dapat memberikan izin Formasi Danau pada lokasi data dan sumber daya Katalog Data kepada prinsipal mana pun (termasuk diri sendiri). Untuk informasi selengkapnya tentang kemampuan administrator data lake, lihatIzin Formasi Danau Implisit. Secara default, Lake Formation memungkinkan Anda membuat hingga 30 administrator danau data.

Anda dapat membuat administrator danau data menggunakan konsol Formasi Danau atau pen `PutDataLakeSettings` goperasian API Formasi Danau.

Izin berikut diperlukan untuk membuat administrator danau data. `Administrator`Pengguna memiliki izin ini secara implisit.

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

Jika Anda memberikan `AWSLakeFormationDataAdmin` kebijakan kepada pengguna, pengguna tersebut tidak akan dapat membuat pengguna administrator Lake Formation tambahan.

Untuk membuat administrator danau data (konsol)

1. Jika pengguna yang akan menjadi administrator data lake belum ada, gunakan konsol IAM untuk membuatnya. Jika tidak, pilih pengguna yang sudah ada yang akan menjadi administrator danau data.

   > **Note**
   >
   > Sebaiknya Anda tidak memilih pengguna administratif IAM (pengguna dengan kebijakan terkel `AdministratorAccess` AWS ola) untuk menjadi administrator danau data.

   Lampirkan kebijakan AWS terkelola berikut ke pengguna:

Kebijakan	Wajib?	Catatan
`AWSLakeFormationDataAdmin`	Wajib	Izin administrator danau data dasar. Kebijakan AWS terkelola ini berisi penolakan eksplisit untuk operasi Lake Formation API, `PutDataLakeSetting` yang membatasi pengguna membuat administrator danau data baru.
`AWSGlueConsoleFullAccess` , `CloudWatchLogsReadOnlyAccess`	Opsional	Lampirkan kebijakan ini jika administrator danau data akan memecahkan masalah alur kerja yang dibuat dari cetak biru Lake Formation. Kebijakan ini memungkinkan administrator danau data untuk melihat informasi pemecahan masalah di AWS Glue konsol dan Amazon CloudWatch Logs konsol. Untuk informasi tentang alur kerja, lihatthe section called “Mengimpor data menggunakan alur kerja”.

Kebijakan	Wajib?	Catatan
`AWSLakeFormationCrossAccountManager`	Opsional	Lampirkan kebijakan ini untuk mengaktifkan administrator danau data untuk memberikan dan mencabut izin lintas akun pada sumber daya Katalog Data. Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation.
`AmazonAthenaFullAccess`	Opsional	Lampirkan kebijakan ini jika administrator danau data akan menjalankan kueri di Amazon Athena.

2. Lampirkan kebijakan sebaris berikut, yang memberikan izin administrator danau data untuk membuat peran terkait layanan Lake Formation. Nama yang disarankan untuk kebijakan tersebut adalah`LakeFormationSLR`.

   Peran terkait layanan memungkinkan administrator danau data untuk lebih mudah mendaftarkan lokasi Amazon S3 dengan Lake Formation. Untuk informasi lebih lanjut tentang peran terkait layanan Formasi Danau, lihat. the section called “Menggunakan Peran Terkait Layanan”

   > **Important**
   >
   > Dalam semua kebijakan berikut, ganti *<account-id>* dengan nomor AWS akun yang valid.

```
{
    "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. (Opsional) Lampirkan kebijakan `PassRole` sebaris berikut ke pengguna. Kebijakan ini memungkinkan administrator danau data untuk membuat dan menjalankan alur kerja. Iz `iam:PassRole` in memungkinkan alur kerja untuk mengambil peran `LakeFormationWorkflowRole` untuk membuat crawler dan pekerjaan, dan untuk melampirkan peran ke crawler dan pekerjaan yang dibuat. Nama yang disarankan untuk kebijakan tersebut adalah`UserPassRole`.

> **Important**
>
> Ganti *<account-id>* dengan nomor re AWS kening yang valid.

JSON

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

```
}
```

4. (Opsional) Lampirkan kebijakan inline tambahan ini jika akun Anda akan memberikan atau menerima izin Formasi Danau lintas akun. Kebijakan ini memungkinkan administrator danau data untuk melihat dan menerima undangan berbagi sumber daya AWS Resource Access Manager (AWS RAM). Selain itu, untuk administrator danau data di akun AWS Organizations manajemen, kebijakan menyertakan izin untuk mengaktifkan hibah lintas akun kepada organisasi. Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation.

   Nama yang disarankan untuk kebijakan tersebut adalah`RAMAccess`.

   JSON

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

5. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/ dan masuk sebagai pengguna dengan kebijakan ter `AdministratorAccess` AWS kelola.
6. Jika jendela Welcome to Lake Formation muncul, pilih pengguna IAM yang Anda buat atau pilih di Langkah 1, lalu pilih Mulai.
7. Jika Anda tidak melihat jendela Welcome to Lake Formation, maka lakukan langkah-langkah berikut untuk mengonfigurasi Administrator Formasi Danau.

   a. Di panel navigasi, di bawah Administrasi, pilih Per an dan tugas administratif. Di bagian Administrator danau data pada halaman konsol, pilih Tam bah.

   b. Di kotak dialog Tambah administrator, di bawah Jenis akses, pilih Administrator danau data.

c. Untuk pengguna dan peran IAM, pilih pengguna IAM yang Anda buat atau pilih di Langkah 1, lalu pilih Simpan.

Ubah model izin default atau gunakan mode akses hybrid

Formasi Danau dimulai dengan pengaturan “Gunakan hanya kontrol akses IAM” yang diaktifkan untuk kompatibilitas dengan AWS Glue Data Catalog perilaku yang ada. Pengaturan ini memungkinkan Anda mengelola akses ke data Anda di danau data dan metadatanya melalui kebijakan IAM dan kebijakan bucket Amazon S3.

Untuk memudahkan transisi izin danau data dari model IAM dan Amazon S3 ke izin Lake Formation, kami sarankan Anda menggunakan mode akses hibrida untuk Katalog Data. Dengan mode akses hybrid, Anda memiliki jalur tambahan di mana Anda dapat mengaktifkan izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu pengguna atau beban kerja lain yang ada.

Untuk informasi selengkapnya, lihat Mode akses hibrid.

Nonaktifkan pengaturan default untuk memindahkan semua pengguna tabel yang ada ke Lake Formation dalam satu langkah.

Important

Jika Anda memiliki AWS Glue Data Catalog database dan tabel yang ada, jangan ikuti instruksi di bagian ini. Sebagai gantinya, ikuti instruksi dithe section called “Memutakhirkan izin AWS Glue data ke model Lake Formation”.

Warning

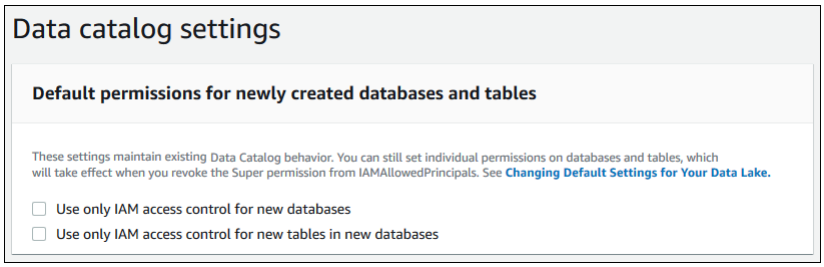
Jika Anda memiliki otomatisasi yang membuat database dan tabel di Katalog Data, langkah-langkah berikut dapat menyebabkan otomatisasi dan pekerjaan ekstraksi, transformasi, dan muat (ETL) hilir gagal. Lanjutkan hanya setelah Anda mengubah proses yang ada atau memberikan izin Formasi Danau eksplisit kepada prinsipal yang diperlukan. Untuk informasi tentang izin Formasi Danau, lihatthe section called “Referensi izin Lake Formation”.

Untuk mengubah pengaturan Katalog Data default

1. Lanjutkan di konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/. Pastikan Anda masuk sebagai pengguna dengan kebijakan ter `AdministratorAccess` AWS kelola.
2. Ubah pengaturan Katalog Data:

   a. Di panel navigasi, di bawah Administrasi, pilih pengaturan Katalog Data.

   b. Kosongkan kedua kotak centang dan pilih Simpan.



3. Cabut `IAMAllowedPrincipals` izin untuk pembuat database.

   a. Di panel navigasi, di bawah Administrasi, pilih Per an dan tugas administratif.

   b. Di halaman konsol Peran dan tugas administratif, di bagian Pembuat database, pilih `IAMAllowedPrincipals` grup, lalu pilih Cab ut.

      Ko tak dialog Cabut izin muncul, menunjukkan bahwa `IAMAllowedPrincipals` memiliki izin Buat database.

   c. Pilih Cabut.

Tetapkan izin ke pengguna Lake Formation

Buat pengguna untuk memiliki akses ke danau data di AWS Lake Formation. Pengguna ini memiliki hak istimewa terkecil untuk menanyakan data lake.

Untuk informasi selengkapnya tentang membuat pengguna atau grup, lihat Identi tas IAM di Panduan Pengguna IAM.

Untuk melampirkan izin ke pengguna non-administrator untuk mengakses data Lake Formation

1. Buka konsol IAM di https://console.aws.amazon.com/iam dan masuk sebagai pengguna dengan kebijakan terkel `AdministratorAccess` AWS ola.
2. Pilih Pengguna atau Grup Pengguna.

3. Dalam daftar, pilih nama pengguna atau grup untuk menyematkan kebijakan.

   Pilih Izin.

4. Pilih Tambah izin, dan pilih Lamp irkan kebijakan secara langsung. Masukkan `Athena` di bidang teks kebijakan filter. Dalam daftar hasil, centang kotak untuk`AmazonAthenaFullAccess`.

5. Pilih tombol Buat kebijakan. Di halaman Buat kebijakan, pilih tab JSON. Salin dan tempel kode berikut ke editor kebijakan.

   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. Pilih Berikutnya hingga Anda melihat halaman Kebijakan Tinjauan. Masukkan nama untuk kebijakan, misalnya,`DatalakeUserBasic`. Pilih Buat kebijakan, lalu tutup tab Kebijakan atau jendela browser.

Konfigurasikan lokasi Amazon S3 untuk danau data Anda

Untuk menggunakan Lake Formation untuk mengelola dan mengamankan data di danau data Anda, Anda harus terlebih dahulu mendaftarkan lokasi Amazon S3. Saat Anda mendaftarkan lokasi, jalur Amazon S3 dan semua folder di bawah jalur tersebut terdaftar, yang memungkinkan Lake Formation menerapkan izin tingkat penyimpanan. Ketika pengguna meminta data dari mesin terintegrasi seperti Amazon Athena, Lake Formation menyediakan akses data daripada menggunakan izin pengguna.

Saat mendaftarkan lokasi, Anda menentukan peran IAM yang memberikan read/write izin di lokasi tersebut. Lake Formation mengambil peran itu ketika menyediakan kredenSIAL sementara ke AWS layanan terintegrasi yang meminta akses ke data di lokasi Amazon S3 terdaftar. Anda dapat menentukan peran terkait layanan Lake Formation (SLR) atau membuat peran Anda sendiri.

Gunakan peran khusus dalam situasi berikut:

- Anda berencana untuk mempublikasikan metrik di Amazon CloudWatch Logs. Peran yang ditentukan pengguna harus menyertakan kebijakan untuk menambahkan CloudWatch log di Log dan metrik penerbitan selain izin SLR. Untuk contoh kebijakan sebaris yang memberikan CloudWatch izin yang diperlukan, lihatPersyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.
- Lokasi Amazon S3 ada di akun yang berbeda. Lihat perinciannya di the section called “Mendaftarkan lokasi Amazon S3 di tempat lain AWS akun”.
- Lokasi Amazon S3 berisi data yang dienkripsi dengan Kunci yang dikelola AWS. Untuk detailnya, lihat Mendaftarkan lokasi Amazon S3 terenkripsi dan Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening.
- Anda berencana untuk mengakses lokasi Amazon S3 menggunakan Amazon EMR. Untuk informasi selengkapnya tentang persyaratan peran, lihat peran IAM untuk Formasi Danau di Panduan Manajemen EMR Amazon.

Peran yang Anda pilih harus memiliki izin yang diperlukan, seperti yang dijelaskan dalamPersyaratan untuk peran yang digunakan untuk mendaftarkan lokasi. Untuk petunjuk tentang cara mendaftarkan lokasi Amazon S3, lihatMenambahkan lokasi Amazon S3 ke danau data Anda.

(Opsional) Pengaturan penyaringan data eksternal

Jika Anda bermaksud menganalisis dan memproses data di danau data Anda menggunakan mesin kueri pihak ketiga, Anda harus memilih untuk mengizinkan mesin eksternal mengakses data yang

dikelola oleh Lake Formation. Jika Anda tidak ikut serta, mesin eksternal tidak akan dapat mengakses data di lokasi Amazon S3 yang terdaftar di Lake Formation.

Lake Formation mendukung izin tingkat kolom untuk membatasi akses ke kolom tertentu dalam tabel. Layanan analitik terintegrasi seperti Amazon Athena, Amazon Redshift Spectrum, dan Amazon EMR mengambil metadata tabel yang tidak difilter dari. AWS Glue Data Catalog Pemfilteran kolom yang sebenarnya dalam tanggapan kueri adalah tanggung jawab layanan terintegrasi. Adalah tanggung jawab administrator pihak ketiga untuk menangani izin dengan benar untuk menghindari akses tidak sah ke data.

Untuk memilih untuk mengizinkan mesin pihak ketiga mengakses dan memfilter data (konsol)

1. Lanjutkan di konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/. Pastikan Anda masuk sebagai kepala sekolah yang memiliki izin IAM pada operasi `PutDataLakeSettings` API Formasi Danau. Pengguna administrator IAM yang Anda buat Mendaftar untuk Akun AWS memiliki izin ini.
2. Di panel navigasi, di bawah Administrasi, pilih Pengaturan integrasi aplikasi.
3. Pada halaman Pengaturan integrasi aplikasi, lakukan hal berikut:
   a. Centang kotak Iz inkan mesin eksternal memfilter data di lokasi Amazon S3 yang terdaftar di Lake Formation.
   b. Masukkan nilai tag sesi yang ditentukan untuk mesin pihak ketiga.
   c. Untuk ID AWS akun, masukkan ID akun tempat mesin pihak ketiga diizinkan mengakses lokasi yang terdaftar di Lake Formation. Tek an Enter setelah setiap ID akun.
   d. Pilih Simpan.

Untuk mengizinkan mesin eksternal mengakses data tanpa validasi tag sesi, lihat Integrasi aplikasi untuk akses tabel penuh

(Opsional) Berikan akses ke kunci enkripsi Katalog Data

Jika AWS Glue Data Catalog dienkripsi, berikan izin AWS Identity and Access Management (IAM) pada AWS KMS kunci kepada kepala sekolah mana pun yang perlu memberikan izin Lake Formation pada database dan tabel Katalog Data.

Lihat informasi selengkapnya di Panduan Developer AWS Key Management Service .

(Opsional) Membuat peran IAM untuk alur kerja

Dengan AWS Lake Formation, Anda dapat mengimpor data Anda menggunakan alur kerja yang dieksekusi oleh AWS Glue crawler. Alur kerja menentukan sumber data dan jadwal untuk mengimpor data ke danau data Anda. Anda dapat dengan mudah menentukan alur kerja menggunakan cetak biru, atau templat yang disediakan Lake Formation.

Saat Anda membuat alur kerja, Anda harus menetapkan peran AWS Identity and Access Management (IAM) yang memberikan Lake Formation izin yang diperlukan untuk menyerap data.

Prosedur berikut mengasumsikan keakraban dengan IAM.

Untuk membuat peran IAM untuk alur kerja

1. Buka konsol IAM di https://console.aws.amazon.com/iam dan masuk sebagai pengguna dengan kebijakan terkel `AdministratorAccess` AWS ola.
2. Di panel navigasi, pilih Per an, lalu Buat peran.
3. Pada halaman Buat peran, pilih AWS layanan, lalu pilih Lem. Pilih Berikutnya.
4. Pada halaman Tambah izin, cari kebijakan ter AWSGlueServiceRole kelola, dan pilih kotak centang di sebelah nama kebijakan dalam daftar. Kemudian selesaikan wizard Buat peran, beri nama peran`LFWorkflowRole`. Untuk menyelesaikan, pilih Buat peran.
5. Kembali ke halaman Peran, cari`LFWorkflowRole`, dan pilih nama peran.
6. Pada halaman Ring kasan peran, di bawah tab Izin, pilih Buat kebijakan sebaris. Pada layar Buat kebijakan, navigasikan ke tab JSON, dan tambahkan kebijakan sebaris berikut. Nama yang disarankan untuk kebijakan tersebut adalah`LakeFormationWorkflow`.

   > **Important**
   >
   > Dalam kebijakan berikut, ganti *<account-id>* dengan Akun AWS nomor yang valid.

   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"
            ]
        }
    ]
}
```

Berikut ini adalah deskripsi singkat tentang izin dalam kebijakan ini:

- `lakeformation:GetDataAccess`memungkinkan pekerjaan yang dibuat oleh alur kerja untuk menulis ke lokasi target.
- `lakeformation:GrantPermissions`memungkinkan alur kerja untuk memberikan `SELECT` izin pada tabel target.
- `iam:PassRole`memungkinkan layanan untuk mengambil peran `LakeFormationWorkflowRole` untuk membuat crawler dan pekerjaan (contoh alur kerja), dan untuk melampirkan peran ke crawler dan pekerjaan yang dibuat.

7. Pastikan peran tersebut `LakeFormationWorkflowRole` memiliki dua kebijakan yang terlampir.
8. Jika Anda menelan data yang berada di luar lokasi danau data, tambahkan kebijakan sebaris yang memberikan izin untuk membaca data sumber.

Memutakhirkan izin AWS Glue data ke model AWS Lake Formation

AWS Lake Formation izin mengaktifkan kontrol akses berbutir halus untuk data di danau data Anda. Anda dapat menggunakan model izin Lake Formation untuk mengelola AWS Glue Data Catalog objek dan lokasi data yang ada di Amazon Simple Storage Service (Amazon S3).

Model izin Lake Formation menggunakan izin kasar AWS Identity and Access Management (IAM) untuk akses layanan API. Lake Formation menggunakan Pemfilteran data dan keamanan tingkat sel

di Lake Formation fungsionalitas untuk membatasi akses tabel di kolom, baris, dan tingkat sel untuk pengguna dan aplikasinya. Sebagai perbandingan, AWS Glue model memberikan akses data melalui kebijakan IAM berbasis Identitas dan sumber daya.

Untuk beralih, ikuti langkah-langkah dalam panduan ini.

Untuk informasi selengkapnya, lihat Ikhtisar izin Lake Formation.

Tentang izin default

Untuk mempertahankan kompatibilitas mundur denganAWS Glue, secara default, AWS Lake Formation memberikan `Super` izin kepada `IAMAllowedPrincipals` grup pada semua sumber daya Katalog AWS Glue Data yang ada, dan memberikan `Super` izin pada sumber daya Katalog Data baru jika pengaturan kontrol akses Use only IAM diaktifkan. Hal ini secara efektif menyebabkan akses ke sumber daya Katalog Data dan lokasi Amazon S3 dikendalikan semata-mata oleh kebijakan AWS Identity and Access Management (IAM). `IAMAllowedPrincipals`Grup ini mencakup setiap pengguna IAM dan peran yang diizinkan mengakses objek Katalog Data Anda oleh kebijakan IAM Anda. `Super`Izin memungkinkan kepala sekolah untuk melakukan setiap operasi Lake Formation yang didukung pada database atau tabel yang diberikan.

Anda dapat mulai menggunakan Lake Formation untuk mengelola akses ke data Anda dengan mendaftarkan lokasi sumber daya Katalog Data yang ada di Lake Formation atau dengan menggunakan mode akses hybrid. Saat mendaftarkan lokasi Amazon S3 dalam mode akses hybrid, Anda dapat mengaktifkan izin Lake Formation dengan memilih prinsipal untuk database dan tabel di bawah lokasi tersebut.

Untuk memudahkan transisi izin data lake dari model IAM dan Amazon S3 ke izin Lake Formation, kami sarankan Anda untuk menggunakan mode akses hybrid untuk Katalog Data. Dengan mode akses hybrid, Anda memiliki jalur tambahan di mana Anda dapat mengaktifkan izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu pengguna atau beban kerja lain yang ada.

Untuk informasi selengkapnya, lihat Mode akses hibrid.

Nonaktifkan pengaturan Katalog Data default untuk memindahkan semua pengguna tabel yang ada ke Lake Formation dalam satu langkah.

Untuk mulai menggunakan izin Lake Formation dengan database dan tabel Katalog AWS Glue Data yang ada, Anda harus melakukan hal berikut:

1. Tentukan izin IAM pengguna Anda yang ada untuk setiap database dan tabel.

2. Replikasi izin ini di Lake Formation.
3. Untuk setiap lokasi Amazon S3 yang berisi data:
   a. Cabut `Super` izin dari `IAMAllowedPrincipals` grup pada setiap sumber daya Katalog Data yang mereferensikan lokasi tersebut.
   b. Daftarkan lokasi dengan Lake Formation.
4. Bersihkan kebijakan IAM kontrol akses berbutir halus yang ada.

> **Important**
>
> Untuk menambahkan pengguna baru saat dalam proses transisi Katalog Data Anda, Anda harus menyiapkan AWS Glue izin granular di IAM seperti sebelumnya. Anda juga harus mereplikasi izin tersebut di Lake Formation seperti yang dijelaskan di bagian ini. Jika pengguna baru memiliki kebijakan IAM berbutir kasar yang dijelaskan dalam panduan ini, mereka dapat mencantumkan database atau tabel apa pun yang memiliki izin yang diberikan. `Super` `IAMAllowedPrincipals` Mereka juga dapat melihat metadata untuk sumber daya tersebut.

Ikuti langkah-langkah di bagian ini untuk meningkatkan ke model izin Lake Formation.

Topik

- Langkah 1: Daftar izin pengguna dan peran yang ada
- Langkah 2: Siapkan izin Lake Formation yang setara
- Langkah 3: Berikan izin IAM kepada pengguna untuk menggunakan Lake Formation
- Langkah 4: Alihkan penyimpanan data Anda ke model izin Lake Formation
- Langkah 5: Amankan sumber daya Katalog Data baru
- Langkah 6: Beri pengguna kebijakan IAM baru untuk akses data lake future
- Langkah 7: Bersihkan kebijakan IAM yang ada

Langkah 1: Daftar izin pengguna dan peran yang ada

Untuk mulai menggunakan AWS Lake Formation izin dengan AWS Glue database dan tabel yang ada, Anda harus terlebih dahulu menentukan izin pengguna yang ada.

> **Important**
>
> Sebelum Anda mulai, pastikan bahwa Anda telah menyelesaikan tugas diMemulai.

Topik

- Menggunakan operasi API
- Menggunakan Konsol Manajemen AWS
- Menggunakan AWS CloudTrail

Menggunakan operasi API

Gunakan operasi ListPoliciesGrantingServiceAccessAPI AWS Identity and Access Management (IAM) untuk menentukan kebijakan IAM yang dilampirkan pada setiap prinsipal (pengguna atau peran). Dari kebijakan yang ditampilkan dalam hasil, Anda dapat menentukan izin IAM yang diberikan kepada prinsipal. Anda harus menjalankan API untuk setiap prinsipal secara terpisah.

Example

AWS CLI Contoh berikut mengembalikan kebijakan yang dilampirkan ke penggunaglue_user1.

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

Perintah mengembalikan hasil yang mirip dengan berikut ini.

```
{
    "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
}
```

Menggunakan Konsol Manajemen AWS

Anda juga dapat melihat informasi ini di konsol AWS Identity and Access Management (IAM), di tab Access Advisor di halaman Ringkasan pengguna atau peran:

1. Buka konsol IAM di https://console.aws.amazon.com/iam/.
2. Di panel navigasi, pilih Pengguna atau Peran.
3. Pilih nama dalam daftar untuk membuka halaman Ringkasannya, dan pilih tab Access Advisor.
4. Periksa setiap kebijakan untuk menentukan kombinasi database, tabel, dan tindakan yang setiap pengguna memiliki izin.

   Ingatlah untuk memeriksa peran selain pengguna selama proses ini karena pekerjaan pemrosesan data Anda mungkin mengambil peran untuk mengakses data.

Menggunakan AWS CloudTrail

Cara lain untuk menentukan izin yang ada adalah dengan mencari panggilan AWS Glue API di mana `additionaleventdata` bidang log berisi `insufficientLakeFormationPermissions` entri. AWS CloudTrail Entri ini mencantumkan database dan tabel tempat pengguna memerlukan izin Lake Formation untuk mengambil tindakan yang sama.

Ini adalah log akses data, sehingga mereka tidak dijamin untuk menghasilkan daftar lengkap pengguna dan izin mereka. Sebaiknya pilih rentang waktu yang luas untuk menangkap sebagian besar pola akses data pengguna Anda, misalnya, beberapa minggu atau bulan.

Untuk informasi selengkapnya, lihat Melihat CloudTrail Acara dengan Riwayat Acara di Panduan AWS CloudTrail Pengguna.

Selanjutnya, Anda dapat mengatur izin Lake Formation agar sesuai dengan izin. AWS Glue Lihat Langkah 2: Siapkan izin Lake Formation yang setara.

Langkah 2: Siapkan izin Lake Formation yang setara

Menggunakan informasi yang dikumpulkanLangkah 1: Daftar izin pengguna dan peran yang ada, berikan AWS Lake Formation izin untuk mencocokkan AWS Glue izin. Gunakan salah satu metode berikut untuk melakukan hibah:

- Gunakan konsol Lake Formation atau AWS CLI.

  Lihat the section called “Memberikan izin data”.
- Gunakan operasi `GrantPermissions` atau `BatchGrantPermissions` API.

  Lihat Izin APIs.

Untuk informasi selengkapnya, lihat Ikhtisar izin Lake Formation.

Setelah mengatur izin Lake Formation, lanjutkan keLangkah 3: Berikan izin IAM kepada pengguna untuk menggunakan Lake Formation.

Langkah 3: Berikan izin IAM kepada pengguna untuk menggunakan Lake Formation

Untuk menggunakan model AWS Lake Formation izin, kepala sekolah harus memiliki izin AWS Identity and Access Management (IAM) di Lake Formation. APIs

Buat kebijakan berikut di IAM dan lampirkan ke setiap pengguna yang membutuhkan akses ke data lake Anda. Sebutkan kebijakan `LakeFormationDataAccess`.

JSON

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

```
            "Resource": "*"
        }
    ]
}
```

Selanjutnya, tingkatkan ke Lake Formation mengizinkan satu lokasi data pada satu waktu. Lihat Langkah 4: Alihkan penyimpanan data Anda ke model izin Lake Formation.

Langkah 4: Alihkan penyimpanan data Anda ke model izin Lake Formation

Tingkatkan ke Lake Formation mengizinkan satu lokasi data pada satu waktu. Untuk melakukan itu, ulangi seluruh bagian ini sampai Anda telah mendaftarkan semua jalur Amazon Simple Storage Service (Amazon S3) yang direferensikan oleh Katalog Data Anda.

Topik

- Verifikasi izin Lake Formation
- Mengamankan sumber daya Katalog Data yang ada
- Aktifkan izin Lake Formation untuk lokasi Amazon S3

Verifikasi izin Lake Formation

Sebelum mendaftarkan lokasi, lakukan langkah verifikasi untuk memastikan bahwa kepala sekolah yang benar memiliki izin Lake Formation yang diperlukan, dan bahwa tidak ada izin Lake Formation yang diberikan kepada kepala sekolah yang seharusnya tidak memilikinya. Menggunakan operasi Lake Formation `GetEffectivePermissionsForPath` API, identifikasi sumber daya Katalog Data yang mereferensikan lokasi Amazon S3, bersama dengan prinsipal yang memiliki izin pada sumber daya tersebut.

AWS CLI Contoh berikut menampilkan database dan tabel Katalog Data yang mereferensikan bucket Amazon `products` S3.

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

Perhatikan `profile` opsinya. Kami menyarankan Anda menjalankan perintah sebagai administrator danau data.

Berikut ini adalah kutipan dari hasil yang dikembalikan.

```
{
        "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

Jika Katalog AWS Glue Data Anda dienkripsi, hanya `GetEffectivePermissionsForPath` mengembalikan database dan tabel yang dibuat atau dimodifikasi setelah ketersediaan umum Lake Formation.

Mengamankan sumber daya Katalog Data yang ada

Selanjutnya, cabut `Super` izin dari `IAMAllowedPrincipals` setiap tabel dan database yang Anda identifikasi untuk lokasi tersebut.

Warning

Jika Anda memiliki otomatisasi yang membuat database dan tabel di Katalog Data, langkah-langkah berikut dapat menyebabkan pekerjaan otomatisasi dan hilir ekstrak, transformasi, dan pemuatan (ETL) gagal. Lanjutkan hanya setelah Anda memodifikasi proses yang ada

atau memberikan izin Formasi Danau eksplisit ke kepala sekolah yang diperlukan. Untuk informasi tentang izin Lake Formation, lihatthe section called “Referensi izin Lake Formation”.

Untuk mencabut **`Super`** dari **`IAMAllowedPrincipals`** atas meja

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, pilih Tabel.
3. Pada halaman Tabel, pilih tombol radio di sebelah tabel yang diinginkan.
4. Pada menu Tindakan, pilih Cabut.
5. Dalam kotak dialog Cabut izin, di daftar pengguna dan peran IAM, gulir ke bawah ke judul Grup, dan pilih Prinsip. IAMAllowed
6. Di bawah izin Tabel, pastikan Super dipilih, lalu pilih Batalkan.

Untuk mencabut **`Super`** dari **`IAMAllowedPrincipals`** database

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, pilih Basis data.
3. Pada halaman Database, pilih tombol radio di sebelah database yang diinginkan.
4. Di menu Tindakan, pilih Edit.
5. Pada halaman Edit database, hapus Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini, lalu pilih Simpan.
6. Kembali ke halaman Database, pastikan bahwa database masih dipilih, dan kemudian pada menu Tindakan, pilih Cabut.
7. Dalam kotak dialog Cabut izin, di daftar pengguna dan peran IAM, gulir ke bawah ke judul Grup, dan pilih Prinsip. IAMAllowed
8. Di bawah Izin database, pastikan Super dipilih, lalu pilih Cabut.

Aktifkan izin Lake Formation untuk lokasi Amazon S3

Selanjutnya, daftarkan lokasi Amazon S3 dengan Lake Formation. Untuk melakukan ini, Anda dapat menggunakan proses yang dijelaskan dalamMenambahkan lokasi Amazon S3 ke danau data Anda.

Atau, gunakan operasi `RegisterResource` API seperti yang dijelaskan dalamPenjual kredensi APIs.

> **Note**
>
> Jika lokasi induk terdaftar, Anda tidak perlu mendaftarkan lokasi anak.

Setelah Anda menyelesaikan langkah-langkah ini dan menguji apakah pengguna Anda dapat mengakses data mereka, Anda telah berhasil meningkatkan ke izin Lake Formation. Lanjutkan dengan langkah selanjutnya,Langkah 5: Amankan sumber daya Katalog Data baru.

Langkah 5: Amankan sumber daya Katalog Data baru

Selanjutnya, amankan semua sumber daya Katalog Data baru dengan mengubah pengaturan Katalog Data default. Matikan opsi untuk menggunakan kontrol akses hanya AWS Identity and Access Management (IAM) untuk database dan tabel baru.

> **Warning**
>
> Jika Anda memiliki otomatisasi yang membuat database dan tabel di Katalog Data, langkah-langkah berikut dapat menyebabkan pekerjaan otomatisasi dan hilir ekstrak, transformasi, dan pemuatan (ETL) gagal. Lanjutkan hanya setelah Anda memodifikasi proses yang ada atau memberikan izin Formasi Danau eksplisit ke kepala sekolah yang diperlukan. Untuk informasi tentang izin Lake Formation, lihatthe section called “Referensi izin Lake Formation”.

Untuk mengubah pengaturan Katalog Data default

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai pengguna administratif IAM (pengguna `Administrator` atau pengguna lain dengan kebijakan `AdministratorAccess` AWS terkelola).
2. Pada panel navigasi, silakan pilih Pengaturan.
3. Pada halaman Pengaturan katalog data, kosongkan kedua kotak centang, lalu pilih Simpan.

Langkah selanjutnya adalah memberi pengguna akses ke database atau tabel tambahan di masa depan. Lihat Langkah 6: Beri pengguna kebijakan IAM baru untuk akses data lake future.

Langkah 6: Beri pengguna kebijakan IAM baru untuk akses data lake future

Untuk memberi pengguna akses ke database atau tabel Katalog Data tambahan di masa mendatang, Anda harus memberi mereka kebijakan inline berbutir kasar AWS Identity and Access Management (IAM) berikut. Sebutkan kebijakan `GlueFullReadAccess`.

> **Important**
>
> Jika Anda melampirkan kebijakan ini ke pengguna sebelum mencabut `Super` dari setiap database dan tabel `IAMAllowedPrincipals` di Katalog Data Anda, pengguna tersebut dapat melihat semua metadata untuk sumber daya apa pun yang `Super` diberikan. `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**
>
> Kebijakan inline yang ditetapkan dalam langkah ini dan langkah sebelumnya berisi izin IAM minimal. Untuk kebijakan yang disarankan untuk administrator data lake, analis data, dan persona lainnya, lihat. the section called “Referensi personas Lake Formation dan izin IAM”

Selanjutnya, lanjutkan keLangkah 7: Bersihkan kebijakan IAM yang ada.

Langkah 7: Bersihkan kebijakan IAM yang ada

Setelah menyiapkan AWS Lake Formation izin dan membuat serta melampirkan kebijakan kontrol akses kasar AWS Identity and Access Management (IAM), selesaikan langkah terakhir berikut:

- Hapus dari pengguna, grup, dan peran kebijakan IAM kontrol akses berbutir halus lama yang Anda replikasi di Lake Formation.

Dengan melakukan ini, Anda memastikan bahwa prinsipal tersebut tidak lagi memiliki akses langsung ke data di Amazon Simple Storage Service (Amazon S3). Anda kemudian dapat mengelola akses data lake untuk prinsipal tersebut sepenuhnya melalui Lake Formation.

AWS Lake Formation dan antarmuka titik akhir VPC ()AWS PrivateLink

Amazon VPC adalah AWS layanan yang dapat Anda gunakan untuk meluncurkan AWS sumber daya di jaringan virtual yang Anda tentukan. Dengan VPC, Anda memiliki kendali terhadap pengaturan jaringan, seperti rentang alamat IP, subnet, tabel rute, dan pintu masuk jaringan.

Jika Anda menggunakan Amazon Virtual Private Cloud (Amazon VPC) untuk meng-host AWS sumber daya Anda, Anda dapat membuat koneksi pribadi antara VPC dan Lake Formation. Anda menggunakan koneksi ini sehingga Lake Formation dapat berkomunikasi dengan sumber daya di VPC Anda tanpa melalui internet publik.

Anda dapat membuat koneksi pribadi antara VPC Anda dan AWS Lake Formation dengan membuat antarmuka VPC endpoint. Endpoint antarmuka didukung oleh AWS PrivateLink, teknologi yang memungkinkan Anda mengakses Lake Formation secara pribadi APIs tanpa gateway internet, perangkat NAT, koneksi VPN, atau koneksi. Direct Connect Instans di VPC Anda tidak memerlukan

alamat IP publik untuk berkomunikasi dengan Lake Formation. APIs Lalu lintas antara VPC dan Lake Formation Anda tidak meninggalkan jaringan Amazon.

Setiap titik akhir antarmuka diwakili oleh satu atau beberapa Antarmuka Jaringan Elastis di subnet Anda.

Untuk informasi selengkapnya, lihat Antarmuka VPC endpoint (AWS PrivateLink) dalam Panduan Pengguna Amazon VPC.

Pertimbangan untuk titik akhir VPC Lake Formation

Sebelum menyiapkan titik akhir VPC antarmuka untuk Lake Formation, pastikan Anda meninjau properti dan batasan titik akhir Antarmuka di Panduan Pengguna Amazon VPC.

Lake Formation mendukung panggilan ke semua tindakan API-nya dari VPC Anda. Anda dapat menggunakan Lake Formation dengan titik akhir VPC di semua Wilayah AWS yang mendukung titik akhir Lake Formation dan Amazon VPC.

Membuat titik akhir VPC antarmuka untuk Lake Formation

Anda dapat membuat titik akhir VPC untuk layanan Lake Formation menggunakan konsol Amazon VPC atau (). AWS Command Line Interface AWS CLI Untuk informasi selengkapnya, lihat Membuat titik akhir antarmuka dalam Panduan Pengguna Amazon VPC.

Buat titik akhir VPC untuk Lake Formation menggunakan nama layanan berikut:

- com.amazonaws. *region*.lakeformasi

Jika Anda mengaktifkan DNS pribadi untuk titik akhir, Anda dapat membuat permintaan API ke Lake Formation menggunakan nama DNS default untuk Wilayah, misalnya,. `lakeformation.us-east-1.amazonaws.com`

Untuk informasi selengkapnya, lihat Mengakses layanan melalui titik akhir antarmuka dalam Panduan Pengguna Amazon VPC.

Membuat kebijakan titik akhir VPC untuk Lake Formation

Lake Formation mendukung kebijakan titik akhir VPC. Kebijakan endpoint adalah kebijakan berbasis sumber daya yang Anda lampirkan ke titik akhir VPC untuk mengontrol AWS prinsipal mana yang dapat menggunakan titik akhir untuk mengakses layanan. AWS

Anda dapat melampirkan kebijakan titik akhir ke titik akhir VPC Anda yang mengontrol akses ke Lake Formation. Kebijakan titik akhir menentukan informasi berikut:

- Prinsipal yang dapat melakukan tindakan.
- Tindakan yang dapat dilakukan.
- Sumber daya yang menjadi target tindakan.

Untuk informasi selengkapnya, lihat Mengontrol Akses ke Layanan dengan titik akhir VPC dalam Panduan Pengguna Amazon VPC.

Contoh: Kebijakan titik akhir VPC untuk tindakan Lake Formation

Contoh kebijakan titik akhir VPC berikut untuk Lake Formation memungkinkan penjual kredenal menggunakan izin Lake Formation. Anda dapat menggunakan kebijakan ini untuk menjalankan kueri menggunakan izin Lake Formation dari klaster Amazon Redshift atau cluster Amazon EMR yang terletak di subnet pribadi.

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

> **Note**
>
> Jika Anda tidak melampirkan kebijakan saat membuat titik akhir, kebijakan default yang mengizinkan akses penuh ke layanan akan dilampirkan.

Untuk informasi selengkapnya, lihat topik ini di dokumentasi Amazon VPC:

- Apa yang Dimaksud Amazon VPC?
- Buat Endpoint Antarmuka
- Gunakan kebijakan titik akhir VPC

AWS Lake Formation tutorial

Tutorial berikut disusun menjadi tiga trek dan memberikan step-by-step instruksi tentang cara membangun data lake, menelan data, berbagi, dan mengamankan data lake menggunakan AWS Lake Formation:

1. Membangun data lake dan menelan data: Belajar membangun data lake dan menggunakan cetak biru untuk memindahkan, menyimpan, membuat katalog, membersihkan, dan mengatur data Anda. Anda juga akan belajar mengatur tabel yang diatur. Tabel yang diatur adalah jenis tabel Amazon S3 baru yang mendukung transaksi atom, konsisten, terisolasi, dan tahan lama (ACID).

   Sebelum Anda mulai, pastikan Anda telah menyelesaikan langkah-langkahnyaMemulai dengan Formasi Danau.

   - Membuat danau data dari sebuah AWS CloudTrail sumber

     Buat dan muat data lake pertama Anda dengan menggunakan CloudTrail log Anda sendiri sebagai sumber data.

   - Membuat data lake dari sumber JDBC di Lake Formation

     Buat data lake dengan menggunakan salah satu penyimpanan data yang dapat diakses JDBC Anda, seperti database relasional, sebagai sumber data.

2. Mengamankan data lake: Pelajari cara menggunakan kontrol akses berbasis tag dan tingkat baris untuk mengamankan dan mengelola akses ke danau data Anda secara efektif.

   - Menyiapkan izin untuk format penyimpanan tabel terbuka di Lake Formation

     Tutorial ini menunjukkan cara mengatur izin untuk format tabel transaksional open source (Apache Iceberg, Apache Hudi, dan tabel Linux Foundation Delta Lake) di Lake Formation.

   - Mengelola data lake menggunakan kontrol akses berbasis tag Lake Formation

     Pelajari cara mengelola akses ke data dalam data lake menggunakan kontrol akses berbasis tag di Lake Formation.

   - Mengamankan data lake dengan kontrol akses tingkat baris

     Pelajari cara menyiapkan izin tingkat baris yang memungkinkan Anda membatasi akses ke baris tertentu berdasarkan kepatuhan data dan kebijakan tata kelola di Lake Formation.

3. Berbagi data: Pelajari cara berbagi data dengan aman Akun AWS menggunakan kontrol akses berbasis tag (TBAC) dan mengelola izin terperinci pada kumpulan data yang dibagikan di antaranya. Akun AWS
   - Berbagi data lake menggunakan kontrol akses berbasis tag Lake Formation dan sumber daya bernama

     Dalam tutorial ini, Anda mempelajari cara membagikan data Anda dengan aman Akun AWS menggunakan Lake Formation.
   - Berbagi data lake menggunakan kontrol akses berbutir halus Lake Formation

     Dalam tutorial ini, Anda mempelajari cara berbagi kumpulan data dengan cepat dan mudah menggunakan Lake Formation saat mengelola beberapa Akun AWS dengan. AWS Organizations

Topik

- Membuat danau data dari sebuah AWS CloudTrail sumber
- Membuat data lake dari sumber JDBC di Lake Formation
- Menyiapkan izin untuk format penyimpanan tabel terbuka di Lake Formation
- Mengelola data lake menggunakan kontrol akses berbasis tag Lake Formation
- Mengamankan data lake dengan kontrol akses tingkat baris
- Berbagi data lake menggunakan kontrol akses berbasis tag Lake Formation dan sumber daya bernama
- Berbagi data lake menggunakan kontrol akses berbutir halus Lake Formation

Membuat danau data dari sebuah AWS CloudTrail sumber

Tutorial ini memandu Anda melalui tindakan yang harus diambil pada konsol Lake Formation untuk membuat dan memuat data lake pertama Anda dari AWS CloudTrail sumber.

High-level langkah-langkah untuk membuat danau data

1. Daftarkan jalur Amazon Simple Storage Service (Amazon S3) sebagai data lake.
2. Berikan izin Lake Formation untuk menulis ke Katalog Data dan ke lokasi Amazon S3 di data lake.
3. Buat database untuk mengatur tabel metadata dalam Katalog Data.

4. Gunakan cetak biru untuk membuat alur kerja. Jalankan alur kerja untuk menyerap data dari sumber data.
5. Siapkan izin Lake Formation Anda untuk memungkinkan orang lain mengelola data di Katalog Data dan data lake.
6. Siapkan Amazon Athena untuk menanyakan data yang Anda impor ke danau data Amazon S3 Anda.
7. Untuk beberapa jenis penyimpanan data, siapkan Amazon Redshift Spectrum untuk menanyakan data yang Anda impor ke data lake Amazon S3 Anda.

Topik

- Audiens yang dituju
- Prasyarat
- Langkah 1: Buat pengguna analis data
- Langkah 2: Tambahkan izin untuk membaca AWS CloudTrail log ke peran alur kerja
- Langkah 3: Buat bucket Amazon S3 untuk data lake
- Langkah 4: Daftarkan jalur Amazon S3
- Langkah 5: Berikan izin lokasi data
- Langkah 6: Buat database di Katalog Data
- Langkah 7: Berikan izin data
- Langkah 8: Gunakan cetak biru untuk membuat alur kerja
- Langkah 9: Jalankan alur kerja
- Langkah 10: Berikan SELECT pada tabel
- Langkah 11: Kueri data lake Menggunakan Amazon Athena

Audiens yang dituju

Tabel berikut mencantumkan peran yang digunakan dalam tutorial ini untuk membuat danau data.

Audiens yang dituju

Peran	Deskripsi
Administrator IAM	Memiliki kebijakan AWS terkelola:`AdministratorAccess` . Dapat membuat peran IAM dan bucket Amazon S3.
Administrator danau data	Pengguna yang dapat mengakses katalog data, membuat database, dan memberikan izin Lake Formation kepada pengguna lain. Memiliki izin IAM lebih sedikit daripada administrator IAM, tetapi cukup untuk mengelola data lake.
Analis data	Pengguna yang dapat menjalankan kueri terhadap data lake. Hanya memiliki izin yang cukup untuk menjalankan kueri.
Peran alur kerja	Berperan dengan kebijakan IAM yang diperlukan untuk menjalankan alur kerja. Untuk informasi selengkapnya, lihat (Opsional) Membuat peran IAM untuk alur kerja.

Prasyarat

Sebelum Anda memulai:

- Pastikan Anda telah menyelesaikan tugas diPenyiapan AWS Lake Formation.
- Ketahui lokasi CloudTrail log Anda.
- Athena mengharuskan persona analis data untuk membuat bucket Amazon S3 untuk menyimpan hasil kueri sebelum menggunakan Athena.

Keakraban dengan AWS Identity and Access Management (IAM) diasumsikan. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Langkah 1: Buat pengguna analis data

Pengguna ini memiliki set izin minimum untuk menanyakan data lake.

1. Buka konsol IAM di https://console.aws.amazon.com/iam. Masuk sebagai pengguna dengan kebijakan `AdministratorAccess` AWS terkelola.
2. Buat pengguna bernama `datalake_user` dengan pengaturan berikut:
   - Aktifkan Konsol Manajemen AWS akses.
   - Tetapkan kata sandi dan tidak memerlukan pengaturan ulang kata sandi.
   - Lampirkan kebijakan `AmazonAthenaFullAccess` AWS terkelola.
   - Lampirkan kebijakan inline berikut. Sebutkan kebijakan `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": "*"
        }
    ]
}
```

Langkah 2: Tambahkan izin untuk membaca AWS CloudTrail log ke peran alur kerja

1. Lampirkan kebijakan inline berikut ke peran`LakeFormationWorkflowRole`. Kebijakan memberikan izin untuk membaca AWS CloudTrail log Anda. Sebutkan kebijakan `DatalakeGetCloudTrail`.

Untuk membuat `LakeFormationWorkflowRole` peran, lihat(Opsional) Membuat peran IAM untuk alur kerja.

> **Important**
>
> Ganti *<your-s3-cloudtrail-bucket>* dengan lokasi Amazon S3 data Anda CloudTrail .

JSON

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

2. Verifikasi bahwa ada tiga kebijakan yang melekat pada peran tersebut.

Langkah 3: Buat bucket Amazon S3 untuk data lake

Buat bucket Amazon S3 yang akan menjadi lokasi root danau data Anda.

1. Buka konsol Amazon S3 di https://console.aws.amazon.com/s3/dan masuk sebagai pengguna dengan akses administratif.
2. Pilih Buat ember, dan buka wizard untuk membuat bucket bernama*<yourName>*`-datalake-cloudtrail`, di *<yourName>* mana nama awal dan belakang pertama Anda. Sebagai contoh: `jdoe-datalake-cloudtrail`.

   Untuk petunjuk mendetail tentang cara membuat bucket Amazon S3, lihat Membuat bucket.

Langkah 4: Daftarkan jalur Amazon S3

Daftarkan jalur Amazon S3 sebagai lokasi root danau data Anda.

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Daftar dan konsumsi, pilih Lokasi danau data.
3. Pilih Daftar lokasi dan kemudian Jelajahi.
4. Pilih *<yourName>*`-datalake-cloudtrail` bucket yang Anda buat sebelumnya, terima peran IAM default`AWSServiceRoleForLakeFormationDataAccess`, lalu pilih Daftar lokasi.

   Untuk informasi selengkapnya tentang mendaftarkan lokasi, lihatMenambahkan lokasi Amazon S3 ke danau data Anda.

Langkah 5: Berikan izin lokasi data

Prinsipal harus memiliki izin lokasi data pada lokasi data lake untuk membuat tabel Katalog Data atau database yang mengarah ke lokasi tersebut. Anda harus memberikan izin lokasi data ke peran IAM untuk alur kerja sehingga alur kerja dapat menulis ke tujuan konsumsi data.

1. Di panel navigasi, di bawah Izin, pilih Lokasi data.
2. Pilih Hibah, dan di kotak dialog Hibah izin, buat pilihan ini:

   a. Untuk pengguna dan peran IAM, pilih`LakeFormationWorkflowRole`.

   b. Untuk lokasi Penyimpanan, pilih *<yourName>*`-datalake-cloudtrail` bucket Anda.
3. PilihIzin.

Untuk informasi selengkapnya tentang izin lokasi data, lihatUnderlying data access control.

Langkah 6: Buat database di Katalog Data

Tabel metadata dalam Katalog Data Lake Formation disimpan dalam database.

1. Di panel navigasi, di bawah Katalog data, pilih Database.
2. Pilih Buat database, dan di bawah rincian Database, masukkan nama`lakeformation_cloudtrail`.
3. Biarkan bidang lainnya kosong, dan pilih Buat database.

Langkah 7: Berikan izin data

Anda harus memberikan izin untuk membuat tabel metadata di Katalog Data. Karena alur kerja akan berjalan dengan peran`LakeFormationWorkflowRole`, Anda harus memberikan izin ini ke peran tersebut.

1. Di konsol Lake Formation, di panel navigasi, di bawah Katalog data, pilih Database.
2. Pilih `lakeformation_cloudtrail` database, lalu, dari daftar drop-down Tindakan, pilih Hibah di bawah judul Izin.
3. Di kotak dialog Berikan izin data, buat pilihan ini:
   a. Di bawah Prinsipal, untuk pengguna dan peran IAM, pilih. `LakeFormationWorkflowRole`
   b. Di bawah LF-Tags atau sumber katalog, pilih Sumber daya Katalog Data Bernama.
   c. Untuk Database, Anda harus melihat bahwa `lakeformation_cloudtrail` database sudah ditambahkan.
   d. Di bawah Izin database, pilih Buat tabel, Ubah, dan Jatuhkan, dan hapus Super jika dipilih.
4. PilihIzin.

Untuk informasi selengkapnya tentang pemberian izin Lake Formation, lihat. Mengelola izin Lake Formation

Langkah 8: Gunakan cetak biru untuk membuat alur kerja

Untuk membaca CloudTrail log, memahami strukturnya, membuat tabel yang sesuai di Katalog Data, kita perlu menyiapkan alur kerja yang terdiri dari AWS Glue crawler, pekerjaan, pemicu, dan alur kerja. Cetak biru Lake Formation menyederhanakan proses ini.

Alur kerja menghasilkan pekerjaan, crawler, dan pemicu yang menemukan dan menelan data ke dalam data lake Anda. Anda membuat alur kerja berdasarkan salah satu cetak biru Lake Formation yang telah ditentukan sebelumnya.

1. Di konsol Lake Formation, di panel navigasi, pilih Blueprints di bawah Tertelan, lalu pilih Gunakan cetak biru.
2. Pada halaman Gunakan cetak biru, di bawah Jenis cetak biru, pilih. AWS CloudTrail
3. Di bawah Impor sumber, pilih CloudTrail sumber dan tanggal mulai.
4. Di bawah target Impor, tentukan parameter ini:

Basis data target	`lakeformation_cloudtrail`
Target lokasi penyimpanan	`s3://<yourName> -datalake-cloudtrail`
Format data	Parquet

5. Untuk frekuensi impor, pilih Jalankan sesuai permintaan.
6. Di bawah opsi Impor, tentukan parameter ini:

Nama alur kerja	`lakeformationcloudtrailtest`
Peran IAM	`LakeFormationWorkflowRole`
Awalan tabel	`cloudtrailtest` **Note** Harus huruf kecil.

7. Pilih Buat, dan tunggu konsol melaporkan bahwa alur kerja berhasil dibuat.

> **Tip**
>
> Apakah Anda mendapatkan pesan kesalahan berikut?
> `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...`
> Jika demikian, periksa apakah Anda mengganti *<account-id>* dalam kebijakan inline untuk pengguna administrator data lake dengan nomor AWS akun yang valid.

Langkah 9: Jalankan alur kerja

Karena Anda menetapkan bahwa alur kerja adalah run-on-demand, Anda harus memulai alur kerja secara manual.

- Pada halaman Blueprints, pilih alur kerja, dan pada menu Tindakan **`lakeformationcloudtrailtest`**, pilih Mulai.

  Saat alur kerja berjalan, Anda dapat melihat kemajuannya di kolom Status Last run. Pilih tombol refresh sesekali.

  Status berubah dari RUNNING, ke Discovering, ke Importing, ke COMPLETED.

  Saat alur kerja selesai:

  - Katalog Data akan memiliki tabel metadata baru.
  - CloudTrail Log Anda akan tertelan ke dalam danau data.

  Jika alur kerja gagal, lakukan hal berikut:

  a. Pilih alur kerja, dan pada menu Tindakan, pilih Lihat grafik.

     Alur kerja terbuka di AWS Glue konsol.

  b. Pastikan bahwa alur kerja sudah dipilih, dan pilih tab Riwayat.

  c. Di bawah Riwayat, pilih proses terbaru dan pilih Lihat detail jalankan.

  d. Pilih job atau crawler yang gagal dalam grafik dinamis (runtime), dan tinjau pesan galatnya. Node yang gagal berwarna merah atau kuning.

Langkah 10: Berikan SELECT pada tabel

Anda harus memberikan `SELECT` izin pada tabel Katalog Data baru sehingga analis data dapat melakukan kueri data yang ditunjukkan tabel.

> **Note**
>
> Alur kerja secara otomatis memberikan `SELECT` izin pada tabel yang dibuatnya kepada pengguna yang menjalankannya. Karena administrator data lake menjalankan alur kerja ini, Anda harus memberikan `SELECT` kepada analis data.

1. Di konsol Lake Formation, di panel navigasi, di bawah Katalog data, pilih Database.

2. Pilih `lakeformation_cloudtrail` database, lalu, dari daftar drop-down Tindakan, pilih Hibah di bawah judul Izin.
3. Di kotak dialog Berikan izin data, buat pilihan ini:
   a. Di bawah Prinsipal, untuk pengguna dan peran IAM, pilih. `datalake_user`
   b. Di bawah LF-Tags atau sumber katalog, pilih Sumber daya katalog data bernama.
   c. Untuk Database, `lakeformation_cloudtrail` database harus sudah dipilih.
   d. Untuk Tabel, pilih`cloudtrailtest-cloudtrail`.
   e. Di bawah Izin tabel dan kolom, pilih Pilih.
4. PilihIzin.

Langkah selanjutnya dilakukan sebagai analis data.

Langkah 11: Kueri data lake Menggunakan Amazon Athena

Gunakan Amazon Athena konsol untuk menanyakan CloudTrail data di danau data Anda.

1. Buka konsol Athena di https://console.aws.amazon.com/athena/dan masuk sebagai analis data, pengguna. `datalake_user`
2. Jika perlu, pilih Mulai untuk melanjutkan ke editor kueri Athena.
3. Untuk Sumber Data, pilih AwsDataCatalog.
4. Untuk Database, pilih`lakeformation_cloudtrail`.

   Daftar Tabel terisi.
5. Pada menu overflow (3 titik disusun secara horizontal) di samping tabel, pilih tabel Pratinjau **`cloudtrailtest-cloudtrail`**, lalu pilih Jalankan.

   Kueri berjalan dan menampilkan 10 baris data.

   Jika Anda belum pernah menggunakan Athena sebelumnya, Anda harus terlebih dahulu mengonfigurasi lokasi Amazon S3 di konsol Athena untuk menyimpan hasil kueri. `datalake_user`Harus memiliki izin yang diperlukan untuk mengakses bucket Amazon S3 yang Anda pilih.

> **Note**
>
> Sekarang setelah Anda menyelesaikan tutorial, berikan izin data dan izin lokasi data ke kepala sekolah di organisasi Anda.

Membuat data lake dari sumber JDBC di Lake Formation

Tutorial ini memandu Anda melalui langkah-langkah yang harus diambil pada AWS Lake Formation konsol untuk membuat dan memuat data lake pertama Anda dari sumber JDBC menggunakan Lake Formation.

Topik

- Audiens yang dituju
- Prasyarat tutorial JDBC
- Langkah 1: Buat pengguna analis data
- Langkah 2: Buat koneksi di AWS Glue
- Langkah 3: Buat bucket Amazon S3 untuk data lake
- Langkah 4: Daftarkan jalur Amazon S3
- Langkah 5: Berikan izin lokasi data
- Langkah 6: Buat database di Katalog Data
- Langkah 7: Berikan izin data
- Langkah 8: Gunakan cetak biru untuk membuat alur kerja
- Langkah 9: Jalankan alur kerja
- Langkah 10: Berikan SELECT pada tabel
- Langkah 11: Kueri data lake menggunakan Amazon Athena
- Langkah 12: Kueri data di danau data menggunakan Amazon Redshift Spectrum
- Langkah 13: Berikan atau cabut izin Lake Formation menggunakan Amazon Redshift Spectrum

Audiens yang dituju

Tabel berikut mencantumkan peran yang digunakan dalam tutorial AWS Lake Formation JDBC ini.

Peran	Deskripsi
Administrator IAM	Pengguna yang dapat membuat pengguna dan peran AWS Identity and Access Managemen t (IAM) serta bucket Amazon Simple Storage Service (Amazon S3). Memiliki kebijakan yang `AdministratorAccess` AWS dikelola.
Administrator danau data	Pengguna yang dapat mengakses Katalog Data, membuat database, dan memberikan izin Lake Formation kepada pengguna lain. Memiliki izin IAM lebih sedikit daripada administrator IAM, tetapi cukup untuk mengelola data lake.
Analis data	Pengguna yang dapat menjalankan kueri terhadap data lake. Hanya memiliki izin yang cukup untuk menjalankan kueri.
Peran alur kerja	Peran dengan kebijakan IAM yang diperlukan untuk menjalankan alur kerja.

Untuk informasi tentang prasyarat untuk menyelesaikan tutorial, lihat. Prasyarat tutorial JDBC

Prasyarat tutorial JDBC

Sebelum Anda memulai tutorial AWS Lake Formation JDBC, pastikan bahwa Anda telah melakukan hal berikut:

- Selesaikan tugas dalam Memulai dengan Formasi Danau.
- Tentukan penyimpanan JDBC-accessible data yang ingin Anda gunakan untuk tutorial.
- Kumpulkan informasi yang diperlukan untuk membuat AWS Glue koneksi tipe JDBC. Objek Katalog Data ini menyertakan URL ke penyimpanan data, kredensi login, dan jika penyimpanan data dibuat di Amazon Virtual Private Cloud (Amazon VPC), informasi konfigurasi tambahan. VPC-specific Untuk informasi selengkapnya, lihat Mendefinisikan Koneksi di Katalog AWS Glue Data di Panduan AWS Glue Pengembang.

Tutorial mengasumsikan bahwa Anda sudah familiar dengan AWS Identity and Access Management (IAM). Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Untuk memulai, lanjutkan kethe section called “Langkah 1: Buat pengguna analis data”.

Langkah 1: Buat pengguna analis data

Pada langkah ini, Anda membuat pengguna AWS Identity and Access Management (IAM) untuk menjadi analis data untuk data lake Anda. AWS Lake Formation

Pengguna ini memiliki set izin minimum untuk menanyakan data lake.

1. Buka konsol IAM di https://console.aws.amazon.com/iam. Masuk sebagai pengguna dengan kebijakan `AdministratorAccess` AWS terkelola.
2. Buat pengguna bernama `datalake_user` dengan pengaturan berikut:
   - Aktifkan Konsol Manajemen AWS akses.
   - Tetapkan kata sandi dan tidak memerlukan pengaturan ulang kata sandi.
   - Lampirkan kebijakan `AmazonAthenaFullAccess` AWS terkelola.
   - Lampirkan kebijakan inline berikut. Sebutkan kebijakan `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": "*"
```

```
        }
    ]
}
```

Langkah 2: Buat koneksi di AWS Glue

> **Note**
>
> Lewati langkah ini jika Anda sudah memiliki AWS Glue koneksi ke sumber data JDBC Anda.

AWS Lake Formation mengakses sumber data JDBC melalui koneksi. AWS Glue Koneksi adalah objek Katalog Data yang berisi semua informasi yang diperlukan untuk terhubung ke sumber data. Anda dapat membuat koneksi menggunakan AWS Glue konsol.

Untuk membuat koneksi

1. Buka AWS Glue konsol dihttps://console.aws.amazon.com/glue/, dan masuk sebagai pengguna dengan akses administratif.
2. Pada panel navigasi, di Katalog data, pilih Koneksi.
3. Pada halaman Konektor, pilih Buat koneksi.
4. Pada halaman Pilih sumber data, pilih JDBC sebagai jenis koneksi. Lalu pilih Selanjutnya.
5. Lanjutkan melalui wizard koneksi dan simpan koneksi.

   Untuk informasi tentang membuat sambungan, lihat properti koneksi AWS Glue JDBC di Panduan AWS Glue Pengembang.

Langkah 3: Buat bucket Amazon S3 untuk data lake

Pada langkah ini, Anda membuat bucket Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) yang akan menjadi lokasi root danau data Anda.

1. Buka konsol Amazon S3 di https://console.aws.amazon.com/s3/dan masuk sebagai pengguna dengan akses administratif.
2. Pilih Buat ember, dan buka wizard untuk membuat bucket bernama`<yourName>-datalake-tutorial`, di `<yourName>` mana nama awal dan belakang pertama Anda. Sebagai contoh: `jdoe-datalake-tutorial`.

Untuk petunjuk mendetail tentang cara membuat bucket Amazon S3, lihat Bagaimana Cara Membuat Bucket S3? di Panduan Pengguna Layanan Penyimpanan Sederhana Amazon.

Langkah 4: Daftarkan jalur Amazon S3

Pada langkah ini, Anda mendaftarkan jalur Amazon Simple Storage Service (Amazon S3) sebagai lokasi root danau data Anda.

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
3. Pilih Daftarkan lokasi, lalu pilih Jelajahi.
4. Pilih *<yourName>*`-datalake-tutorial` bucket yang Anda buat sebelumnya, terima peran IAM default`AWSServiceRoleForLakeFormationDataAccess`, lalu pilih Daftar lokasi.

   Untuk informasi selengkapnya tentang mendaftarkan lokasi, lihatMenambahkan lokasi Amazon S3 ke danau data Anda.

Langkah 5: Berikan izin lokasi data

Prinsipal harus memiliki izin lokasi data pada lokasi data lake untuk membuat tabel Katalog Data atau database yang mengarah ke lokasi tersebut. Anda harus memberikan izin lokasi data ke peran IAM untuk alur kerja sehingga alur kerja dapat menulis ke tujuan konsumsi data.

1. Di konsol Lake Formation, di panel navigasi, di bawah Izin, pilih Lokasi data.
2. Pilih Hibah, dan di kotak dialog Hibah izin, lakukan hal berikut:

   a. Untuk pengguna dan peran IAM, pilih`LakeFormationWorkflowRole`.

   b. Untuk lokasi Penyimpanan, pilih *<yourName>*`-datalake-tutorial` bucket Anda.
3. PilihIzin.

Untuk informasi selengkapnya tentang izin lokasi data, lihatUnderlying data access control.

Langkah 6: Buat database di Katalog Data

Tabel metadata dalam Katalog Data Lake Formation disimpan dalam database.

1. Di konsol Lake Formation, di panel navigasi, di bawah Katalog data, pilih Database.
2. Pilih Buat database, dan di bawah rincian Database, masukkan nama`lakeformation_tutorial`.
3. Biarkan bidang lainnya kosong, dan pilih Buat database.

Langkah 7: Berikan izin data

Anda harus memberikan izin untuk membuat tabel metadata di Katalog Data. Karena alur kerja berjalan dengan peran`LakeFormationWorkflowRole`, Anda harus memberikan izin ini ke peran tersebut.

1. Di konsol Lake Formation, di panel navigasi, di bawah Izin, pilih Izin danau data.
2. Pilih Hibah, dan di kotak dialog Hibah izin data, lakukan hal berikut:
   a. Di bawah Prinsipal, untuk pengguna dan peran IAM, pilih. `LakeFormationWorkflowRole`
   b. Di bawah LF-Tags atau sumber katalog, pilih Sumber daya katalog data bernama.
   c. Untuk Database, pilih database yang Anda buat sebelumnya. `lakeformation_tutorial`
   d. Di bawah Izin database, pilih Buat tabel, Ubah, dan Jatuhkan, dan hapus Super jika dipilih.
3. PilihIzin.

Untuk informasi selengkapnya tentang pemberian izin Lake Formation, lihat. Ikhtisar izin Lake Formation

Langkah 8: Gunakan cetak biru untuk membuat alur kerja

AWS Lake Formation Alur kerja menghasilkan AWS Glue pekerjaan, crawler, dan pemicu yang menemukan dan menelan data ke dalam data lake Anda. Anda membuat alur kerja berdasarkan salah satu cetak biru Lake Formation yang telah ditentukan sebelumnya.

1. Di konsol Lake Formation, di panel navigasi, pilih Blueprints, lalu pilih Use blueprint.
2. Pada halaman Gunakan cetak biru, di bawah Jenis cetak biru, pilih snapshot Database.
3. Di bawah Impor sumber, untuk koneksi Database, pilih koneksi yang baru saja Anda buat`datalake-tutorial`, atau pilih sambungan yang ada untuk sumber data Anda.
4. Untuk jalur data Sumber, masukkan jalur untuk menelan data, dalam formulir*<database>*/*<schema>*/*<table>*.

Anda dapat mengganti wildcard persen (%) untuk skema atau tabel. Untuk database yang mendukung skema, masukkan*<database>*/*<schema>*/% untuk mencocokkan semua tabel di dalamnya. *<schema>* *<database>* Oracle Database dan MySQL tidak mendukung skema di jalur; sebagai gantinya, masukkan /%. *<database>* Untuk Oracle Database, *<database>* adalah pengenal sistem (SID).

Misalnya, jika database Oracle memiliki `orcl` SID-nya, masukkan `orcl/%` untuk mencocokkan semua tabel yang pengguna ditentukan dalam koneksi JDCB memiliki akses ke.

> **Important**
>
> Bidang ini peka terhadap huruf besar dan kecil.

5. Di bawah target Impor, tentukan parameter ini:

Basis data target	`lakeformation_tutorial`
Target lokasi penyimpanan	`s3://`*<yourName>* `-datalake-tutorial`
Format data	(Pilih Parket atau CSV)

6. Untuk frekuensi impor, pilih Jalankan sesuai permintaan.
7. Di bawah opsi Impor, tentukan parameter ini:

Nama alur kerja	`lakeformationjdbctest`
Peran IAM	`LakeFormationWorkflowRole`
Awalan tabel	`jdbctest` **Note** Harus huruf kecil.

8. Pilih Buat, dan tunggu konsol melaporkan bahwa alur kerja berhasil dibuat.

> **Tip**
>
> Apakah Anda mendapatkan pesan kesalahan berikut?
> `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...`
> Jika demikian, periksa apakah Anda mengganti *<account-id>* dalam kebijakan inline untuk pengguna administrator data lake dengan nomor AWS akun yang valid.

Langkah 9: Jalankan alur kerja

Karena Anda menetapkan bahwa alur kerja adalah run-on-demand, Anda harus memulai alur kerja secara manual. AWS Lake Formation

1. Di konsol Lake Formation, pada halaman Blueprints, pilih alur kerja. `lakeformationjdbctest`
2. Pilih Tindakan, lalu pilih Mulai.
3. Saat alur kerja berjalan, lihat kemajuannya di kolom Status Last run. Pilih tombol refresh sesekali.

   Status berubah dari RUNNING, ke Discovering, ke Importing, ke COMPLETED.

   Saat alur kerja selesai:

   - Katalog Data memiliki tabel metadata baru.
   - Data Anda tertelan ke danau data.

   Jika alur kerja gagal, lakukan hal berikut:

   a. Pilih alur kerja. Pilih Tindakan, lalu pilih Lihat grafik.

      Alur kerja terbuka di AWS Glue konsol.

   b. Pilih alur kerja dan pilih tab History.

   c. Pilih run terbaru dan pilih View run details.

d. Pilih job atau crawler yang gagal dalam grafik dinamis (runtime), dan tinjau pesan galatnya. Node yang gagal berwarna merah atau kuning.

Langkah 10: Berikan SELECT pada tabel

Anda harus memberikan `SELECT` izin pada tabel Katalog Data baru AWS Lake Formation agar analis data dapat melakukan kueri data yang ditunjukkan tabel.

> **Note**
>
> Alur kerja secara otomatis memberikan `SELECT` izin pada tabel yang dibuatnya kepada pengguna yang menjalankannya. Karena administrator data lake menjalankan alur kerja ini, Anda harus memberikan `SELECT` kepada analis data.

1. Di konsol Lake Formation, di panel navigasi, di bawah Izin, pilih Izin danau data.
2. Pilih Hibah, dan di kotak dialog Hibah izin data, lakukan hal berikut:
   a. Di bawah Prinsipal, untuk pengguna dan peran IAM, pilih. `datalake_user`
   b. Di bawah LF-Tags atau sumber katalog, pilih Sumber daya katalog data bernama.
   c. Untuk Database, pilih`lakeformation_tutorial`.

      Daftar Tabel terisi.
   d. Untuk Tabel, pilih satu atau beberapa tabel dari sumber data Anda.
   e. Di bawah Izin tabel dan kolom, pilih Pilih.
3. PilihIzin.

Langkah selanjutnya dilakukan sebagai analis data.

Langkah 11: Kueri data lake menggunakan Amazon Athena

Gunakan Amazon Athena konsol untuk menanyakan data di danau data Anda.

1. Buka konsol Athena di https://console.aws.amazon.com/athena/, dan masuk sebagai analis data, pengguna. `datalake_user`
2. Jika perlu, pilih Mulai untuk melanjutkan ke editor kueri Athena.

3. Untuk Sumber Data, pilih AwsDataCatalog.
4. Untuk Database, pilih`lakeformation_tutorial`.

   Daftar Tabel terisi.
5. Di menu pop-up di samping salah satu tabel, pilih tabel Pratinjau.

   Kueri berjalan dan menampilkan 10 baris data.

Langkah 12: Kueri data di danau data menggunakan Amazon Redshift Spectrum

Anda dapat mengatur Amazon Redshift Spectrum untuk menanyakan data yang Anda impor ke Amazon Simple Storage Service (Amazon S3) data lake. Pertama, buat peran AWS Identity and Access Management (IAM) yang digunakan untuk meluncurkan cluster Amazon Redshift dan untuk menanyakan data Amazon S3. Kemudian, berikan peran ini `Select` izin pada tabel yang ingin Anda kueri. Kemudian, berikan izin pengguna untuk menggunakan editor kueri Amazon Redshift. Terakhir, buat cluster Amazon Redshift dan jalankan kueri.

Anda membuat klaster sebagai administrator, dan menanyakan klaster sebagai analis data.

Untuk informasi selengkapnya tentang Amazon Redshift Spectrum, lihat Menggunakan Amazon Redshift Spectrum untuk Menanyakan Data Eksternal di Panduan Pengembang Database Amazon Redshift.

Untuk mengatur izin untuk menjalankan kueri Amazon Redshift

1. Buka konsol IAM di https://console.aws.amazon.com/iam/. Masuk sebagai pengguna dengan kebijakan `AdministratorAccess` AWS terkelola.
2. Di panel navigasi, pilih Kebijakan.

   Jika ini pertama kalinya Anda memilih Kebijakan, akan muncul halaman Selamat Datang di Kebijakan Terkelola. Pilih Memulai.
3. Pilih Buat kebijakan.
4. Pilih tab JSON.
5. Tempel di dokumen kebijakan JSON berikut.

```
{
```

```
    "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. Setelah selesai, pilih Tinjau untuk meninjau kebijakan. Validator kebijakan melaporkan kesalahan sintaksis.
7. Pada halaman Kebijakan tinjau, masukkan Nama **`RedshiftLakeFormationPolicy`** untuk kebijakan yang Anda buat. Masukkan Deskripsi (opsional). Ulas Ringkasan kebijakan untuk melihat izin yang diberikan oleh kebijakan Anda. Kemudian pilih Buat kebijakan untuk menyimpan pekerjaan Anda.
8. Di panel navigasi konsol IAM, pilih Peran, dan lalu pilih Buat peran.
9. Untuk Pilih entitas tepercaya, pilih AWS layanan.
10. Pilih layanan Amazon Redshift untuk mengambil peran ini.
11. Pilih kasus penggunaan Redshift Customizable untuk layanan Anda. Kemudian, pilih Selanjutnya: Izin.
12. Cari kebijakan izin yang Anda buat`RedshiftLakeFormationPolicy`, dan pilih kotak centang di samping nama kebijakan dalam daftar.
13. Pilih Berikutnya: Tanda.
14. Pilih Berikutnya: Tinjau.
15. Untuk nama Peran, masukkan nama**`RedshiftLakeFormationRole`**.

16. (Opsional) Untuk Deskripsi peran, masukkan deskripsi.
17. Tinjau peran lalu pilih Buat peran.

Untuk memberikan izin **`Pilih`** pada tabel yang akan ditanyakan dalam database Lake Formation

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, pilih Izin danau data, lalu pilih Hibah.
3. Saat diminta, berikan informasi berikut:
   - Untuk pengguna dan peran IAM, pilih peran IAM yang Anda buat. `RedshiftLakeFormationRole` Saat Anda menjalankan Amazon Redshift Query Editor, ia menggunakan peran IAM ini untuk izin ke data.
   - Untuk Database, pilih`lakeformation_tutorial`.

     Daftar tabel terisi.
   - Untuk Tabel, pilih tabel dalam sumber data untuk kueri.
   - Pilih izin Pilih tabel.
4. PilihIzin.

Untuk mengatur Amazon Redshift Spectrum dan menjalankan kueri

1. Buka konsol Amazon Redshift di. https://console.aws.amazon.com/redshift Masuk sebagai penggunaAdministrator.
2. Pilih Buat klaster.
3. Pada halaman Create cluster, masukkan `redshift-lakeformation-demo` untuk pengenal Cluster.
4. Untuk tipe Node, pilih dc2.large.
5. Gulir ke bawah, dan di bawah konfigurasi Database, masukkan atau terima parameter ini:
   - Nama pengguna admin: `awsuser`
   - Kata sandi pengguna admin: (*`Choose a password`*)
6. Perluas izin Cluster, dan untuk peran IAM yang Tersedia, pilih. RedshiftLakeFormationRole Kemudian pilih Tambahkan peran IAM.

7. Jika Anda harus menggunakan port yang berbeda dari nilai default 5439, di samping Konfigurasi tambahan, matikan opsi Gunakan default. Perluas bagian untuk konfigurasi Database, dan masukkan nomor port Database baru.
8. Pilih Buat klaster.

   Halaman Clusters dimuat.
9. Tunggu hingga status klaster menjadi Tersedia. Pilih ikon penyegaran secara berkala.
10. Berikan izin analis data untuk menjalankan kueri terhadap cluster. Untuk melakukannya, selesaikan langkah-langkah berikut.

    a. Buka konsol IAM di https://console.aws.amazon.com/iam/, dan masuk sebagai `Administrator` pengguna.

    b. Di panel navigasi, pilih Pengguna, dan lampirkan kebijakan terkelola berikut ke penggunadatalake_user.

    - `AmazonRedshiftQueryEditor`
    - `AmazonRedshiftReadOnlyAccess`
11. Keluar dari konsol Amazon Redshift dan masuk kembali sebagai pengguna. `datalake_user`
12. Di bilah alat vertikal kiri, pilih ikon EDITOR untuk membuka editor kueri dan terhubung ke cluster. Jika kotak dialog Connect to database muncul, pilih nama cluster`redshift-lakeformation-demo`, dan masukkan nama database**`dev`**, nama pengguna**`awsuser`**, dan kata sandi yang Anda buat. Kemudian pilih Connect to database.

    > **Note**
    >
    > Jika Anda tidak diminta untuk parameter koneksi dan cluster lain sudah dipilih di editor kueri, pilih Ubah Koneksi untuk membuka kotak dialog Connect to database.
13. Di kotak teks New Query 1, masukkan dan jalankan pernyataan berikut untuk memetakan database `lakeformation_tutorial` di Lake Formation ke nama skema Amazon Redshift: `redshift_jdbc`

    > **Important**
    >
    > Ganti *`<account-id>`* dengan nomor AWS akun yang valid, dan *`<region>`* dengan nama AWS Region yang valid (misalnya,`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. Dalam daftar skema di bawah Pilih skema, pilih redshift_jdbc.

    Daftar tabel terisi. Editor kueri hanya menampilkan tabel di mana Anda diberikan izin danau data Lake Formation.

15. Pada menu pop-up di samping nama tabel, pilih Pratinjau data.

    Amazon Redshift mengembalikan 10 baris pertama.

    Anda sekarang dapat menjalankan kueri terhadap tabel dan kolom yang Anda memiliki izin.

Langkah 13: Berikan atau cabut izin Lake Formation menggunakan Amazon Redshift Spectrum

Amazon Redshift mendukung kemampuan untuk memberikan dan mencabut izin Lake Formation pada database dan tabel menggunakan pernyataan SQL yang dimodifikasi. Pernyataan ini mirip dengan pernyataan Amazon Redshift yang ada. Untuk informasi selengkapnya, lihat GRANT dan REVOKE di Panduan Pengembang Database Amazon Redshift.

Menyiapkan izin untuk format penyimpanan tabel terbuka di Lake Formation

AWS Lake Formation mendukung pengelolaan izin akses untuk Open Table Format (OTF) seperti Apache Iceberg, ApacheHudi, dan Linux foundation Delta Lake. Dalam tutorial ini, Anda akan belajar cara membuat Iceberg, Hudi, dan Delta Lake dengan tabel manifes symlink dalam AWS Glue Data Catalog penggunaan AWS Glue, mengatur izin berbutir halus menggunakan Lake Formation, dan kueri data menggunakan Amazon Athena.

> **Note**
>
> AWS Layanan analisis tidak mendukung semua format tabel transaksional. Untuk informasi selengkapnya, lihat Bekerja dengan AWS layanan lain. Tutorial ini secara manual mencakup

pembuatan database baru dan tabel di Katalog Data menggunakan AWS Glue pekerjaan saja.

Tutorial ini mencakup AWS CloudFormation template untuk pengaturan cepat. Anda dapat meninjau dan menyesuaikannya sesuai dengan kebutuhan Anda.

Topik

- Audiens yang dituju
- Prasyarat
- Langkah 1: Menyediakan sumber daya Anda
- Langkah 2: Siapkan izin untuk tabel Iceberg
- Langkah 3: Siapkan izin untuk tabel Hudi
- Langkah 4: Siapkan izin untuk tabel Delta Lake
- Langkah 5: Bersihkan AWS sumber daya

Audiens yang dituju

Tutorial ini ditujukan untuk administrator IAM, administrator data lake, dan analis bisnis. Tabel berikut mencantumkan peran yang digunakan dalam tutorial ini untuk membuat tabel yang diatur menggunakan Lake Formation.

Peran	Deskripsi
Administrator IAM	Pengguna yang dapat membuat pengguna dan peran IAM dan bucket Amazon S3. Memiliki kebijakan yang `AdministratorAccess` AWS dikelola.
Administrator danau data	Pengguna yang dapat mengakses Katalog Data, membuat database, dan memberikan izin Lake Formation kepada pengguna lain. Memiliki izin IAM lebih sedikit daripada administrator IAM, tetapi cukup untuk mengelola data lake.

Peran	Deskripsi
Analis bisnis	Pengguna yang dapat menjalankan kueri terhadap data lake. Memiliki izin untuk menjalankan kueri.

Prasyarat

Sebelum Anda memulai tutorial ini, Anda harus memiliki Akun AWS yang dapat Anda masuk sebagai pengguna dengan izin yang benar. Untuk informasi selengkapnya, lihat Mendaftar untuk Akun AWS.

Tutorial mengasumsikan bahwa Anda terbiasa dengan peran dan kebijakan IAM. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Anda perlu mengatur AWS sumber daya berikut untuk menyelesaikan tutorial ini:

- Pengguna administrator danau data
- Pengaturan danau data Lake Formation
- Mesin Amazon Athena versi 3

Untuk membuat administrator data lake

1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/sebagai pengguna administrator. Anda akan membuat sumber daya di Wilayah AS Timur (Virginia N.) untuk tutorial ini.
2. Di konsol Lake Formation, di panel navigasi, di bawah Izin, pilih Peran dan tugas administratif.
3. Pilih Pilih Administrator di bawah Administrator danau data.
4. Di jendela pop-up, Kelola administrator danau data, di bawah pengguna dan peran IAM, pilih pengguna admin IAM.
5. Pilih Simpan.

Untuk mengaktifkan pengaturan data lake

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/. Di panel navigasi, di bawah Katalog data, pilih Pengaturan. Hapus centang pada hal berikut:

- Gunakan hanya kontrol akses IAM untuk database baru.
- Gunakan hanya kontrol akses IAM untuk tabel baru di database baru.

2. Di bawah Pengaturan versi Cross account, pilih Versi 3 sebagai versi lintas akun.
3. Pilih Simpan.

Untuk meningkatkan mesin Amazon Athena ke versi 3

1. Buka konsol Athena di. https://console.aws.amazon.com/athena/
2. Pilih Workgroup dan pilih workgroup utama.
3. Pastikan bahwa workgroup berada pada versi minimal 3. Jika tidak, edit workgroup, pilih Manual for Upgrade query engine, dan pilih versi 3.
4. Pilih Simpan perubahan.

Langkah 1: Menyediakan sumber daya Anda

Bagian ini menunjukkan cara mengatur AWS sumber daya menggunakan CloudFormation templat.

Untuk membuat sumber daya Anda menggunakan CloudFormation templat

1. Masuk ke AWS CloudFormation konsol di https://console.aws.amazon.com/cloudformationsebagai administrator IAM di Wilayah AS Timur (Virginia N.).
2. Pilih Launch Stack.
3. Pilih Berikutnya di layar Buat tumpukan.
4. Masukkan Nama tumpukan.
5. Pilih Berikutnya.
6. Di halaman berikutnya, pilih Berikutnya.
7. Tinjau detail di halaman akhir dan pilih Saya akui yang AWS CloudFormation mungkin membuat sumber daya IAM.
8. Pilih Buat.

   Pembuatan tumpukan bisa memakan waktu hingga dua menit.

Meluncurkan tumpukan formasi cloud menciptakan sumber daya berikut:

- lf-otf-datalake-123456789012 - Amazon S3 bucket untuk menyimpan data

> **Note**
>
> Id akun yang ditambahkan ke nama bucket Amazon S3 diganti dengan id akun Anda.

- lf-otf-tutorial-123456789012 — Amazon S3 bucket untuk menyimpan hasil query dan skrip pekerjaan AWS Glue
- lficebergdb - Database Gunung Es AWS Glue
- lfhudidb — Database Hudi AWS Glue
- lfdeltadb - Database Delta AWS Glue
- native-iceberg-create — AWS Glue pekerjaan yang membuat tabel Iceberg di Katalog Data
- native-hudi-create — AWS Glue pekerjaan yang membuat tabel Hudi di Katalog Data
- native-delta-create — AWS Glue pekerjaan yang membuat tabel Delta di Katalog Data
- LF-OTF-GlueServiceRole — Peran IAM yang Anda berikan AWS Glue untuk menjalankan pekerjaan. Peran ini memiliki kebijakan yang diperlukan yang dilampirkan untuk mengakses sumber daya seperti Katalog Data, bucket Amazon S3, dll.
- LF-OTF-RegisterRole — Peran IAM untuk mendaftarkan lokasi Amazon S3 dengan Lake Formation. Peran ini `LF-Data-Lake-Storage-Policy` melekat pada peran tersebut.
- lf-consumer-analystuser - pengguna IAM untuk menanyakan data menggunakan Athena
- lf-consumer-analystuser-credentials - Kata sandi untuk pengguna analis data yang disimpan di AWS Secrets Manager

Setelah pembuatan tumpukan selesai, navigasikan ke tab output dan catat nilai untuk:

- AthenaQueryResultLocation — Lokasi Amazon S3 untuk output kueri Athena
- BusinessAnalystUserCredentials — Kata sandi untuk pengguna analis data

  Untuk mengambil nilai kata sandi:

  1. Pilih `lf-consumer-analystuser-credentials` nilainya dengan menavigasi ke konsol Secrets Manager.
  2. Di bagian Nilai rahasia, pilih Ambil nilai rahasia.
  3. Catat nilai rahasia untuk kata sandi.

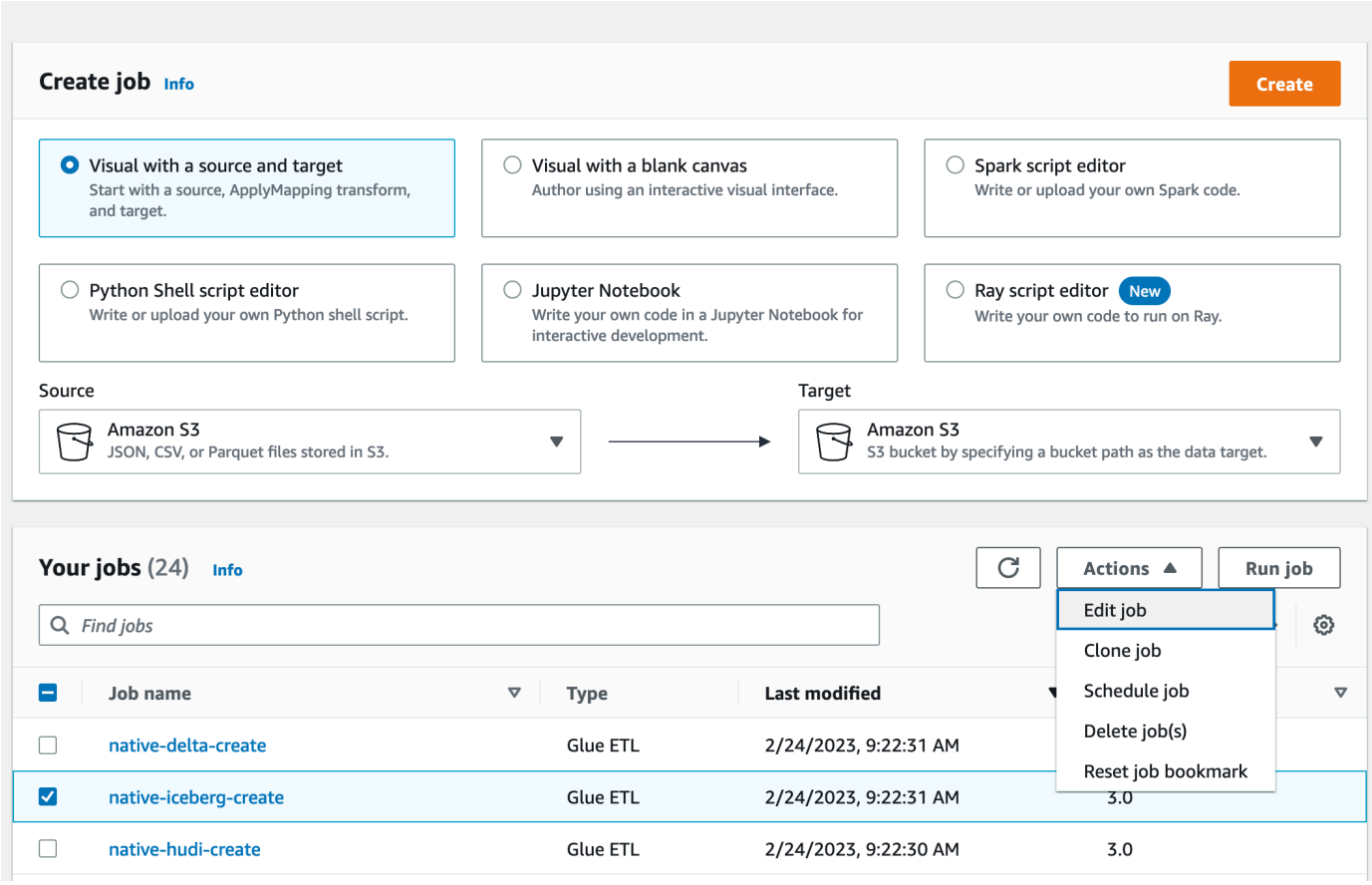
Langkah 2: Siapkan izin untuk tabel Iceberg

Di bagian ini, Anda akan mempelajari cara membuat tabel Gunung Es di AWS Glue Data Catalog, mengatur izin data di AWS Lake Formation, dan kueri data menggunakan Amazon Athena.

Untuk membuat tabel Iceberg

Pada langkah ini, Anda akan menjalankan AWS Glue pekerjaan yang membuat tabel transaksional Iceberg di Katalog Data.

1. Buka AWS Glue konsol https://console.aws.amazon.com/glue/di Wilayah AS Timur (Virginia N.) sebagai pengguna administrator danau data.
2. Pilih pekerjaan dari panel navigasi kiri.
3. Pilih `native-iceberg-create`.



4. Di bawah Tindakan, pilih Edit pekerjaan.
5. Di bawah Job details, perluas properti Advanced, dan centang kotak di samping Use AWS Glue Data Catalog as the Hive metastore untuk menambahkan metadata tabel di. AWS Glue Data Catalog Ini ditentukan AWS Glue Data Catalog sebagai metastore untuk sumber daya Katalog

Data yang digunakan dalam pekerjaan dan memungkinkan izin Lake Formation diterapkan nanti pada sumber daya katalog.

6. Pilih Simpan.
7. Pilih Jalankan. Anda dapat melihat status pekerjaan saat sedang berjalan.

   Untuk informasi selengkapnya tentang AWS Glue lowongan, lihat Bekerja dengan pekerjaan di AWS Glue konsol di Panduan AWS Glue Pengembang.

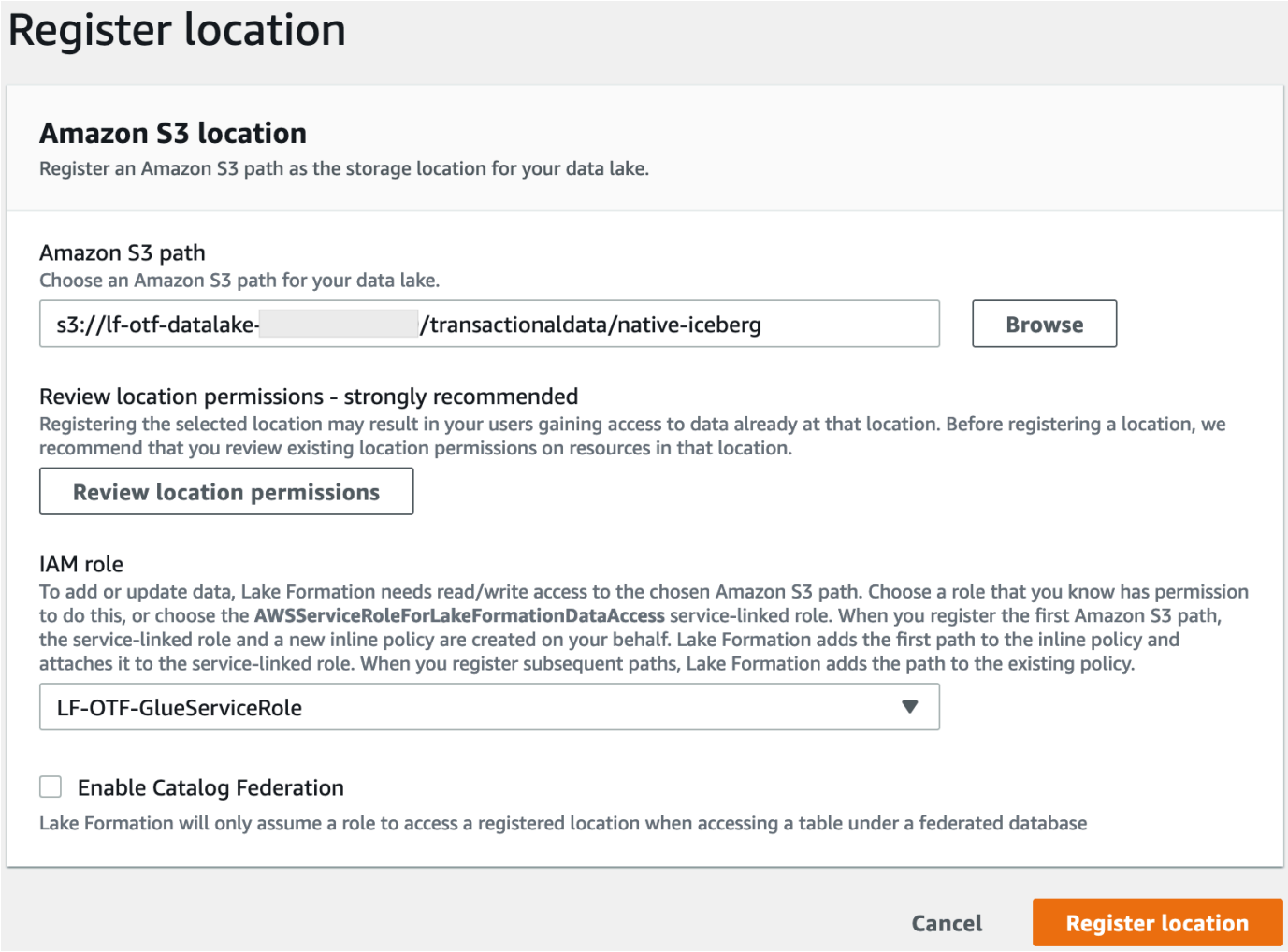
   Pekerjaan ini menciptakan tabel Iceberg bernama `product` dalam database. `lficebergdb` Verifikasi tabel produk di konsol Lake Formation.

Untuk mendaftarkan lokasi data dengan Lake Formation

Selanjutnya, daftarkan jalur Amazon S3 sebagai lokasi danau data Anda.

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/sebagai pengguna administrator danau data.
2. Di panel navigasi, di bawah Daftar dan konsumsi, pilih Lokasi data.
3. Pilih Daftar lokasi.
4. Pada halaman Daftar lokasi, masukkan yang berikut ini:

   - Jalur Amazon S3 - Pilih Jelajahi dan pilih. `lf-otf-datalake-123456789012` Pilih panah perluas di sebelah lokasi root Amazon S3 untuk menavigasi ke lokasi. `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-iceberg`
   - Peran IAM - Pilih **`LF-OTF-RegisterRole`** sebagai peran IAM.
   - Pilih Daftar lokasi.

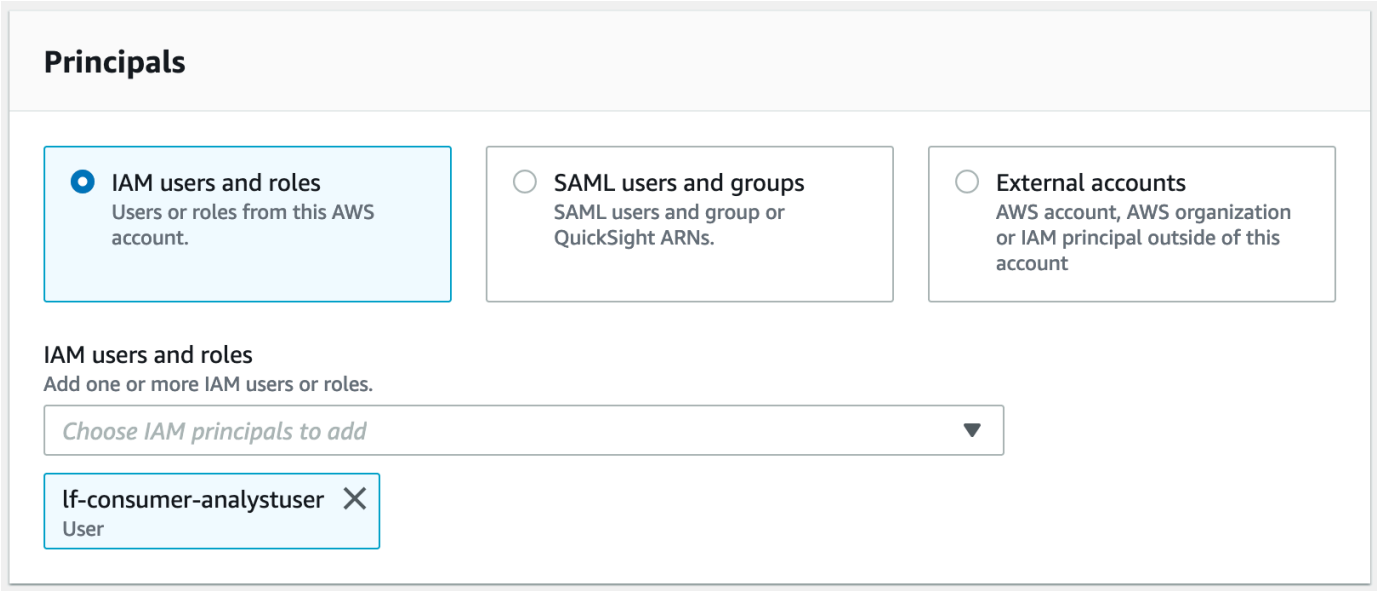


Untuk informasi lebih lanjut tentang mendaftarkan lokasi data dengan Lake Formation, lihat[Menambahkan lokasi Amazon S3 ke danau data Anda](#).

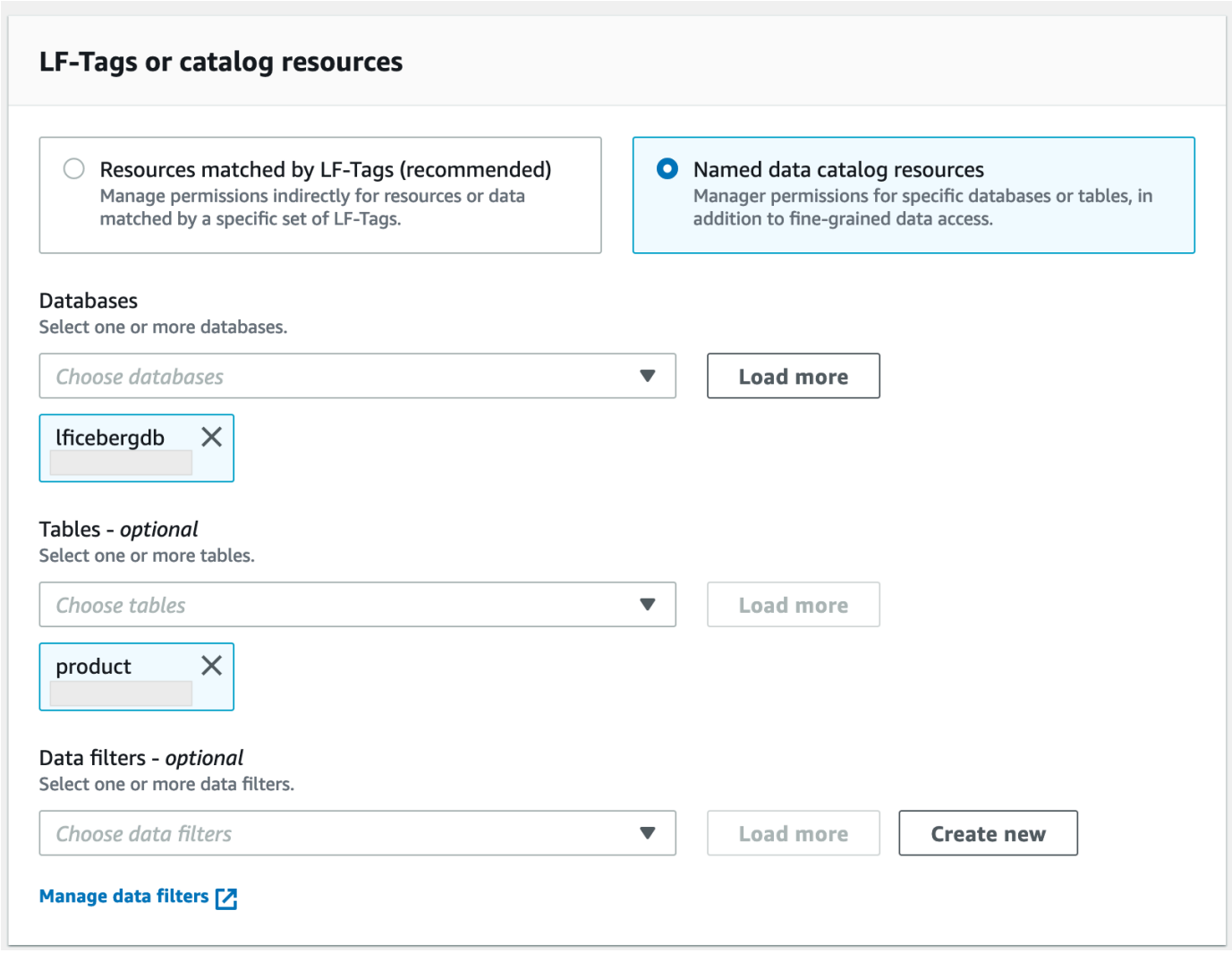
Untuk memberikan izin Lake Formation di tabel Gunung Es

Pada langkah ini, kami akan memberikan izin data lake kepada pengguna analis bisnis.

1. Di bawah Izin data lake, pilih Grant.
2. Di layar Berikan izin data, pilih, pengguna dan peran IAM.
3. Pilih `lf-consumer-analystuser` dari drop down.



4. Pilih Sumber daya katalog data bernama.
5. Untuk Database pilih`lficebergdb`.
6. Untuk Tabel, pilih`product`.



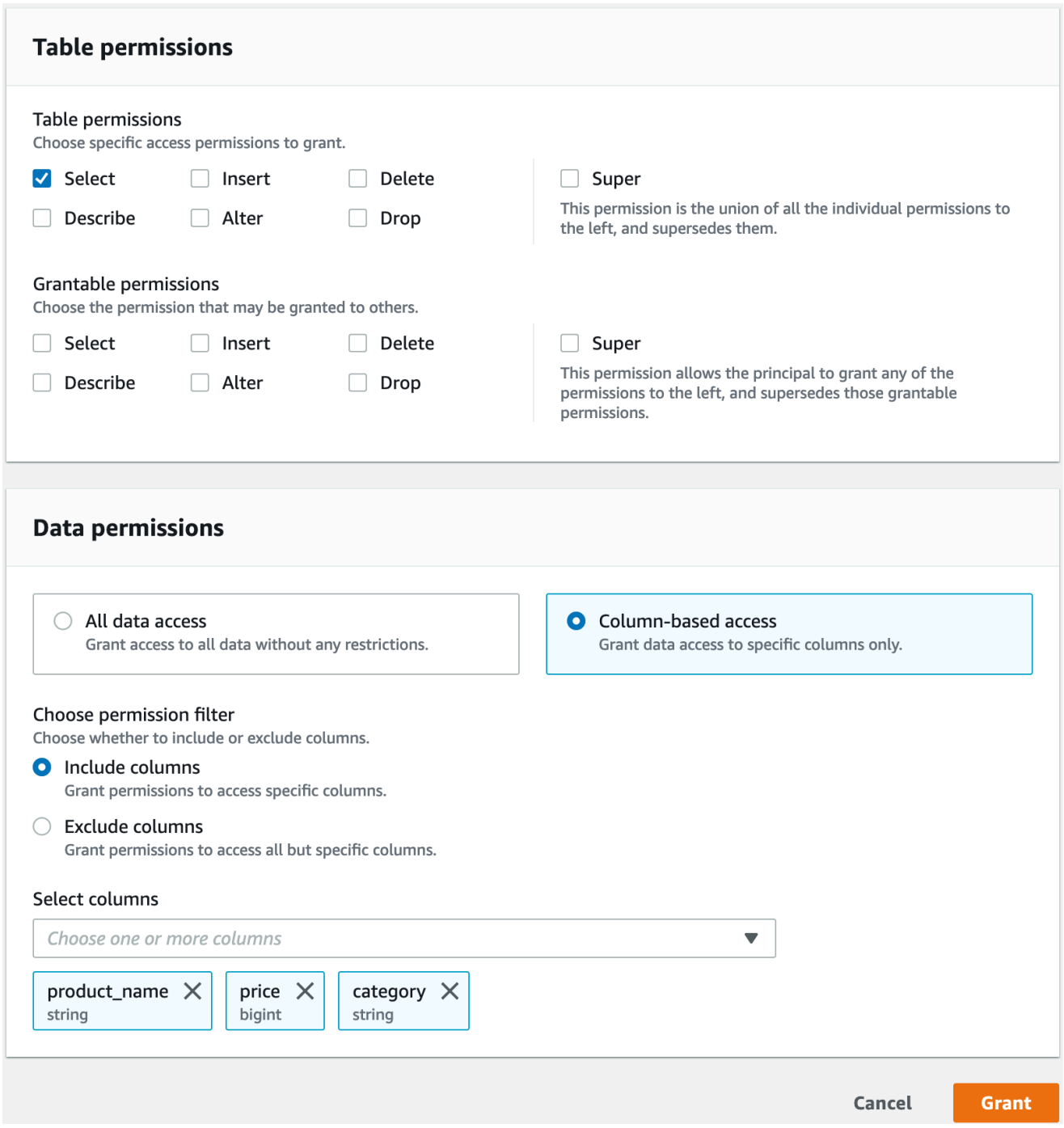
7. Selanjutnya, Anda dapat memberikan akses berbasis kolom dengan menentukan kolom.

   a. Di bawah Izin tabel, pilih Pilih.

   b. Di bawah Izin data, pilih Column-based akses, pilih Sertakan kolom.

   c. Pilih`product_name`,`price`, dan `category` kolom.

   d. PilihIzin.



Untuk menanyakan tabel Iceberg menggunakan Athena

Sekarang Anda dapat mulai menanyakan tabel Gunung Es yang Anda buat menggunakan Athena. Jika ini adalah pertama kalinya Anda menjalankan kueri di Athena, Anda perlu mengonfigurasi lokasi hasil kueri. Untuk informasi selengkapnya, lihat Menentukan lokasi hasil kueri.

1. Keluar sebagai pengguna administrator data lake dan masuk seperti `lf-consumer-analystuser` di Wilayah AS Timur (Virginia N.) menggunakan kata sandi yang disebutkan sebelumnya dari CloudFormation output.
2. Buka konsol Athena di https://console.aws.amazon.com/athena/.
3. Pilih Pengaturan dan pilih Kelola.
4. Di kotak Lokasi hasil kueri, masukkan jalur ke bucket yang Anda buat di CloudFormation output. Salin nilai **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial - 123456789012/athena -results/) dan pilih Simpan.
5. Jalankan kueri berikut untuk melihat pratinjau 10 catatan yang disimpan dalam tabel Iceberg:

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

   Untuk informasi selengkapnya tentang menanyakan tabel Gunung Es menggunakan Athena, lihat Menanyakan tabel Gunung Es di Panduan Pengguna Amazon Athena.

Langkah 3: Siapkan izin untuk tabel Hudi

Di bagian ini, Anda akan mempelajari cara membuat tabel Hudi di AWS Glue Data Catalog, mengatur izin data di AWS Lake Formation, dan kueri data menggunakan Amazon Athena.

Untuk membuat tabel Hudi

Pada langkah ini, Anda akan menjalankan AWS Glue pekerjaan yang membuat tabel transaksional Hudi di Katalog Data.

1. Masuk ke AWS Glue konsol https://console.aws.amazon.com/glue/di Wilayah AS Timur (Virginia Utara)

   sebagai pengguna administrator danau data.
2. Pilih pekerjaan dari panel navigasi kiri.
3. Pilih `native-hudi-create`.
4. Di bawah Tindakan, pilih Edit pekerjaan.
5. Di bawah Job details, perluas properti Advanced, dan centang kotak di samping Use AWS Glue Data Catalog as the Hive metastore untuk menambahkan metadata tabel di. AWS Glue Data Catalog Ini ditentukan AWS Glue Data Catalog sebagai metastore untuk sumber daya Katalog

Data yang digunakan dalam pekerjaan dan memungkinkan izin Lake Formation diterapkan nanti pada sumber daya katalog.

6. Pilih Simpan.
7. Pilih Jalankan. Anda dapat melihat status pekerjaan saat sedang berjalan.

   Untuk informasi selengkapnya tentang AWS Glue lowongan, lihat Bekerja dengan pekerjaan di AWS Glue konsol di Panduan AWS Glue Pengembang.

   Pekerjaan ini membuat tabel Hudi (sapi) di database: lfhudidb. Verifikasi `product` tabel di konsol Lake Formation.

Untuk mendaftarkan lokasi data dengan Lake Formation

Selanjutnya, daftarkan jalur Amazon S3 sebagai lokasi root danau data Anda.

1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/sebagai pengguna administrator danau data.
2. Di panel navigasi, di bawah Daftar dan konsumsi, pilih Lokasi data.
3. Pilih Daftar lokasi.
4. Pada halaman Daftar lokasi, masukkan yang berikut ini:

   - Jalur Amazon S3 - Pilih Jelajahi dan pilih. `lf-otf-datalake-123456789012` Pilih panah perluas di sebelah lokasi root Amazon S3 untuk menavigasi ke lokasi. `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-hudi`
   - Peran IAM - Pilih **`LF-OTF-RegisterRole`** sebagai peran IAM.
   - Pilih Daftar lokasi.

Untuk memberikan izin data lake pada tabel Hudi

Pada langkah ini, kami akan memberikan izin data lake kepada pengguna analis bisnis.

1. Di bawah Izin data lake, pilih Grant.
2. Di layar Berikan izin data, pilih, pengguna dan peran IAM.
3. `lf-consumer-analystuser`dari drop down.
4. Pilih Sumber daya katalog data bernama.
5. Untuk Database pilih`lfhudidb`.

6. Untuk Tabel, pilih`product`.
7. Selanjutnya, Anda dapat memberikan akses berbasis kolom dengan menentukan kolom.
   a. Di bawah Izin tabel, pilih Pilih.
   b. Di bawah Izin data, pilih Column-based akses, pilih Sertakan kolom.
   c. Pilih`product_name`,`price`, dan `category` kolom.
   d. PilihIzin.

Untuk menanyakan tabel Hudi menggunakan Athena

Sekarang mulailah menanyakan tabel Hudi yang Anda buat menggunakan Athena. Jika ini adalah pertama kalinya Anda menjalankan kueri di Athena, Anda perlu mengonfigurasi lokasi hasil kueri. Untuk informasi selengkapnya, lihat Menentukan lokasi hasil kueri.

1. Keluar sebagai pengguna administrator data lake dan masuk seperti `lf-consumer-analystuser` di Wilayah AS Timur (Virginia N.) menggunakan kata sandi yang disebutkan sebelumnya dari CloudFormation output.
2. Buka konsol Athena di https://console.aws.amazon.com/athena/.
3. Pilih Pengaturan dan pilih Kelola.
4. Di kotak Lokasi hasil kueri, masukkan jalur ke bucket yang Anda buat di CloudFormation output. Salin nilai **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial - 123456789012/athena -results/) dan Simpan.
5. Jalankan kueri berikut untuk melihat pratinjau 10 catatan yang disimpan dalam tabel Hudi:

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

   Untuk informasi selengkapnya tentang menanyakan tabel Hudi, lihat bagian Menanyakan tabel Hudi di Panduan Pengguna Amazon Athena.

Langkah 4: Siapkan izin untuk tabel Delta Lake

Di bagian ini, Anda akan mempelajari cara membuat tabel Delta Lake dengan file manifes symlink di AWS Glue Data Catalog, mengatur izin data AWS Lake Formation dan kueri data menggunakan Amazon Athena.

Untuk membuat tabel Delta Lake

Pada langkah ini, Anda akan menjalankan AWS Glue pekerjaan yang membuat tabel transaksional Delta Lake di Katalog Data.

1. Masuk ke AWS Glue konsol https://console.aws.amazon.com/glue/di Wilayah AS Timur (Virginia Utara)

   sebagai pengguna administrator danau data.
2. Pilih pekerjaan dari panel navigasi kiri.
3. Pilih `native-delta-create`.
4. Di bawah Tindakan, pilih Edit pekerjaan.
5. Di bawah Job details, perluas properti Advanced, dan centang kotak di samping Use AWS Glue Data Catalog as the Hive metastore untuk menambahkan metadata tabel di. AWS Glue Data Catalog Ini ditentukan AWS Glue Data Catalog sebagai metastore untuk sumber daya Katalog Data yang digunakan dalam pekerjaan dan memungkinkan izin Lake Formation diterapkan nanti pada sumber daya katalog.
6. Pilih Simpan.
7. Pilih Jalankan di bawah Tindakan.

   Pekerjaan ini menciptakan tabel Delta Lake bernama `product` dalam `lfdeltadb` database. Verifikasi `product` tabel di konsol Lake Formation.

Untuk mendaftarkan lokasi data dengan Lake Formation

Selanjutnya, daftarkan jalur Amazon S3 sebagai lokasi root danau data Anda.

1. Buka konsol Lake Formation https://console.aws.amazon.com/lakeformation/di pengguna administrator danau data.
2. Di panel navigasi, di bawah Daftar dan konsumsi, pilih Lokasi data.
3. Pilih Daftar lokasi.
4. Pada halaman Daftar lokasi, masukkan yang berikut ini:

   - Jalur Amazon S3 - Pilih Jelajahi dan pilih. `lf-otf-datalake-123456789012` Pilih panah perluas di sebelah lokasi root Amazon S3 untuk menavigasi ke lokasi. `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-delta`
   - Peran IAM - Pilih **`LF-OTF-RegisterRole`** sebagai peran IAM.

- Pilih Daftar lokasi.

Untuk memberikan izin data lake pada tabel Delta Lake

Pada langkah ini, kami akan memberikan izin data lake kepada pengguna analis bisnis.

1. Di bawah Izin data lake, pilih Grant.
2. Di layar Berikan izin data, pilih, pengguna dan peran IAM.
3. `lf-consumer-analystuser`dari drop down.
4. Pilih Sumber daya katalog data bernama.
5. Untuk Database pilih`lfdeltadb`.
6. Untuk Tabel, pilih`product`.
7. Selanjutnya, Anda dapat memberikan akses berbasis kolom dengan menentukan kolom.

   a. Di bawah Izin tabel, pilih Pilih.

   b. Di bawah Izin data, pilih Column-based akses, pilih Sertakan kolom.

   c. Pilih`product_name`,`price`, dan `category` kolom.

   d. PilihIzin.

Untuk menanyakan tabel Danau Delta menggunakan Athena

Sekarang mulailah menanyakan tabel Delta Lake yang Anda buat menggunakan Athena. Jika ini adalah pertama kalinya Anda menjalankan kueri di Athena, Anda perlu mengonfigurasi lokasi hasil kueri. Untuk informasi selengkapnya, lihat Menentukan lokasi hasil kueri.

1. Keluar sebagai pengguna administrator data lake dan masuk seperti `BusinessAnalystUser` di Wilayah AS Timur (Virginia N.) menggunakan kata sandi yang disebutkan sebelumnya dari CloudFormation output.
2. Buka konsol Athena di https://console.aws.amazon.com/athena/.
3. Pilih Pengaturan dan pilih Kelola.
4. Di kotak Lokasi hasil kueri, masukkan jalur ke bucket yang Anda buat di CloudFormation output. Salin nilai **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial - 123456789012/athena -results/) dan Simpan.
5. Jalankan kueri berikut untuk melihat pratinjau 10 catatan yang disimpan dalam tabel Delta Lake:

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

Untuk informasi selengkapnya tentang menanyakan tabel Delta Lake, lihat bagian Menanyakan tabel Danau Delta di Panduan Pengguna Amazon Athena.

Langkah 5: Bersihkan AWS sumber daya

Untuk membersihkan sumber daya

Untuk mencegah biaya yang tidak diinginkan ke Anda Akun AWS, hapus AWS sumber daya yang Anda gunakan untuk tutorial ini.

1. Masuk ke CloudFormation konsol di https://console.aws.amazon.com/cloudformationsebagai administrator IAM.
2. Hapus tumpukan formasi cloud. Tabel yang Anda buat secara otomatis dihapus dengan tumpukan.

Mengelola data lake menggunakan kontrol akses berbasis tag Lake Formation

Ribuan pelanggan sedang membangun danau data skala petabyte. AWS Banyak dari pelanggan ini menggunakan AWS Lake Formation untuk dengan mudah membangun dan berbagi data lake mereka di seluruh organisasi. Seiring bertambahnya jumlah tabel dan pengguna, pengelola data dan administrator mencari cara untuk mengelola izin pada data lake dengan mudah dalam skala besar. Lake Formation Tag-based Access Control (LF-TBAC) memecahkan masalah ini dengan memungkinkan data steward untuk membuat LF-tag (berdasarkan klasifikasi data dan ontologi mereka) yang kemudian dapat dilampirkan ke sumber daya.

LF-TBAC adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Dalam Lake Formation, atribut ini disebut LF-tag. Anda dapat melampirkan LF-tag ke sumber daya Katalog Data dan prinsip Lake Formation. Administrator data lake dapat menetapkan dan mencabut izin pada sumber daya Lake Formation menggunakan LF-tag. Untuk informasi lebih lanjut tentang lihat,Kontrol akses berbasis tag Lake Formation.

Tutorial ini menunjukkan cara membuat kebijakan kontrol akses berbasis tag Lake Formation menggunakan dataset AWS publik. Selain itu, ini menunjukkan cara menanyakan tabel, database, dan kolom yang memiliki kebijakan akses berbasis tag Lake Formation yang terkait dengannya.

Anda dapat menggunakan LF-TBAC untuk kasus penggunaan berikut:

- Anda memiliki sejumlah besar tabel dan prinsip yang harus diberikan oleh administrator danau data
- Anda ingin mengklasifikasikan data Anda berdasarkan ontologi dan memberikan izin berdasarkan klasifikasi
- Administrator data lake ingin menetapkan izin secara dinamis, dengan cara yang digabungkan secara longgar

Berikut ini adalah langkah-langkah tingkat tinggi untuk mengonfigurasi izin menggunakan LF-TBAC:

1. Data steward mendefinisikan ontologi tag dengan dua LF-tag: dan. `Confidential` `Sensitive` Data dengan `Confidential=True` memiliki kontrol akses yang lebih ketat. Data dengan `Sensitive=True` membutuhkan analisis spesifik dari analis.
2. Data steward memberikan tingkat izin yang berbeda kepada insinyur data untuk membuat tabel dengan tag LF yang berbeda.
3. Insinyur data membangun dua database: `tag_database` dan. `col_tag_database` Semua tabel di `tag_database` dikonfigurasi dengan`Confidential=True`. Semua tabel di `col_tag_database` dikonfigurasi dengan`Confidential=False`. Beberapa kolom tabel di `col_tag_database` ditandai dengan `Sensitive=True` untuk kebutuhan analisis spesifik.
4. Insinyur data memberikan izin baca kepada analis untuk tabel dengan kondisi ekspresi tertentu `Confidential=True` dan`Confidential=False`,`Sensitive=True`.
5. Dengan konfigurasi ini, analis data dapat fokus melakukan analisis dengan data yang tepat.

Topik

- Audiens yang dituju
- Prasyarat
- Langkah 1: Menyediakan sumber daya Anda
- Langkah 2: Daftarkan lokasi data Anda, buat ontologi LF-tag, dan berikan izin
- Langkah 3: Buat database Lake Formation
- Langkah 4: Berikan izin tabel

- Langkah 5: Jalankan kueri di Amazon Athena untuk memverifikasi izin
- Langkah 6: Bersihkan AWS sumber daya

Audiens yang dituju

Tutorial ini ditujukan untuk pengelola data, insinyur data, dan analis data. Dalam hal mengelola AWS Glue Data Catalog dan mengelola izin di Lake Formation, pengelola data dalam akun penghasil memiliki kepemilikan fungsional berdasarkan fungsi yang mereka dukung, dan dapat memberikan akses ke berbagai konsumen, organisasi eksternal, dan akun.

Tabel berikut mencantumkan peran yang digunakan dalam tutorial ini:

Peran	Deskripsi
Pelayan data (administrator)	`lf-data-steward` Pengguna memiliki akses berikut: • Baca akses ke semua sumber daya di Katalog Data • Dapat membuat LF-tag dan mengasosiasikan ke peran insinyur data untuk izin yang dapat diberikan kepada prinsipal lain
Insinyur data	`lf-data-engineer` pengguna memiliki akses berikut: • Akses baca, tulis, dan perbarui lengkap ke semua sumber daya di Katalog Data • Izin lokasi data di danau data • Dapat mengaitkan LF-tag dan mengasosi asikan ke Katalog Data • Dapat melampirkan tag LF ke sumber daya, yang menyediakan akses ke prinsipal berdasarkan kebijakan apa pun yang dibuat oleh pengelola data

Peran	Deskripsi
Analis data	`lf-data-analyst` Pengguna memiliki akses berikut: • Akses halus ke sumber daya yang dibagikan oleh kebijakan akses berbasis tag Lake Formation

Prasyarat

Sebelum Anda memulai tutorial ini, Anda harus memiliki Akun AWS yang dapat Anda gunakan untuk masuk sebagai pengguna administratif dengan izin yang benar. Untuk informasi selengkapnya, lihat Selesaikan inisial AWS tugas konfigurasi.

Tutorial mengasumsikan bahwa Anda sudah familiar dengan IAM. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Langkah 1: Menyediakan sumber daya Anda

Tutorial ini mencakup AWS CloudFormation template untuk pengaturan cepat. Anda dapat meninjau dan menyesuaikannya sesuai dengan kebutuhan Anda. Template membuat tiga peran berbeda (tercantum dalamAudiens yang dituju) untuk melakukan latihan ini dan menyalin nyc-taxi-data kumpulan data ke bucket Amazon S3 lokal Anda.

- Bucket Amazon S3
- Pengaturan Lake Formation yang sesuai
- Sumber daya Amazon EC2 yang sesuai
- Tiga peran IAM dengan kredensil

Buat sumber daya Anda

1. Masuk ke AWS CloudFormation konsol di https://console.aws.amazon.com/cloudformation di wilayah US East (Virginia N.).
2. Pilih Launch Stack.
3. Pilih Berikutnya.

4. Di bagian Konfigurasi Pengguna, masukkan kata sandi untuk tiga peran:`DataStewardUserPassword`, `DataEngineerUserPassword` dan`DataAnalystUserPassword`.
5. Tinjau detail di halaman akhir dan pilih Saya akui yang AWS CloudFormation mungkin membuat sumber daya IAM.
6. Pilih Buat.

   Pembuatan tumpukan bisa memakan waktu hingga lima menit.

> **Note**
>
> Setelah Anda menyelesaikan tutorial, Anda mungkin ingin menghapus tumpukan CloudFormation untuk menghindari terus dikenakan biaya. Verifikasi bahwa sumber daya berhasil dihapus dalam status acara untuk tumpukan.

Langkah 2: Daftarkan lokasi data Anda, buat ontologi LF-tag, dan berikan izin

Pada langkah ini, pengguna data steward mendefinisikan ontologi tag dengan dua LF-tag: `Confidential` dan`Sensitive`, dan memberikan prinsip-prinsip IAM tertentu kemampuan untuk melampirkan LF-tag yang baru dibuat ke sumber daya.

Daftarkan lokasi data dan tentukan ontologi LF-tag

1. Lakukan langkah pertama sebagai pengguna data steward (`lf-data-steward`) untuk memverifikasi data di Amazon S3 dan Katalog Data di Lake Formation.

   a. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/seperti kata sandi `lf-data-steward` yang digunakan saat menerapkan CloudFormation tumpukan.

   b. Di panel navigasi, di bawah Izin ¸ pilih Peran dan tugas administratif.

   c. Pilih Tambah di bagian Administrator danau data.

   d. Pada halaman Add administrator, untuk pengguna dan peran IAM, pilih penggunalf-data-steward.

   e. Pilih Simpan untuk ditambahkan `lf-data-steward` sebagai administrator Lake Formation.

2. Selanjutnya, perbarui pengaturan Katalog Data untuk menggunakan izin Lake Formation untuk mengontrol sumber daya katalog alih-alih kontrol akses berbasis IAM.

   a. Di panel navigasi, di bawah Administrasi, pilih Pengaturan Katalog Data.

   b. Hapus centang Gunakan hanya kontrol akses IAM untuk database baru.

   c. Hapus centang Gunakan hanya kontrol akses IAM untuk tabel baru di database baru.

   d. Klik Simpan.

3. Selanjutnya, daftarkan lokasi data untuk data lake.

   a. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.

   b. Pilih Daftar lokasi.

   c. Pada halaman Daftar lokasi, untuk jalur Amazon S3, masukkan. `s3://lf-tagbased-demo-`*`Account-ID`*

   d. Untuk peran IAM, biarkan nilai default apa `AWSServiceRoleForLakeFormationDataAccess` adanya.

   e. Pilih Lake Formation sebagai mode izin.

   f. Pilih Daftar lokasi.

4. Selanjutnya, buat ontologi dengan mendefinisikan LF-tag.

   a. Di bawah Izin di panel navigasi, pilih LF-tag dan izin. .

   b. Pilih Tambahkan LF-Tag.

   c. Untuk Kunci, masukkan `Confidential`.

   d. Untuk Nilai, tambahkan `True` dan`False`.

   e. Pilih Tambahkan LF-Tag.

   f. Ulangi langkah-langkah untuk membuat LF-tag `Sensitive` dengan nilai. `True`

   Anda telah membuat semua LF-tag yang diperlukan untuk latihan ini.

Berikan izin kepada pengguna IAM

1. Selanjutnya, berikan kepala sekolah IAM tertentu kemampuan untuk melampirkan tag LF yang baru dibuat ke sumber daya.

   a. Di bawah Izin di panel navigasi, pilih LF-tag dan izin.

b. Di bagian izin LF-tag, pilih Hibah izin.

c. Untuk jenis Izin, pilih izin pasangan nilai kunci LF-tag.

d. Pilih pengguna dan peran IAM.

e. Untuk pengguna dan peran IAM, cari dan pilih `lf-data-engineer` peran.

f. Di bagian LF-tag, tambahkan kunci `Confidential` dengan nilai `True` dan`False`, dan `key Sensitive` dengan nilai. `True`

g. Di bawah Izin, pilih Jelaskan dan Kaitkan untuk Izin dan Izin yang Dapat Diberikan.

h. PilihIzin.

2. Selanjutnya, berikan izin `lf-data-engineer` untuk membuat database di Katalog Data kami dan pada bucket Amazon S3 yang mendasari yang dibuat oleh. AWS CloudFormation

   a. Di bawah Administrasi di panel navigasi, pilih Peran dan tugas administratif.

   b. Di bagian Pembuat basis data, pilih Hibah.

   c. Untuk pengguna dan peran IAM, pilih `lf-data-engineer` peran.

   d. Untuk izin Katalog, pilih Buat database.

   e. PilihIzin.

3. Selanjutnya, berikan izin pada (`s3://lf-tagbased-demo-`*`Account-ID`*`)` bucket Amazon S3 kepada `lf-data-engineer` pengguna.

   a. Di panel navigasi, di bawah Izin, pilih Lokasi data.

   b. PilihIzin.

   c. Pilih Akun saya.

   d. Untuk pengguna dan peran IAM, pilih `lf-data-engineer` peran.

   e. Untuk lokasi Penyimpanan, masukkan bucket Amazon S3 yang dibuat oleh template. CloudFormation (`s3://lf-tagbased-demo-`*`Account-ID`*`)`

   f. PilihIzin.

4. Selanjutnya, berikan **`lf-data-engineer`** izin yang dapat diberikan pada sumber daya yang terkait dengan ekspresi LF-tag. `Confidential=True`

   a. Di panel navigasi, di bawah Izin, pilih Izin danau data.

   b. PilihIzin.

   c. Pilih pengguna dan peran IAM.

   d. Pilih perannya`lf-data-engineer`.

e. Di bagian LF-tag atau sumber katalog, pilih Sumber daya yang cocok dengan LF-tag.

f. Pilih Tambahkan pasangan nilai kunci LF-tag.

g. Tambahkan kunci `Confidential` dengan nilainya`True`.

h. Di bagian Izin database, pilih Jelaskan untuk izin Database dan izin yang dapat diberikan.

i. Di bagian Izin tabel, pilih Jelaskan, Pilih, dan Ubah untuk izin Tabel dan izin yang Dapat Diberikan.

j. PilihIzin.

5. Selanjutnya, berikan `lf-data-engineer` izin yang dapat diberikan pada sumber daya yang terkait dengan ekspresi LF-tag. `Confidential=False`

a. Di panel navigasi, di bawah Izin, pilih Izin danau data.

b. PilihIzin.

c. Pilih pengguna dan peran IAM.

d. Pilih perannya`lf-data-engineer`.

e. Pilih Sumber daya yang cocok dengan LF-tag.

f. Pilih Tambahkan LF-Tag.

g. Tambahkan kunci `Confidential` dengan nilainya`False`.

h. Di bagian Izin database, pilih Jelaskan untuk izin Database dan izin yang dapat diberikan.

i. Di bagian Tabel dan kolom izin, jangan pilih apa pun.

j. PilihIzin.

6. Selanjutnya, kami `lf-data-engineer` memberikan izin yang dapat diberikan pada sumber daya yang terkait dengan pasangan nilai kunci LF-tag dan. `Confidential=False Sensitive=True`

a. Di panel navigasi, di bawah Izin, pilih Izin data.

b. PilihIzin.

c. Pilih pengguna dan peran IAM.

d. Pilih perannya`lf-data-engineer`.

e. Di bagian LF-tag atau sumber katalog, pilih Sumber daya yang cocok dengan LF-tag.

f. Pilih Tambahkan LF-Tag.

g. Tambahkan kunci `Confidential` dengan nilainya`False`.

h. Pilih Tambahkan pasangan nilai kunci LF-tag.

i. Tambahkan kunci `Sensitive` dengan nilainya`True`.

j. Di bagian Izin database, pilih Jelaskan untuk izin Database dan izin yang dapat diberikan.

k. Di bagian Izin tabel, pilih Jelaskan, Pilih, dan Ubah untuk izin Tabel dan izin yang Dapat Diberikan.

l. PilihIzin.

Langkah 3: Buat database Lake Formation

Pada langkah ini, Anda membuat dua database dan melampirkan LF-tag ke database dan kolom tertentu untuk tujuan pengujian.

Buat database dan tabel Anda untuk akses tingkat database

1. Pertama, buat database`tag_database`, tabel`source_data`, dan lampirkan LF-tag yang sesuai.

   a. Pada konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), di bawah Katalog Data, pilih Database.

   b. Pilih Buat basis data.

   c. Untuk Nama, masukkan `tag_database`.

   d. Untuk Lokasi, masukkan lokasi Amazon S3 yang dibuat oleh template. CloudFormation `(s3://lf-tagbased-demo-`*`Account-ID`*`/tag_database/)`

   e. Hapus pilih Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini.

   f. Pilih Buat basis data.

2. Selanjutnya, buat tabel baru di dalamnya`tag_database`.

   a. Pada halaman Database, pilih database`tag_database`.

   b. Pilih Lihat Tabel dan klik Buat tabel.

   c. Untuk Nama, masukkan `source_data`.

   d. Untuk Basis data, pilih basis data `tag_database`.

   e. Untuk format Tabel, pilih AWS Glue Tabel standar.

   f. Untuk Data terletak di, pilih Jalur yang ditentukan di akun saya.

   g. Untuk jalur Sertakan, masukkan jalur yang akan `tag_database` dibuat oleh CloudFormation template`(s3://lf-tagbased-demo`*`Account-ID`*`/tag_database/)`.

h. Untuk format Data, pilih CSV.

i. Di bawah Upload skema, masukkan array JSON berikut dari struktur kolom untuk membuat skema:

```
[
                {
                     "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. Pilih Unggah. Setelah mengunggah skema, skema tabel akan terlihat seperti tangkapan layar berikut:

#	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. Pilih Kirim.

3. Selanjutnya, lampirkan LF-tag di tingkat database.

   a. Pada halaman Database, temukan dan pilih`tag_database`.
   b. Pada menu Tindakan, pilih Edit LF-tag.
   c. Pilih Tetapkan LF-Tag baru.
   d. Untuk kunci yang Ditugaskan ¸ pilih `Confidential` LF-tag yang Anda buat sebelumnya.
   e. Untuk Nilai, pilih`True`.
   f. Pilih Simpan.

   Ini melengkapi penugasan LF-tag ke database tag_database.

Buat database dan tabel Anda untuk akses tingkat kolom

Ulangi langkah-langkah berikut untuk membuat database `col_tag_database` dan tabel`source_data_col_lvl`, dan melampirkan LF-tag pada tingkat kolom.

1. Pada halaman Database, pilih Buat database.
2. Untuk Nama, masukkan `col_tag_database`.
3. Untuk Lokasi, masukkan lokasi Amazon S3 yang dibuat oleh template. CloudFormation (`s3://lf-tagbased-demo-`*`Account-ID`*`/col_tag_database/`)
4. Hapus pilih Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini.
5. Pilih Buat basis data.
6. Pada halaman Database, pilih database (`col_tag_database`) baru Anda.
7. Pilih Lihat tabel dan klik Buat tabel.
8. Untuk Nama, masukkan `source_data_col_lvl`.
9. Untuk Database, pilih database baru Anda(`col_tag_database`).
10. Untuk format Tabel, pilih AWS Glue Tabel standar.
11. Untuk Data terletak di, pilih Jalur yang ditentukan di akun saya.
12. Masukkan jalur Amazon S3 untuk. `col_tag_database` (`s3://lf-tagbased-demo-`*`Account-ID`*`/col_tag_database/`)
13. Untuk format Data, pilih`CSV`.
14. Di bawah`Upload schema`, masukkan skema JSON berikut:

```
[
                {
                     "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. Pilih `Upload`. Setelah mengunggah skema, skema tabel akan terlihat seperti tangkapan layar berikut.

#	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. Pilih Kirim untuk menyelesaikan pembuatan tabel.
17. Sekarang, kaitkan `Sensitive=True` LF-tag ke kolom `vendorid` dan. `fare_amount`

    a. Pada halaman Tabel, pilih tabel yang Anda buat(`source_data_col_lvl`).

    b. Pada menu Tindakan, pilih Skema.

    c. Pilih kolom `vendorid` dan pilih Edit LF-tag.

    d. Untuk kunci yang Ditugaskan, pilih Sensitif.

    e. Untuk Nilai, pilih Benar.

    f. Pilih Simpan.
18. Selanjutnya, kaitkan `Confidential=False` LF-tag ke. `col_tag_database` Ini diperlukan `lf-data-analyst` agar dapat menggambarkan database `col_tag_database` saat masuk dari Amazon Athena.

    a. Pada halaman Database, temukan dan pilih`col_tag_database`.

    b. Pada menu Tindakan, pilih Edit LF-tag.

    c. Pilih Tetapkan LF-Tag baru.

    d. Untuk kunci yang Ditugaskan, pilih `Confidential` LF-tag yang Anda buat sebelumnya.

    e. Untuk Nilai, pilih`False`.

    f. Pilih Simpan.

Langkah 4: Berikan izin tabel

Berikan izin kepada analis data untuk konsumsi database `tag_database` dan tabel `col_tag_database` menggunakan `Confidential` LF-tag dan. `Sensitive`

1. Ikuti langkah-langkah ini untuk memberikan izin kepada `lf-data-analyst` pengguna pada objek yang terkait dengan LF-tag `Confidential=True` (Database:TAG_DATABASE) untuk memiliki database dan izin pada tabel. `Describe` `Select`

    a. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/as`lf-data-engineer`.

    b. Di bawah Izin, pilih Izin data lake.

    c. PilihIzin.

    d. Di bawah Prinsipal, pilih pengguna dan peran IAM.

e. Untuk pengguna dan peran IAM, pilih`lf-data-analyst`.

f. Di bawah LF-tag atau sumber katalog, pilih Sumber daya yang cocok dengan LF-tag.

g. Pilih Tambahkan LF-Tag.

h. Untuk Key, pilih`Confidential`.

i. Untuk Nilai, pilih`True`.

j. Untuk izin Database, pilih`Describe`.

k. Untuk izin Tabel, pilih Pilih dan Jelaskan.

l. PilihIzin.

2. Selanjutnya, ulangi langkah-langkah untuk memberikan izin kepada analis data untuk ekspresi LF-tag untuk. `Confidential=False` LF-tag ini digunakan untuk menggambarkan `col_tag_database` dan tabel `source_data_col_lvl` saat login dari Amazon `lf-data-analyst` Athena.

   a. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/as`lf-data-engineer`.

   b. Pada halaman Database, pilih database`col_tag_database`.

   c. Pilih Action dan Grant.

   d. Di bawah Prinsipal, pilih pengguna dan peran IAM.

   e. Untuk pengguna dan peran IAM, pilih`lf-data-analyst`.

   f. Pilih Sumber daya yang cocok dengan LF-tag.

   g. Pilih Tambahkan LF-Tag.

   h. Untuk Key, pilih`Confidential`.

   i. Untuk Nilai ¸ pilih`False`.

   j. Untuk izin Database, pilih`Describe`.

   k. Untuk izin Tabel, jangan pilih apa pun.

   l. PilihIzin.

3. Selanjutnya, ulangi langkah-langkah untuk memberikan izin kepada analis data untuk ekspresi LF-tag untuk dan. `Confidential=False` `Sensitive=True` Tag LF ini digunakan untuk mendeskripsikan `col_tag_database` dan tabel `source_data_col_lvl` (tingkat kolom) saat masuk sebagai dari Amazon Athena. `lf-data-analyst`

   a. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/as`lf-data-engineer`.

b. Pada halaman Database, pilih databasecol_tag_database.
c. Pilih Action dan Grant.
d. Di bawah Prinsipal, pilih pengguna dan peran IAM.
e. Untuk pengguna dan peran IAM, pilihlf-data-analyst.
f. Pilih Sumber daya yang cocok dengan LF-tag.
g. Pilih Tambahkan LF-Tag.
h. Untuk Key, pilihConfidential.
i. Untuk Nilai ¸ pilihFalse.
j. Pilih Tambahkan LF-Tag.
k. Untuk Key, pilihSensitive.
l. Untuk Nilai ¸ pilihTrue.
m. Untuk izin Database, pilihDescribe.
n. Untuk izin Tabel, pilih Select danDescribe.
o. PilihIzin.

Langkah 5: Jalankan kueri di Amazon Athena untuk memverifikasi izin

Untuk langkah ini, gunakan Amazon Athena untuk menjalankan SELECT kueri terhadap dua tabel. (source_data and source_data_col_lvl) Gunakan jalur Amazon S3 sebagai lokasi hasil kueri. (s3://lf-tagbased-demo-*Account-ID*/athena-results/)

1. Masuk ke konsol Athena di https://console.aws.amazon.com/athena/as. lf-data-analyst
2. Di editor kueri Athena, pilih tag_database di panel kiri.
3. Pilih ikon opsi menu tambahan (tiga titik vertikal) di sebelah source_data dan pilih tabel Pratinjau.
4. Pilih Run query (Jalankan kueri).

   Kueri harus memakan waktu beberapa menit untuk dijalankan. Query menampilkan semua kolom dalam output karena LF-tag dikaitkan pada tingkat database dan source_data tabel secara otomatis mewarisi LF-tag dari database. tag_database
5. Jalankan kueri lain menggunakan col_tag_database dansource_data_col_lvl.

   Query kedua mengembalikan dua kolom yang ditandai sebagai Non-Confidential danSensitive.

6. Anda juga dapat memeriksa untuk melihat perilaku kebijakan akses berbasis tag Lake Formation pada kolom yang Anda tidak memiliki hibah kebijakan. Ketika kolom untagged dipilih dari tabel, `source_data_col_lvl` Athena mengembalikan kesalahan. Misalnya, Anda dapat menjalankan kueri berikut untuk memilih kolom yang tidak ditandai: `geolocationid`

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

Langkah 6: Bersihkan AWS sumber daya

Untuk mencegah biaya yang tidak diinginkan ke Anda Akun AWS, Anda dapat menghapus AWS sumber daya yang Anda gunakan untuk tutorial ini.

1. Masuk ke konsol Lake Formation sebagai `lf-data-engineer` dan hapus database `tag_database` dan`col_tag_database`.
2. Selanjutnya, masuk sebagai `lf-data-steward` dan bersihkan semua Izin LF-Tag, Izin Data, dan Izin Lokasi Data yang diberikan di atas yang diberikan dan. `lf-data-engineer` `lf-data-analyst.`
3. Masuk ke konsol Amazon S3 sebagai pemilik akun menggunakan kredenal IAM yang Anda gunakan untuk menyebarkan tumpukan. CloudFormation
4. Hapus ember berikut:
   - lf-tagbased-demo-accesslogs-*acct-id*
   - lf-tagbased-demo-*acct-id*
5. Masuk ke CloudFormation konsol di https://console.aws.amazon.com/cloudformation, dan hapus tumpukan yang Anda buat. Tunggu status tumpukan berubah menjadi`DELETE_COMPLETE`.

Mengamankan data lake dengan kontrol akses tingkat baris

AWS Lake Formation izin tingkat baris memungkinkan Anda memberikan akses ke baris tertentu dalam tabel berdasarkan kepatuhan data dan kebijakan tata kelola. Jika Anda memiliki tabel besar yang menyimpan miliaran catatan, Anda memerlukan cara untuk memungkinkan pengguna dan tim yang berbeda untuk hanya mengakses data yang diizinkan untuk dilihat. Kontrol akses tingkat baris adalah cara sederhana dan berkinerja baik untuk melindungi data, sekaligus memberi pengguna akses ke data yang mereka butuhkan untuk melakukan pekerjaan mereka. Lake Formation

menyediakan audit terpusat dan pelaporan kepatuhan dengan mengidentifikasi kepala sekolah mana yang mengakses data apa, kapan, dan melalui layanan mana.

Dalam tutorial ini, Anda mempelajari cara kerja kontrol akses tingkat baris di Lake Formation, dan cara mengaturnya.

Tutorial ini mencakup AWS CloudFormation template untuk mengatur sumber daya yang dibutuhkan dengan cepat. Anda dapat meninjau dan menyesuaikannya sesuai dengan kebutuhan Anda.

Topik

- Audiens yang dituju
- Prasyarat
- Langkah 1: Menyediakan sumber daya Anda
- Langkah 2: Kueri tanpa filter data
- Langkah 3: Siapkan filter data dan berikan izin
- Langkah 4: Kueri dengan filter data
- Langkah 5: Bersihkan AWS sumber daya

Audiens yang dituju

Tutorial ini ditujukan untuk pengelola data, insinyur data, dan analis data. Tabel berikut mencantumkan peran dan tanggung jawab pemilik data dan konsumen data.

Peran	Deskripsi
Administrator IAM	Pengguna yang dapat membuat pengguna dan peran serta bucket Amazon Simple Storage Service (Amazon S3). Memiliki kebijakan yang `AdministratorAccess` AWS dikelola.
Administrator danau data	Pengguna yang bertanggung jawab untuk menyiapkan data lake, membuat filter data, dan memberikan izin kepada analis data.
Analis data	Pengguna yang dapat menjalankan kueri terhadap data lake. Analis data yang berada

Peran	Deskripsi
	di berbagai negara (untuk kasus penggunaan kami, AS dan Jepang) hanya dapat menganali sis ulasan produk untuk pelanggan yang berlokasi di negara mereka sendiri dan untuk alasan kepatuhan, seharusnya tidak dapat melihat data pelanggan yang berlokasi di negara lain.

Prasyarat

Sebelum Anda memulai tutorial ini, Anda harus memiliki Akun AWS yang dapat Anda gunakan untuk masuk sebagai pengguna administratif dengan izin yang benar. Untuk informasi selengkapnya, lihat Selesaikan inisial AWS tugas konfigurasi.

Tutorial mengasumsikan bahwa Anda sudah familiar dengan IAM. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Ubah pengaturan Lake Formation

> **Important**
>
> Sebelum meluncurkan CloudFormation template, nonaktifkan opsi Gunakan hanya kontrol akses IAM untuk database/tabel baru di Lake Formation dengan mengikuti langkah-langkah di bawah ini:

1. Masuk ke konsol Lake Formation https://console.aws.amazon.com/lakeformation/di wilayah AS Timur (Virginia N.) atau wilayah AS Barat (Oregon).
2. Di bawah Katalog Data, pilih Pengaturan.
3. Hapus pilih Gunakan hanya kontrol akses IAM untuk database baru dan Gunakan hanya kontrol akses IAM untuk tabel baru di database baru.
4. Pilih Simpan.

Langkah 1: Menyediakan sumber daya Anda

Tutorial ini mencakup CloudFormation template untuk pengaturan cepat. Anda dapat meninjau dan menyesuaikannya sesuai dengan kebutuhan Anda. CloudFormation Template menghasilkan sumber daya berikut:

- Pengguna dan kebijakan untuk:
  - DataLakeAdmin
  - DataAnalystAS
  - DataAnalystJP
- Pengaturan dan izin danau data Lake Formation
- Fungsi Lambda (untuk sumber daya CloudFormation khusus yang didukung Lambda) digunakan untuk menyalin file data sampel dari bucket Amazon S3 publik ke bucket Amazon S3 Anda
- Bucket Amazon S3 untuk berfungsi sebagai danau data kami
- AWS Glue Data Catalog Database, tabel, dan partisi

Buat sumber daya Anda

Ikuti langkah-langkah ini untuk membuat sumber daya Anda menggunakan CloudFormation template.

1. Masuk ke CloudFormation konsol di https://console.aws.amazon.com/cloudformation di wilayah US East (Virginia N.).
2. Pilih Launch Stack.
3. Pilih Berikutnya di layar Buat tumpukan.
4. Masukkan Nama tumpukan.
5. Untuk DatalakeAdminUserNamedan DatalakeAdminUserPassword, masukkan nama pengguna dan kata sandi IAM Anda untuk pengguna admin danau data.
6. Untuk DataAnalystUsUserNamedan DataAnalystUsUserPassword, masukkan nama pengguna dan kata sandi untuk nama pengguna dan kata sandi yang Anda inginkan untuk pengguna analis data yang bertanggung jawab atas pasar AS.
7. Untuk DataAnalystJpUserNamedan DataAnalystJpUserPassword, masukkan nama pengguna dan kata sandi untuk nama pengguna dan kata sandi yang Anda inginkan untuk pengguna analis data yang bertanggung jawab atas pasar Jepang.
8. Untuk DataLakeBucketName, masukkan nama bucket data Anda.

9. Untuk DatabaseName, dan TableNamebiarkan sebagai default.
10. Pilih Berikutnya
11. Di halaman berikutnya, pilih Berikutnya.
12. Tinjau detail di halaman akhir dan pilih Saya akui yang CloudFormation mungkin membuat sumber daya IAM.
13. Pilih Buat.

    Pembuatan tumpukan dapat memakan waktu satu menit untuk diselesaikan.

Langkah 2: Kueri tanpa filter data

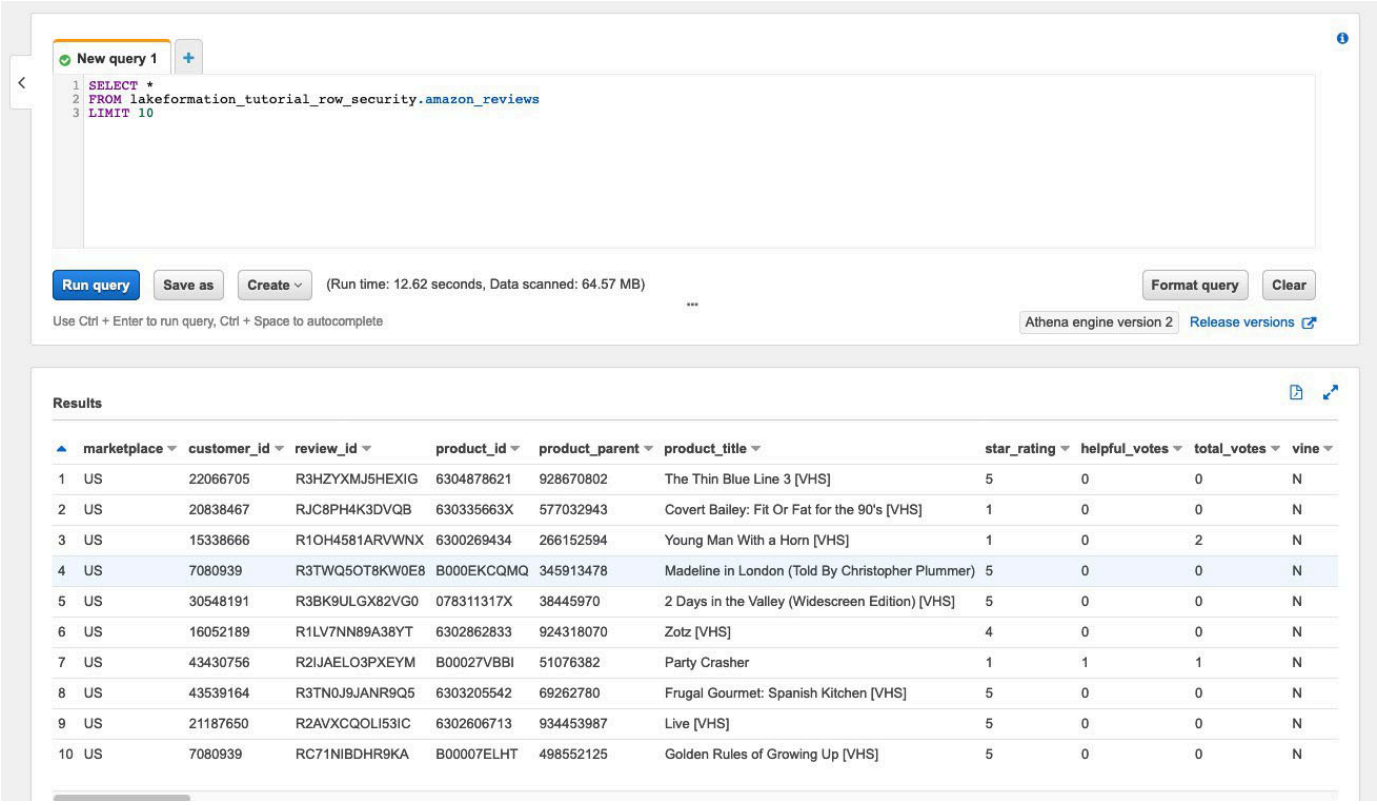
Setelah Anda mengatur lingkungan, Anda dapat menanyakan tabel ulasan produk. Pertama kueri tabel tanpa kontrol akses tingkat baris untuk memastikan Anda dapat melihat data. Jika Anda menjalankan kueri di Amazon Athena untuk pertama kalinya, Anda perlu mengonfigurasi lokasi hasil kueri.

Kueri tabel tanpa kontrol akses tingkat baris

1. Masuk ke Athena konsol di https://console.aws.amazon.com/athena/sebagai `DatalakeAdmin` pengguna, dan jalankan kueri berikut:

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

   Tangkapan layar berikut menunjukkan hasil kueri. Tabel ini hanya memiliki satu partisi`product_category=Video`, sehingga setiap rekaman adalah komentar ulasan untuk produk video.

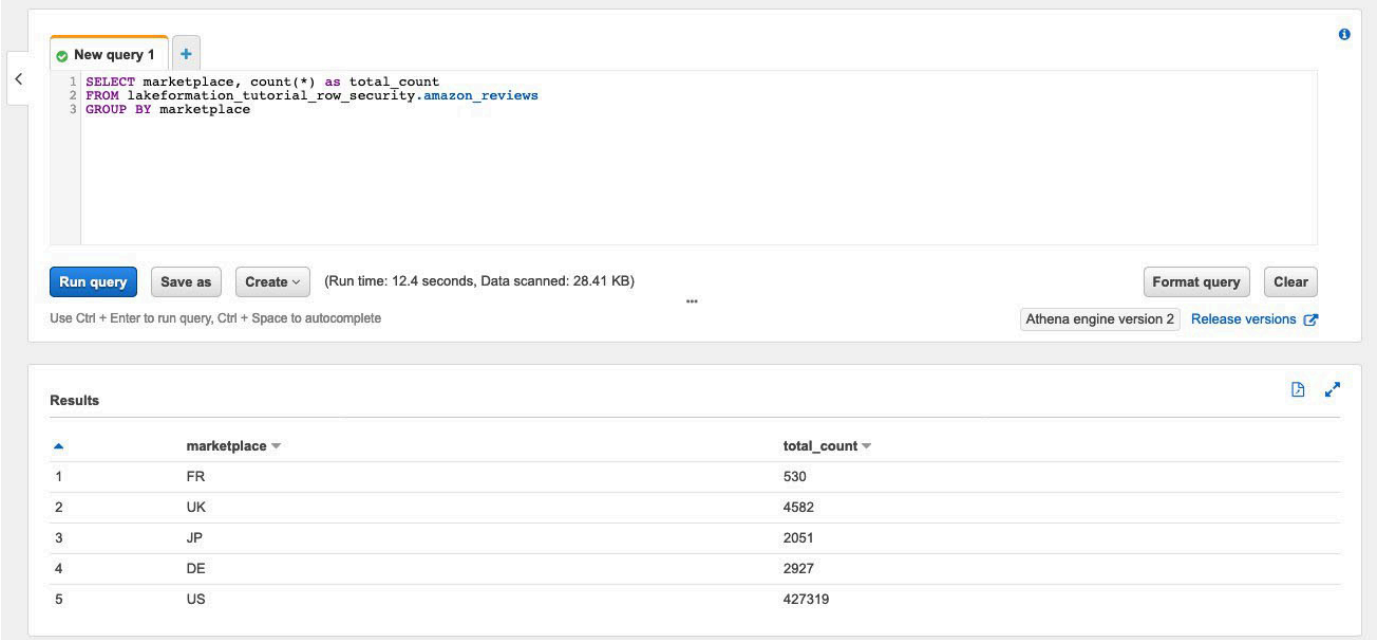


	marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine
1	US	22066705	R3HZYXMJ5HEXIG	6304878621	928670802	The Thin Blue Line 3 [VHS]	5	0	0	N
2	US	20838467	RJC8PH4K3DVQB	630335663X	577032943	Covert Bailey: Fit Or Fat for the 90's [VHS]	1	0	0	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	Zotz [VHS]	4	0	0	N
7	US	43430756	R2IJAELO3PXEYM	B00027VBBI	51076382	Party Crasher	1	1	1	N
8	US	43539164	R3TN0J9JANR9Q5	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. Selanjutnya, jalankan kueri agregasi untuk mengambil jumlah total catatan per. `marketplace`

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

Tangkapan layar berikut menunjukkan hasil kueri. `marketplace`Kolom memiliki lima nilai yang berbeda. Pada langkah selanjutnya, Anda akan mengatur filter berbasis baris menggunakan kolom. `marketplace`



Langkah 3: Siapkan filter data dan berikan izin

Tutorial ini menggunakan dua analis data: satu bertanggung jawab untuk pasar AS dan satu lagi untuk pasar Jepang. Setiap analis menggunakan Athena untuk menganalisis ulasan pelanggan hanya untuk pasar spesifik mereka. Buat dua filter data yang berbeda, satu untuk analis yang bertanggung jawab atas pasar AS, dan satu lagi untuk yang bertanggung jawab atas pasar Jepang. Kemudian, berikan izin masing-masing kepada analis.

Buat filter data dan berikan izin

1. Buat filter untuk membatasi akses ke US `marketplace` data.

   a. Masuk ke konsol Lake Formation https://console.aws.amazon.com/lakeformation/di wilayah US East (Virginia N.) sebagai `DatalakeAdmin` pengguna.

   b. Pilih Filter data.

   c. Pilih Buat filter baru.

   d. Untuk nama filter Data, masukkan`amazon_reviews_US`.

   e. Untuk database Target, pilih database`lakeformation_tutorial_row_security`.

   f. Untuk tabel Target, pilih tabel`amazon_reviews`.

   g. Untuk akses tingkat kolom, biarkan sebagai default.

   h. Untuk ekspresi filter Baris, masukkan`marketplace='US'`.

i. Pilih Buat filter.

2. Buat filter untuk membatasi akses ke `marketplace` data Jepang.

   a. Pada halaman Filter data, pilih Buat filter baru.

   b. Untuk nama filter Data, masukkan`amazon_reviews_JP`.

   c. Untuk database Target, pilih database`lakeformation_tutorial_row_security`.

   d. Untuk tabel Target, pilih`table amazon_reviews`.

   e. Untuk akses tingkat kolom, biarkan sebagai default.

   f. Untuk ekspresi filter Baris, masukkan`marketplace='JP'`.

   g. Pilih Buat filter.

3. Selanjutnya, berikan izin kepada analis data menggunakan filter data ini. Ikuti langkah-langkah berikut untuk memberikan izin kepada analis data AS (`DataAnalystUS`):

   a. Di bawah Izin, pilih Izin data lake.

   b. Di bawah Izin data, pilih Hibah.

   c. Untuk Prinsipal, pilih pengguna dan peran IAM, lalu pilih peran. `DataAnalystUS`

   d. Untuk tag LF atau sumber katalog, pilih Sumber daya katalog data bernama.

   e. Untuk Database, pilih`lakeformation_tutorial_row_security`.

   f. Untuk tabel-opsional, pilih. `amazon_reviews`

   g. Untuk filter Data — opsional ¸ pilih`amazon_reviews_US`.

   h. Untuk izin filter data, pilih Pilih.

   i. PilihIzin.

4. Ikuti langkah-langkah berikut untuk memberikan izin kepada analis data Jepang (`DataAnalystJP`):

   a. Di bawah Izin, pilih Izin data lake.

   b. Di bawah Izin data, pilih Hibah.

   c. Untuk Prinsipal, pilih pengguna dan peran IAM, lalu pilih peran. `DataAnalystJP`

   d. Untuk tag LF atau sumber katalog, pilih Sumber daya katalog data bernama.

   e. Untuk Database, pilih`lakeformation_tutorial_row_security`.

   f. Untuk tabel-opsional, pilih. `amazon_reviews`

   g. Untuk filter Data — opsional ¸ pilih`amazon_reviews_JP`.

h. Untuk izin filter data, pilih Pilih.

i. PilihIzin.

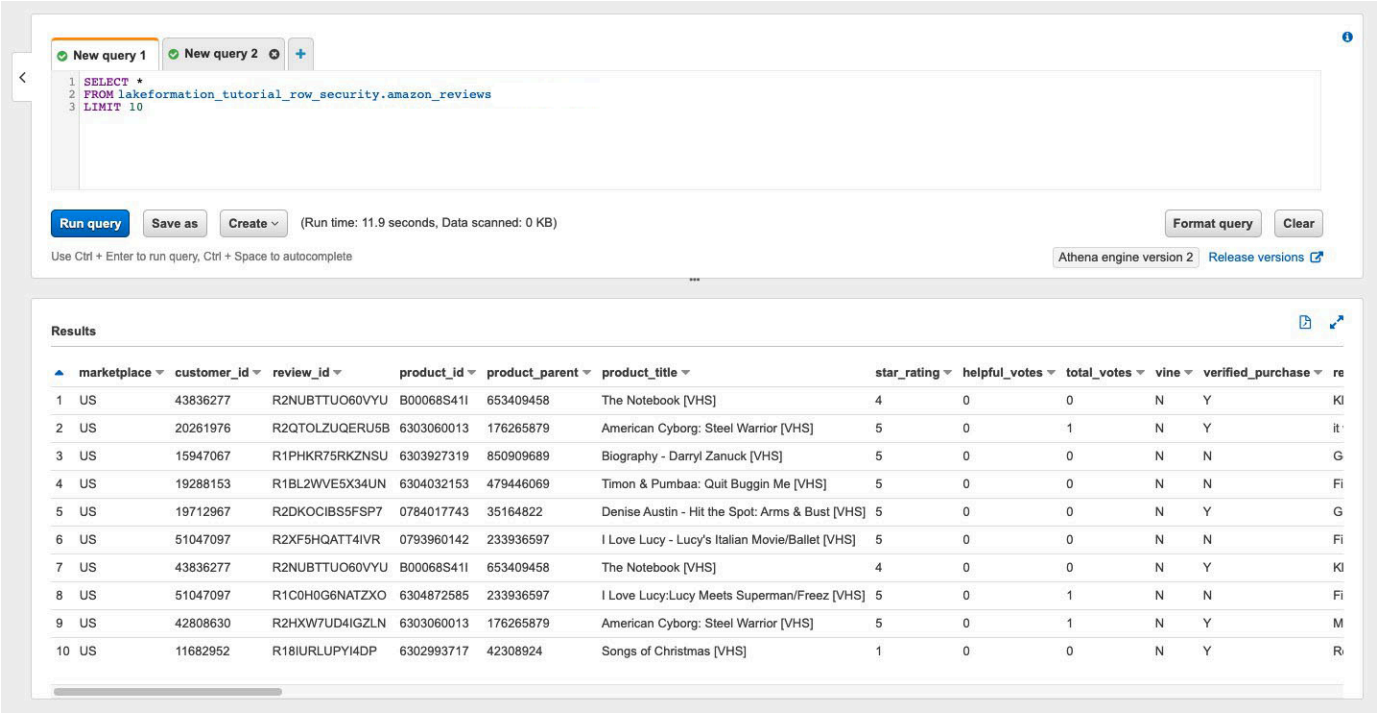
Langkah 4: Kueri dengan filter data

Dengan filter data yang dilampirkan pada tabel ulasan produk, jalankan beberapa kueri dan lihat bagaimana izin diberlakukan oleh Lake Formation.

1. Masuk ke konsol Athena di https://console.aws.amazon.com/athena/sebagai pengguna. `DataAnalystUS`
2. Jalankan kueri berikut untuk mengambil beberapa catatan, yang difilter berdasarkan izin tingkat baris yang kami tentukan:

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

Tangkapan layar berikut menunjukkan hasil kueri.



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

3. Demikian pula, jalankan kueri untuk menghitung jumlah total catatan per pasar.

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

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

Hasil query hanya menunjukkan `marketplace` US dalam hasil. Ini karena pengguna hanya diperbolehkan untuk melihat baris di mana nilai `marketplace` kolom sama denganUS.

4. Beralih ke `DataAnalystJP` pengguna dan jalankan kueri yang sama.

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

Hasil kueri hanya menunjukkan catatan milik `JPmarketplace`.

5. Jalankan kueri untuk menghitung jumlah total catatan per`marketplace`.

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

Hasil kueri hanya menunjukkan baris milik `JPmarketplace`.

Langkah 5: Bersihkan AWS sumber daya

Pembersihan sumber daya

Untuk mencegah biaya yang tidak diinginkan ke Anda Akun AWS, Anda dapat menghapus AWS sumber daya yang Anda gunakan untuk tutorial ini.

- Hapus tumpukan formasi cloud.

Berbagi data lake menggunakan kontrol akses berbasis tag Lake Formation dan sumber daya bernama

Tutorial ini menunjukkan bagaimana Anda dapat mengkonfigurasi AWS Lake Formation untuk aman berbagi data yang disimpan dalam data lake dengan beberapa perusahaan, organisasi, atau unit bisnis, tanpa harus menyalin seluruh database. Ada dua opsi untuk berbagi database dan tabel Anda dengan yang lain Akun AWS dengan menggunakan kontrol akses lintas akun Lake Formation:

- Kontrol akses berbasis tag Lake Formation (disarankan)

  Kontrol akses berbasis tag Lake Formation adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Dalam Lake Formation, atribut ini disebut LF-tag. Untuk detail selengkapnya, lihat Mengelola data lake menggunakan kontrol akses berbasis tag Lake Formation.

- Lake Formation bernama sumber daya

  Metode sumber daya bernama Lake Formation adalah strategi otorisasi yang mendefinisikan izin untuk sumber daya. Sumber daya termasuk database, tabel, dan kolom. Administrator data lake dapat menetapkan dan mencabut izin pada sumber daya Lake Formation. Untuk detail selengkapnya, lihat Cross-account berbagi data di Lake Formation.

  Sebaiknya gunakan sumber daya bernama jika administrator data lake lebih suka memberikan izin secara eksplisit ke sumber daya individu. Saat Anda menggunakan metode sumber daya bernama untuk memberikan izin Lake Formation pada sumber daya Katalog Data ke akun eksternal, Lake Formation menggunakan AWS Resource Access Manager (AWS RAM) untuk membagikan sumber daya.

Topik

- Audiens yang dituju
- Konfigurasikan pengaturan Katalog Data Lake Formation di akun produsen
- Langkah 1: Menyediakan sumber daya Anda menggunakan AWS CloudFormation template
- Langkah 2: Prasyarat berbagi lintas akun Lake Formation
- Langkah 3: Terapkan berbagi lintas akun menggunakan metode kontrol akses berbasis tag
- Langkah 4: Menerapkan metode sumber daya bernama
- Langkah 5: Bersihkan AWS sumber daya

Audiens yang dituju

Tutorial ini ditujukan untuk pengelola data, insinyur data, dan analis data. Dalam hal berbagi tabel Katalog Data dari AWS Glue dan mengelola izin di Lake Formation, pengelola data dalam akun penghasil memiliki kepemilikan fungsional berdasarkan fungsi yang mereka dukung, dan dapat memberikan akses ke berbagai konsumen, organisasi eksternal, dan akun. Tabel berikut mencantumkan peran yang digunakan dalam tutorial ini:

Peran	Deskripsi
DataLakeAdminProducer	Pengguna IAM admin danau data memiliki akses sebagai berikut: • Akses baca, tulis, dan perbarui lengkap ke semua sumber daya di Katalog Data • Kemampuan untuk memberikan izin ke sumber daya • Dapat membuat link sumber daya untuk tabel bersama • Dapat melampirkan tag LF ke sumber daya, yang menyediakan akses ke prinsipal berdasarkan kebijakan apa pun yang dibuat oleh pengelola data
DataLakeAdminConsumer	Pengguna IAM admin danau data memiliki akses sebagai berikut: • Akses baca, tulis, dan perbarui lengkap ke semua sumber daya di Katalog Data • Kemampuan untuk memberikan izin ke sumber daya • Dapat membuat link sumber daya untuk tabel bersama • Dapat melampirkan tag LF ke sumber daya, yang menyediakan akses ke prinsipal berdasarkan kebijakan apa pun yang dibuat oleh pengelola data
DataAnalyst	DataAnalyst Pengguna memiliki akses berikut: • Akses halus ke sumber daya yang dibagikan oleh kebijakan akses berbasis tag Lake Formation atau menggunakan metode sumber daya bernama

Konfigurasikan pengaturan Katalog Data Lake Formation di akun produsen

Sebelum Anda memulai tutorial ini, Anda harus memiliki Akun AWS yang dapat Anda gunakan untuk masuk sebagai pengguna administratif dengan izin yang benar. Untuk informasi selengkapnya, lihat Selesaikan inisial AWS tugas konfigurasi.

Tutorial mengasumsikan bahwa Anda sudah familiar dengan IAM. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Konfigurasikan pengaturan Katalog Data Lake Formation di akun produsen

> **Note**
>
> Dalam tutorial ini, akun yang memiliki tabel sumber disebut akun produser, dan akun yang membutuhkan akses ke tabel sumber disebut akun konsumen.

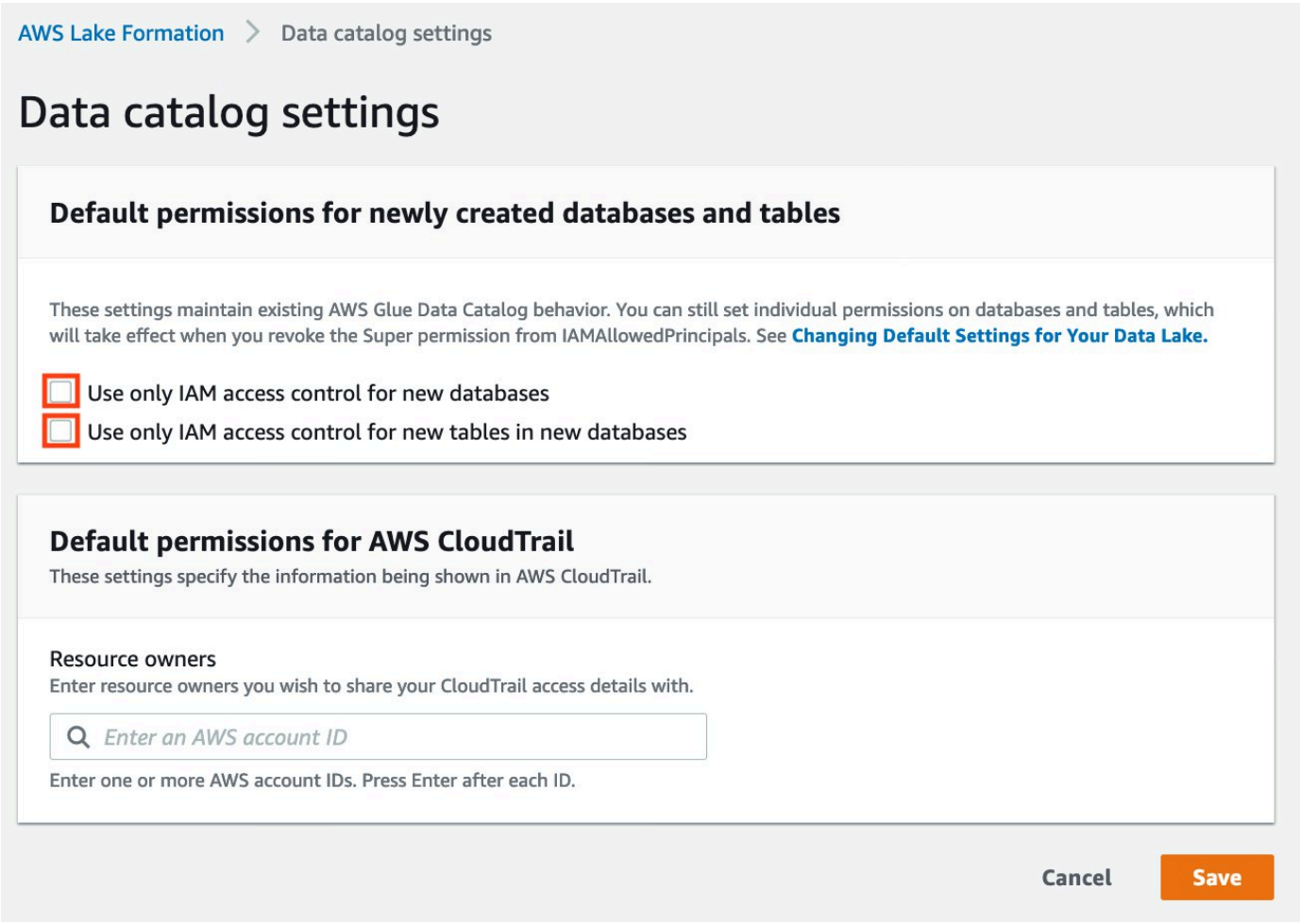
Lake Formation menyediakan model manajemen izinnya sendiri. Untuk mempertahankan kompatibilitas mundur dengan model izin IAM, `Super` izin diberikan kepada grup `IAMAllowedPrincipals` pada semua AWS Glue Data Catalog sumber daya yang ada secara default. Selain itu, Gunakan hanya pengaturan kontrol akses IAM yang diaktifkan untuk sumber daya Katalog Data baru. Tutorial ini menggunakan kontrol akses berbutir halus menggunakan izin Lake Formation dan menggunakan kebijakan IAM untuk kontrol akses berbutir kasar. Lihat Metode untuk kontrol akses berbutir halus untuk detail. Oleh karena itu, sebelum Anda menggunakan AWS CloudFormation templat untuk pengaturan cepat, Anda perlu mengubah pengaturan Katalog Data Formasi Danau di akun produsen.

> **Important**
>
> Pengaturan ini memengaruhi semua database dan tabel yang baru dibuat, jadi kami sangat menyarankan untuk menyelesaikan tutorial ini di akun non-produksi atau di akun baru. Juga, jika Anda menggunakan akun bersama (seperti akun pengembangan perusahaan Anda), pastikan itu tidak memengaruhi sumber daya orang lain. Jika Anda lebih suka mempertahankan pengaturan keamanan default, Anda harus menyelesaikan langkah ekstra saat berbagi sumber daya ke akun lain, di mana Anda mencabut izin Super default dari `IAMAllowedPrincipals` database atau tabel. Kami membahas detailnya nanti dalam tutorial ini.

Untuk mengonfigurasi pengaturan Katalog Data Lake Formation di akun produsen, selesaikan langkah-langkah berikut:

1. Masuk ke Konsol Manajemen AWS menggunakan akun produser sebagai pengguna admin, atau sebagai pengguna dengan izin Lake Formation `PutDataLakeSettings` API.
2. Di konsol Lake Formation, di panel navigasi, di bawah Katalog Data, pilih Pengaturan.
3. Hapus pilih Gunakan hanya kontrol akses IAM untuk database baru dan Gunakan hanya kontrol akses IAM untuk tabel baru di database baru

   Pilih Simpan.



   Selain itu, Anda dapat menghapus `CREATE_DATABASE` izin untuk peran dan **`IAMAllowedPrincipals`** tugas Administratif, pembuat Database. Hanya dengan begitu, Anda dapat mengatur siapa yang dapat membuat database baru melalui izin Lake Formation.

Langkah 1: Menyediakan sumber daya Anda menggunakan AWS CloudFormation template

CloudFormation Template untuk akun produsen menghasilkan sumber daya berikut:

- Bucket Amazon S3 untuk berfungsi sebagai data lake.
- Fungsi Lambda (untuk sumber daya kustom yang didukung Lambda CloudFormation). Kami menggunakan fungsi ini untuk menyalin file data sampel dari bucket Amazon S3 publik ke bucket Amazon S3 Anda.
- Pengguna dan kebijakan IAM: DataLakeAdminProducer.
- Pengaturan dan izin Lake Formation yang sesuai termasuk:
  - Mendefinisikan administrator danau data Lake Formation di akun produsen
  - Mendaftarkan bucket Amazon S3 sebagai lokasi danau data Lake Formation (akun produsen)
- AWS Glue Data Catalog Database, tabel, dan partisi. Karena ada dua opsi untuk berbagi sumber daya Akun AWS, template ini membuat dua set database dan tabel terpisah.

CloudFormation Template untuk akun konsumen menghasilkan sumber daya berikut:

- Pengguna dan kebijakan IAM:
  - DataLakeAdminConsumer
  - DataAnalyst
- AWS Glue Data Catalog Database. Database ini untuk membuat tautan sumber daya ke sumber daya bersama.

Buat sumber daya Anda di akun produsen

1. Masuk ke AWS CloudFormation konsol di https://console.aws.amazon.com/cloudformation di wilayah US East (Virginia N.).
2. Pilih Launch Stack.
3. Pilih Berikutnya.
4. Untuk nama Stack, masukkan nama tumpukan, seperti`stack-producer`.
5. Di bagian Konfigurasi Pengguna, masukkan nama pengguna dan kata sandi untuk `ProducerDatalakeAdminUserName` dan`ProducerDatalakeAdminUserPassword`.

6. Untuk DataLakeBucketName, masukkan nama bucket danau data Anda. Nama ini harus unik secara global.
7. Untuk DatabaseNamedan TableName, tinggalkan nilai default.
8. Pilih Berikutnya.
9. Di halaman berikutnya, pilih Berikutnya.
10. Tinjau detail di halaman akhir dan pilih Saya akui yang AWS CloudFormation mungkin membuat sumber daya IAM.
11. Pilih Buat.

    Pembuatan tumpukan bisa memakan waktu hingga satu menit.

Buat sumber daya Anda di akun konsumen

1. Masuk ke AWS CloudFormation konsol di https://console.aws.amazon.com/cloudformation di wilayah US East (Virginia N.).
2. Pilih Launch Stack.
3. Pilih Berikutnya.
4. Untuk nama Stack, masukkan nama tumpukan, seperti`stack-consumer`.
5. Di bagian Konfigurasi Pengguna, masukkan nama pengguna dan kata sandi untuk `ConsumerDatalakeAdminUserName` dan`ConsumerDatalakeAdminUserPassword`.
6. Untuk `DataAnalystUserName` dan`DataAnalystUserPassword`, masukkan nama pengguna dan kata sandi yang Anda inginkan untuk pengguna IAM analis data.
7. Untuk DataLakeBucketName, masukkan nama bucket danau data Anda. Nama ini harus unik secara global.
8. Untuk DatabaseName, tinggalkan nilai default.
9. Untuk`AthenaQueryResultS3BucketName`, masukkan nama bucket Amazon S3 yang menyimpan hasil kueri Amazon Athena. Jika Anda tidak memilikinya, buat ember Amazon S3.
10. Pilih Berikutnya.
11. Di halaman berikutnya, pilih Berikutnya.
12. Tinjau detail di halaman akhir dan pilih Saya akui yang AWS CloudFormation mungkin membuat sumber daya IAM.
13. Pilih Buat.

    Pembuatan tumpukan dapat memakan waktu hingga satu menit.

> **Note**
>
> Setelah menyelesaikan tutorial, hapus tumpukan CloudFormation untuk menghindari biaya yang dikenakan. Verifikasi bahwa sumber daya berhasil dihapus dalam status acara untuk tumpukan.

Langkah 2: Prasyarat berbagi lintas akun Lake Formation

Sebelum berbagi sumber daya dengan Lake Formation, ada prasyarat untuk metode kontrol akses berbasis tag dan metode sumber daya bernama.

Prasyarat berbagi data lintas akun kontrol akses berbasis tag lengkap

- Untuk informasi selengkapnya tentang persyaratan berbagi data lintas akun, lihat Prasyarat bagian di bagian berbagi data lintas akun.

  Untuk membagikan sumber daya Katalog Data dengan versi 3 atau lebih tinggi dari setelan versi Cross account, pemberi harus memiliki izin IAM yang ditentukan dalam kebijakan AWS `AWSLakeFormationCrossAccountManager` terkelola di akun Anda.

  Jika Anda menggunakan versi 1 atau versi 2 dari setelan versi Cross account, sebelum Anda dapat menggunakan metode kontrol akses berbasis tag untuk memberikan akses lintas akun ke sumber daya, Anda harus menambahkan objek `JSON` izin berikut ke kebijakan sumber daya Katalog Data di akun produsen. Ini memberikan izin akun konsumen untuk mengakses Katalog Data ketika `glue:EvaluatedByLakeFormationTags` benar. Selain itu, kondisi ini menjadi benar untuk sumber daya yang Anda berikan izin menggunakan tag izin Lake Formation ke akun konsumen. Kebijakan ini diperlukan Akun AWS untuk setiap yang Anda berikan izin.

  Kebijakan berikut harus berada dalam `Statement` elemen. Kami membahas kebijakan IAM lengkap di bagian selanjutnya.

  ```
  {
      "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
        }
    }
}
```

Prasyarat berbagi lintas akun metode sumber daya bernama lengkap

1. Jika tidak ada kebijakan sumber daya Katalog Data di akun Anda, hibah lintas akun Lake Formation yang Anda lakukan melanjutkan seperti biasa. Namun, jika kebijakan sumber daya Katalog Data ada, Anda harus menambahkan pernyataan berikut untuk mengizinkan hibah lintas akun Anda berhasil jika dibuat dengan metode sumber daya bernama. Jika Anda berencana untuk hanya menggunakan metode sumber daya bernama, atau hanya metode kontrol akses berbasis tag, Anda dapat melewati langkah ini. Dalam tutorial ini, kita mengevaluasi kedua metode, dan kita perlu menambahkan kebijakan berikut.

   Kebijakan berikut harus berada dalam `Statement` elemen. Kami membahas kebijakan IAM lengkap di bagian selanjutnya.

```
{
          "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. Selanjutnya, tambahkan kebijakan AWS Glue Data Catalog sumber daya menggunakan AWS Command Line Interface (AWS CLI).

   Jika Anda memberikan izin lintas akun dengan menggunakan metode kontrol akses berbasis tag dan metode sumber daya bernama, Anda harus menetapkan `EnableHybrid` argumen ke 'true' saat menambahkan kebijakan sebelumnya. Karena opsi ini saat ini tidak didukung di konsol, dan Anda harus menggunakan `glue:PutResourcePolicy` API dan AWS CLI.

   Pertama, buat dokumen kebijakan (seperti policy.json) dan tambahkan dua kebijakan sebelumnya. Ganti *`consumer-account-id`* dengan Akun AWS penerima hibah, *`region`* dengan Wilayah Katalog Data yang berisi database dan tabel tempat Anda memberikan izin, dan *`account-id`* dengan ID produsen. *`account ID`* Akun AWS

   Masukkan AWS CLI perintah berikut. Ganti *`glue-resource-policy`* dengan nilai yang benar (seperti file: //policy.json).

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

   Untuk informasi lebih lanjut, lihat put-resource-policy.

Langkah 3: Terapkan berbagi lintas akun menggunakan metode kontrol akses berbasis tag

Di bagian ini, kami memandu Anda melalui langkah-langkah tingkat tinggi berikut:

1. Tentukan LF-tag.
2. Tetapkan LF-tag ke sumber daya target.
3. Berikan izin LF-tag ke akun konsumen.
4. Berikan izin data ke akun konsumen.
5. Secara opsional, cabut izin untuk `IAMAllowedPrincipals` pada database, tabel, dan kolom.
6. Buat tautan sumber daya ke tabel bersama.
7. Buat LF-tag dan tetapkan ke database target.
8. Berikan izin data LF-tag ke akun konsumen.

Mendefinisikan LF-tag

> **Note**
>
> Jika Anda masuk ke akun produser Anda, keluar sebelum menyelesaikan langkah-langkah berikut.

1. Masuk ke akun produser sebagai administrator danau data di https://console.aws.amazon.com/lakeformation/. Gunakan nomor akun produsen, nama pengguna IAM (defaultnya adalah`DatalakeAdminProducer`), dan kata sandi yang Anda tentukan selama pembuatan CloudFormation tumpukan.
2. Di konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), di panel navigasi, di bawah Izin, pilih LF-tag dan Izin.
3. Pilih Tambahkan LF-Tag.

Tetapkan LF-tag ke sumber daya target

Tetapkan LF-tag ke sumber daya target dan berikan izin data ke akun lain

Sebagai administrator data lake, Anda dapat melampirkan tag ke sumber daya. Jika Anda berencana untuk menggunakan peran terpisah, Anda mungkin harus memberikan izin menjelaskan dan melampirkan ke peran terpisah.

1. Di panel navigasi, di bawah Katalog Data, pilih Database.
2. Pilih database target (`lakeformation_tutorial_cross_account_database_tbac`) dan pada menu Tindakan, pilih Edit LF-tag.

   Untuk tutorial ini, Anda menetapkan LF-tag ke database, tetapi Anda juga dapat menetapkan LF-tag ke tabel dan kolom.
3. Pilih Tetapkan LF-Tag baru.
4. Tambahkan kunci `Confidentiality` dan nilai`public`.
5. Pilih Simpan.

Berikan izin LF-tag ke akun konsumen

Masih di akun produsen, berikan izin ke akun konsumen untuk mengakses LF-tag.

1. Di panel navigasi, di bawah Izin, pilih LF-tag dan izin.
2. Pilih tab LF-tag, dan pilih kunci dan nilai LF-tag yang sedang dibagikan dengan akun konsumen (kunci **`Confidentiality`** dan nilai). `public`
3. Pilih Berikan izin.
4. Untuk jenis Izin, pilih izin pasangan nilai kunci LF-tag.
5. Untuk Prinsipal, pilih Akun eksternal.
6. Masukkan Akun AWS ID target.

   Akun AWS dalam organisasi yang sama muncul secara otomatis. Jika tidak, Anda harus memasukkan Akun AWS ID secara manual.
7. Di bawah Izin, pilih Jelaskan.

   Ini adalah izin yang diberikan ke akun konsumen. Izin yang dapat diberikan adalah izin yang dapat diberikan oleh akun konsumen kepada prinsipal lain.
8. PilihIzin.

   Pada titik ini, administrator danau data konsumen harus dapat menemukan tag kebijakan yang dibagikan melalui konsol Lake Formation akun konsumen, di bawah Izin, LF-tag, dan izin.

Berikan izin data ke akun konsumen

Kami sekarang akan menyediakan akses data ke akun konsumen dengan menentukan ekspresi LF-tag dan memberikan akses akun konsumen ke tabel atau database apa pun yang cocok dengan ekspresi..

1. Di panel navigasi, di bawah Izin, Izin danau data, pilih Hibah.
2. Untuk Prinsipal, pilih Akun eksternal, dan masukkan ID target. Akun AWS
3. Untuk LF-tag atau sumber katalog, pilih kunci dan nilai LF-tag yang sedang dibagikan dengan akun konsumen (kunci **`Confidentiality`** dan nilai). `public`
4. Untuk Izin, di bawah Sumber daya yang cocok dengan LF-tag (disarankan) pilih Tambahkan LF-tag.
5. Pilih kunci dan nilai tag yang dibagikan dengan akun konsumen (kunci `Confidentiality` dan nilai`public`).
6. Untuk izin Database, pilih Jelaskan di bawah Izin database untuk memberikan izin akses di tingkat database.

7. Administrator danau data konsumen harus dapat menemukan tag kebijakan yang dibagikan melalui akun konsumen di konsol Lake Formation di, di bawah Izin https://console.aws.amazon.com/lakeformation/, peran dan tugas Administratif, LF-tag.
8. Pilih Jelaskan di bawah Izin yang dapat diberikan sehingga akun konsumen dapat memberikan izin tingkat database kepada penggunanya.
9. Untuk izin Tabel dan kolom, pilih Pilih dan Jelaskan di bawah Izin tabel.
10. Pilih Pilih dan Jelaskan di bawah Izin yang dapat diberikan.
11. PilihIzin.

Mencabut izin untuk **`IAMAllowedPrincipals`** database, tabel, dan kolom (Opsional).

Di awal tutorial ini, Anda mengubah pengaturan Katalog Data Lake Formation. Jika Anda melewatkan bagian itu, langkah ini diperlukan. Jika Anda mengubah pengaturan Katalog Data Lake Formation, Anda dapat melewati langkah ini.

Pada langkah ini, kita perlu mencabut izin Super default dari `IAMAllowedPrincipals` database atau tabel. Lihat Langkah 4: Alihkan penyimpanan data Anda ke model izin Lake Formation untuk detail.

Sebelum mencabut izin`IAMAllowedPrincipals`, pastikan bahwa Anda memberikan kepala sekolah IAM yang ada dengan izin yang diperlukan melalui Lake Formation. Ini termasuk tiga langkah:

1. Tambahkan izin IAM ke pengguna IAM target atau peran dengan `GetDataAccess` tindakan Lake Formation (dengan kebijakan IAM).
2. Berikan pengguna atau peran IAM target dengan izin data Lake Formation (ubah, pilih, dan sebagainya).
3. Kemudian, cabut izin untuk. `IAMAllowedPrincipals` Jika tidak, setelah mencabut izin untuk`IAMAllowedPrincipals`, prinsipal IAM yang ada mungkin tidak lagi dapat mengakses database target atau Katalog Data.

   Pencabutan izin Super untuk `IAMAllowedPrincipals` diperlukan saat Anda ingin menerapkan model izin Lake Formation (bukan model kebijakan IAM) untuk mengelola akses pengguna dalam satu akun atau di antara beberapa akun menggunakan model izin Lake Formation. Anda tidak perlu mencabut izin `IAMAllowedPrincipals` untuk tabel lain di mana Anda ingin mempertahankan model kebijakan IAM tradisional.

Pada titik ini, administrator danau data akun konsumen harus dapat menemukan database dan tabel yang dibagikan melalui akun konsumen di konsol Lake Formation di https://console.aws.amazon.com/lakeformation/, di bawah Data Catalog, database. Jika tidak, konfirmasikan apakah yang berikut ini dikonfigurasi dengan benar:

1. Tag kebijakan dan nilai yang benar ditetapkan ke database dan tabel target.
2. Izin tag dan izin data yang benar ditetapkan ke akun konsumen.
3. Cabut izin super default dari `IAMAllowedPrincipals` database atau tabel.

Buat tautan sumber daya ke tabel bersama

Ketika sumber daya dibagi antar akun, dan sumber daya bersama tidak dimasukkan ke dalam Katalog Data akun konsumen. Untuk membuatnya tersedia, dan menanyakan data dasar tabel bersama menggunakan layanan seperti Athena, kita perlu membuat tautan sumber daya ke tabel bersama. Tautan sumber daya adalah objek Katalog Data yang merupakan tautan ke database atau tabel lokal atau bersama. Lihat perinciannya di Membuat tautan sumber daya. Dengan membuat tautan sumber daya, Anda dapat:

- Tetapkan nama yang berbeda ke database atau tabel yang sejajar dengan kebijakan penamaan sumber daya Katalog Data Anda.
- Gunakan layanan seperti Athena dan Redshift Spectrum untuk menanyakan database atau tabel bersama.

Untuk membuat tautan sumber daya, selesaikan langkah-langkah berikut:

1. Jika Anda masuk ke akun konsumen Anda, keluar.
2. Masuk sebagai administrator danau data akun konsumen. Gunakan ID akun konsumen, nama pengguna IAM (default DatalakeAdminConsumer) dan kata sandi yang Anda tentukan selama pembuatan CloudFormation tumpukan.
3. Di konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), di panel navigasi, di bawah Katalog Data, Database, pilih database bersama. `lakeformation_tutorial_cross_account_database_tbac`

   Jika Anda tidak melihat database, kunjungi kembali langkah-langkah sebelumnya untuk melihat apakah semuanya sudah dikonfigurasi dengan benar.
4. Pilih Lihat Tabel.

5. Pilih tabel bersamaamazon_reviews_table_tbac.
6. Pada menu Tindakan, pilih Buat tautan sumber daya.
7. Untuk nama link Resource, masukkan nama (untuk tutorial ini,amazon_reviews_table_tbac_resource_link).
8. Di bawah Database, pilih database tempat tautan sumber daya dibuat (untuk posting ini, tumpukan CloudFormation n membuat databaselakeformation_tutorial_cross_account_database_consumer).
9. Pilih Buat.

   Tautan sumber daya muncul di bawah Katalog data, Tabel.

Buat LF-tag dan tetapkan ke database target

Tag Lake Formation berada di Katalog Data yang sama dengan sumber daya. Ini berarti bahwa tag yang dibuat di akun produsen tidak tersedia untuk digunakan saat memberikan akses ke tautan sumber daya di akun konsumen. Anda perlu membuat satu set tag LF terpisah di akun konsumen untuk menggunakan kontrol akses berbasis tag LF saat membagikan tautan sumber daya di akun konsumen.

1. Tentukan LF-tag di akun konsumen. Untuk tutorial ini, kita menggunakan kunci Division dan nilaisales,marketing, dananalyst.
2. Tetapkan kunci Division dan nilai LF-tag analyst ke databaselakeformation_tutorial_cross_account_database_consumer, tempat tautan sumber daya dibuat.

Berikan izin data LF-tag kepada konsumen

Sebagai langkah terakhir, berikan izin data LF-tag kepada konsumen.

1. Di panel navigasi, di bawah Izin, Izin danau data, pilih Hibah.
2. Untuk Prinsipal, pilih pengguna dan peran IAM, dan pilih pengguna. DataAnalyst
3. Untuk tag LF atau sumber katalog, pilih Sumber daya yang cocok dengan LF-tag (disarankan).
4. Pilih Divisi kunci dan analis nilai.
5. Untuk izin Database, pilih Jelaskan di bawah Izin database.
6. Untuk izin Tabel dan kolom, pilih Pilih dan Jelaskan di bawah Izin tabel.

7. PilihIzin.
8. Ulangi langkah-langkah ini untuk penggunaDataAnalyst, di mana kunci LF-tag Confidentiality dan nilainya. public

   Pada titik ini, pengguna analis data di akun konsumen harus dapat menemukan database dan tautan sumber daya, dan menanyakan tabel bersama melalui konsol Athena di. https://console.aws.amazon.com/athena/ Jika tidak, konfirmasikan apakah yang berikut ini dikonfigurasi dengan benar:

   - Tautan sumber daya dibuat untuk tabel bersama
   - Anda memberi pengguna akses ke LF-tag yang dibagikan oleh akun produsen
   - Anda memberi pengguna akses ke LF-tag yang terkait dengan tautan sumber daya dan database tempat tautan sumber daya dibuat
   - Periksa apakah Anda menetapkan LF-tag yang benar ke tautan sumber daya, dan ke database tempat tautan sumber daya dibuat

Langkah 4: Menerapkan metode sumber daya bernama

Untuk menggunakan metode sumber daya bernama, kami memandu Anda melalui langkah-langkah tingkat tinggi berikut:

1. Secara opsional, cabut izin untuk IAMAllowedPrincipals pada database, tabel, dan kolom.
2. Berikan izin data ke akun konsumen.
3. Terima pembagian sumber daya dari AWS Resource Access Manager.
4. Buat tautan sumber daya untuk tabel bersama.
5. Berikan izin data untuk tabel bersama kepada konsumen.
6. Berikan izin data untuk tautan sumber daya ke konsumen.

Mencabut izin untuk **IAMAllowedPrincipals** database, tabel, dan kolom (Opsional)

- Di awal tutorial ini, kami mengubah pengaturan Katalog Data Lake Formation. Jika Anda melewatkan bagian itu, langkah ini diperlukan. Untuk petunjuk, lihat langkah opsional di bagian sebelumnya.

Berikan izin data ke akun konsumen

1. > **Note**
   >
   > Jika Anda masuk ke akun produser sebagai pengguna lain, keluar terlebih dahulu.

   Masuk ke konsol Lake Formation saat https://console.aws.amazon.com/lakeformation/menggunakan administrator danau data akun produser menggunakan Akun AWS ID, nama pengguna IAM (default adalah`DatalakeAdminProducer`), dan kata sandi yang ditentukan selama pembuatan CloudFormation tumpukan.
2. Pada halaman Izin, di bawah Izin Danau data pilih Hibah.
3. Di bawah Prinsipal, pilih Akun eksternal, dan masukkan satu atau lebih Akun AWS IDs atau organisasi. AWS IDs Untuk informasi lebih lanjut, lihat: AWS Organizations.

   Organizations yang menjadi milik akun produsen dan Akun AWS dalam organisasi yang sama muncul secara otomatis. Jika tidak, masukkan ID akun atau ID organisasi secara manual.
4. Untuk tag LF atau sumber daya katalog, pilih. `Named data catalog resources`
5. Di bawah Database, pilih database`lakeformation_tutorial_cross_account_database_named_resource`.
6. Pilih Tambahkan LF-Tag.
7. Di bawah Tabel, pilih Semua tabel.
8. Untuk izin kolom Tabel ¸ pilih Pilih, dan Jelaskan di bawah Izin tabel.
9. Pilih Pilih dan Jelaskan, di bawah Izin yang Dapat Diberikan.
10. Secara opsional, untuk izin Data, pilih Akses berbasis kolom sederhana jika manajemen izin tingkat kolom diperlukan.
11. PilihIzin.

Jika Anda belum mencabut izin`IAMAllowedPrincipals`, Anda mendapatkan kesalahan gagal izin Hibah. Pada titik ini, Anda akan melihat tabel target yang AWS RAM dibagikan melalui akun konsumen di bawah Izin, Izin data.

Terima pembagian sumber daya dari AWS RAM

> **Note**
>
> Langkah ini diperlukan hanya untuk berbagi Akun AWS berbasis, bukan untuk berbagi berbasis organisasi.

1. Masuk ke AWS konsol saat https://console.aws.amazon.com/connect/menggunakan administrator danau data akun konsumen menggunakan nama pengguna IAM (default adalah DatalakeAdminConsumer) dan kata sandi yang ditentukan selama pembuatan CloudFormation tumpukan.
2. Di AWS RAM konsol, di panel navigasi, di bawah Berbagi dengan saya, Sumber daya berbagi, pilih sumber daya Lake Formation bersama. Status harus Tertunda.
3. Pilih Action dan Grant.
4. Konfirmasikan detail sumber daya, dan pilih Terima berbagi sumber daya.

   Pada titik ini, administrator danau data akun konsumen harus dapat menemukan sumber daya bersama di konsol Lake Formation (https://console.aws.amazon.com/lakeformation/) di bawah Katalog Data, Database.

Buat tautan sumber daya untuk tabel bersama

- Ikuti petunjuk di Langkah 3: Terapkan berbagi lintas akun menggunakan metode kontrol akses berbasis tag (langkah 6) untuk membuat tautan sumber daya untuk tabel bersama. Beri nama tautan sumber dayaamazon_reviews_table_named_resource_resource_link. Buat tautan sumber daya dalam databaselakeformation_tutorial_cross_account_database_consumer.

Berikan izin data untuk tabel bersama kepada konsumen

Untuk memberikan izin data untuk tabel bersama kepada konsumen, selesaikan langkah-langkah berikut:

1. Di Lake Formationconsole (https://console.aws.amazon.com/lakeformation/), di bawah Izin, izin danau data, pilih Hibah.
2. Untuk Prinsipal, pilih pengguna dan peran IAM, dan pilih pengguna. DataAnalyst

3. Untuk LF-tag atau sumber katalog, pilih Sumber daya katalog data bernama.
4. Di bawah Database, pilih databaselakeformation_tutorial_cross_account_database_named_resource. Jika Anda tidak melihat database pada daftar drop-down, pilih Muat lebih banyak.
5. Di bawah Tabel, pilih tabelamazon_reviews_table_named_resource.
6. Untuk izin Tabel dan kolom, pilih Pilih dan Jelaskan di bawah Izin tabel.
7. PilihIzin.

Berikan izin data untuk tautan sumber daya ke konsumen

Selain memberikan izin pengguna data lake untuk mengakses tabel bersama, Anda juga perlu memberikan izin pengguna data lake untuk mengakses tautan sumber daya.

1. Di konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), di bawah Izin, izin danau data, pilih Hibah.
2. Untuk Prinsipal, pilih pengguna dan peran IAM, dan pilih pengguna. DataAnalyst
3. Untuk LF-tag atau sumber katalog, pilih Sumber daya katalog data bernama.
4. Di bawah Database, pilih databaselakeformation_tutorial_cross_account_database_consumer. Jika Anda tidak melihat database pada daftar drop-down, pilih Muat lebih banyak.
5. Di bawah Tabel, pilih tabelamazon_reviews_table_named_resource_resource_link.
6. Untuk izin tautan Sumber daya, pilih Jelaskan di bawah Izin tautan sumber daya.
7. PilihIzin.

   Pada titik ini, pengguna analis data di akun konsumen harus dapat menemukan database dan tautan sumber daya, dan menanyakan tabel bersama melalui konsol Athena.

   Jika tidak, konfirmasikan apakah yang berikut ini dikonfigurasi dengan benar:

   - Tautan sumber daya dibuat untuk tabel bersama
   - Anda memberi pengguna akses ke tabel yang dibagikan oleh akun produsen
   - Anda memberi pengguna akses ke tautan sumber daya dan database tempat tautan sumber daya dibuat

Langkah 5: Bersihkan AWS sumber daya

Untuk mencegah biaya yang tidak diinginkan ke Anda Akun AWS, Anda dapat menghapus AWS sumber daya yang Anda gunakan untuk tutorial ini.

1. Masuk ke konsol Lake Formation saat https://console.aws.amazon.com/lakeformation/menggunakan akun produser dan hapus atau ubah yang berikut ini:
   - AWS Resource Access Manager berbagi sumber daya
   - Tag Lake Formation
   - CloudFormation tumpukan
   - Pengaturan Lake Formation
   - AWS Glue Data Catalog
2. Masuk ke konsol Lake Formation saat https://console.aws.amazon.com/lakeformation/menggunakan akun konsumen dan hapus atau ubah yang berikut ini:
   - Tag Lake Formation
   - CloudFormation tumpukan

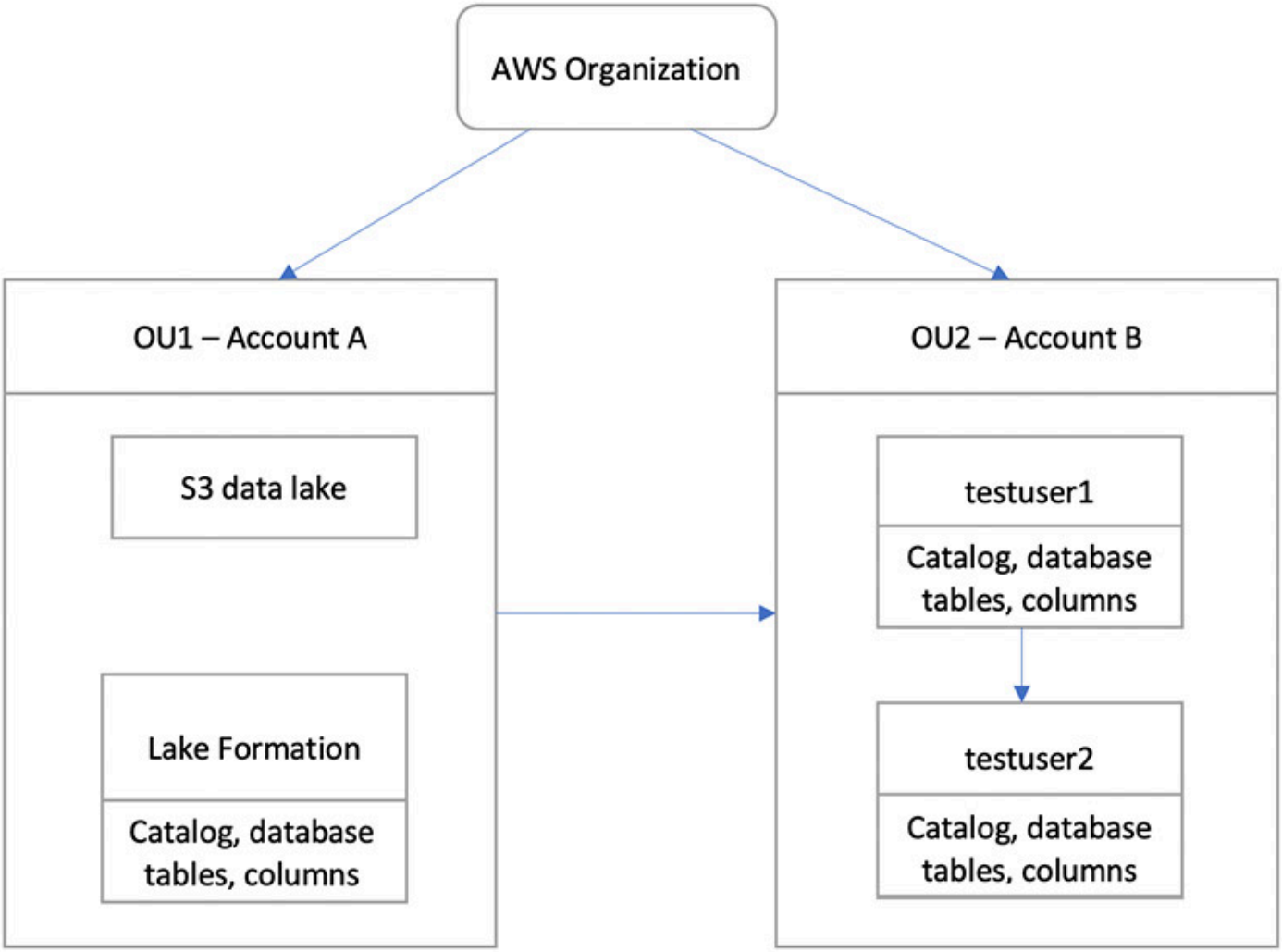
Berbagi data lake menggunakan kontrol akses berbutir halus Lake Formation

Tutorial ini memberikan step-by-step petunjuk tentang bagaimana Anda dapat dengan cepat dan mudah berbagi kumpulan data menggunakan Lake Formation saat mengelola beberapa Akun AWS dengan. AWS Organizations Anda menentukan izin granular untuk mengontrol akses ke data sensitif.

Prosedur berikut juga menunjukkan bagaimana administrator data lake Akun A dapat memberikan akses halus untuk Akun B, dan bagaimana pengguna di Akun B, yang bertindak sebagai pengelola data, dapat memberikan akses halus ke tabel bersama untuk pengguna lain di akun mereka. Pengurus data dalam setiap akun dapat secara independen mendelegasikan akses ke pengguna mereka sendiri, memberikan otonomi masing-masing tim atau lini bisnis (LOB).

Kasus penggunaan mengasumsikan Anda menggunakan AWS Organizations untuk mengelola Anda Akun AWS. Pengguna Akun A dalam satu unit organisasi (OU1) memberikan akses ke pengguna Akun B di OU2. Anda dapat menggunakan pendekatan yang sama ketika tidak menggunakan Organizations, seperti ketika Anda hanya memiliki beberapa akun. Diagram berikut menggambarkan

kontrol akses berbutir halus dari dataset di danau data. Data lake tersedia di Akun A. Administrator data lake Akun A menyediakan akses halus untuk Akun B. Diagram juga menunjukkan bahwa pengguna Akun B menyediakan akses tingkat kolom tabel data lake Akun A ke pengguna lain di Akun B.



Topik

- Audiens yang dituju
- Prasyarat
- Langkah 1: Berikan akses halus ke akun lain
- Langkah 2: Berikan akses halus ke pengguna di akun yang sama

Audiens yang dituju

Tutorial ini ditujukan untuk pengelola data, insinyur data, dan analis data. Tabel berikut mencantumkan peran yang digunakan dalam tutorial ini:

Peran	Deskripsi
Administrator IAM	Pengguna yang memiliki kebijakan AWS terkelola:`AdministratorAccess` .
Administrator danau data	Pengguna yang memiliki kebijakan AWS terkelola: `AWSLakeFormationDataAdmin` melekat pada peran.
Analis data	Pengguna yang memiliki kebijakan AWS terkelola: `AmazonAthenaFullAccess` terlampir.

Prasyarat

Sebelum Anda memulai tutorial ini, Anda harus memiliki Akun AWS yang dapat Anda gunakan untuk masuk sebagai pengguna administratif dengan izin yang benar. Untuk informasi selengkapnya, lihat Selesaikan inisial AWS tugas konfigurasi.

Tutorial mengasumsikan bahwa Anda sudah familiar dengan IAM. Untuk informasi tentang IAM, lihat Panduan Pengguna IAM.

Anda memerlukan sumber daya berikut untuk tutorial ini:

- Dua unit organisasi:
  - OU1 — Berisi Akun A
  - OU2 — Berisi Akun B
- Lokasi danau data Amazon S3 (bucket) di Akun A.
- Pengguna administrator data lake di Akun A. Anda dapat membuat administrator danau data menggunakan konsol Lake Formation (https://console.aws.amazon.com/lakeformation/) atau `PutDataLakeSettings` pengoperasian Lake Formation API.
- Lake Formation dikonfigurasi di Akun A, dan lokasi danau data Amazon S3 terdaftar dengan Lake Formation di Akun A.
- Dua pengguna di Akun B dengan kebijakan terkelola IAM berikut:

  - testuser1 - memiliki kebijakan AWS `AWSLakeFormationDataAdmin` terkelola terlampir.
  - testuser2 - Memiliki kebijakan AWS `AmazonAthenaFullAccess` terkelola terlampir.
- Database testdb dalam database Lake Formation untuk Akun B.

Langkah 1: Berikan akses halus ke akun lain

Pelajari cara administrator data lake Akun A menyediakan akses halus untuk Akun B.

Berikan akses halus ke akun lain

1. Masuk ke Konsol Manajemen AWS https://console.aws.amazon.com/connect/dalam Akun A sebagai administrator danau data.
2. Buka konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), dan pilih Memulai.
3. di panel navigasi, pilih Databases.
4. Pilih Buat basis data.
5. Di bagian Detail database, pilih Database.
6. Untuk Nama, masukkan nama (untuk tutorial ini, kita gunakan`sampledb01`).
7. Pastikan bahwa Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini tidak dipilih. Membiarkan ini tidak dipilih memungkinkan kita untuk mengontrol akses dari Lake Formation.
8. Pilih Buat basis data.
9. Pada halaman Database, pilih database `sampledb01` Anda.
10. Pada menu Actions, pilih Grant.
11. Di bagian Hibah izin, pilih Akun eksternal.
12. Untuk Akun AWS ID atau ID AWS organisasi, masukkan ID akun untuk Akun B di OU2.
13. Untuk Tabel, pilih tabel yang ingin Anda akses ke Akun B (untuk posting ini, kami menggunakan tabel`acc_a_area`). Secara opsional, Anda dapat memberikan akses ke kolom dalam tabel, yang kami lakukan di posting ini.
14. Untuk Sertakan kolom ¸ pilih kolom yang ingin diakses Akun B (untuk posting ini, kami memberikan izin untuk mengetik, nama, dan pengidentifikasi).
15. Untuk Kolom, pilih Sertakan kolom.
16. Untuk izin Tabel, pilih Pilih.

17. Untuk izin yang Dapat Diberikan, pilih Pilih. Izin yang dapat diberikan diperlukan agar pengguna admin di Akun B dapat memberikan izin kepada pengguna lain di Akun B.
18. PilihIzin.
19. Di panel navigasi, pilih Tabel.
20. Anda dapat melihat satu koneksi aktif di Akun AWS dan AWS organisasi dengan bagian akses.

Buat tautan sumber daya

Layanan terintegrasi seperti Amazon Athena tidak dapat langsung mengakses database atau tabel di seluruh akun. Oleh karena itu, Anda perlu membuat tautan sumber daya sehingga Athena dapat mengakses tautan sumber daya di akun Anda ke database dan tabel di akun lain. Buat tautan sumber daya ke tabel (`acc_a_area`) sehingga pengguna Akun B dapat melakukan kueri datanya dengan Athena.

1. Masuk ke AWS konsol https://console.aws.amazon.com/connect/di Akun B sebagai`testuser1`.
2. Pada konsol Lake Formation (https://console.aws.amazon.com/lakeformation/), di panel navigasi, pilih Tabel. Anda akan melihat tabel yang telah disediakan akses oleh Akun A.
3. Pilih tabel `acc_a_area`.
4. Pada menu Tindakan, pilih Buat tautan sumber daya.
5. Untuk nama link Resource, masukkan nama (untuk tutorial ini,`acc_a_area_rl`).
6. Untuk Database, pilih database Anda (`testdb`).
7. Pilih Buat.
8. Di panel navigasi, pilih Tabel.
9. Pilih tabel `acc_b_area_rl`.
10. Pada menu Tindakan, pilih Lihat data.

    Anda diarahkan ke konsol Athena, di mana Anda akan melihat database dan tabel.

    Sekarang Anda dapat menjalankan kueri pada tabel untuk melihat nilai kolom yang aksesnya diberikan ke testuser1 dari Akun B.

Langkah 2: Berikan akses halus ke pengguna di akun yang sama

Bagian ini menunjukkan bagaimana pengguna di Akun B (`testuser1`), bertindak sebagai data steward, menyediakan akses halus ke pengguna lain di akun yang sama (`testuser2`) ke nama kolom dalam tabel bersama. `aac_b_area_rl`

Berikan akses halus ke pengguna di akun yang sama

1. Masuk ke AWS konsol https://console.aws.amazon.com/connect/di Akun B sebagai`testuser1`.
2. Di konsol Lake Formation, di panel navigasi, pilih Tabel.

   Anda dapat memberikan izin pada tabel melalui tautan sumber dayanya. Untuk melakukannya, pada halaman Tabel, pilih tautan sumber daya`aac_b_area_rl`, dan pada menu Tindakan, pilih Hibah sesuai target.
3. Di bagian Hibah izin, pilih Akun saya.
4. Untuk pengguna dan peran IAM ¸ pilih penggun`atestuser2`.
5. Untuk Kolom, pilih nama kolom.
6. Untuk izin Tabel, pilih Pilih.
7. PilihIzin.

   Saat Anda membuat tautan sumber daya, hanya Anda yang dapat melihat dan mengaksesnya. Untuk mengizinkan pengguna lain di akun Anda mengakses tautan sumber daya, Anda perlu memberikan izin pada tautan sumber daya itu sendiri. Anda harus memberikan izin DESCRIPTION atau DROP. Pada halaman Tabel, pilih tabel Anda lagi dan pada menu Tindakan, pilih Grant.
8. Di bagian Hibah izin, pilih Akun saya.
9. Untuk pengguna dan peran IAM, pilih penggun`atestuser2`.
10. Untuk izin tautan Sumber daya ¸ pilih Jelaskan.
11. PilihIzin.
12. Masuk ke AWS konsol di Akun B sebagai`testuser2`.

    Pada konsol Athena (https://console.aws.amazon.com/athena/), Anda akan melihat database dan tabel. `acc_b_area_rl` Anda sekarang dapat menjalankan query pada tabel untuk melihat nilai kolom yang `testuser2` memiliki akses ke.

Orientasi ke izin Lake Formation

AWS Lake Formation menggunakan AWS Glue Data Catalog (Katalog Data) untuk menyimpan metadata untuk danau data Amazon S3 dan sumber data eksternal seperti Amazon Redshift dalam bentuk katalog, database, dan tabel. Metadata dalam Katalog Data diatur dalam hierarki data tiga tingkat yang terdiri dari katalog, database, dan tabel. Ini mengatur data dari berbagai sumber ke dalam wadah logis yang disebut katalog. Database adalah kumpulan tabel. Katalog Data juga berisi tautan sumber daya, yang merupakan tautan ke database dan tabel bersama di akun eksternal, dan digunakan untuk akses lintas akun ke data di danau data. Setiap AWS akun memiliki satu Katalog Data per AWS Wilayah.

Lake Formation menyediakan model izin sistem manajemen basis data relasional (RDBMS) untuk memberikan atau mencabut akses ke katalog, database, tabel, dan kolom di Katalog Data dengan data dasar di Amazon S3.

Sebelum Anda mempelajari tentang detail model izin Lake Formation, akan sangat membantu untuk meninjau informasi latar belakang berikut:

- Danau data yang dikelola oleh Lake Formation berada di lokasi yang ditentukan di Amazon Simple Storage Service (Amazon S3). Katalog Data juga berisi objek katalog. Setiap katalog mewakili data dari sumber seperti gudang data Amazon Redshift, Amazon DynamoDB database, dan sumber data pihak ketiga seperti Snowflake, MySQL, dan lebih dari 30 sumber data eksternal, yang terintegrasi melalui konektor federasi.
- Lake Formation memelihara Katalog Data yang berisi metadata tentang data sumber yang akan diimpor ke danau data Anda, seperti data dalam log dan database relasional, dan tentang data di danau data Anda di Amazon S3. Katalog Data juga berisi metadata tentang data dari sumber data eksternal selain Amazon S3. Metadata diatur sebagai katalog, database, dan tabel. Tabel metadata berisi skema, lokasi, partisi, dan informasi lain tentang data yang mereka wakili. Database metadata adalah kumpulan tabel.
- Katalog Data Lake Formation adalah Katalog Data yang sama yang digunakan olehAWS Glue. Anda dapat menggunakan AWS Glue crawler untuk membuat tabel Katalog Data, dan Anda dapat menggunakan pekerjaan AWS Glue ekstrak, transformasi, dan muat (ETL) untuk mengisi data yang mendasarinya di data lake Anda.
- Katalog, database, dan tabel dalam Katalog Data disebut sebagai sumber daya Katalog Data. Tabel dalam Katalog Data disebut sebagai tabel metadata untuk membedakannya dari tabel di

sumber data atau data tabular di Amazon S3. Data yang ditunjukkan tabel metadata di Amazon S3 atau dalam sumber data disebut sebagai data yang mendasari.

- Prinsipal adalah pengguna atau peran, pengguna atau grup Amazon Quick, pengguna atau grup yang mengautentikasi dengan Lake Formation melalui penyedia SAMB, atau untuk kontrol akses lintas akun, ID AWS akun, ID organisasi, atau ID unit organisasi.
- AWS Gluecrawler membuat tabel metadata, tetapi Anda juga dapat membuat tabel metadata secara manual dengan konsol Lake Formation, API, atau (). AWS Command Line Interface AWS CLI Saat membuat tabel metadata, Anda harus menentukan lokasi. Ketika Anda membuat database, lokasi adalah opsional. Lokasi tabel dapat berupa lokasi Amazon S3 atau lokasi sumber data seperti database Amazon Relational Database Service (Amazon RDS). Lokasi database selalu merupakan lokasi Amazon S3.
- Layanan yang terintegrasi dengan Lake Formation, seperti Amazon Athena dan Amazon Redshift, dapat mengakses Katalog Data untuk mendapatkan metadata dan untuk memeriksa otorisasi untuk menjalankan kueri. Untuk daftar lengkap layanan terintegrasi, lihatAWS integrasi layanan dengan Lake Formation.

Topik

- Ikhtisar izin Lake Formation
- Referensi personas Lake Formation dan izin IAM
- Mengubah pengaturan default untuk data lake
- Izin Formasi Danau Implisit
- Referensi izin Lake Formation
- Mengintegrasikan Pusat Identitas IAM
- Menambahkan lokasi Amazon S3 ke danau data Anda
- Mode akses hibrid
- Membuat objek di AWS Glue Data Catalog
- Mengimpor data menggunakan alur kerja di Lake Formation

Ikhtisar izin Lake Formation

Ada dua jenis izin utama di AWS Lake Formation:

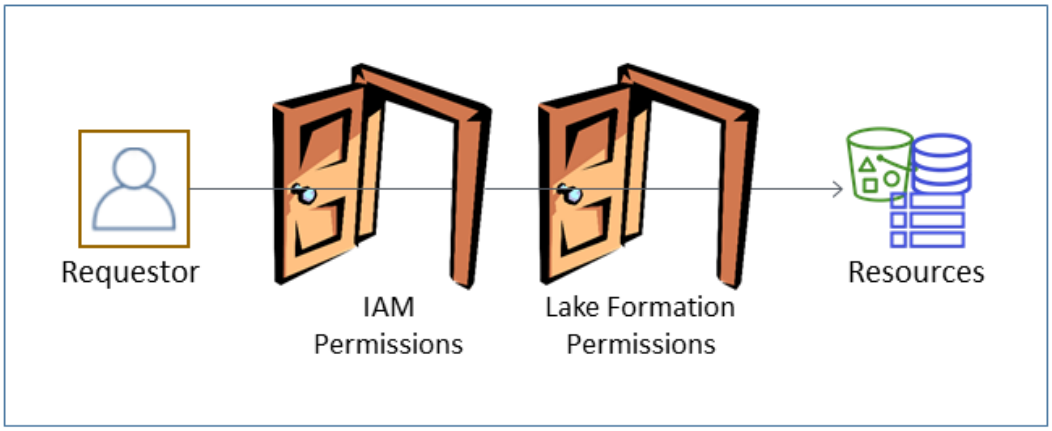
- Akses metadata — Izin pada sumber daya Katalog Data (Izin Katalog Data).

Izin ini memungkinkan prinsipal untuk membuat, membaca, memperbarui, dan menghapus database dan tabel metadata di Katalog Data.

- Akses data yang mendasari — Izin pada lokasi di Amazon Simple Storage Service (Amazon S3) (izin akses data dan izin lokasi data).
  - Izin data lake memungkinkan prinsipal untuk membaca dan menulis data ke lokasi Amazon S3 yang mendasarinya—data yang ditunjukkan oleh sumber daya Katalog Data.
  - Izin lokasi data memungkinkan prinsipal untuk membuat dan mengubah database dan tabel metadata yang mengarah ke lokasi Amazon S3 tertentu.

Untuk kedua area tersebut, Lake Formation menggunakan kombinasi izin Lake Formation dan AWS Identity and Access Management (IAM). Model izin IAM terdiri dari kebijakan IAM. Model izin Lake Formation diimplementasikan sebagai DBMS-style GRANT/REVOKE perintah, sepertiGrant SELECT on *tableName* to *userName*.

Ketika kepala sekolah membuat permintaan untuk mengakses sumber daya Katalog Data atau data yang mendasarinya, agar permintaan berhasil, ia harus lulus pemeriksaan izin oleh IAM dan Lake Formation.



Izin Lake Formation mengontrol akses ke sumber daya Katalog Data, lokasi Amazon S3, dan data dasar di lokasi tersebut. Izin IAM mengontrol akses ke Lake Formation dan AWS Glue API serta sumber daya. Jadi meskipun Anda mungkin memiliki izin Lake Formation untuk membuat tabel metadata di Data Catalog (CREATE_TABLE), operasi Anda gagal jika Anda tidak memiliki izin IAM pada API. glue:CreateTable (Mengapa glue: izin? Karena Lake Formation menggunakan Katalog AWS Glue Data.)

> **Note**
>
> Izin Lake Formation hanya berlaku di Wilayah di mana mereka diberikan.

AWS Lake Formation mengharuskan setiap kepala sekolah (pengguna atau peran) diberi wewenang untuk melakukan tindakan pada sumber daya yang dikelola Lake Formation. Seorang kepala sekolah diberikan otorisasi yang diperlukan oleh administrator danau data atau kepala sekolah lain dengan izin untuk memberikan izin Lake Formation.

Ketika Anda memberikan izin Lake Formation kepada kepala sekolah, Anda dapat secara opsional memberikan kemampuan untuk memberikan izin itu kepada kepala sekolah lain.

Anda dapat menggunakan Lake Formation API, halaman AWS Command Line Interface (AWS CLI), atau izin Data dan lokasi Data pada konsol Lake Formation untuk memberikan dan mencabut izin Lake Formation.

Metode untuk kontrol akses berbutir halus

Dengan data lake, tujuannya adalah untuk memiliki kontrol akses halus ke data. Di Lake Formation, ini berarti kontrol akses berbutir halus ke sumber daya Katalog Data dan lokasi Amazon S3. Anda dapat mencapai kontrol akses berbutir halus dengan salah satu metode berikut.

Metode	Izin Lake Formation	Izin IAM	Komentar
Metode 1	Buka	Fine-grained	Ini adalah metode default untuk kompatibi litas mundur dengan AWS Glue. • Open berarti bahwa izin khusus `Super` diberikan kepada grup `IAMAllowe dPrincipals` , di mana `IAMAllowe dPrincipals` secara otomatis dibuat dan mencakup setiap pengguna dan peran IAM yang diizinkan mengakses sumber daya Katalog Data Anda oleh kebijakan IAM Anda, dan `Super` izin tersebut memungkinkan prinsipal

Metode	Izin Lake Formation	Izin IAM	Komentar
			untuk melakukan setiap operasi Lake Formation yang didukung pada database atau tabel yang diberikan . Hal ini secara efektif menyebabkan akses ke sumber daya Katalog Data dan lokasi Amazon S3 dikendalikan semata-mata oleh kebijakan IAM. Untuk informasi selengkapnya, lihat Mengubah pengaturan default untuk data lake dan Memutakhirkan izin AWS Glue data ke model AWS Lake Formation. • Fine-grainedberarti bahwa kebijakan IAM mengontrol semua akses ke sumber daya Katalog Data dan ke bucket Amazon S3 individual. Pada konsol Lake Formation, metode ini muncul sebagai Gunakan hanya kontrol akses IAM.

Metode	Izin Lake Formation	Izin IAM	Komentar
Metode 2	Fine-grained	Coarse-grained	Ini adalah metode yang direkomendasikan. • Fine-grainedakses berarti memberikan izin Lake Formation terbatas kepada masing-masing kepala sekolah pada sumber daya Katalog Data, lokasi Amazon S3, dan data dasar di lokasi tersebut. • Coarse-grainedberarti izin yang lebih luas pada operasi individual dan akses ke lokasi Amazon S3. Misalnya, kebijakan IAM berbutir kasar mungkin menyertakan `"glue:*"` atau lebih `"glue:Create*"` tepatnya, membiarkan izin `"glue:CreateTables"` Lake Formation untuk mengontrol apakah prinsipal dapat membuat objek katalog atau tidak. Ini juga berarti memberi kepala sekolah akses ke API yang mereka butuhkan untuk melakukan pekerjaan mereka, tetapi mengunci API dan sumber daya lainnya. Misalnya, Anda dapat membuat kebijakan IAM yang memungkinkan prinsipal untuk membuat sumber daya Katalog Data dan membuat serta menjalankan alur kerja, tetapi tidak mengaktifkan pembuatan AWS Glue koneksi atau fungsi yang ditentukan pengguna. Lihat contoh nanti di bagian ini.

> **Important**
>
> Ketahui hal-hal berikut:
>
> - Secara default, Lake Formation memiliki pengaturan kontrol akses Use only IAM yang diaktifkan untuk kompatibilitas dengan perilaku Katalog AWS Glue Data yang ada. Kami menyarankan Anda menonaktifkan pengaturan ini setelah Anda beralih menggunakan izin Lake Formation. Untuk informasi selengkapnya, lihat Mengubah pengaturan default untuk data lake.
> - Administrator data lake dan pembuat database memiliki izin Lake Formation implisit yang harus Anda pahami. Untuk informasi selengkapnya, lihat Izin Formasi Danau Implisit.

Kontrol akses metadata

Untuk kontrol akses sumber daya Katalog Data, diskusi berikut mengasumsikan kontrol akses berbutir halus dengan izin Lake Formation dan kontrol akses berbutir kasar dengan kebijakan IAM.

Ada dua metode berbeda untuk memberikan izin Lake Formation pada sumber daya Katalog Data:

- Kontrol akses sumber daya bernama — Dengan metode ini, Anda memberikan izin pada database atau tabel tertentu dengan menentukan nama database atau tabel. Hibah memiliki formulir ini:

  Berikan izin kepada kepala sekolah tentang sumber daya [dengan opsi hibah].

  Dengan opsi hibah, Anda dapat mengizinkan penerima hibah untuk memberikan izin kepada kepala sekolah lainnya.
- Tag-based kontrol akses — Dengan metode ini, Anda menetapkan satu atau beberapa LF-Tags database, tabel, dan kolom Katalog Data, dan memberikan izin pada satu atau beberapa LF-Tags prinsipal. Masing-masing LF-Tag adalah pasangan kunci-nilai, seperti. `department=sales` Prinsipal LF-Tags yang cocok dengan sumber daya Katalog Data dapat mengakses sumber daya tersebut. LF-Tags Metode ini direkomendasikan untuk danau data dengan sejumlah besar database dan tabel. Ini dijelaskan secara rinci dalamKontrol akses berbasis tag Lake Formation.

Izin yang dimiliki kepala sekolah pada sumber daya adalah gabungan izin yang diberikan oleh kedua metode.

Tabel berikut merangkum izin Lake Formation yang tersedia pada sumber daya Katalog Data. Judul kolom menunjukkan sumber daya tempat izin diberikan.

Katalog	Basis Data	Tabel
CREATE_DATABASE	CREATE_TABLE	ALTER
	ALTER	DROP
	DROP	DESCRIBE
	DESCRIBE	SELECT*
		INSERT*
		DELETE*

Misalnya, CREATE_TABLE izin diberikan pada database. Ini berarti bahwa prinsipal diperbolehkan untuk membuat tabel dalam database itu.

Izin dengan tanda bintang (*) diberikan pada sumber daya Katalog Data, tetapi berlaku untuk data yang mendasarinya. Misalnya, DROP izin pada tabel metadata memungkinkan Anda untuk menjatuhkan tabel dari Katalog Data. Namun, DELETE izin yang diberikan pada tabel yang sama memungkinkan Anda untuk menghapus data dasar tabel di Amazon S3, menggunakan, misalnya, pernyataan SQLDELETE. Dengan izin ini, Anda juga dapat melihat tabel di konsol Lake Formation dan mengambil informasi tentang tabel dengan API. AWS Glue Dengan demikianSELECT,INSERT,, dan DELETE keduanya merupakan izin Katalog Data dan izin akses data.

Saat memberikan SELECT pada tabel, Anda dapat menambahkan filter yang menyertakan atau mengecualikan satu atau beberapa kolom. Ini memungkinkan kontrol akses berbutir halus pada kolom tabel metadata, membatasi kolom yang dapat dilihat pengguna layanan terintegrasi saat menjalankan kueri. Kemampuan ini tidak tersedia hanya dengan menggunakan kebijakan IAM.

Ada juga izin khusus bernamaSuper. SuperIzin memungkinkan kepala sekolah untuk melakukan setiap operasi Lake Formation yang didukung pada database atau tabel yang diberikan. Izin ini dapat hidup berdampingan dengan izin Lake Formation lainnya. Misalnya, Anda dapat memberikanSuper,SELECT, dan INSERT pada tabel metadata. Prinsipal dapat melakukan semua tindakan yang didukung di atas meja, dan ketika Anda mencabutSuper, INSERT izin SELECT dan tetap ada.

Untuk detail tentang setiap izin, lihatReferensi izin Lake Formation.

> **Important**
>
> Untuk dapat melihat tabel Katalog Data yang dibuat oleh pengguna lain, Anda harus diberikan setidaknya satu izin Lake Formation di atas meja. Jika Anda diberikan setidaknya satu izin di atas meja, Anda juga dapat melihat tabel yang berisi database.

Anda dapat memberikan atau mencabut izin Katalog Data menggunakan konsol Lake Formation, API, atau (). AWS Command Line Interface AWS CLI Berikut ini adalah contoh AWS CLI perintah yang memberikan `datalake_user1` izin pengguna untuk membuat tabel dalam `retail` database.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Berikut ini adalah contoh kebijakan IAM kontrol akses kasar yang melengkapi kontrol akses berbutir halus dengan izin Lake Formation. Ini memungkinkan semua operasi pada database atau tabel metadata apa pun.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "glue:*Database*",
                "glue:*Table*",
                "glue:*Partition*"
            ],
            "Resource": "*"
        }
    ]
}
```

Contoh selanjutnya juga berbutir kasar tetapi agak lebih membatasi. Ini memungkinkan operasi hanya-baca pada semua database dan tabel metadata di Katalog Data di akun dan Wilayah yang ditunjuk.

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:*"
        }
    ]
}
```

Bandingkan kebijakan ini dengan kebijakan berikut, yang menerapkan kontrol akses IAM-based berbutir halus. Ini memberikan izin hanya pada subset tabel dalam database metadata manajemen hubungan pelanggan (CRM) di akun dan Wilayah yang ditunjuk.

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*"
            ]
        }
    ]
}
```

Untuk lebih banyak contoh kebijakan kontrol akses berbutir kasar, lihat. Referensi personas Lake Formation dan izin IAM

Kontrol akses data yang mendasari

Ketika AWS layanan terintegrasi meminta akses ke data di lokasi Amazon S3 yang dikendalikan oleh akses, Lake AWS Lake Formation Formation menyediakan kredensi sementara untuk mengakses data.

Untuk mengaktifkan Lake Formation mengontrol akses ke data dasar di lokasi Amazon S3, Anda mendaftarkan lokasi tersebut dengan Lake Formation.

Setelah mendaftarkan lokasi Amazon S3, Anda dapat mulai memberikan izin Lake Formation berikut:

- Izin akses data (`SELECT`,`INSERT`, dan `DELETE`) pada tabel Katalog Data yang mengarah ke lokasi tersebut.
- Izin lokasi data di lokasi tersebut.

Izin lokasi data Lake Formation mengontrol kemampuan untuk membuat sumber daya Katalog Data yang mengarah ke lokasi Amazon S3 tertentu. Izin lokasi data memberikan lapisan keamanan ekstra ke lokasi di dalam danau data. Ketika Anda memberikan `CREATE_TABLE` atau `ALTER` izin kepada kepala sekolah, Anda juga memberikan izin lokasi data untuk membatasi lokasi yang prinsipal dapat membuat atau mengubah tabel metadata.

Lokasi Amazon S3 adalah ember atau awalan di bawah ember, tetapi bukan objek Amazon S3 individual.

Anda dapat memberikan izin lokasi data kepada prinsipal menggunakan konsol Lake Formation, API, atau. AWS CLI Bentuk umum hibah adalah sebagai berikut:

```
grant DATA_LOCATION_ACCESS to principal on S3 location [with grant option]
```

Jika Anda menyertakan`with grant option`, penerima hibah dapat memberikan izin kepada prinsipal lain.

Ingatlah bahwa izin Lake Formation selalu bekerja dalam kombinasi dengan izin AWS Identity and Access Management (IAM) untuk kontrol akses berbutir halus. Untuk read/write izin pada data Amazon S3 yang mendasarinya, izin IAM diberikan sebagai berikut:

Saat mendaftarkan lokasi, Anda menentukan peran IAM yang memberikan read/write izin di lokasi tersebut. Lake Formation mengasumsikan peran itu ketika memasok kredensi sementara ke layanan terintegrasi. AWS Peran tipikal mungkin memiliki kebijakan berikut yang dilampirkan, di mana lokasi terdaftar adalah 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 menyediakan peran terkait layanan yang dapat Anda gunakan saat pendaftaran untuk membuat kebijakan seperti ini secara otomatis. Untuk informasi selengkapnya, lihat Menggunakan peran terkait layanan untuk Lake Formation.

Oleh karena itu, mendaftarkan lokasi Amazon S3 memberikan `s3:` izin IAM yang diperlukan di lokasi tersebut, di mana izin ditentukan oleh peran yang digunakan untuk mendaftarkan lokasi.

> **Important**
>
> Hindari mendaftarkan bucket Amazon S3 yang mengaktifkan Requester pay. Untuk ember yang terdaftar di Lake Formation, peran yang digunakan untuk mendaftarkan ember selalu dipandang sebagai pemohon. Jika bucket diakses oleh AWS akun lain, pemilik bucket akan dikenakan biaya untuk akses data jika peran tersebut milik akun yang sama dengan pemilik bucket.

Untuk read/write akses ke data yang mendasarinya, selain izin Lake Formation, kepala sekolah juga memerlukan izin IAM. `lakeformation:GetDataAccess` Dengan izin ini, Lake Formation memberikan permintaan kredensil sementara untuk mengakses data.

JSON

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

Dalam kebijakan di atas, Anda harus mengatur parameter Resource ke '*' (all. Menentukan sumber daya lain untuk izin ini tidak didukung. Konfigurasi ini memastikan bahwa Lake Formation dapat mengelola akses data di seluruh lingkungan danau data Anda secara efisien.

> **Note**
>
> Amazon Athena mengharuskan pengguna untuk memiliki izin. `lakeformation:GetDataAccess` Layanan terintegrasi lainnya memerlukan peran eksekusi yang mendasarinya untuk memiliki `lakeformation:GetDataAccess` izin.

Izin ini termasuk dalam kebijakan yang disarankan di[Referensi personas Lake Formation dan izin IAM](#).

Untuk meringkas, untuk mengaktifkan kepala sekolah Lake Formation membaca dan menulis data yang mendasarinya dengan akses yang dikendalikan oleh izin Lake Formation:

- Daftarkan lokasi Amazon S3 yang berisi data dengan Lake Formation.
- Prinsipal yang membuat tabel Katalog Data yang mengarah ke lokasi data yang mendasarinya harus memiliki izin lokasi data.
- Kepala sekolah yang membaca dan menulis data dasar harus memiliki izin akses data Lake Formation pada tabel Katalog Data yang mengarah ke lokasi data yang mendasarinya.
- Kepala sekolah yang membaca dan menulis data dasar harus memiliki izin `lakeformation:GetDataAccess` IAM ketika lokasi data yang mendasarinya terdaftar di Lake Formation.

> **Note**
>
> Model izin Lake Formation tidak mencegah akses ke lokasi Amazon S3 melalui Amazon S3 API atau konsol jika Anda memiliki akses ke lokasi tersebut melalui kebijakan IAM atau Amazon S3. Anda dapat melampirkan kebijakan IAM ke kepala sekolah untuk memblokir akses ini.

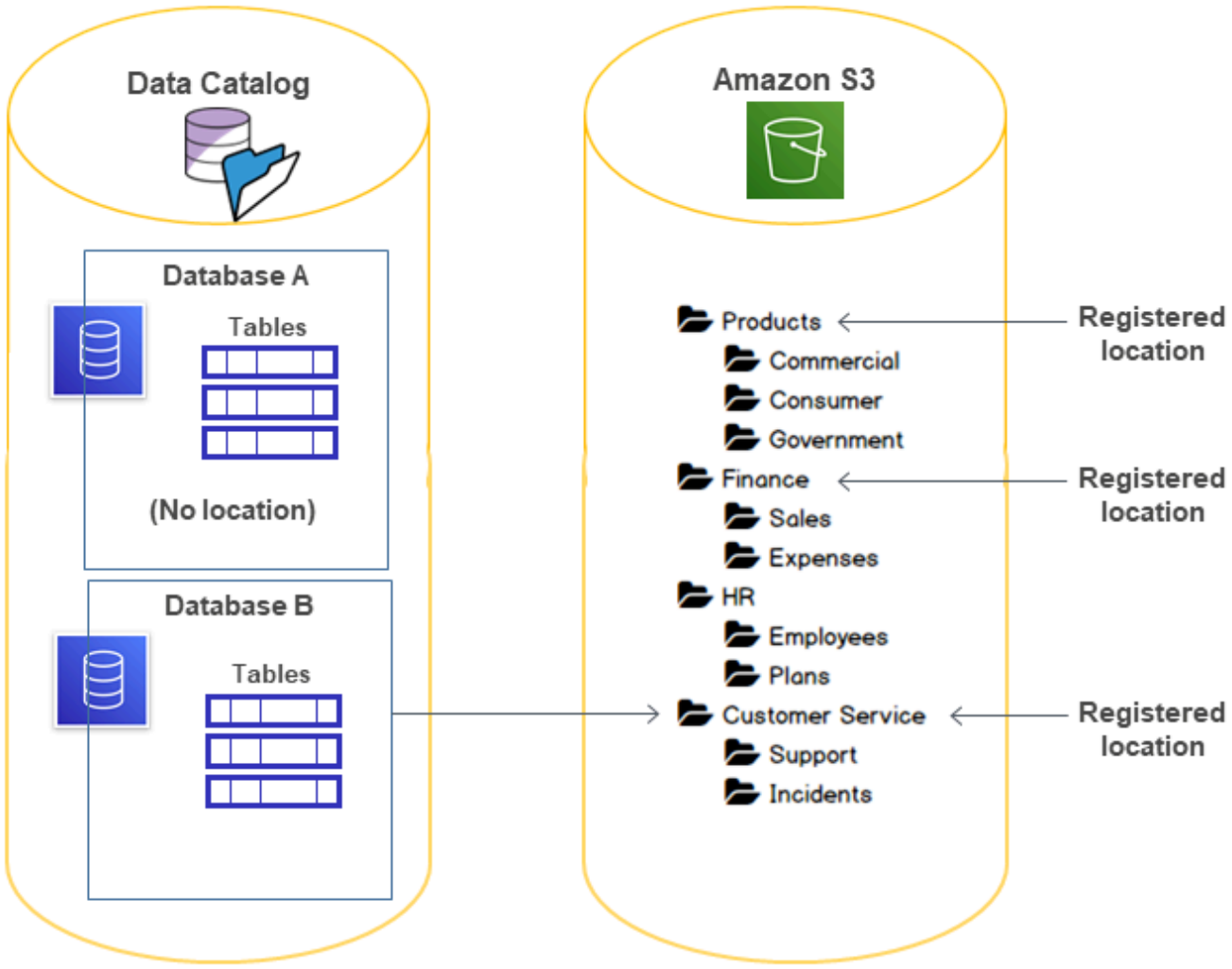
Lebih lanjut tentang izin lokasi data

Izin lokasi data mengatur hasil dari membuat dan memperbarui operasi pada database dan tabel Katalog Data. Aturannya adalah sebagai berikut:

- Prinsipal harus memiliki izin lokasi data eksplisit atau implisit di lokasi Amazon S3 untuk membuat atau memperbarui database atau tabel yang menentukan lokasi tersebut.

- Izin eksplisit `DATA_LOCATION_ACCESS` diberikan menggunakan konsol, API, atau AWS CLI.
- Izin implisit diberikan ketika database memiliki properti lokasi yang menunjuk ke lokasi terdaftar, kepala sekolah memiliki `CREATE_TABLE` izin pada database, dan prinsipal mencoba membuat tabel di lokasi tersebut atau lokasi anak.
- Jika prinsipal diberikan izin lokasi data pada suatu lokasi, kepala sekolah memiliki izin lokasi data di semua lokasi turunan.
- Seorang prinsipal tidak memerlukan izin lokasi data untuk melakukan read/write operasi pada data yang mendasarinya. Hal ini cukup untuk memiliki `SELECT` atau izin akses `INSERT` data. Izin lokasi data hanya berlaku untuk membuat sumber daya Katalog Data yang mengarah ke lokasi.
- Operasi partisi (`CreatePartitionBatchCreatePartition`,,`UpdatePartition`) tidak memerlukan izin lokasi data ketika lokasi partisi adalah anak dari lokasi tabel yang sudah terdaftar dengan Lake Formation.

Pertimbangkan skenario yang ditunjukkan pada diagram berikut.



Dalam diagram ini:

- Bucket Amazon S3`Products`,`Finance`, dan `Customer Service` terdaftar di Lake Formation.
- `Database A` tidak memiliki properti lokasi, dan `Database B` memiliki properti lokasi yang menunjuk ke `Customer Service` ember.
- Pengguna `datalake_user` memiliki `CREATE_TABLE` pada kedua database.
- Pengguna `datalake_user` telah diberikan izin lokasi data hanya pada `Products` bucket.

Berikut ini adalah hasil ketika pengguna `datalake_user` mencoba untuk membuat tabel katalog dalam database tertentu di lokasi tertentu.

Lokasi di mana **`datalake_user`** mencoba membuat tabel

Database dan Lokasi	Berhasil atau Gagal	Alasan
Database A di `Finance/Sales`	Gagal	Tidak ada izin lokasi data
Database A di `Products`	Berhasil	Memiliki izin lokasi data
Database A di `HR/Plans`	Berhasil	Lokasi tidak terdaftar
Database B di `Customer Service/Incidents`	Berhasil	Database memiliki properti lokasi di `Customer Service`

Untuk informasi selengkapnya, lihat informasi berikut:

- Menambahkan lokasi Amazon S3 ke danau data Anda
- Referensi izin Lake Formation
- Referensi personas Lake Formation dan izin IAM

Referensi personas Lake Formation dan izin IAM

Bagian ini mencantumkan beberapa persona Lake Formation yang disarankan dan izin yang disarankan AWS Identity and Access Management (IAM) mereka. Untuk informasi tentang izin Lake Formation, lihatthe section called “Referensi izin Lake Formation”.

AWS Lake Formation persona

Tabel berikut mencantumkan AWS Lake Formation persona yang disarankan.

Personas Lake Formation

Persona	Deskripsi
Administrator IAM (pengguna super)	(Wajib) Pengguna yang dapat membuat pengguna dan peran IAM. Memiliki kebijakan yang `AdministratorAccess` AWS dikelola. Memiliki semua izin pada semua sumber daya Lake Formation. Dapat menambahkan administrator danau data.

Persona	Deskripsi
	Tidak dapat memberikan izin Lake Formation jika tidak juga ditunjuk sebagai administrator danau data.
Administrator danau data	(Wajib) Pengguna yang dapat mendaftarkan lokasi Amazon S3, mengakses Katalog Data, membuat database, membuat dan menjalankan alur kerja, memberikan izin Lake Formation kepada pengguna lain, dan melihat log. AWS CloudTrail Memiliki izin IAM lebih sedikit daripada administrator IAM, tetapi cukup untuk mengelola data lake. Tidak dapat menambahkan administrator danau data lainnya.
Hanya baca administrator	(Opsional) Pengguna yang dapat melihat prinsipal, sumber daya Katalog Data, izin, dan AWS CloudTrail log, tanpa izin untuk melakukan pembaruan.
Insinyur data	(Opsional) Pengguna yang dapat membuat database, membuat dan menjalankan crawler dan alur kerja, serta memberikan izin Lake Formation pada tabel Katalog Data yang dibuat oleh crawler dan alur kerja. Kami menyarankan Anda membuat semua pembuat database insinyur data. Untuk informasi selengkapnya, lihat Membuat basis data.
Analis data	(Opsional) Pengguna yang dapat menjalankan kueri terhadap data lake menggunakan, misalnya, Amazon Athena. Hanya memiliki izin yang cukup untuk menjalankan kueri.
Peran alur kerja	(Wajib) Peran yang menjalankan alur kerja atas nama pengguna. Anda menentukan peran ini saat membuat alur kerja dari cetak biru.

> **Note**
>
> Di Lake Formation, administrator data lake yang ditambahkan setelah pembuatan database dapat memberikan izin tetapi tidak secara otomatis memiliki izin akses data seperti SELECT atau DESCRIPTE. Administrator yang membuat database menerima SUPER izin pada database tersebut. Perilaku ini disengaja—sementara semua administrator dapat

memberikan izin yang diperlukan, izin ini tidak diterapkan secara otomatis ke sumber daya yang sudah ada sebelumnya. Oleh karena itu, administrator harus secara eksplisit memberikan diri mereka akses ke database yang ada sebelum mereka diberi hak istimewa admin.

AWS kebijakan terkelola untuk Lake Formation

Anda dapat memberikan izin AWS Identity and Access Management (IAM) yang diperlukan untuk bekerja AWS Lake Formation dengan menggunakan kebijakan AWS terkelola dan kebijakan inline. Kebijakan AWS terkelola berikut tersedia untuk Lake Formation.

AWS kebijakan terkelola: AWSLake FormationDataAdmin

AWSLakeFormationDataAdminkebijakan memberikan akses administratif ke AWS Lake Formation dan layanan terkait seperti AWS Glue untuk mengelola danau data.

Anda dapat melampirkan `AWSLakeFormationDataAdmin` ke pengguna, grup, dan peran Anda.

Detail izin

- `CloudTrail`— Memungkinkan kepala sekolah untuk melihat log. AWS CloudTrail Ini diperlukan untuk meninjau kesalahan apa pun dalam pengaturan data lake.
- `Glue`— Memungkinkan prinsipal untuk melihat, membuat, dan memperbarui tabel metadata dan database dalam Katalog Data. Ini termasuk operasi API yang dimulai dengan`Get`,`List`,`Create`,`Update`,`Delete`, dan`Search`. Ini diperlukan untuk mengelola metadata tabel data lake.
- `IAM`— Memungkinkan kepala sekolah untuk mengambil informasi tentang pengguna IAM, peran, dan kebijakan yang dilampirkan pada peran. Ini diperlukan agar admin data meninjau dan mencantumkan pengguna dan peran IAM untuk memberikan izin Lake Formation.
- `Lake Formation`— Memberikan admin data lake memerlukan izin Lake Formation untuk mengelola data lake.
- `S3`— Memungkinkan kepala sekolah untuk mengambil informasi tentang bucket Amazon S3 dan lokasinya untuk mengatur lokasi data untuk 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**
>
> `AWSLakeFormationDataAdmin`Kebijakan ini tidak memberikan setiap izin yang diperlukan untuk administrator data lake. Izin tambahan diperlukan untuk membuat dan menjalankan alur kerja dan mendaftarkan lokasi dengan peran terkait layanan. `AWSServiceRoleForLakeFormationDataAccess` Untuk informasi selengkapnya, lihat Buat administrator danau data dan Menggunakan peran terkait layanan untuk Lake Formation.

AWS kebijakan terkelola: AWSLake FormationCrossAccountManager

AWSLakeFormationCrossAccountManagerkebijakan menyediakan akses lintas akun ke AWS Glue sumber daya melalui Lake Formation, dan memberikan akses baca ke layanan lain yang diperlukan seperti AWS Organizations dan AWS RAM.

Anda dapat melampirkan `AWSLakeFormationCrossAccountManager` ke pengguna, grup, dan peran Anda.

Detail izin

Kebijakan ini mencakup izin berikut.

- `Glue`— Memungkinkan prinsipal untuk mengatur atau menghapus kebijakan sumber daya Katalog Data untuk kontrol akses.
- `Organizations`— Memungkinkan kepala sekolah untuk mengambil informasi akun dan unit organisasi (OU) untuk suatu organisasi.
- `ram:CreateResourceShare`— Memungkinkan kepala sekolah untuk membuat pembagian sumber daya.

- `ram:UpdateResourceShare`—Memungkinkan prinsipal untuk memodifikasi beberapa properti dari pembagian sumber daya yang ditentukan.
- `ram:DeleteResourceShare`— Memungkinkan prinsipal untuk menghapus pembagian sumber daya yang ditentukan.
- `ram:AssociateResourceShare`— Memungkinkan prinsipal untuk menambahkan daftar prinsipal dan daftar sumber daya yang ditentukan ke pembagian sumber daya.
- `ram:DisassociateResourceShare`— Memungkinkan prinsipal untuk menghapus prinsip atau sumber daya yang ditentukan dari berpartisipasi dalam pembagian sumber daya yang ditentukan.
- `ram:GetResourceShares`— Memungkinkan kepala sekolah untuk mengambil rincian tentang pembagian sumber daya yang Anda miliki atau yang dibagikan dengan Anda.
- `ram:RequestedResourceType`— Memungkinkan prinsipal untuk mengambil jenis sumber daya (database, tabel atau katalog).
- `AssociateResourceSharePermission`— Memungkinkan prinsipal untuk menambah atau mengganti AWS RAM izin untuk jenis sumber daya yang disertakan dalam pembagian sumber daya. Anda dapat memiliki persis satu izin yang terkait dengan setiap jenis sumber daya dalam pembagian sumber daya.

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 kebijakan terkelola: AWSGlue ConsoleFullAccess

AWSGlueConsoleFullAccesskebijakan memberikan akses penuh ke AWS Glue sumber daya ketika identitas yang dilampirkan kebijakan menggunakan. Konsol Manajemen AWS Jika Anda mengikuti konvensi penamaan untuk sumber daya yang ditentukan dalam kebijakan ini, maka pengguna memiliki kemampuan konsol penuh. Kebijakan ini biasanya dilampirkan ke pengguna AWS Glue konsol.

Selain itu, AWS Glue Lake Formation mengambil peran layanan `AWSGlueServiceRole` untuk memungkinkan akses ke layanan terkait, termasuk Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3), dan Amazon. CloudWatch

AWS managed policy:LakeFormationDataAccessServiceRolePolicy

Kebijakan ini dilampirkan ke peran terkait layanan bernama `ServiceRoleForLakeFormationDataAccess` yang memungkinkan layanan melakukan tindakan pada sumber daya atas permintaan Anda. Anda tidak dapat melampirkan kebijakan ini ke identitas IAM Anda.

Kebijakan ini memungkinkan AWS layanan terintegrasi Lake Formation seperti Amazon Athena atau Amazon Redshift menggunakan peran terkait layanan untuk menemukan sumber daya Amazon S3.

Untuk informasi selengkapnya, lihat Menggunakan peran terkait layanan untuk Lake Formation.

Detail izin

Kebijakan ini mencakup izin berikut.

- `s3:ListAllMyBuckets`— Mengembalikan daftar semua bucket yang dimiliki oleh pengirim permintaan yang diautentikasi.

JSON

```
{
 "Version":"2012-10-17",
 "Statement": [
  {
   "Sid": "LakeFormationDataAccessServiceRolePolicy",
   "Effect": "Allow",
   "Action": [
    "s3:ListAllMyBuckets"
   ],
   "Resource": [
    "arn:aws:s3:::*"
   ]
  }
 ]
}
```

Lake Formation memperbarui kebijakan AWS terkelola

Lihat detail tentang pembaruan kebijakan AWS terkelola untuk Lake Formation sejak layanan ini mulai melacak perubahan ini.

Ubah	Deskripsi	Date
`AWSLakeFormationCrossAccountManager` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationCrossAccountManagerkebijakan dengan mengganti operator `StringLike` kondisi dengan `ArnLike` operator yang memungkin	Januari, 2025

Ubah	Deskripsi	Date
	kan IAM melakukan pemeriksaan format ARN.	
`AWSLakeFormationDataAdmin` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationDataAdminkebijakan dengan menambahkan AWS Glue Data Catalog CRUD berikut APIs sebagai bagian dari fitur multi-katalog. • lem: CreateCatalog • lem: UpdateCatalog • lem: DeleteCatalog • lem: GetCatalog • lem: GetCatalogs Perubahan kebijakan terkelola ini adalah untuk memastikan bahwa persona administrator Lake Formation secara default memiliki izin IAM pada operasi baru ini.	Desember, 2024
`AWSLakeFormationCrossAccountManager` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationCrossAccountManagerkebijakan dengan menambahkan elemen Sid ke dalam pernyataan kebijakan.	Maret, 2024
`AWSLakeFormationDataAdmin` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationDataAdminkebijakan dengan menambahkan elemen Sid ke pernyataan kebijakan dan menghapus tindakan yang berlebihan.	Maret, 2024

Ubah	Deskripsi	Date
`LakeFormationDataAccessServiceRolePolicy` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan LakeFormationDataAccessServiceRolePolicykebijakan dengan menambahkan elemen Sid ke pernyataan kebijakan.	Februari, 2024
`AWSLakeFormationCrossAccountManager` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationCrossAccountManagerkebijakan dengan menambahkan izin baru untuk mengaktifkan berbagi data lintas akun dalam mode akses hibrida.	Oktober, 2023
`AWSLakeFormationCrossAccountManager` Kebijakan yang diperbarui Lake Formation.	Lake Formation menyempurnakan AWSLakeFormationCrossAccountManagerkebijakan untuk membuat hanya satu pembagian sumber daya per akun penerima saat sumber daya pertama kali dibagikan. Semua sumber daya yang dibagikan setelahnya dengan akun yang sama dilampirkan ke pembagian sumber daya yang sama.	6 Mei 2022
Lake Formation mulai melacak perubahan.	Lake Formation mulai melacak perubahan untuk kebijakan yang AWS dikelola.	6 Mei 2022

Personas menyarankan izin

Berikut ini adalah izin yang disarankan untuk setiap persona. Administrator IAM tidak disertakan karena pengguna tersebut memiliki semua izin pada semua sumber daya.

Topik

- Izin administrator danau data

- Baca hanya izin administrator
- Izin insinyur data
- Izin analis data
- Izin peran alur kerja

Izin administrator danau data

> **Important**
>
> Dalam kebijakan berikut, ganti *<account-id>* dengan nomor AWS akun yang valid, dan ganti *<workflow_role>* dengan nama peran yang memiliki izin untuk menjalankan alur kerja, seperti yang didefinisikan dalam. Izin peran alur kerja

Jenis Kebijakan	Kebijakan
AWS kebijakan terkelola	• `AWSLakeFormationDataAdmin` • `LakeFormationDataAccessServiceRolePolicy` (kebijakan peran terkait layanan) • `AWSGlueConsoleFullAccess` (Opsional) • `CloudWatchLogsReadOnlyAccess` (Opsional) • `AWSLakeFormationCrossAccountManager` (Opsional) • `AmazonAthenaFullAccess` (Opsional) Untuk informasi tentang kebijakan AWS terkelola opsional, lihatthe section called “Buat administrator danau data”.

Kebijakan sebaris (untuk membuat peran terkait layanan Lake Formation)	<pre>{
	    "Version": "2012-10-17",
	    "Statement": [
	        {
	            "Effect": "Allow",
	            "Action": "iam:CreateServiceLinkedRole",</pre>

Jenis Kebijakan	Kebijakan

	"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"
	}
	]
	}

(Opsional) Kebijakan sebaris (kebijakan peran sandi untuk peran alur kerja). Ini diperlukan hanya jika administrator data lake membuat dan menjalankan alur kerja.	**{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *111122223333* :role/*<workflow_role>* "
	]
	}
	]
	}**

Jenis Kebijakan	Kebijakan

(Opsional) Kebijakan sebaris (jika akun Anda memberikan atau menerima izin Lake Formation lintas akun). Kebijakan ini untuk menerima atau menolak undangan berbagi AWS RAM sumber daya, dan untuk memungkinkan pemberian izin lintas akun kepada organisasi. `ram:EnableSharingWithAwsOrganization` diperlukan hanya untuk administrator danau data di akun AWS Organizations manajemen.	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"ram:AcceptResourceShareInvitation",
	"ram:RejectResourceShareInvitation",
	"ec2:DescribeAvailabilityZones",
	"ram:EnableSharingWithAwsOrganization"
	],
	"Resource": "*"
	}
	]
	}

Baca hanya izin administrator

Tipe kebijakan	Kebijakan
Kebijakan inline (dasar)	{ "Version":"2012-10-17", "Statement":[{ "Effect":"Allow", "Action":["lakeformation:GetEffectivePermissionsForPath", "lakeformation:ListPermissions", "lakeformation:ListDataCellsFilter", "lakeformation:GetDataCellsFilter", "lakeformation:SearchDatabasesByLFTags",

Tipe kebijakan	Kebijakan

```
            "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":"*"
      }
   ]
```

Tipe kebijakan	Kebijakan
	`}`

Izin insinyur data

> **Important**
>
> Dalam kebijakan berikut, ganti *<account-id>* dengan nomor AWS akun yang valid, dan ganti *<workflow_role>* dengan nama peran alur kerja.

Jenis Kebijakan	Kebijakan
AWS kebijakan terkelola	`AWSGlueConsoleFullAccess`
Kebijakan inline (dasar)	`{ "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",`

Jenis Kebijakan	Kebijakan
	"lakeformation:SearchDatabasesByLFTa gs", "lakeformation:GetWorkUnits", "lakeformation:GetWorkUnitResults", "lakeformation:StartQueryPlanning", "lakeformation:GetQueryState", "lakeformation:GetQueryStatistics"], "Resource": "*" }] }

Kebijakan inline (untuk operasi pada tabel yang diatur, termasuk operasi dalam transaksi)	{
	"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": "*"
	}
	]
	}

Jenis Kebijakan	Kebijakan

Kebijakan inline (untuk kontrol akses metadata menggunakan metode kontrol akses berbasis tag Lake Formation (LF-TBAC))	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"lakeformation:AddLFTagsToResource",
	"lakeformation:RemoveLFTagsFromResource",
	"lakeformation:GetResourceLFTags",
	"lakeformation:ListLFTags",
	"lakeformation:GetLFTag",
	"lakeformation:SearchTablesByLFTags",
	"lakeformation:SearchDatabasesByLFTags"
	],
	"Resource": "*"
	}
	]
	}

Kebijakan inline (kebijakan passrole untuk peran alur kerja)	{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *111122223333* :role/*<workflow_role>* "
	]
	}
	]
	}

Izin analis data

Jenis Kebijakan	Kebijakan
AWS kebijakan terkelola	`AmazonAthenaFullAccess`
Kebijakan inline (dasar)	`{ "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": "*" }] }`
(Opsional) Kebijakan inline (untuk operasi pada tabel yang diatur, termasuk operasi dalam transaksi)	`{ "Version":"2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "lakeformation:StartTransaction", "lakeformation:CommitTransaction", "lakeformation:CancelTransaction", "lakeformation:ExtendTransaction",`

Jenis Kebijakan	Kebijakan

	<pre>"lakeformation:DescribeTransaction",
	"lakeformation:ListTransactions",
	"lakeformation:GetTableObjects",
	"lakeformation:UpdateTableObjects",
	"lakeformation:DeleteObjectsOnCancel
	"
	],
	"Resource": "*"
	}
	]
	}</pre>

Izin peran alur kerja

Peran ini memiliki izin yang diperlukan untuk menjalankan alur kerja. Anda menentukan peran dengan izin ini saat membuat alur kerja.

> **Important**
>
> Dalam kebijakan berikut, ganti *<region>* dengan pengenal AWS Wilayah yang valid (misalnya`us-east-1`), *<account-id>* dengan nomor AWS akun yang valid, *<workflow_role>* dengan nama peran alur kerja, dan *<your-s3-cloudtrail-bucket>* dengan jalur Amazon S3 ke log Anda. AWS CloudTrail

Jenis Kebijakan	Kebijakan
AWS kebijakan terkelola	`AWSGlueServiceRole`

Kebijakan inline (akses data)	<pre>{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "Lakeformation",
	"Effect": "Allow",
	"Action": [
	"lakeformation:GetDataAccess",</pre>

Jenis Kebijakan	Kebijakan

	<pre>"lakeformation:GrantPermissions"
	],
	"Resource": "*"
	}
	]
	}</pre>

Kebijakan inline (kebijakan passrole untuk peran alur kerja)	<pre>{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *111122223333* :role/*<workflow_role>* "
	]
	}
	]
	}</pre>

Kebijakan sebaris (untuk menelan data di luar data lake, misalnya, AWS CloudTrail log)	<pre>{
	"Version":"2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": ["s3:GetObject", "s3:ListBucket"],
	"Resource": ["arn:aws:s3::: *<your-s3-cloudtrail-bucket>* /*"]
	}
	]
	}</pre>

Mengubah pengaturan default untuk data lake

Untuk menjaga kompatibilitas mundur denganAWS Glue, AWS Lake Formation memiliki pengaturan keamanan awal berikut:

- `Super`Izin diberikan kepada grup `IAMAllowedPrincipals` pada semua sumber daya Katalog AWS Glue Data yang ada.
- Pengaturan “Gunakan hanya kontrol akses IAM” diaktifkan untuk sumber daya Katalog Data baru.

Pengaturan ini secara efektif menyebabkan akses ke sumber daya Katalog Data dan lokasi Amazon S3 dikendalikan semata-mata oleh kebijakan AWS Identity and Access Management (IAM). Izin Lake Formation individu tidak berlaku.

`IAMAllowedPrincipals`Grup ini mencakup setiap pengguna IAM dan peran yang diizinkan mengakses sumber daya Katalog Data Anda oleh kebijakan IAM Anda. `Super`Izin memungkinkan kepala sekolah untuk melakukan setiap operasi Lake Formation yang didukung pada database atau tabel yang diberikan.

Untuk mengubah pengaturan keamanan sehingga akses ke sumber Data Catalog (database dan tabel) dikelola oleh izin Lake Formation, lakukan hal berikut:

1. Ubah pengaturan keamanan default untuk sumber daya baru. Untuk petunjuk, lihat Ubah model izin default atau gunakan mode akses hybrid.
2. Ubah pengaturan untuk sumber daya Katalog Data yang ada. Untuk petunjuk, lihat Memutakhirkan izin AWS Glue data ke model AWS Lake Formation.

Mengubah setelan keamanan default menggunakan operasi Lake Formation **`PutDataLakeSettings`** API

Anda juga dapat mengubah pengaturan keamanan default dengan menggunakan operasi Lake Formation PutDataLakeSettingsAPI. Tindakan ini mengambil argumen ID katalog opsional dan DataLakeSettingsstruktur.

Untuk menerapkan metadata dan kontrol akses data yang mendasari oleh Lake Formation pada database dan tabel baru, kodekan struktur sebagai `DataLakeSettings` berikut.

> **Note**
>
> Ganti *<AccountID>* dengan ID AWS akun yang valid dan *<Username>* dengan nama pengguna IAM yang valid. Anda dapat menentukan lebih dari satu pengguna sebagai administrator danau data.

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ],
        "CreateDatabaseDefaultPermissions": [],
        "CreateTableDefaultPermissions": []
    }
}
```

Anda juga dapat mengkodekan struktur sebagai berikut. Menghilangkan `CreateTableDefaultPermissions` parameter `CreateDatabaseDefaultPermissions` or sama dengan melewatkan daftar kosong.

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ]
    }
}
```

Tindakan ini secara efektif mencabut semua izin Lake Formation dari `IAMAllowedPrincipals` grup pada database dan tabel baru. Saat membuat database, Anda dapat mengganti pengaturan ini.

Untuk menegakkan metadata dan kontrol akses data yang mendasarinya hanya oleh IAM pada database dan tabel baru, kodekan struktur sebagai berikut. `DataLakeSettings`

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

Ini memberikan izin `Super` Lake Formation ke `IAMAllowedPrincipals` grup pada database dan tabel baru. Saat membuat database, Anda dapat mengganti pengaturan ini.

> **Note**
>
> Dalam `DataLakeSettings` struktur sebelumnya, satu-satunya nilai yang diizinkan `DataLakePrincipalIdentifier` adalah`IAM_ALLOWED_PRINCIPALS`, dan satu-satunya nilai yang diizinkan adalah. `Permissions` `ALL`

Izin Formasi Danau Implisit

AWS Lake Formation memberikan izin implisit berikut kepada administrator data lake, pembuat database, dan pembuat tabel.

Administrator data lake

- Memiliki `Describe` akses ke semua sumber daya dalam Katalog Data kecuali untuk sumber daya yang dibagikan dari akun lain secara langsung ke prinsipal yang berbeda. Akses ini tidak dapat dicabut dari administrator.
- Memiliki izin lokasi data di mana-mana di danau data.
- Dapat memberikan atau mencabut akses ke sumber daya apa pun dalam Katalog Data kepada prinsipal mana pun (termasuk mandiri). Akses ini tidak dapat dicabut dari administrator.
- Dapat membuat database di Katalog Data.
- Dapat memberikan izin untuk membuat database ke pengguna lain.

> **Note**
>
> Administrator data lake dapat mendaftarkan lokasi Amazon S3 hanya jika mereka memiliki izin IAM untuk melakukannya. Kebijakan administrator data lake yang disarankan dalam panduan ini memberikan izin tersebut. Selain itu, administrator data lake tidak memiliki izin implisit untuk menjatuhkan database atau alter/drop tabel yang dibuat oleh orang lain. Namun, mereka dapat memberikan izin kepada diri mereka sendiri untuk melakukannya.

Untuk informasi selengkapnya tentang administrator data lake, lihatBuat administrator danau data.

Pembuat katalog

- Memiliki semua izin katalog pada katalog yang mereka buat, memiliki izin pada database dan tabel yang mereka buat di katalog, dan dapat memberikan izin kepada prinsipal lain di AWS akun yang sama untuk membuat database dan tabel di katalog. Pembuat katalog yang juga memiliki kebijakan `AWSLakeFormationCrossAccountManager` AWS terkelola dapat memberikan izin pada katalog ke AWS akun atau organisasi lain.

  Administrator data lake dapat menggunakan konsol Lake Formation atau API untuk menunjuk pembuat katalog.

> **Note**
>
> Pembuat katalog tidak secara implisit memiliki izin pada database dan tabel yang dibuat orang lain dalam katalog.

Untuk informasi selengkapnya tentang membuat katalog, lihat. Membawa data Anda ke AWS Glue Data Catalog

Pembuat basis data

- Memiliki semua izin database pada database yang mereka buat, memiliki izin pada tabel yang mereka buat dalam database, dan dapat memberikan prinsipal lain dalam izin AWS akun yang sama untuk membuat tabel dalam database. Pembuat database yang juga memiliki kebijakan `AWSLakeFormationCrossAccountManager` AWS terkelola dapat memberikan izin pada database ke AWS akun atau organisasi lain.

  Administrator data lake dapat menggunakan konsol Lake Formation atau API untuk menunjuk pembuat database.

  > **Note**
  >
  > Pembuat database tidak secara implisit memiliki izin pada tabel yang dibuat orang lain dalam database.

  Untuk informasi selengkapnya, lihat Membuat basis data.

Pembuat tabel

- Memiliki semua izin pada tabel yang mereka buat.
- Dapat memberikan izin pada semua tabel yang mereka buat ke kepala sekolah di akun yang sama. AWS
- Dapat memberikan izin pada semua tabel yang dibuat ke AWS akun atau organisasi lain jika memiliki kebijakan `AWSLakeFormationCrossAccountManager` AWS terkelola.
- Dapat melihat database yang berisi tabel yang mereka buat.

Referensi izin Lake Formation

Untuk melakukan AWS Lake Formation operasi, kepala sekolah memerlukan izin Lake Formation dan AWS Identity and Access Management (IAM). Anda biasanya memberikan izin IAM menggunakan kebijakan kontrol akses berbutir kasar, seperti yang dijelaskan dalam. the section called “Ikhtisar izin Lake Formation” Anda dapat memberikan izin Lake Formation dengan menggunakan konsol, API, atau AWS Command Line Interface (AWS CLI).

Untuk mempelajari cara memberikan atau mencabut izin Lake Formation, lihat dan. the section called “Memberikan izin data” the section called “Memberikan izin lokasi data”

> **Note**
>
> Contoh di bagian ini menunjukkan cara memberikan izin kepada kepala sekolah di akun yang sama. AWS Untuk contoh hibah lintas akun, lihat. the section called “Cross-account berbagi data”

Izin Lake Formation per jenis sumber daya

Berikut ini adalah izin Lake Formation valid yang tersedia untuk setiap jenis sumber daya:

Sumber daya	Izin
`Catalog`	`ALL(Super)`, Pengguna super
	`ALTER`
	`CREATE_DATABASE`
	`DESCRIBE`
	`DROP`
`Database`	`ALL (Super)`
	`ALTER`
	`CREATE_TABLE`

Sumber daya	Izin	
	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	
	GrantWithLFTagExpression	

Sumber daya	Izin
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

Topik

- Lake Formation memberikan dan mencabut AWS CLI perintah

- Izin Lake Formation

Lake Formation memberikan dan mencabut AWS CLI perintah

Setiap deskripsi izin di bagian ini mencakup contoh pemberian izin menggunakan AWS CLI perintah. Berikut ini adalah sinopsis dari Formasi dan perintah Lake. grant-permissions revoke-permissions AWS CLI

```
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>]
```

Untuk deskripsi terperinci tentang perintah ini, lihat izin pemberian dan pencabutan izin di Referensi Perintah.AWS CLI Bagian ini memberikan informasi tambahan tentang `--principal` opsi.

Nilai `--principal` opsi adalah salah satu dari yang berikut:

- Nama Sumber Daya Amazon (ARN) untuk pengguna atau peran AWS Identity and Access Management (IAM)
- ARN untuk pengguna atau grup yang mengautentikasi melalui penyedia SAFL, seperti Microsoft Active Directory Federation Service (AD FS)
- ARN untuk pengguna atau grup Amazon Quick
- Untuk izin lintas akun, ID AWS akun, ID organisasi, atau ID unit organisasi
- Untuk pengguna atau grup IAM Identity Center, pengguna IAM Identity Center atau grup ARN.

Berikut ini adalah sintaks dan contoh untuk semua `--principal` jenis.

Principal adalah pengguna IAM

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1
```

Principal adalah peran IAM

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/workflowrole
```

Principal adalah pengguna yang mengautentikasi melalui penyedia SAFL

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:user/<user-name>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
idp1:user/datalake_user1
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
AthenaLakeFormationOkta:user/athena-user@example.com
```

Principal adalah grup yang mengautentikasi melalui penyedia SAFL

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:group/<group-name>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
idp1:group/data-scientists
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
AthenaLakeFormationOkta:group/my-group
```

Principal adalah pengguna Amazon Quick Enterprise Edition

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-
id>:user/<namespace>/<user-name>
```

> **Note**
>
> Untuk *<namespace>*, Anda harus menentukan `default`.

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-
east-1:111122223333:user/default/bi_user1
```

Principal adalah grup Amazon Quick Enterprise Edition

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-
id>:group/<namespace>/<group-name>
```

> **Note**
>
> Untuk *<namespace>*, Anda harus menentukan `default`.

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-
east-1:111122223333:group/default/data_scientists
```

Principal adalah AWS akun

Sintaksis:

```
--principal DataLakePrincipalIdentifier=<account-id>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=111122223333
```

Principal adalah sebuah organisasi

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-
id>:organization/<organization-id>
```

Contoh:

```
--principal
 DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/o-
abcdefghijkl
```

Principal adalah unit organisasi

Sintaksis:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-
id>:ou/<organization-id>/<organizational-unit-id>
```

Contoh:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:ou/o-
abcdefghijkl/ou-ab00-cdefghij
```

Principal adalah pengguna atau grup identitas IAM Identity Center

Example:User

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserID>
```

Example:Group:

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::group/<GroupID>
```

Principal adalah grup IAM - **`IAMAllowedPrincipals`**

Lake Formation menetapkan `Super` izin pada semua database dan tabel dalam Katalog Data ke grup yang dipanggil secara `IAMAllowedPrincipals` default. Jika izin grup ini ada pada database atau tabel, semua prinsipal di akun Anda akan memiliki akses ke sumber daya melalui kebijakan utama IAM untuk. AWS Glue Ini memberikan kompatibilitas mundur saat Anda mulai menggunakan izin Lake Formation untuk mengamankan sumber daya Katalog Data yang sebelumnya dilindungi oleh kebijakan IAM. AWS Glue

Saat Anda menggunakan Lake Formation untuk mengelola izin untuk sumber daya Katalog Data Anda, Anda harus terlebih dahulu mencabut `IAMAllowedPrincipals` izin pada sumber daya, atau memilih prinsip dan sumber daya ke mode akses hibrid agar izin Lake Formation berfungsi.

Contoh:

```
--principal DataLakePrincipalIdentifier=IAM_Allowed_Principals
```

Principal adalah grup IAM - **`ALLIAMPrincipals`**

Saat Anda memberikan izin untuk `ALLIAMPrincipals` mengelompokkan sumber daya Katalog Data, setiap prinsipal di akun mendapatkan akses ke sumber daya Katalog Data menggunakan izin Lake Formation dan izin IAM.

Contoh:

```
--principal DataLakePrincipalIdentifier=123456789012:IAMPrincipals
```

Izin Lake Formation

Bagian ini berisi izin Lake Formation yang tersedia yang dapat Anda berikan kepada kepala sekolah.

MENGUBAH

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
ALTER	DATABASE	glue:UpdateDatabase
ALTER	TABLE	glue:UpdateTable
ALTER	LF-Tag	lakeformation:UpdateLFTag

Prinsipal dengan izin ini dapat mengubah metadata untuk database atau tabel di Katalog Data. Untuk tabel, Anda dapat mengubah skema kolom dan menambahkan parameter kolom. Anda tidak dapat mengubah kolom dalam data dasar yang ditunjukkan oleh tabel metadata.

Jika properti yang sedang diubah adalah lokasi Amazon Simple Storage Service (Amazon S3) terdaftar, prinsipal harus memiliki izin lokasi data di lokasi baru.

Example

Contoh berikut memberikan ALTER izin kepada pengguna datalake_user1 pada database retail di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "ALTER" --resource '{ "Database": {"Name":"retail"}}'
```

Example

Contoh berikut memberikan ALTER kepada pengguna datalake_user1 pada tabel inventory dalam databaseretail.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
```

```
 --permissions "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

CREATE_DATABASE

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
CREATE_DATABASE	Katalog Data	glue:CreateDatabase

Prinsipal dengan izin ini dapat membuat database metadata atau tautan sumber daya di Katalog Data. Prinsipal juga dapat membuat tabel dalam database.

Example

Contoh berikut memberikan CREATE_DATABASE kepada pengguna datalake_user1 di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

Saat prinsipal membuat database di Katalog Data, tidak ada izin untuk data dasar yang diberikan. Izin metadata tambahan berikut diberikan (bersama dengan kemampuan untuk memberikan izin ini kepada orang lain):

- CREATE_TABLEdalam database
- ALTERbasis data
- DROPbasis data

Saat membuat database, prinsipal dapat secara opsional menentukan lokasi Amazon S3. Bergantung pada apakah prinsipal memiliki izin lokasi data, CREATE_DATABASE izin mungkin tidak cukup untuk membuat database dalam semua kasus. Penting untuk mengingat tiga kasus berikut.

Buat kasus penggunaan database	Izin diperlukan
Properti lokasi tidak ditentukan.	CREATE_DATABASE sudah cukup.

Buat kasus penggunaan database	Izin diperlukan
Properti lokasi ditentukan, dan lokasi tidak dikelola oleh Lake Formation (tidak terdaftar).	`CREATE_DATABASE` sudah cukup.
Properti lokasi ditentukan, dan lokasi dikelola oleh Lake Formation (terdaftar).	`CREATE_DATABASE` diperlukan ditambah izin lokasi data pada lokasi yang ditentukan.

`CREATE_TABLE`

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
`CREATE_TABLE`	`DATABASE`	`glue:CreateTable`

Prinsipal dengan izin ini dapat membuat tabel metadata atau tautan sumber daya di Katalog Data dalam database yang ditentukan.

Example

Contoh berikut memberikan `datalake_user1` izin pengguna untuk membuat tabel dalam `retail` database di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Saat prinsipal membuat tabel di Katalog Data, semua izin Lake Formation pada tabel diberikan kepada kepala sekolah, dengan kemampuan untuk memberikan izin ini kepada orang lain.

Cross-account Hibah

Jika akun pemilik database memberikan `CREATE_TABLE` akun penerima, dan pengguna di akun penerima berhasil membuat tabel di database akun pemilik, aturan berikut berlaku:

- Administrator pengguna dan data lake di akun penerima memiliki semua izin Lake Formation di atas meja. Mereka dapat memberikan izin di atas meja ke kepala sekolah lain di akun mereka. Mereka tidak dapat memberikan izin kepada kepala sekolah di akun pemilik atau akun lainnya.

- Administrator data lake di akun pemilik dapat memberikan izin di atas meja ke kepala sekolah lain di akun mereka.

Izin Lokasi Data

Saat Anda mencoba membuat tabel yang mengarah ke lokasi Amazon S3, tergantung pada apakah Anda memiliki izin lokasi data, `CREATE_TABLE` izin tersebut mungkin tidak cukup untuk membuat tabel. Penting untuk mengingat tiga kasus berikut.

Buat kasus penggunaan tabel	Izin diperlukan
Lokasi yang ditentukan tidak dikelola oleh Lake Formation (tidak terdaftar).	`CREATE_TABLE` sudah cukup.
Lokasi yang ditentukan dikelola oleh Lake Formation (terdaftar), dan database yang berisi tidak memiliki properti lokasi atau memiliki properti lokasi yang bukan awalan Amazon S3 dari lokasi tabel.	`CREATE_TABLE` diperlukan ditambah izin lokasi data pada lokasi yang ditentukan.
Lokasi yang ditentukan dikelola oleh Lake Formation (terdaftar), dan database yang berisi memiliki properti lokasi yang menunjuk ke lokasi yang terdaftar dan merupakan awalan Amazon S3 dari lokasi tabel.	`CREATE_TABLE` sudah cukup.

`DATA_LOCATION_ACCESS`

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
`DATA_LOCATION_ACCESS`	Lokasi Amazon S3	(Izin Amazon S3 di lokasi, yang harus ditentukan oleh peran yang digunakan untuk mendaftarkan lokasi.)

Ini adalah satu-satunya izin lokasi data. Prinsipal dengan izin ini dapat membuat database metadata atau tabel yang menunjuk ke lokasi Amazon S3 yang ditentukan. Lokasi harus terdaftar. Kepala sekolah yang memiliki izin lokasi data di lokasi juga memiliki izin lokasi pada lokasi anak.

Example

Contoh berikut memberikan izin lokasi data `s3://products/retail` ke pengguna di AWS akun `datalake_user1` 1111-2222-3333.

```
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`tidak diperlukan untuk menanyakan atau memperbarui data yang mendasarinya. Izin ini hanya berlaku untuk membuat sumber daya Katalog Data.

Untuk informasi selengkapnya tentang izin lokasi data, lihat[Underlying data access control](#).

DELETE

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
`DELETE`	`TABLE`	(Tidak diperlukan izin IAM tambahan jika lokasi terdaftar.)

Prinsipal dengan izin ini dapat menyisipkan, memperbarui, dan membaca data yang mendasarinya di lokasi Amazon S3 yang ditentukan oleh tabel. Kepala sekolah juga dapat melihat tabel di konsol Lake Formation dan mengambil informasi tentang tabel dengan AWS Glue API.

Example

Contoh berikut memberikan `DELETE` izin kepada pengguna `datalake_user1` pada tabel `inventory` dalam database `retail` di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DELETE" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Izin ini hanya berlaku untuk data di Amazon S3, dan bukan untuk data di penyimpanan data lain seperti Amazon Relational Database Service (Amazon RDS).

MENJELASKAN

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
DESCRIBE	Tautan sumber daya tabel Tautan sumber daya basis data	glue:GetTable glue:GetDatabase
DESCRIBE	DATABASE	glue:GetDatabase
DESCRIBE	TABLE	glue:GetTable
DESCRIBE	LF-Tag	glue:GetTable glue:GetDatabase lakeformation:GetResourceLFTags lakeformation:ListLFTags lakeformation:GetLFTag lakeformation:SearchTablesByLFTags lakeformation:SearchDatabasesByLFTags

Prinsipal dengan izin ini dapat melihat database, tabel, atau tautan sumber daya yang ditentukan. Tidak ada izin Katalog Data lainnya yang diberikan secara implisit, dan tidak ada izin akses data yang diberikan secara implisit. Database dan tabel muncul di editor kueri layanan terintegrasi, tetapi tidak

ada kueri yang dapat dibuat terhadapnya kecuali izin Lake Formation lainnya (misalnya,SELECT) diberikan.

Misalnya, pengguna yang memiliki DESCRIBE database dapat melihat database dan semua metadata database (deskripsi, lokasi, dan sebagainya). Namun, pengguna tidak dapat mengetahui tabel mana yang berisi database, dan tidak dapat menjatuhkan, mengubah, atau membuat tabel dalam database. Demikian pula, pengguna yang memiliki DESCRIBE tabel dapat melihat metadata tabel dan tabel (deskripsi, skema, lokasi, dan sebagainya), tetapi tidak dapat menjatuhkan, mengubah, atau menjalankan kueri terhadap tabel.

Berikut ini adalah beberapa aturan tambahan untukDESCRIBE:

- Jika pengguna memiliki izin Lake Formation lainnya pada database, tabel, atau tautan sumber daya, secara implisit DESCRIBE diberikan.
- Jika pengguna hanya memiliki SELECT subset kolom untuk tabel (sebagianSELECT), pengguna dibatasi untuk hanya melihat kolom tersebut.
- Anda tidak dapat memberikan DESCRIBE kepada pengguna yang memiliki pilihan sebagian pada tabel. Sebaliknya, Anda tidak dapat menentukan penyertaan kolom atau daftar pengecualian untuk tabel yang DESCRIBE diberikan pada.

Example

Contoh berikut memberikan DESCRIBE izin kepada pengguna datalake_user1 pada tautan inventory-link sumber daya tabel di database retail di AWS akun 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"}}'
```

MENJATUHKAN

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
DROP	DATABASE	glue:DeleteDatabase
DROP	TABLE	glue:DeleteTable

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
DROP	LF-Tag	lakeformation:DeleteLFTag
DROP	Tautan sumber daya basis data Tautan sumber daya tabel	glue:DeleteDatabase glue:DeleteTable

Prinsipal dengan izin ini dapat menjatuhkan database, tabel, atau tautan sumber daya di Katalog Data. Anda tidak dapat memberikan DROP pada database ke akun atau organisasi eksternal.

> **Warning**
>
> Menjatuhkan database menjatuhkan semua tabel dalam database.

Example

Contoh berikut memberikan DROP izin kepada pengguna pada database datalake_user1 di AWS akun retail 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Database": {"Name":"retail"}}'
```

Example

Contoh berikut memberikan DROP kepada pengguna datalake_user1 pada tabel inventory dalam databaseretail.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Example

Contoh berikut memberikan DROP kepada pengguna `datalake_user1` pada link sumber daya tabel `inventory-link` dalam 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

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
INSERT	TABLE	(Tidak diperlukan izin IAM tambahan jika lokasi terdaftar.)

Prinsipal dengan izin ini dapat menyisipkan, memperbarui, dan membaca data yang mendasarinya di lokasi Amazon S3 yang ditentukan oleh tabel. Kepala sekolah juga dapat melihat tabel di konsol Lake Formation dan mengambil informasi tentang tabel dengan AWS Glue API.

Example

Contoh berikut memberikan `INSERT` izin kepada pengguna `datalake_user1` pada tabel `inventory` dalam database `retail` di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "INSERT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Izin ini hanya berlaku untuk data di Amazon S3, dan bukan untuk data di penyimpanan data lain seperti Amazon RDS.

SELECT

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
SELECT	• TABLE	(Tidak diperlukan izin IAM tambahan jika lokasi terdaftar.)

Prinsipal dengan izin ini dapat melihat tabel di Katalog Data, dan dapat menanyakan data yang mendasarinya di Amazon S3 di lokasi yang ditentukan oleh tabel. Kepala sekolah dapat melihat tabel di konsol Lake Formation dan mengambil informasi tentang tabel dengan AWS Glue API. Jika pemfilteran kolom diterapkan saat izin ini diberikan, prinsipal dapat melihat metadata hanya untuk kolom yang disertakan dan hanya dapat menanyakan data dari kolom yang disertakan.

> **Note**
>
> Merupakan tanggung jawab layanan analitik terintegrasi untuk menerapkan pemfilteran kolom saat memproses kueri.

Example

Contoh berikut memberikan SELECT izin kepada pengguna `datalake_user1` pada tabel `inventory` dalam database `retail` di AWS akun 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "SELECT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Izin ini hanya berlaku untuk data di Amazon S3, dan bukan untuk data di penyimpanan data lain seperti Amazon RDS.

Anda dapat memfilter (membatasi akses ke) kolom tertentu dengan daftar inklusi opsional atau daftar pengecualian. Daftar inklusi menentukan kolom yang dapat diakses. Daftar pengecualian menentukan kolom yang tidak dapat diakses. Dengan tidak adanya daftar inklusi atau pengecualian, semua kolom tabel dapat diakses.

Hasil `glue:GetTable` pengembalian hanya kolom yang pemanggil memiliki izin untuk melihat. Layanan terintegrasi seperti Amazon Athena dan Amazon Redshift menghormati inklusi kolom dan daftar pengecualian.

Example

Contoh berikut memberikan `SELECT` kepada pengguna `datalake_user1` pada tabel `inventory` menggunakan daftar inklusi.

```
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

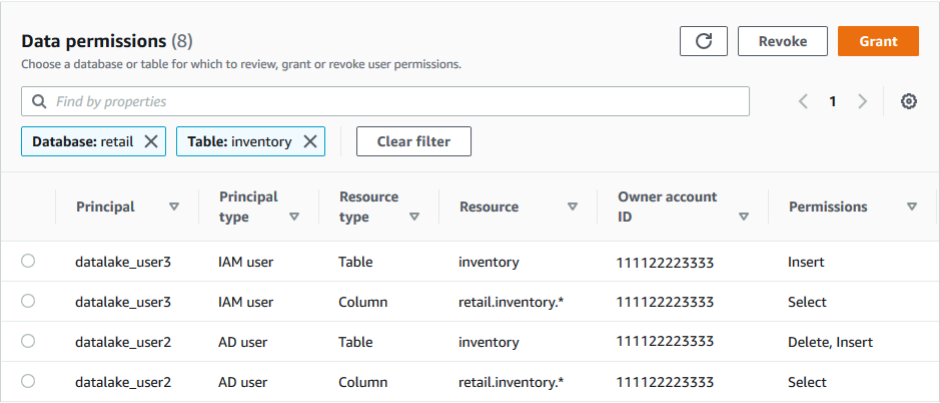
Contoh berikutnya ini memberikan `SELECT` pada `inventory` tabel menggunakan daftar pengecualian.

```
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"]}}}'
```

Pembatasan berikut berlaku untuk `SELECT` izin:

- Saat memberikan`SELECT`, Anda tidak dapat menyertakan opsi hibah jika pemfilteran kolom diterapkan.
- Anda tidak dapat membatasi kontrol akses pada kolom yang merupakan kunci partisi.
- Seorang kepala sekolah dengan `SELECT` izin pada subset kolom dalam tabel tidak dapat diberikan`ALTER`,, `DROP``DELETE`, atau `INSERT` izin pada tabel itu. Demikian pula, kepala sekolah dengan`ALTER`,`DROP`,`DELETE`, atau `INSERT` izin di atas meja tidak dapat diberikan `SELECT` izin dengan pemfilteran kolom.

`SELECT`Izin selalu muncul di halaman izin Data konsol Lake Formation sebagai baris terpisah. Gambar berikut ini menunjukkan bahwa `SELECT` diberikan kepada pengguna `datalake_user2` dan `datalake_user3` pada semua kolom dalam `inventory` tabel.



Super

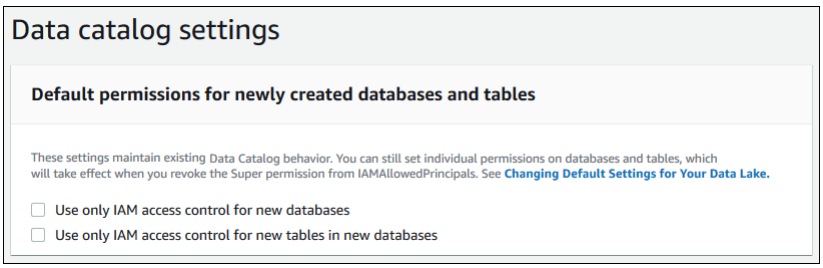
Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
`Super`	`DATABASE`	`glue:*Database*`
`Super`	`TABLE`	`glue:*Table*, glue:*Partition*`

Izin ini memungkinkan kepala sekolah untuk melakukan setiap operasi Lake Formation yang didukung pada database atau tabel. Anda tidak dapat memberikan `Super` pada database ke akun eksternal.

Izin ini dapat hidup berdampingan dengan izin Lake Formation lainnya. Misalnya, Anda dapat memberikan `INSERT` izin`Super`,`SELECT`, dan pada tabel metadata. Kepala sekolah kemudian dapat melakukan semua operasi yang didukung di atas meja. Saat Anda mencabut`Super`, `INSERT` izin `SELECT` dan tetap ada, dan prinsipal hanya dapat melakukan operasi pilih dan sisipkan.

Alih-alih memberikan `Super` kepada kepala sekolah individu, Anda dapat memberikannya kepada grup`IAMAllowedPrincipals`. `IAMAllowedPrincipals`Grup dibuat secara otomatis dan mencakup semua pengguna dan peran IAM yang diizinkan mengakses sumber daya Katalog Data Anda oleh kebijakan IAM Anda. Ketika `Super` diberikan kepada `IAMAllowedPrincipals` sumber daya Katalog Data, akses ke sumber daya dikendalikan secara efektif hanya oleh kebijakan IAM.

Anda dapat memiliki `Super` izin untuk secara otomatis diberikan kepada `IAMAllowedPrincipals` sumber daya katalog baru dengan memanfaatkan opsi di halaman Pengaturan konsol Lake Formation.



- `SuperIAMAllowedPrincipals`Untuk memberikan semua database baru, pilih Gunakan hanya kontrol akses IAM untuk database baru.
- `SuperIAMAllowedPrincipals`Untuk memberikan semua tabel baru dalam database baru, pilih Gunakan hanya kontrol akses IAM untuk tabel baru di database baru.

> **Note**
>
> Opsi ini menyebabkan kotak centang Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini di kotak dialog Buat database yang akan dipilih secara default. Itu tidak lebih dari itu. Ini adalah kotak centang di kotak dialog Buat database yang memungkinkan pemberian `Super` to`IAMAllowedPrincipals`.

Opsi halaman Pengaturan ini diaktifkan secara default. Untuk informasi selengkapnya, lihat berikut ini:

- the section called “Mengubah pengaturan default untuk data lake”
- the section called “Memutakhirkan izin AWS Glue data ke model Lake Formation”

SUPER_PENGGUNA

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan
`Super user`	`Catalog`	`glue:GetCatalog`

Anda dapat memberikan `Super user` izin hanya kepada prinsipal tertentu pada katalog dalam Katalog Data default. Anda tidak dapat memberikan `Super user` izin pada katalog default atau jenis sumber daya lain seperti database dan tabel atau kepada prinsipal di akun eksternal. `Super`

`user`Izin izin memungkinkan kepala sekolah untuk melakukan setiap operasi Lake Formation yang didukung pada database dan tabel dalam katalog yang diberikan.

Dengan `Super user` izin, kepala sekolah (penerima hibah) dapat melakukan tindakan berikut pada sumber daya (katalog, database, dan tabel) dalam katalog:

- `CREATE_DATABASE`, `DESCRIBE` izin pada katalog.
- `DROP`,`ALTER`,`CREATE_TABLE`, `DESCRIBE` (efektif`SUPER`) izin pada semua database dalam katalog.
- `DROP`,`ALTER`,`DESCRIBE`,`SELECT`,`INSERT`, `DELETE` (efektif`SUPER`) izin pada semua tabel dalam semua database dalam katalog.
- `All`(efektif SUPER) izin pada katalog dalam katalog.
- Izin yang dapat diberikan (kemampuan untuk memberikan izin ini kepada prinsipal lain) pada semua katalog, database, dan tabel dalam katalog.

Dengan `Super user` izin pada sumber daya katalog, penerima hibah tidak diizinkan untuk melakukan atau mendelegasikan `ALTER` dan `DROP` tindakan pada katalog.

REKAN

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan

`ASSOCIATE`	`LF-Tag`	`glue:GetDatabase`
		`glue:GetTable`
		`lakeformation:AddLFTagsToResource"`
		`lakeformation:RemoveLFTagsFromResource"`
		`lakeformation:GetResourceLFTags`
		`lakeformation:ListLFTags`
		`lakeformation:GetLFTag`

Izin	Diberikan pada sumber daya ini	Penerima hibah juga membutuhkan

		`lakeformation:SearchTablesByLFTags`
		
		`lakeformation:SearchDatabasesByLFTags`

Seorang kepala sekolah dengan izin ini pada LF-Tag dapat menetapkan LF-Tag ke sumber daya Katalog Data. Memberikan hibah `ASSOCIATE` implisit. `DESCRIBE`

Example

Contoh ini memberikan `ASSOCIATE` izin `datalake_user1` kepada pengguna LF-Tag dengan kunci`module`. Ini memberikan izin untuk melihat dan menetapkan semua nilai untuk kunci itu, seperti yang ditunjukkan oleh tanda bintang (*)..

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Mengintegrasikan Pusat Identitas IAM

Dengan AWS IAM Identity Center, Anda dapat terhubung ke penyedia identitas (IdPs) dan mengelola akses secara terpusat untuk pengguna dan grup di seluruh layanan AWS analitik. Anda dapat mengintegrasikan penyedia identitas seperti Okta, Ping, dan Microsoft Entra ID (sebelumnya Azure Active Directory) dengan Pusat Identitas IAM bagi pengguna di organisasi Anda untuk mengakses data menggunakan pengalaman masuk tunggal. Pusat Identitas IAM juga mendukung menghubungkan penyedia identitas pihak ketiga tambahan.

Untuk informasi selengkapnya lihat, Penyedia identitas yang didukung di Panduan AWS IAM Identity Center Pengguna.

Anda dapat mengonfigurasi AWS Lake Formation sebagai aplikasi yang diaktifkan di Pusat Identitas IAM, dan administrator danau data dapat memberikan izin berbutir halus kepada pengguna dan grup resmi pada sumber daya. AWS Glue Data Catalog

Pengguna dari organisasi Anda dapat masuk ke aplikasi yang diaktifkan Pusat Identitas menggunakan penyedia identitas organisasi Anda, dan menanyakan kumpulan data yang menerapkan izin Lake Formation. Dengan integrasi ini, Anda dapat mengelola akses ke AWS layanan, tanpa membuat beberapa peran IAM.

Propagasi identitas tepercaya adalah AWS IAM Identity Center fitur yang Layanan AWS dapat digunakan administrator yang terhubung untuk memberikan dan mengaudit akses ke data layanan. Akses ke data ini didasarkan pada atribut pengguna seperti asosiasi grup. Menyiapkan propagasi identitas tepercaya memerlukan kolaborasi antara administrator yang terhubung Layanan AWS dan administrator Pusat Identitas IAM. Untuk informasi selengkapnya, lihat Pras yarat dan pertimbangan.

Untuk batasan, lihat Keterbatasan integrasi Pusat Identitas IAM.

Topik

- Prasyarat untuk integrasi Pusat Identitas IAM dengan Formasi Danau
- Menghubungkan Formasi Danau dengan Pusat Identitas IAM
- Memperbarui integrasi Pusat Identitas IAM
- Menghapus koneksi Lake Formation dengan IAM Identity Center
- Memberikan izin kepada pengguna dan grup
- Termasuk konteks pengguna Pusat Identitas IAM dalam log CloudTrail

Prasyarat untuk integrasi Pusat Identitas IAM dengan Formasi Danau

Berikut ini adalah prasyarat untuk mengintegrasikan Pusat Identitas IAM dengan Formasi Danau.

1. Aktifkan Pusat Identitas IAM — Mengaktifkan Pusat Identitas IAM merupakan prasyarat untuk mendukung otentikasi dan propagasi identitas.
2. Pilih sumber identitas Anda — Setelah mengaktifkan Pusat Identitas IAM, Anda harus memiliki penyedia identifikasi untuk mengelola pengguna dan grup. Anda dapat menggunakan direktori Pusat Identitas bawaan sebagai sumber identitas atau menggunakan IdP eksternal, seperti Microsoft Entra ID atau Okta.

Untuk informasi selengkapnya, lihat Mengel ola sumber identitas Anda dan Menghubun gkan ke penyedia identitas eksternal di Panduan AWS IAM Identity Center Pengguna.

3. Buat peran IAM — Peran yang membuat koneksi Pusat Identitas IAM memerlukan izin untuk membuat dan memodifikasi konfigurasi aplikasi di Lake Formation dan Pusat Identitas IAM seperti dalam kebijakan sebaris berikut.

   Anda perlu menambahkan izin sesuai praktik terbaik IAM. Izin khusus dirinci dalam prosedur berikut. Untuk informasi selengkapnya, lihat Memulai dengan Pusat Identitas IAM.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:CreateLakeFormationIdentityCenterConfiguration",
                "sso:CreateApplication",
                "sso:PutApplicationAssignmentConfiguration",
                "sso:PutApplicationAuthenticationMethod",
                "sso:PutApplicationGrant",
                "sso:PutApplicationAccessScope"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

   Jika Anda berbagi sumber daya Katalog Data dengan eksternal Akun AWS atau organisasi, Anda harus memiliki izin AWS Resource Access Manager (AWS RAM) untuk membuat berbagi sumber daya. Untuk informasi selengkapnya tentang izin yang diperlukan untuk berbagi sumber daya, lihat prasyarat berbagi Cross-account data.

Kebijakan sebaris berikut berisi izin khusus yang diperlukan untuk melihat, memperbarui, dan menghapus properti integrasi Lake Formation dengan Pusat Identitas IAM.

- Gunakan kebijakan sebaris berikut untuk mengizinkan peran IAM untuk melihat integrasi Lake Formation dengan Pusat Identitas IAM.

  JSON

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

 "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- Gunakan kebijakan sebaris berikut untuk mengizinkan peran IAM memperbarui integrasi Lake Formation dengan Pusat Identitas IAM. Kebijakan ini juga mencakup izin opsional yang diperlukan untuk berbagi sumber daya dengan akun eksternal.

  JSON

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

 "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication",
                "sso:UpdateApplication"
            ],
            "Resource": [
                "*"
            ]
```

```
        }
    ]
}
```

- Gunakan kebijakan sebaris berikut untuk mengizinkan peran IAM menghapus integrasi Lake Formation dengan Pusat Identitas IAM.

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:DeleteLakeFormationIdentityCenterConfiguration",
                "sso:DeleteApplication"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- Untuk izin IAM yang diperlukan untuk memberikan atau mencabut izin danau data untuk pengguna dan grup Pusat Identitas IAM, lihat. Izin IAM diperlukan untuk memberikan atau mencabut izin Lake Formation

Deskripsi izin

- `lakeformation:CreateLakeFormationIdentityCenterConfiguration`— Menciptakan konfigurasi iDC Formasi Danau.
- `lakeformation:DescribeLakeFormationIdentityCenterConfiguration`— Menjelaskan konfigurasi iDC yang ada.
- `lakeformation:DeleteLakeFormationIdentityCenterConfiguration`- Memberikan kemampuan untuk menghapus konfigurasi iDC Formasi Danau yang ada.
- `lakeformation:UpdateLakeFormationIdentityCenterConfiguration`— Digunakan untuk mengubah konfigurasi Formasi Danau yang ada.
- `sso:CreateApplication`— Digunakan untuk membuat aplikasi Pusat Identitas IAM.

- `sso:DeleteApplication`— Digunakan untuk menghapus aplikasi Pusat Identitas IAM.
- `sso:UpdateApplication`— Digunakan untuk memperbarui aplikasi Pusat Identitas IAM.
- `sso:PutApplicationGrant`— Digunakan untuk mengubah informasi penerbit token tepercaya.
- `sso:PutApplicationAuthenticationMethod`— Memberikan akses otentikasi Formasi Danau.
- `sso:GetApplicationGrant`— Digunakan untuk mencantumkan informasi penerbit token tepercaya.
- `sso:DeleteApplicationGrant`— Menghapus informasi penerbit token kepercayaan.
- `sso:PutApplicationAccessScope`— Menambahkan atau memperbarui daftar target resmi untuk cakupan akses Pusat Identitas IAM untuk aplikasi.
- `sso:PutApplicationAssignmentConfiguration`— Digunakan untuk mengkonfigurasi bagaimana pengguna mendapatkan akses ke aplikasi.

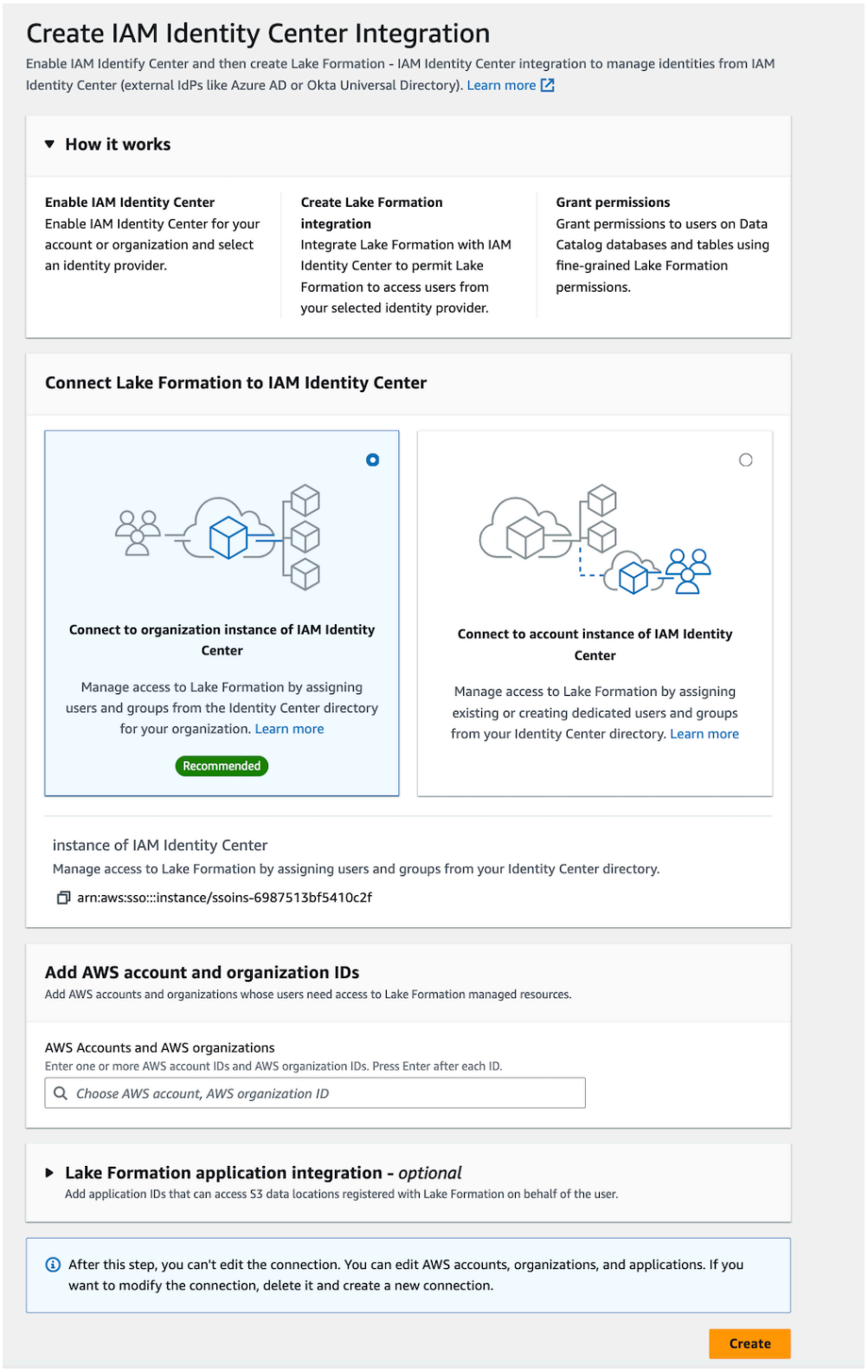
Menghubungkan Formasi Danau dengan Pusat Identitas IAM

Sebelum Anda dapat menggunakan Pusat Identitas IAM untuk mengelola identitas guna memberikan akses ke sumber daya Katalog Data menggunakan Lake Formation, Anda harus menyelesaikan langkah-langkah berikut. Anda dapat membuat integrasi Pusat Identitas IAM menggunakan konsol Lake Formation atau AWS CLI.

Konsol Manajemen AWS

Untuk menghubungkan Formasi Danau dengan Pusat Identitas IAM

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi kiri, pilih integrasi Pusat Identi tas IAM.



3. (Opsional) Masukkan satu atau beberapa Akun AWS ID yang valid, ID and/or organisasi, ID unit organisasi untuk memungkinkan akun eksternal mengakses sumber daya Katalog Data. Ketika pengguna atau grup IAM Identity Center mencoba mengakses sumber daya

Katalog Data yang dikelola Lake Formation, Lake Formation mengambil peran IAM untuk mengotorisasi akses metadata. Jika peran IAM milik akun eksternal yang tidak memiliki kebijakan AWS Glue sumber daya dan pembagian AWS RAM sumber daya, pengguna dan grup Pusat Identitas IAM tidak akan dapat mengakses sumber daya meskipun mereka memiliki izin Lake Formation.

Lake Formation menggunakan layanan AWS Resource Access Manager (AWS RAM) untuk berbagi sumber daya dengan akun dan organisasi eksternal. AWS RAM mengirimkan undangan ke akun penerima hibah untuk menerima atau menolak pembagian sumber daya.

Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

> **Note**
>
> Lake Formation mengizinkan peran IAM dari akun eksternal untuk bertindak sebagai peran operator atas nama pengguna dan grup Pusat Identitas IAM untuk mengakses sumber daya Katalog Data, tetapi izin hanya dapat diberikan pada sumber daya Katalog Data dalam akun yang dimiliki. Jika Anda mencoba memberikan izin kepada pengguna dan grup Pusat Identitas IAM pada sumber daya Katalog Data di akun eksternal, Lake Formation menampilkan kesalahan berikut - "Cross-account hibah tidak didukung untuk prinsipal."

4. (Opsional) Pada layar integrasi Create Lake Formation, tentukan ARN aplikasi pihak ketiga yang dapat mengakses data di lokasi Amazon S3 yang terdaftar di Lake Formation. Lake Formation menjual kredenSIAL sementara dalam bentuk AWS STS token ke lokasi Amazon S3 terdaftar berdasarkan izin efektif, sehingga aplikasi resmi dapat mengakses data atas nama pengguna.
5. (Opsional) Pada layar integrasi Create Lake Formation, centang kotak centang Amazon Redshift Connect di Propagasi Identitas Tepercaya untuk mengaktifkan penemuan Izin Federasi Amazon Redshift melalui IDC. Lake Formation menyebarkan identitas ke hilir berdasarkan izin yang efektif, sehingga aplikasi resmi dapat mengakses data atas nama pengguna.
6. Pilih Kirim.

Setelah administrator Danau Formasi menyelesaikan langkah-langkah dan membuat integrasi, properti Pusat Identitas IAM muncul di konsol Lake Formation. Menyelesaikan tugas-tugas ini menjadikan Lake Formation aplikasi yang diaktifkan Pusat Identitas IAM.

Properti di konsol menyertakan status integrasi. Status integrasi mengatakan `Success` kapan selesai. Status ini menunjukkan jika konfigurasi Pusat Identitas IAM selesai.

AWS CLI

- Contoh berikut menunjukkan cara membuat integrasi Lake Formation dengan IAM Identity Center. Anda juga dapat menentukan `Status` (`ENABLED`,`DISABLED`) dari aplikasi.

```
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"}'
```

- Contoh berikut menunjukkan cara melihat integrasi Danau Formasi dengan Pusat Identitas IAM.

```
aws lakeformation describe-lake-formation-identity-center-configuration
 --catalog-id <123456789012>
```

- Contoh berikut menunjukkan cara mengaktifkan Otor `Redshift:Connect` isasi. Otorisasi dapat DIAKTIFKAN atau DINONAKTIFKAN.

```
aws lakeformation  create-lake-formation-identity-center-configuration \
--instance-arn <arn:aws:sso:::instance/ssoins-112111f12ca1122p> \
--service-integrations '[{
  "Redshift": [{
    "RedshiftConnect": {
      "Authorization": "ENABLED"
    }
  }]
}]'
```

- Gunakan `describe-lake-formation-identity-center-configuration` perintah untuk menggambarkan aplikasi pusat identitas formasi danau. `Redshift:Connect`integrasi layanan sangat penting untuk propagasi identitas iDC lintas layanan dan lintas cluster:

```
aws lakeformation describe-lake-formation-identity-center-configuration --catalog-
id <123456789012>
```

Respons:

```
{
    "CatalogId": "CATALOG ID",
    "InstanceArn": "INSTANCE ARN",
    "ApplicationArn": "APPLICATION ARN",
    "ShareRecipients": [],
    "ServiceIntegrations": [
        {
            "Redshift": [
                {
                    "RedshiftConnect": {
                        "Authorization": "ENABLED"
                    }
                }
            ]
        }
    ]
}
```

Menggunakan Pusat Identitas IAM di beberapa Wilayah AWS

Lake Formation mendukung Pusat Identitas IAM dalam beberapa Wilayah AWS hal. Anda dapat memperluas Pusat Identitas IAM dari Wilayah utama Wilayah AWS ke Wilayah tambahan untuk meningkatkan kinerja melalui kedekatan dengan pengguna dan keandalan. Ketika Wilayah baru ditambahkan di Pusat Identitas IAM, Anda dapat membuat aplikasi Lake Formation Identity Center di Wilayah baru tanpa mereplikasi identitas dari Wilayah utama. Untuk detail selengkapnya untuk memulai dengan Pusat Identitas IAM di beberapa Wilayah, lihat Pusat Identi Multi-Region tas IAM di Panduan Pengguna Pusat Identi tas IAM.

Memperbarui integrasi Pusat Identitas IAM

Setelah membuat koneksi, Anda dapat menambahkan aplikasi pihak ketiga untuk integrasi Pusat Identitas IAM untuk diintegrasikan dengan Lake Formation, dan mendapatkan akses ke data Amazon S3 atas nama pengguna. Anda juga dapat menghapus aplikasi yang ada dari integrasi

Pusat Identitas IAM. Anda dapat menambah atau menghapus aplikasi menggunakan konsol Lake Formation, AWS CLI, dan menggunakan UpdateLakeFormationIdentityCenterConfiguration operasi.

> **Note**
>
> Setelah membuat integrasi Pusat Identitas IAM, Anda tidak dapat memperbarui instanceARN.

Konsol Manajemen AWS

Untuk memperbarui koneksi Pusat Identitas IAM yang ada dengan Formasi Danau

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi kiri, pilih integrasi Pusat Identi tas IAM.
3. Pilih Tambah pada halaman integrasi Pusat Identitas IAM.
4. Masukkan satu atau beberapa Akun AWS ID yang valid, ID and/or organisasi, ID unit organisasi untuk memungkinkan akun eksternal mengakses sumber daya Katalog Data.
5. Pada layar Tambah aplikasi, masukkan ID aplikasi dari aplikasi pihak ketiga yang ingin Anda integrasikan dengan Lake Formation.
6. Pilih Tambahkan.
7. (Opsional) Pada halaman integrasi Pusat Identitas IAM, Anda dapat mengaktifkan propagasi identitas tepercaya untuk Amazon Redshift connect atau menonaktifkannya. Lake Formation menyebarkan identitas ke hilir berdasarkan izin yang efektif, sehingga aplikasi resmi dapat mengakses data atas nama pengguna.

AWS CLI

Anda dapat menambahkan atau menghapus aplikasi pihak ketiga untuk integrasi Pusat Identitas IAM dengan menjalankan AWS CLI perintah berikut. Saat Anda menyetel status pemfilteran eksternal keENABLED, itu memungkinkan Pusat Identitas IAM untuk menyediakan manajemen identitas bagi aplikasi pihak ketiga untuk mengakses data yang dikelola oleh Lake Formation. Anda juga dapat mengaktifkan atau menonaktifkan integrasi Pusat Identitas IAM dengan mengatur status aplikasi.

```
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
```

Jika Anda memiliki aplikasi LF IDC yang sudah ada, tetapi ingin menambahkan `Redshift:Connect` otorisasi, Anda dapat menggunakan yang berikut ini untuk memperbarui Aplikasi IDC Lake Formation Anda. Otorisasi dapat DIAKTIFKAN atau DINONAKTIFKAN.

```
aws lakeformation update-lake-formation-identity-center-configuration \
--service-integrations '[{

  "Redshift": [{
    "RedshiftConnect": {
      "Authorization": "ENABLED"
    }
  }]
}]'
```

Menghapus koneksi Lake Formation dengan IAM Identity Center

Jika Anda ingin menghapus integrasi Pusat Identitas IAM yang ada, Anda dapat melakukannya menggunakan konsol Lake Formation, AWS CLI, atau DeleteLakeFormationIdentityCenterConfiguration operasi.

Konsol Manajemen AWS

Untuk menghapus koneksi Pusat Identitas IAM yang ada dengan Formasi Danau

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi kiri, pilih integrasi Pusat Identi tas IAM.
3. Pilih Hapus pada halaman integrasi Pusat Identitas IAM.
4. Pada layar Kon firmasi integrasi, konfirmasikan tindakan, dan pilih H apus.

AWS CLI

Anda dapat menghapus integrasi Pusat Identitas IAM dengan menjalankan AWS CLI perintah berikut.

```
aws lakeformation delete-lake-formation-identity-center-configuration \
    --catalog-id <123456789012>
```

Memberikan izin kepada pengguna dan grup

Administrator danau data Anda dapat memberikan izin kepada pengguna dan grup Pusat Identitas IAM pada sumber daya Katalog Data (database, tabel, dan tampilan) untuk memungkinkan akses data yang mudah. Untuk memberikan atau mencabut izin danau data, pemberi memerlukan izin untuk tindakan Pusat Identitas IAM berikut.

- DescribeUser
- DescribeGroup
- DescribeInstance

Anda dapat memberikan izin dengan menggunakan konsol Lake Formation, API, atau AWS CLI.

Untuk informasi selengkapnya tentang pemberian izin, lihatthe section called “Memberikan izin data”.

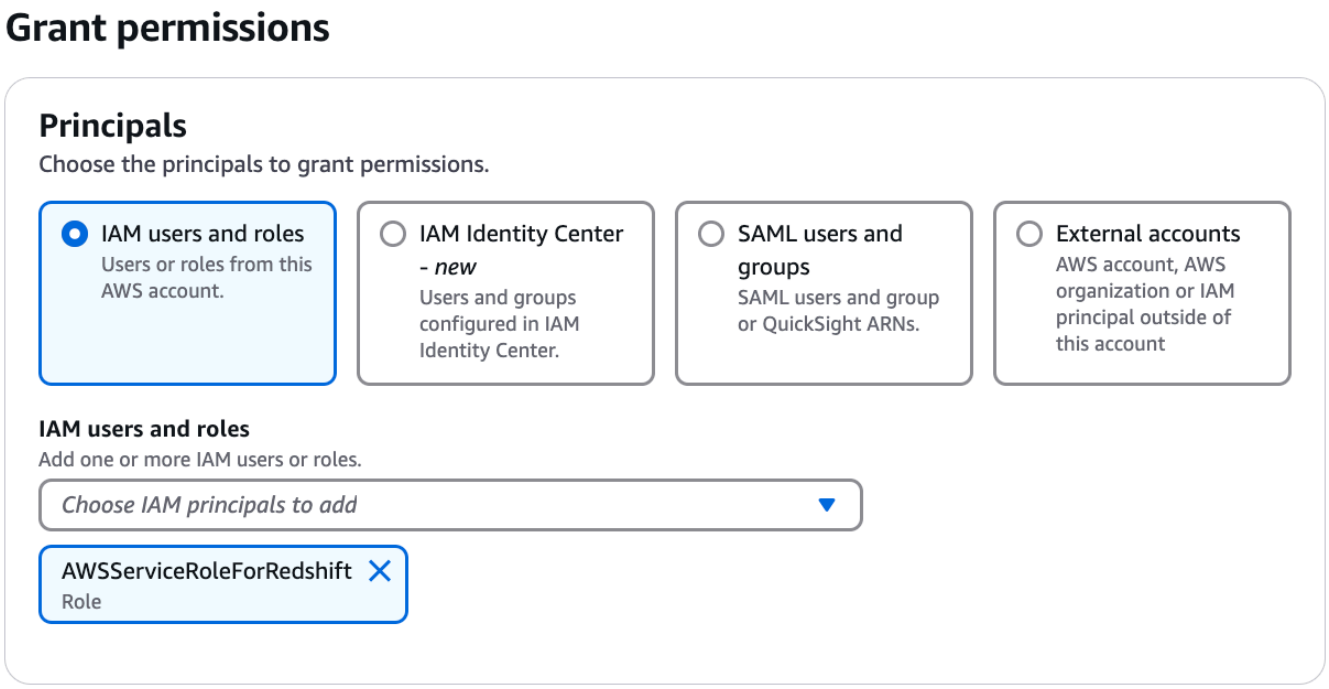
> **Note**
>
> Anda hanya dapat memberikan izin pada sumber daya di akun Anda. Untuk mengalokasikan izin ke pengguna dan grup pada sumber daya yang dibagikan dengan Anda, Anda harus menggunakan AWS RAM berbagi sumber daya.

Konsol Manajemen AWS

Untuk memberikan izin kepada pengguna dan grup

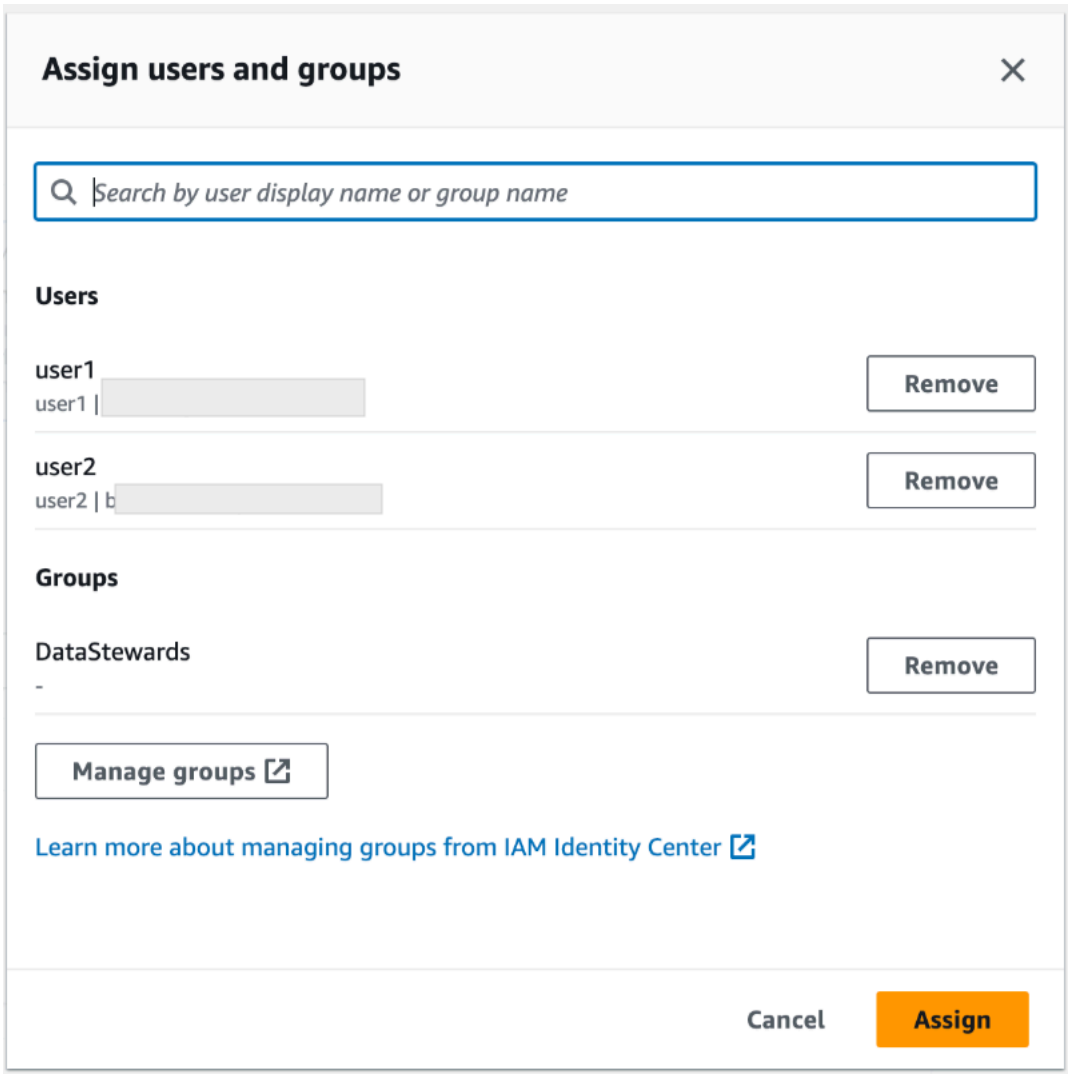
1. Masuk ke Konsol Manajemen AWS, dan buka konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Pilih izin danau data di bawah Iz in di konsol Formasi Danau.

3. Pilih Grant.
4. Pada halaman Berikan izin danau data, pilih, pengguna dan grup Pusat Identitas IAM.
5. Pilih Tambah untuk memilih pengguna dan grup untuk memberikan izin.



6. Pada layar Tetapkan pengguna dan grup, pilih and/or grup pengguna untuk memberikan izin.

   Pilih Tetapkan.



7. Selanjutnya, pilih metode untuk memberikan izin.

   Untuk petunjuk tentang pemberian izin menggunakan metode sumber daya bernama, lihatMemberikan izin data menggunakan metode sumber daya bernama.

   Untuk petunjuk tentang pemberian izin menggunakan LF-Tags, lihatMemberikan izin data lake menggunakan metode LF-TBAC.

8. Pilih sumber daya Katalog Data yang ingin Anda berikan izin.
9. Pilih izin Katalog Data untuk diberikan.
10. Pilih Grant.

AWS CLI

Contoh berikut menunjukkan cara memberikan SELECT izin pengguna Pusat Identitas IAM pada tabel.

```
aws lakeformation grant-permissions \
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserId> \
--permissions "SELECT" \
--resource '{ "Table": { "DatabaseName": "retail", "TableWildcard": {} } }'
```

Untuk mengambil UserId dari Pusat Identitas IAM, lihat GetUserId operasi di Referensi API Pusat Identitas IAM.

Termasuk konteks pengguna Pusat Identitas IAM dalam log CloudTrail

Lake Formation menggunakan fungsi pen jual otomatis kredensi untuk menyediakan akses sementara ke data Amazon S3. Secara default, ketika pengguna Pusat Identitas IAM mengirimkan kueri ke layanan analitik terintegrasi, CloudTrail log hanya menyertakan peran IAM yang diasumsikan oleh layanan untuk menyediakan akses jangka pendek. Jika Anda menggunakan peran yang ditentukan pengguna untuk mendaftarkan lokasi data Amazon S3 dengan Lake Formation, Anda dapat memilih untuk menyertakan konteks pengguna Pusat Identitas IAM dalam CloudTrail acara, lalu melacak pengguna yang mengakses sumber daya Anda.

> **Important**
>
> Untuk menyertakan permintaan API Amazon S3 tingkat objek di dalam CloudTrail, Anda perlu mengaktifkan pencatatan CloudTrail peristiwa untuk bucket dan objek Amazon S3. Untuk informasi selengkapnya, lihat Meng aktifkan pencatatan CloudTrail peristiwa untuk bucket dan objek Amazon S3 di Panduan Pengguna Amazon S3.

Untuk mengaktifkan audit penjual otomatis kredenSIAL pada lokasi danau data yang terdaftar dengan peran yang ditentukan pengguna

1. Masuk ke konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Di navigasi sisi kiri, perluas Administrasi, dan pilih pengaturan Katalog Data.

3. Di bawah Audit yang disempurnakan, pilih Menyebarkan konteks yang disediakan.
4. Pilih Simpan.

Anda juga dapat mengaktifkan opsi audit yang disempurnakan dengan menyet `Parameters` el atribut dalam PutDataLakeSettings operasi. Secara default, nilai `SET_CONTEXT"` parameter diatur ke “true”.

```
{
    "DataLakeSettings": {
        "Parameters": {"SET_CONTEXT": "true"},
    }
}
```

Berikut ini adalah kutipan dari CloudTrail acara dengan opsi audit yang disempurnakan. Log ini mencakup konteks sesi pengguna Pusat Identitas IAM dan peran IAM yang ditentukan pengguna yang diasumsikan oleh Lake Formation untuk mengakses lokasi data Amazon S3. Lihat `onBehalfOf` parameter dalam kutipan berikut.

```
{
         "eventVersion":"1.09",
         "userIdentity":{
            "type":"AssumedRole",
            "principalId":"AROAW7F7MOX4OYE6FLIFN: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":"ASIAW7F7MOX4CQLD4JIZN",
            "sessionContext":{
               "sessionIssuer":{
                  "type":"Role",
                  "principalId":"AROAW7F7MOX4OYE6FLIFN",
                  "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",
    ....
```

Menambahkan lokasi Amazon S3 ke danau data Anda

Untuk menambahkan lokasi data sebagai penyimpanan di danau data Anda, Anda mendaftarkan lokasi (Lokasi danau data) dengan AWS Lake Formation. Anda kemudian dapat menggunakan izin Lake Formation untuk kontrol akses berbutir halus ke AWS Glue Data Catalog objek yang mengarah ke lokasi ini dan ke data yang mendasarinya di lokasi.

Lake Formation juga memungkinkan untuk mendaftarkan lokasi data dalam mode akses hibrida dan memberi Anda fleksibilitas untuk mengaktifkan izin Lake Formation secara selektif untuk database dan tabel di Katalog Data Anda. Dengan mode akses Hybrid, Anda memiliki jalur tambahan yang memungkinkan Anda mengatur izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu kebijakan izin pengguna atau beban kerja lain yang ada.

Untuk informasi selengkapnya tentang pengaturan mode akses hybrid, lihat Mode akses hibrid

Saat Anda mendaftarkan lokasi, jalur Amazon S3 dan semua folder di bawah jalur itu terdaftar.

Misalnya, Anda memiliki organisasi jalur Amazon S3 seperti berikut ini:

`/mybucket/accounting/sales/`

Jika Anda mendaftar`S3://mybucket/accounting`, `sales` folder tersebut juga terdaftar dan berada di bawah manajemen Lake Formation.

Untuk informasi selengkapnya tentang mendaftarkan lokasi, lihatUnderlying data access control.

> **Note**
>
> Izin Lake Formation direkomendasikan untuk data terstruktur (disusun dalam tabel dengan baris dan kolom). Jika data Anda berisi data tidak terstruktur berbasis objek, pertimbangkan untuk menggunakan hibah akses Amazon S3 untuk mengelola akses data.

Topik

- Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi
- Mendaftarkan lokasi Amazon S3
- Mendaftarkan lokasi Amazon S3 terenkripsi
- Mendaftarkan lokasi Amazon S3 di tempat lain AWS akun
- Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening
- Membatalkan pendaftaran lokasi Amazon S3
- Mengakses lokasi Amazon S3

Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi

Anda harus menentukan peran AWS Identity and Access Management (IAM) saat mendaftarkan lokasi Amazon Simple Storage Service (Amazon S3). AWS Lake Formation mengasumsikan peran itu saat mengakses data di lokasi itu.

Anda dapat menggunakan salah satu jenis peran berikut untuk mendaftarkan lokasi:

- Peran terkait layanan Lake Formation. Peran ini memberikan izin yang diperlukan di lokasi. Menggunakan peran ini adalah cara paling sederhana untuk mendaftarkan lokasi. Untuk informasi selengkapnya, lihat Menggunakan peran terkait layanan untuk Lake Formation dan Batasan peran terkait layanan.
- Peran yang ditentukan pengguna. Gunakan peran yang ditentukan pengguna saat Anda perlu memberikan lebih banyak izin daripada yang disediakan oleh peran terkait layanan.

  Anda harus menggunakan peran yang ditentukan pengguna dalam keadaan berikut:

  - Saat mendaftarkan lokasi di akun lain.

Untuk informasi selengkapnya, lihat the section called “Mendaftarkan lokasi Amazon S3 di tempat lain AWS akun” dan the section called “Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening”.

- Jika Anda menggunakan CMK (`aws/s3`) AWS terkelola untuk mengenkripsi lokasi Amazon S3.

  Untuk informasi selengkapnya, lihat Mendaftarkan lokasi Amazon S3 terenkripsi.

- Jika Anda berencana untuk mengakses lokasi menggunakan Amazon EMR.

  Jika Anda sudah mendaftarkan lokasi dengan peran terkait layanan dan ingin mulai mengakses lokasi dengan Amazon EMR, Anda harus membatalkan pendaftaran lokasi dan mendaftarkannya kembali dengan peran yang ditentukan pengguna. Untuk informasi selengkapnya, lihat the section called “Membatalkan pendaftaran lokasi Amazon S3”.

Menggunakan peran terkait layanan untuk Lake Formation

AWS Lake Formation menggunakan peran AWS Identity and Access Management terkait layanan (IAM). Peran terkait layanan adalah jenis unik peran IAM yang terkait langsung dengan Lake Formation. Peran terkait layanan telah ditentukan sebelumnya oleh Lake Formation dan mencakup semua izin yang diperlukan layanan untuk memanggil AWS layanan lain atas nama Anda.

Peran terkait layanan membuat pengaturan Lake Formation menjadi lebih mudah karena Anda tidak perlu membuat peran dan menambahkan izin yang diperlukan secara manual. Lake Formation mendefinisikan izin dari peran terkait layanannya, dan kecuali ditentukan lain, hanya Lake Formation yang dapat mengambil perannya. Izin yang ditentukan mencakup kebijakan kepercayaan dan kebijakan izin, dan kebijakan izin tersebut tidak dapat dilampirkan ke entitas IAM lainnya.

Peran terkait layanan ini mempercayai layanan berikut untuk mengambil peran:

- `lakeformation.amazonaws.com`

Bila Anda menggunakan peran terkait layanan di akun A untuk mendaftarkan lokasi Amazon S3 yang dimiliki oleh akun B, kebijakan bucket Amazon S3 (kebijakan berbasis sumber daya) di akun B harus memberikan izin akses ke peran terkait layanan di akun A.

Untuk informasi tentang menggunakan peran terkait layanan untuk mendaftarkan lokasi data, lihat. Batasan peran terkait layanan

> **Note**
>
> Kebijakan kontrol layanan (SCPs) tidak memengaruhi peran terkait layanan. Untuk informasi selengkapnya, lihat Kebijakan kontrol layanan (SCPs) di panduan AWS Organizations pengguna.

Izin peran terkait layanan untuk Lake Formation

Lake Formation menggunakan peran terkait layanan bernama. `AWSServiceRoleForLakeFormationDataAccess` Peran ini menyediakan satu set izin Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) yang memungkinkan layanan terintegrasi Lake Formation (Amazon Athena seperti) untuk mengakses lokasi terdaftar. Saat mendaftarkan lokasi data lake, Anda harus memberikan peran yang memiliki read/write izin Amazon S3 yang diperlukan di lokasi tersebut. Alih-alih membuat peran dengan izin Amazon S3 yang diperlukan, Anda dapat menggunakan peran terkait layanan ini.

Pertama kali Anda menamai peran terkait layanan sebagai peran yang digunakan untuk mendaftarkan jalur, peran terkait layanan dan kebijakan IAM baru dibuat atas nama Anda. Lake Formation menambahkan jalur ke kebijakan inline dan menempelkannya ke peran terkait layanan. Saat Anda mendaftarkan jalur berikutnya dengan peran terkait layanan, Lake Formation menambahkan jalur ke kebijakan yang ada.

Saat masuk sebagai administrator danau data, daftarkan lokasi danau data. Kemudian, di konsol IAM, cari peran `AWSServiceRoleForLakeFormationDataAccess` dan lihat kebijakan terlampirnya.

Misalnya, setelah Anda mendaftarkan lokasi`s3://my-kinesis-test/logs`, Lake Formation membuat kebijakan inline berikut dan melampirkannya. `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"
            ]
        }
    ]
}
```

Membuat peran terkait layanan untuk Lake Formation

Anda tidak perlu membuat peran terkait layanan secara manual. Saat Anda mendaftarkan lokasi Amazon S3 dengan Lake Formation di Konsol Manajemen AWS, the, atau AWS API AWS CLI, Lake Formation membuat peran terkait layanan untuk Anda.

> **Important**
>
> Peran tertaut layanan ini dapat muncul di akun Anda jika Anda menyelesaikan tindakan di layanan lain yang menggunakan fitur yang disupport oleh peran ini. Untuk mempelajari lebih lanjut, lihat Peran baru muncul di akun IAM saya.

Jika Anda menghapus peran terkait layanan ini, dan ingin membuatnya lagi, Anda dapat mengulangi proses yang sama untuk membuat kembali peran tersebut di akun Anda. Saat Anda mendaftarkan

lokasi Amazon S3 dengan Lake Formation, Lake Formation membuat peran terkait layanan untuk Anda lagi.

Anda juga dapat menggunakan konsol IAM untuk membuat peran terkait layanan dengan kasus penggunaan Lake Formation. Di AWS CLI atau AWS API, buat peran terkait layanan dengan nama `lakeformation.amazonaws.com` layanan. Untuk informasi selengkapnya, lihat Membuat peran tertaut layanan dalam Panduan Pengguna IAM. Jika Anda menghapus peran tertaut layanan ini, Anda dapat mengulang proses yang sama untuk membuat peran tersebut lagi.

Mengedit peran terkait layanan untuk Lake Formation

Lake Formation tidak memungkinkan Anda untuk mengedit peran `AWSServiceRoleForLakeFormationDataAccess` terkait layanan. Setelah membuat peran terkait layanan, Anda tidak dapat mengubah nama peran karena berbagai entitas mungkin merujuk peran tersebut. Namun, Anda dapat mengedit penjelasan peran menggunakan IAM. Untuk informasi selengkapnya, lihat Mengedit Peran Tertaut Layanan dalam Panduan Pengguna IAM.

Menghapus peran terkait layanan untuk Lake Formation

Jika Anda tidak perlu lagi menggunakan fitur atau layanan yang memerlukan peran terkait layanan, sebaiknya hapus peran tersebut. Dengan begitu, Anda tidak memiliki entitas yang tidak digunakan yang tidak dipantau atau dipelihara secara aktif. Tetapi, Anda harus membersihkan sumber daya peran yang terhubung dengan layanan sebelum menghapusnya secara manual.

> **Note**
>
> Jika layanan Lake Formation menggunakan peran saat Anda mencoba menghapus sumber daya, maka penghapusan mungkin gagal. Jika hal itu terjadi, tunggu beberapa menit dan coba mengoperasikannya lagi.

Untuk menghapus sumber daya Lake Formation yang digunakan oleh Lake Formation

- Jika Anda telah menggunakan peran terkait layanan untuk mendaftarkan lokasi Amazon S3 dengan Lake Formation, sebelum menghapus peran terkait layanan, Anda harus membatalkan pendaftaran lokasi dan mendaftarkannya kembali menggunakan peran khusus.

Untuk menghapus peran tertaut layanan secara manual menggunakan IAM

Gunakan konsol IAM, the AWS CLI, atau AWS API untuk menghapus peran `AWSServiceRoleForLakeFormationDataAccess` terkait layanan. Untuk informasi lebih lanjut, lihat Menghapus Peran Tertaut Layanan dalam Panduan Pengguna IAM.

Berikut ini adalah persyaratan untuk peran yang ditentukan pengguna:

- Saat membuat peran baru, pada halaman Buat peran konsol IAM, pilih AWS layanan, lalu di bawah Pilih kasus penggunaan, pilih Lake Formation.

  Jika Anda membuat peran menggunakan jalur yang berbeda, pastikan bahwa peran tersebut memiliki hubungan kepercayaan dengan`lakeformation.amazonaws.com`. Untuk informasi selengkapnya, lihat Memodifikasi kebijakan kepercayaan peran (Konsol).
- Peran harus memiliki kebijakan sebaris yang memberikan izin Amazon read/write S3 di lokasi. Berikut ini adalah kebijakan yang khas.

  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"
            ]
        }
    ]
}
```

- Tambahkan kebijakan kepercayaan berikut ke peran IAM untuk memungkinkan layanan Lake Formation mengambil peran dan menjual kredensi sementara ke mesin analitik terintegrasi.

  Untuk menyertakan konteks pengguna Pusat Identitas IAM dalam CloudTrail log, kebijakan kepercayaan harus memiliki izin untuk `sts:SetContext` tindakan tersebut.

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "DataCatalogViewDefinerAssumeRole1",
            "Effect": "Allow",
            "Principal": {
               "Service": [
                    "lakeformation.amazonaws.com"
                ]
            },
            "Action": [
                "sts:AssumeRole",
                "sts:SetContext"
            ]
        }
    ]
}
```

- Administrator data lake yang mendaftarkan lokasi harus memiliki `iam:PassRole` izin pada peran tersebut.

  Berikut ini adalah kebijakan inline yang memberikan izin ini. Ganti *<account-id>* dengan nomor AWS akun yang valid, dan ganti *<role-name>* dengan nama peran.

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "PassRolePermissions",
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
```

```
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/<role-name>"
            ]
        }
    ]
}
```

- Untuk mengizinkan Lake Formation menambahkan CloudWatch log di Log dan menerbitkan metrik, tambahkan kebijakan sebaris berikut.

> **Note**
>
> Menulis ke CloudWatch Log menimbulkan biaya.

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:*"
            ]
        }
    ]
}
```

Mendaftarkan lokasi Amazon S3

Anda harus menentukan peran AWS Identity and Access Management (IAM) saat mendaftarkan lokasi Amazon Simple Storage Service (Amazon S3). Lake Formation mengasumsikan peran itu ketika memberikan kredensi sementara untuk AWS layanan terintegrasi yang mengakses data di lokasi tersebut.

> **Important**
>
> Hindari mendaftarkan bucket Amazon S3 yang mengaktifkan Requester pay. Untuk ember yang terdaftar di Lake Formation, peran yang digunakan untuk mendaftarkan ember selalu dipandang sebagai pemohon. Jika bucket diakses oleh AWS akun lain, pemilik bucket akan dikenakan biaya untuk akses data jika peran tersebut milik akun yang sama dengan pemilik bucket.

Anda dapat menggunakan AWS Lake Formation konsol, Lake Formation API, atau AWS Command Line Interface (AWS CLI) untuk mendaftarkan lokasi Amazon S3.

Sebelum Anda mulai

Tinjau persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.

Untuk mendaftarkan lokasi (konsol)

> **Important**
>
> Prosedur berikut mengasumsikan bahwa lokasi Amazon S3 berada di AWS akun yang sama dengan Katalog Data dan bahwa data di lokasi tidak dienkripsi. Bagian lain dalam Bab ini mencakup pendaftaran lintas akun dan pendaftaran lokasi terenkripsi.

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna dengan izin `lakeformation:RegisterResource` IAM.
2. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
3. Pilih Daftarkan lokasi, lalu pilih Browse untuk memilih jalur Amazon Simple Storage Service (Amazon S3).

4. (Opsional, tetapi sangat disarankan) Pilih Tinjau izin lokasi untuk melihat daftar semua sumber daya yang ada di lokasi Amazon S3 yang dipilih dan izinnya.

   Mendaftarkan lokasi yang dipilih dapat mengakibatkan pengguna Lake Formation Anda mendapatkan akses ke data yang sudah ada di lokasi tersebut. Melihat daftar ini membantu Anda memastikan bahwa data yang ada tetap aman.

5. Untuk peran IAM, pilih peran `AWSServiceRoleForLakeFormationDataAccess` terkait layanan (default) atau peran IAM khusus yang memenuhi persyaratan di. the section called "Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi"

   Anda dapat memperbarui lokasi terdaftar atau detail lainnya hanya ketika Anda mendaftarkannya menggunakan peran IAM kustom. Untuk mengedit lokasi yang terdaftar menggunakan peran terkait layanan, Anda harus membatalkan pendaftaran lokasi dan mendaftarkannya kembali.

6. Pilih opsi Aktifkan Federasi Katalog Data untuk mengizinkan Lake Formation mengambil peran dan menjual kredensi sementara ke AWS layanan terintegrasi untuk mengakses tabel di bawah database federasi. Jika lokasi terdaftar di Lake Formation, dan Anda ingin menggunakan lokasi yang sama untuk tabel di bawah database federasi, Anda harus mendaftarkan lokasi yang sama dengan opsi Aktifkan Federasi Katalog Data.

7. Pilih mode akses Hybrid untuk tidak mengaktifkan izin Lake Formation secara default. Saat mendaftarkan lokasi Amazon S3 dalam mode akses hybrid, Anda dapat mengaktifkan izin Lake Formation dengan memilih prinsipal untuk database dan tabel di bawah lokasi tersebut.

   Untuk informasi selengkapnya tentang pengaturan mode akses hybrid, lihatMode akses hibrid.

8. Pilih jalur Daftarkan S3 dengan Akun Pemilik dan berikan akun pemilik bucket S3 untuk mengizinkan Lake Formation memverifikasi akses peran pendaftaran di lokasi Amazon S3 dan mengaktifkan Lake Formation untuk menjual kredensi sementara untuk file Amazon S3 di bawah lokasi terdaftar.

9. Pilih Daftarkan lokasi.

Untuk mendaftarkan lokasi (AWS CLI)

1. Daftarkan lokasi baru dengan Lake Formation

   Contoh ini menggunakan peran terkait layanan untuk mendaftarkan lokasi. Anda dapat menggunakan `--role-arn` argumen sebagai gantinya untuk menyediakan peran Anda sendiri.

Ganti *<s3-path>* dengan jalur Amazon S3 yang valid, nomor akun dengan AWS akun yang valid, dan *<s3-access-role>* dengan peran IAM yang memiliki izin untuk mendaftarkan lokasi data.

> **Note**
>
> Anda tidak dapat mengedit properti lokasi terdaftar jika terdaftar menggunakan peran terkait layanan.

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --use-service-linked-role
```

Contoh berikut menggunakan peran kustom untuk mendaftarkan lokasi.

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>
```

2. Untuk memperbarui lokasi yang terdaftar di Lake Formation

Anda dapat mengedit lokasi terdaftar hanya jika terdaftar menggunakan peran IAM kustom. Untuk lokasi yang terdaftar dengan peran terkait layanan, Anda harus membatalkan pendaftaran lokasi dan mendaftarkannya lagi. Untuk informasi selengkapnya, lihat the section called “Membatalkan pendaftaran lokasi 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. Daftarkan lokasi data dalam mode akses hybrid dengan federasi

```
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
```

4. Daftarkan lokasi dengan akun pemilik yang diharapkan

   Gunakan `--expected-resource-owner-account` bendera untuk mengizinkan Lake Formation memverifikasi akses peran pendaftaran di lokasi Amazon S3 dan menjual kredensi sementara untuk file Amazon S3 di bawah lokasi terdaftar.

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \
 --expected-resource-owner-account <123456789012>
```

```
aws lakeformation update-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role> \
 --expected-resource-owner-account <123456789012>
```

Untuk informasi selengkapnya, lihat Operasi RegisterResourceAPI.

> **Note**
>
> Setelah Anda mendaftarkan lokasi Amazon S3, AWS Glue tabel apa pun yang menunjuk ke lokasi (atau lokasi turunannya) akan mengembalikan nilai untuk `IsRegisteredWithLakeFormation` parameter seperti `true` dalam panggilan. `GetTable` Ada batasan yang diketahui bahwa operasi API Katalog Data seperti `GetTables` dan `SearchTables` tidak memperbarui nilai untuk `IsRegisteredWithLakeFormation` parameter, dan mengembalikan default, yang salah. Disarankan untuk menggunakan `GetTable` API untuk melihat nilai yang benar untuk `IsRegisteredWithLakeFormation` parameter.

Mendaftarkan lokasi Amazon S3 terenkripsi

Lake Formation terintegrasi dengan AWS Key Management Service(AWS KMS) untuk memungkinkan Anda mengatur layanan terintegrasi lainnya dengan lebih mudah untuk mengenkripsi dan mendekripsi data di lokasi Amazon Simple Storage Service (Amazon S3).

Baik pelanggan dikelola AWS KMS keys dan Kunci yang dikelola AWS didukung. Saat ini, sisi klien hanya encryption/decryption didukung dengan Athena.

Anda harus menentukan peran AWS Identity and Access Management (IAM) saat mendaftarkan lokasi Amazon S3. Untuk lokasi Amazon S3 terenkripsi, peran harus memiliki izin untuk mengenkripsi dan mendekripsi data dengan AWS KMS key, atau kebijakan kunci KMS harus memberikan izin pada kunci peran tersebut.

> **Important**
>
> Hindari mendaftarkan bucket Amazon S3 yang mengaktifkan Requester pay. Untuk ember yang terdaftar di Lake Formation, peran yang digunakan untuk mendaftarkan ember selalu dipandang sebagai pemohon. Jika bucket diakses oleh AWS akun lain, pemilik bucket akan dikenakan biaya untuk akses data jika peran tersebut milik akun yang sama dengan pemilik bucket.

Lake Formation menggunakan peran terkait layanan untuk mendaftarkan lokasi data Anda. Namun, peran ini memiliki beberapa keterbatasan. Karena kendala ini, kami sarankan untuk membuat dan menggunakan peran IAM khusus sebagai gantinya untuk lebih banyak fleksibilitas dan kontrol. Peran

kustom yang Anda buat untuk mendaftarkan lokasi harus memenuhi persyaratan yang ditentukan dalamthe section called “Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi”.

> **Important**
>
> Jika Anda menggunakan Kunci yang dikelola AWS untuk mengenkripsi lokasi Amazon S3, Anda tidak dapat menggunakan peran terkait layanan Lake Formation. Anda harus menggunakan peran khusus dan menambahkan izin IAM pada kunci peran. Detail diberikan nanti di bagian ini.

Prosedur berikut menjelaskan cara mendaftarkan lokasi Amazon S3 yang dienkripsi dengan kunci yang dikelola pelanggan atau. Kunci yang dikelola AWS

- Mendaftarkan lokasi yang dienkripsi dengan kunci yang dikelola pelanggan
- Mendaftarkan lokasi yang dienkripsi dengan Kunci yang dikelola AWS

Sebelum Anda Memulai

Tinjau persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.

Untuk mendaftarkan lokasi Amazon S3 yang dienkripsi dengan kunci yang dikelola pelanggan

> **Note**
>
> Jika kunci KMS atau lokasi Amazon S3 tidak berada di akun yang AWS sama dengan Katalog Data, ikuti petunjuknya the section called “Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening” sebagai gantinya.

1. Buka AWS KMS konsol di https://console.aws.amazon.com/kmsdan masuk sebagai pengguna administratif AWS Identity and Access Management (IAM) atau sebagai pengguna yang dapat memodifikasi kebijakan kunci kunci KMS yang digunakan untuk mengenkripsi lokasi.
2. Di panel navigasi, pilih Kunci yang dikelola pelanggan, lalu pilih nama kunci KMS yang diinginkan.
3. Pada halaman detail kunci KMS, pilih tab Kebijakan kunci, lalu lakukan salah satu hal berikut untuk menambahkan peran kustom Anda atau peran terkait layanan Lake Formation sebagai pengguna kunci KMS:

- Jika tampilan default ditampilkan (dengan administrator Kunci, Penghapusan kunci, Pengguna kunci, dan bagian AWS Akun lainnya) — Di bawah bagian Pengguna kunci, tambahkan peran kustom Anda atau peran terkait layanan Lake Formation. `AWSServiceRoleForLakeFormationDataAccess`
- Jika kebijakan kunci (JSON) ditampilkan — Edit kebijakan untuk menambahkan peran kustom Anda atau peran terkait layanan Lake Formation `AWSServiceRoleForLakeFormationDataAccess` ke objek “Izinkan penggunaan kunci,” seperti yang ditunjukkan pada contoh berikut.

> **Note**
>
> Jika objek itu hilang, tambahkan dengan izin yang ditunjukkan dalam contoh. Contoh menggunakan peran terkait layanan.

```
        ...
        {
            "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. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna dengan izin `lakeformation:RegisterResource` IAM.
5. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
6. Pilih Daftarkan lokasi, lalu pilih Browse untuk memilih jalur Amazon Simple Storage Service (Amazon S3).
7. (Opsional, tetapi sangat disarankan) Pilih Tinjau izin lokasi untuk melihat daftar semua sumber daya yang ada di lokasi Amazon S3 yang dipilih dan izinnya.

   Mendaftarkan lokasi yang dipilih dapat mengakibatkan pengguna Lake Formation Anda mendapatkan akses ke data yang sudah ada di lokasi tersebut. Melihat daftar ini membantu Anda memastikan bahwa data yang ada tetap aman.
8. Untuk peran IAM, pilih peran `AWSServiceRoleForLakeFormationDataAccess` terkait layanan (default) atau peran kustom Anda yang memenuhi. the section called “Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi”
9. Pilih Daftar lokasi.

Untuk informasi selengkapnya tentang peran tertaut layanan, lihat Izin peran terkait layanan untuk Lake Formation.

Untuk mendaftarkan lokasi Amazon S3 yang dienkripsi dengan Kunci yang dikelola AWS

> **Important**
>
> Jika lokasi Amazon S3 tidak berada di AWS akun yang sama dengan Katalog Data, ikuti petunjuknyathe section called “Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening”.

1. Buat peran IAM yang akan digunakan untuk mendaftarkan lokasi. Pastikan memenuhi persyaratan yang tercantum dithe section called “Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi”.
2. Tambahkan kebijakan inline berikut ke peran. Ini memberikan izin pada kunci peran. `Resource`Spesifikasi harus menunjuk Nama Sumber Daya Amazon (ARN) dari. Kunci yang dikelola AWS Anda dapat memperoleh ARN dari konsol. AWS KMS Untuk mendapatkan ARN

yang benar, pastikan Anda masuk ke AWS KMS konsol dengan AWS akun dan Wilayah yang sama dengan Kunci yang dikelola AWS yang digunakan untuk mengenkripsi lokasi.

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"
    }
  ]
}
```

Anda dapat menggunakan alias kunci KMS alih-alih ID kunci - `arn:aws:kms:region:account-id:key/alias/your-key-alias`

Untuk informasi selengkapnya, lihat Alias di AWS KMS bagian Panduan AWS Key Management Service Pengembang.

3. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna dengan izin `lakeformation:RegisterResource` IAM.
4. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
5. Pilih Daftarkan lokasi, lalu pilih Jelajahi untuk memilih jalur Amazon S3.
6. (Opsional, tetapi sangat disarankan) Pilih Tinjau izin lokasi untuk melihat daftar semua sumber daya yang ada di lokasi Amazon S3 yang dipilih dan izinnya.

Mendaftarkan lokasi yang dipilih dapat mengakibatkan pengguna Lake Formation Anda mendapatkan akses ke data yang sudah ada di lokasi tersebut. Melihat daftar ini membantu Anda memastikan bahwa data yang ada tetap aman.

7. Untuk peran IAM, pilih peran yang Anda buat di Langkah 1.
8. Pilih Daftar lokasi.

Mendaftarkan lokasi Amazon S3 di tempat lain AWS akun

AWS Lake Formation memungkinkan Anda mendaftarkan lokasi Amazon Simple Storage Service (Amazon S3) di seluruh akun. AWS Misalnya, jika AWS Glue Data Catalog ada di akun A, pengguna di akun A dapat mendaftarkan bucket Amazon S3 di akun B.

Mendaftarkan bucket Amazon S3 di AWS akun B menggunakan peran AWS Identity and Access Management (IAM) di AWS akun A memerlukan izin berikut:

- Peran di akun A harus memberikan izin pada bucket di akun B.
- Kebijakan bucket di akun B harus memberikan izin akses ke peran di Akun A.

> **Important**
>
> Hindari mendaftarkan bucket Amazon S3 yang mengaktifkan Requester pay. Untuk ember yang terdaftar di Lake Formation, peran yang digunakan untuk mendaftarkan ember selalu dipandang sebagai pemohon. Jika bucket diakses oleh AWS akun lain, pemilik bucket akan dikenakan biaya untuk akses data jika peran tersebut milik akun yang sama dengan pemilik bucket.
>
> Anda tidak dapat menggunakan peran terkait layanan Lake Formation untuk mendaftarkan lokasi di akun lain. Anda harus menggunakan peran yang ditentukan pengguna sebagai gantinya. Peran tersebut harus memenuhi persyaratan dithe section called “Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi”. Untuk informasi selengkapnya tentang peran tertaut layanan, lihat Izin peran terkait layanan untuk Lake Formation.

Sebelum Anda mulai

Tinjau persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.

Untuk mendaftarkan lokasi di tempat lain AWS akun

> **Note**
>
> Jika lokasi dienkripsi, ikuti instruksi sebagai gantinya. the section called “Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening”

Prosedur berikut mengasumsikan bahwa prinsipal di akun 1111-2222-3333, yang berisi Katalog Data, ingin mendaftarkan bucket Amazon S3, yang ada di akun 1234-5678-9012. `awsexamplebucket1`

1. Di akun 1111-2222-3333, masuk ke Konsol Manajemen AWS dan buka konsol IAM di. https://console.aws.amazon.com/iam/
2. Buat peran baru atau lihat peran yang ada yang memenuhi persyaratan dithe section called “Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi”. Pastikan peran tersebut memberikan izin Amazon S3 aktif. `awsexamplebucket1`
3. Buka konsol Amazon S3 di. https://console.aws.amazon.com/s3/ Masuk dengan akun 1234-5678-9012.
4. Dalam daftar nama Bucket, pilih nama bucket,`awsexamplebucket1`.
5. Pilih Izin.
6. Pada halaman Izin, pilih Kebijakan Bucket.
7. Di editor kebijakan Bucket, tempel kebijakan berikut. Ganti *<role-name>* dengan nama peran Anda.

   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. Pilih Simpan.
9. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk ke akun 1111-2222-3333 sebagai administrator danau data atau sebagai pengguna dengan izin yang cukup untuk mendaftarkan lokasi.
10. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
11. Pada halaman lokasi danau Data, pilih Daftar lokasi.
12. Pada halaman Daftar lokasi, untuk jalur Amazon S3, masukkan nama bucket. `s3://awsexamplebucket1`

> **Note**
>
> Anda harus mengetikkan nama bucket karena bucket lintas akun tidak muncul dalam daftar saat memilih Browse.

13. Untuk peran IAM, pilih peran Anda.
14. Pilih Daftar lokasi.

Mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening

AWS Lake Formation terintegrasi dengan AWS Key Management Service(AWS KMS) untuk memungkinkan Anda mengatur layanan terintegrasi lainnya dengan lebih mudah untuk mengenkripsi dan mendekripsi data di lokasi Amazon Simple Storage Service (Amazon S3).

Kedua kunci yang dikelola pelanggan dan Kunci yang dikelola AWS didukung. Client-side encryption/decryption tidak didukung.

> **Important**
>
> Hindari mendaftarkan bucket Amazon S3 yang mengaktifkan Requester pay. Untuk ember yang terdaftar di Lake Formation, peran yang digunakan untuk mendaftarkan ember selalu dipandang sebagai pemohon. Jika bucket diakses oleh AWS akun lain, pemilik bucket akan dikenakan biaya untuk akses data jika peran tersebut milik akun yang sama dengan pemilik bucket.

Bagian ini menjelaskan cara mendaftarkan lokasi Amazon S3 dalam keadaan berikut:

- Data di lokasi Amazon S3 dienkripsi dengan kunci KMS yang dibuat di. AWS KMS
- Lokasi Amazon S3 tidak berada di AWS akun yang sama dengan. AWS Glue Data Catalog
- Kunci KMS berada atau tidak berada di AWS akun yang sama dengan Katalog Data.

Mendaftarkan AWS KMS bucket Amazon S3 yang dienkripsi di AWS akun B menggunakan peran AWS Identity and Access Management (IAM) di AWS akun A memerlukan izin berikut:

- Peran di akun A harus memberikan izin pada bucket di akun B.
- Kebijakan bucket di akun B harus memberikan izin akses ke peran di Akun A.
- Jika kunci KMS ada di akun B, kebijakan kunci harus memberikan akses ke peran di akun A, dan peran dalam akun A harus memberikan izin pada kunci KMS.

Dalam prosedur berikut, Anda membuat peran dalam AWS akun yang berisi Katalog Data (akun A dalam diskusi sebelumnya). Kemudian, Anda menggunakan peran ini untuk mendaftarkan lokasi. Lake Formation mengasumsikan peran ini saat mengakses data yang mendasarinya di Amazon S3. Peran yang diasumsikan memiliki izin yang diperlukan pada kunci KMS. Akibatnya, Anda tidak perlu memberikan izin pada kunci KMS kepada kepala sekolah yang mengakses data dasar dengan pekerjaan ETL atau dengan layanan terintegrasi seperti. Amazon Athena

> **Important**
>
> Anda tidak dapat menggunakan peran terkait layanan Lake Formation untuk mendaftarkan lokasi di akun lain. Anda harus menggunakan peran yang ditentukan pengguna sebagai gantinya. Peran tersebut harus memenuhi persyaratan dithe section called “Persyaratan

untuk peran yang digunakan untuk mendaftarkan lokasi". Untuk informasi selengkapnya tentang peran tertaut layanan, lihat Izin peran terkait layanan untuk Lake Formation.

Sebelum Anda Memulai

Tinjau persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.

Untuk mendaftarkan lokasi Amazon S3 terenkripsi di seluruh AWS rekening

1. Di AWS akun yang sama dengan Katalog Data, masuk ke Konsol Manajemen AWS dan buka konsol IAM dihttps://console.aws.amazon.com/iam/.
2. Buat peran baru atau lihat peran yang ada yang memenuhi persyaratan dithe section called "Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi". Pastikan peran tersebut menyertakan kebijakan yang memberikan izin Amazon S3 di lokasi.
3. Jika kunci KMS tidak berada di akun yang sama dengan Katalog Data, tambahkan ke peran kebijakan inline yang memberikan izin yang diperlukan pada kunci KMS. Berikut ini adalah contoh kebijakan . Ganti Wilayah dan ID akun dengan wilayah dan nomor akun kunci KMS. Ganti *<key-id>* dengan ID kunci.

   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. Di konsol Amazon S3, tambahkan kebijakan bucket yang memberikan izin Amazon S3 yang diperlukan ke peran tersebut. Berikut ini adalah contoh kebijakan bucket. Ganti ID AWS akun dengan nomor akun Katalog Data, *<role-name>* dengan nama peran Anda, dan *<bucket-name>* dengan nama 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. Di AWS KMS, tambahkan peran sebagai pengguna kunci KMS.

   a. Buka AWS KMS konsol di https://console.aws.amazon.com/kms. Kemudian, masuk sebagai pengguna administrator atau sebagai pengguna yang dapat mengubah kebijakan kunci kunci KMS yang digunakan untuk mengenkripsi lokasi.

   b. Di panel navigasi, pilih Kunci yang dikelola pelanggan, lalu pilih nama kunci KMS.

   c. Pada halaman detail kunci KMS, di bawah tab Kebijakan kunci, jika tampilan JSON dari kebijakan kunci tidak ditampilkan, pilih Beralih ke tampilan kebijakan.

d. Di bagian Kebijakan kunci, pilih Edit, dan tambahkan Nama Sumber Daya Amazon (ARN) peran ke `Allow use of the key` objek, seperti yang ditunjukkan pada contoh berikut.

> **Note**
>
> Jika objek itu hilang, tambahkan dengan izin yang ditunjukkan dalam contoh.

```
...
{
    "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": "*"
},
...
```

Untuk informasi selengkapnya, lihat Mengizinkan Pengguna di Akun Lain Menggunakan kunci KMS di Panduan AWS Key Management Service Pengembang.

6. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk ke AWS akun Katalog Data sebagai administrator danau data.
7. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
8. Pilih Daftar lokasi.
9. Pada halaman Daftar lokasi, untuk jalur Amazon S3, masukkan jalur lokasi sebagai. **s3://*<bucket>*/*<prefix>*** Ganti *<bucket>* dengan nama ember dan *<prefix>* dengan sisa jalur untuk lokasi.

> **Note**
>
> Anda harus mengetik jalur karena bucket lintas akun tidak muncul dalam daftar saat Anda memilih Browse.

10. Untuk peran IAM, pilih peran dari Langkah 2.
11. Pilih Daftar lokasi.

Membatalkan pendaftaran lokasi Amazon S3

Anda dapat membatalkan pendaftaran lokasi Amazon Simple Storage Service (Amazon S3) Simple Storage Service (Amazon S3) jika Anda tidak ingin lagi dikelola oleh Lake Formation. Pencabutan pendaftaran lokasi tidak memengaruhi izin lokasi data Lake Formation yang diberikan pada lokasi tersebut. Anda dapat mendaftarkan ulang lokasi yang dideregistrasi, dan izin lokasi data tetap berlaku. Anda dapat menggunakan peran yang berbeda untuk mendaftarkan ulang lokasi.

Untuk membatalkan pendaftaran lokasi (konsol)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna dengan izin `lakeformation:RegisterResource` IAM.
2. Di panel navigasi, di bawah Administrasi, pilih Lokasi danau data.
3. Pilih lokasi, dan pada menu Tindakan, pilih Hapus.
4. Saat diminta konfirmasi, pilih Hapus.

Mengakses lokasi Amazon S3

AWS Lake Formation memungkinkan Anda membaca dan menulis file data yang mendasarinya di Amazon Simple Storage Service (Amazon S3) untuk tabel yang terdaftar di AWS Glue Data Catalog (Katalog Data). Ini memberi Anda satu set izin untuk kueri SQL dan akses file langsung menggunakan hibah tabel Lake Formation yang ada. Lake Formation memperluas fungsionalitas penjual kredensialnya ke lokasi Amazon S3 yang terdaftar di Katalog Data. Saat aplikasi Anda meminta akses ke file Amazon S3 yang mendasari tabel Katalog Data, Lake Formation mengevaluasi izin tingkat tabel pemanggil yang ada dan, jika diotorisasi, mengembalikan kredensi Amazon S3 yang berumur pendek dan tercakup untuk lokasi terdaftar tabel tersebut.

Aktifkan akses langsung ke lokasi Amazon S3

Untuk mengaktifkan fitur ini, Anda harus menyelesaikan langkah-langkah berikut.

Prasyarat

Lengkapi prasyarat berikut sebelum mengaktifkan flag fitur:

- Tentukan lokasi tabel - Tentukan jalur ke data Anda di Amazon S3 menggunakan properti LOCATION di atribut tabel. Untuk informasi selengkapnya tentang format lokasi, lihat Lokasi tabel di Amazon S3 di Panduan Pengguna Amazon Athena.
- Daftarkan lokasi Amazon S3 — Katalog tabel Anda di Katalog Data untuk lokasi Amazon S3 Anda dan daftarkan ke Lake Formation dengan menyediakan akun pemilik bucket Amazon S3. Saat Anda mendaftarkan lokasi, jalur Amazon S3 dan semua folder di bawah jalur itu terdaftar. Untuk informasi selengkapnya, lihat Mendaftarkan lokasi Amazon S3.
- Berikan izin Lake Formation — Berikan izin Lake Formation `SELECT` (atau`SUPER`) pada tabel kepada ilmuwan data dan aplikasi Anda menggunakan konsol atau API Lake Formation. Jika Anda sudah menggunakan Lake Formation untuk akses tabel dengan layanan seperti Athena atau Amazon EMR, ini sudah dikonfigurasi.
- Aktifkan Integrasi Aplikasi dengan Akses Tabel Penuh - Aktifkan Integrasi Aplikasi dengan Akses Tabel Penuh untuk memungkinkan penjual kredensi ke lokasi tabel terdaftar. Untuk informasi selengkapnya, lihat Integrasi aplikasi untuk akses tabel penuh.

Aktifkan bendera fitur

Aktifkan `fs.s3a.lakeformation.access.grants.enabled` bendera pada mesin yang didukung:

```
fs.s3a.lakeformation.access.grants.enabled = true
```

- Amazon EMR di Amazon EC2 - EMR 7.13 atau yang lebih baru
- Amazon EMR di EKS - EMR 7.13 atau yang lebih baru
- Amazon EMR Tanpa Server - EMR 7.13 atau yang lebih baru

Setelah dikonfigurasi, Anda dapat langsung membaca atau menulis file data dari EMR menggunakan API standar.

Untuk informasi selengkapnya tentang integrasi EMR Amazon, lihat penjual kredensi berbasis jalur Lake Formation di Panduan Manajemen EMR Amazon.

Cara kerjanya

Akses lokasi Danau Formation-based Amazon S3 mengikuti aliran ini:

1. Prinsipal atau peran meminta akses ke file data Amazon S3 melalui layanan seperti pekerjaan Amazon EMR Spark, atau pipeline pemrosesan data.
2. Plugin AWS SDK yang terintegrasi dengan EMR mencegat permintaan Amazon S3 dan memanggil Lake Formation API. GetTemporaryDataLocationCredentials
3. Lake Formation memeriksa apakah jalur Amazon S3 sesuai dengan lokasi terdaftar dan tabel katalog, serta mengevaluasi izin pemanggil pada tabel terkait. AWS Glue
4. Jika pemanggil memiliki `SELECT` atau `SUPER` izin pada tabel yang sesuai dengan akses tabel penuh, Lake Formation memberikan kredensyal sementara yang dicakup ke lokasi Amazon S3 tersebut. `SELECT`memberikan kredensyal BACA; memberikan kredensyal **`SUPER`** READWRITE.
5. Plugin menggunakan kredensyal ini untuk menyelesaikan permintaan Amazon S3, dan Amazon S3 menyediakan data ke aplikasi.
6. Semua operasi penjual kredensyal dicatat CloudTrail, menyediakan jejak audit untuk akses data.

Proses ini transparan bagi Anda dan Anda cukup menggunakan izin Lake Formation yang ada dengan API berbasis file standar.

```
# Read raw data (Lake Formation-based S3 Location access)
raw_df = spark.read.json("s3://finance-datalake/raw/transactions/dt=2024-03-21/")

# Read governed data (Lake Formation-based S3 Location access)
transactions_df = spark.read.parquet("s3://data-lake/transactions/year=2026/")

# Write processed data (Lake Formation-based S3 Location access)
processed_df.write \
    .mode("append") \
    .partitionBy("transaction_date") \
    .parquet("s3://finance-datalake/processed/transactions/")

print("ETL complete. Records written:", processed_df.count())
```

Persyaratan izin

Fitur saat ini menjual kredensyal hanya ketika pemanggil memiliki akses tabel penuh, artinya `SELECT` pada semua kolom tanpa filter baris atau kolom. Jika tabel memiliki filter tingkat baris atau tingkat kolom yang diterapkan, penelepon harus terus menggunakan mesin tepercaya seperti Athena, Amazon EMR, AWS Glue atau Amazon Redshift yang dapat menerapkan filter tersebut. Ini memastikan batas-batas keamanan tetap konsisten.

Cross-account akses

Akses yang disederhanakan Lake Formation ke lokasi Amazon S3 berfungsi dengan kemampuan berbagi lintas akun. Saat Anda membagikan tabel dengan AWS akun lain melalui Lake Formation, penerima dapat mengakses file data Amazon S3 yang mendasarinya menggunakan izin Lake Formation mereka, tunduk pada persyaratan akses tabel lengkap yang sama.

Fitur ini mendukung mekanisme tautan sumber daya Lake Formation untuk akses lintas akun. Jika akun konsumen memiliki izin di tabel bersama, Lake Formation memberikan kredensi yang dicakup ke lokasi Amazon S3 akun terdaftar, memungkinkan akses data lintas akun yang mulus tanpa memerlukan kebijakan bucket Amazon S3 terpisah atau peran IAM lintas akun.

Lokasi Amazon S3 bersarang

Ketika beberapa tabel menunjuk ke lokasi bersarang di bucket yang sama, Lake Formation menerapkan perilaku berikut:

- Saat mengaksess`3://bucket`, Anda menerima izin yang sesuai dengan tabel yang terdaftar di tingkat bucket.
- Saat mengaksess`3://bucket/folder1`, Anda menerima izin yang sesuai dengan tabel yang terdaftar di jalur tertentu.
- Untuk akses ke folder tanpa tabel terdaftar (misalnya,`s3://bucket/folder2`), Anda menerima izin dari lokasi terdaftar induk terdekat.
- Jika beberapa tabel terdaftar di lokasi yang sama, Lake Formation mengembalikan kesalahan karena izin yang bertentangan.

Audit dan kepatuhan

Semua operasi penjual kredensyal dicatat CloudTrail, menyediakan jejak audit untuk akses data. Saat Lake Formation menjual kredensyal melalui `GetTemporaryDataLocationCredentials` API, catat: CloudTrail

- Prinsipal (pengguna atau peran)

- Stempel waktu
- Lokasi Amazon S3
- AWS GlueTabel terkait

Panggilan API Amazon S3 berikutnya yang dilakukan dengan kredensyal tersebut juga dicatat sebagai peristiwa data Amazon CloudTrail S3, dengan konteks yang menautkannya kembali ke hibah izin Lake Formation. Ini memberi auditor visibilitas lengkap tentang siapa yang mengakses data apa, kapan, dan melalui jalur izin mana — semuanya dalam satu aliran log. CloudTrail

Didukung AWS layanan

Akses yang disederhanakan Lake Formation ke lokasi Amazon S3 berfungsi dengan:

- Amazon EMR di Amazon EC2 (EMR 7.13 atau lebih baru)
- Amazon EMR di EKS (EMR 7.13 atau lebih baru)
- Amazon EMR Tanpa Server (EMR 7.13 atau lebih baru)

Plugin open source untuk layanan pihak ketiga

Anda juga dapat mengintegrasikan Trino aplikasi Anda Apache Spark atau menggunakan API atau melalui plugin open source yang disediakan oleh AWS. Untuk informasi selengkapnya, lihat aws-lakeformation-accessgrants-plugin-java-v2 di. GitHub

Pertimbangan-pertimbangan

Perhatikan pertimbangan berikut saat menggunakan penjual kredensyal untuk lokasi Amazon S3:

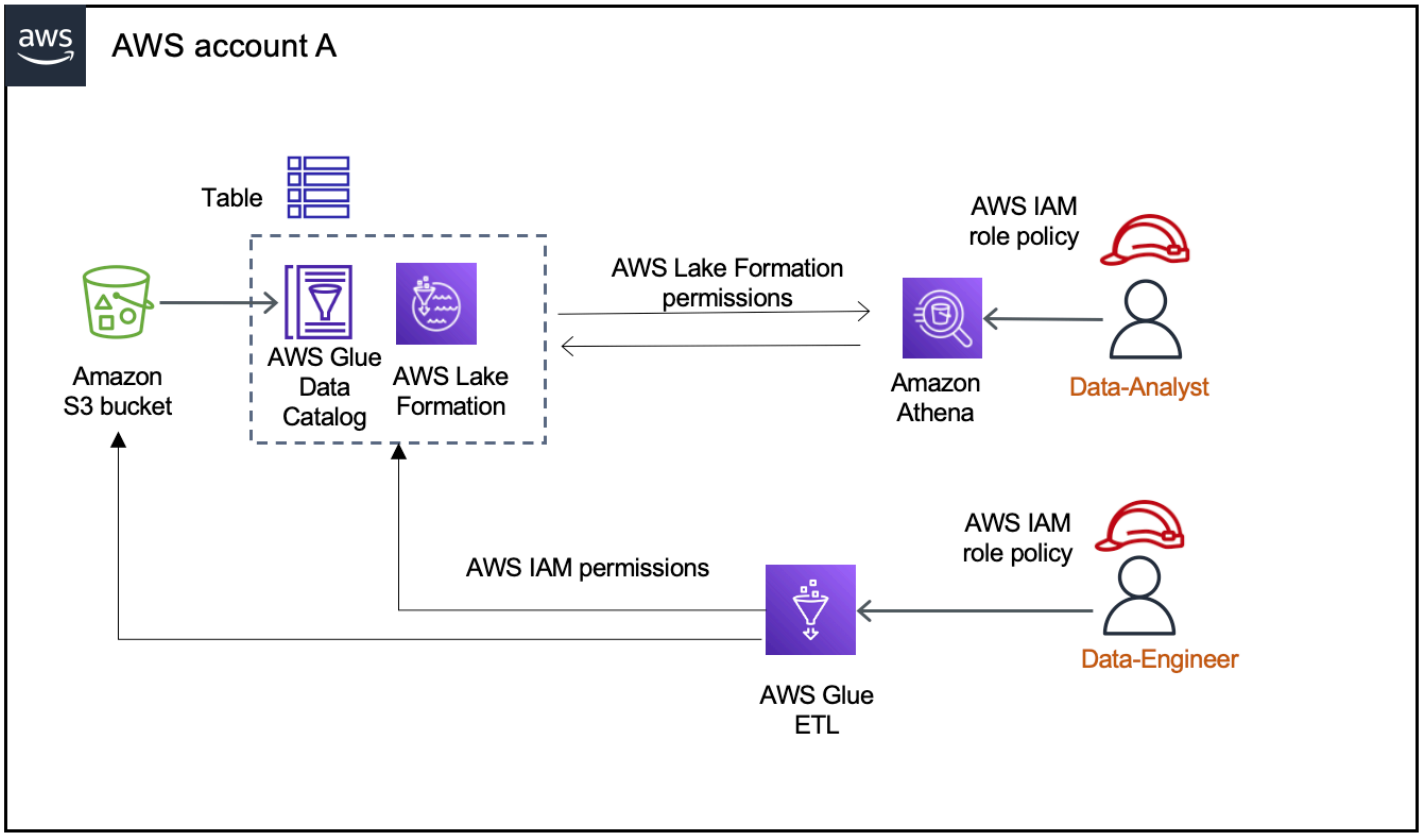
- Penjual kredensyal untuk lokasi Amazon S3 tidak didukung lintas wilayah.
- Penjual kredensyal didukung untuk lokasi Amazon S3 yang disertakan sebagai lokasi utama tabel.
- Plugin mendukungApache Hive,Apache Hudi, dan format Delta Lake tabel. Apache Icebergsaat ini tidak didukung.
- Plugin saat ini tidak didukung dengan mode Amazon EMR Spark Fine-Grained Access Control (FGAC).

Mode akses hibrid

AWS Lake Formation mode akses hybrid mendukung dua jalur izin ke AWS Glue Data Catalog objek yang sama.
Di jalur pertama, Lake Formation memungkinkan Anda memilih prinsip tertentu, dan memberi mereka izin Lake Formation untuk mengakses katalog, database, tabel, dan tampilan dengan ikut serta. Jalur kedua memungkinkan semua prinsipal lain untuk mengakses sumber daya ini melalui kebijakan utama IAM default untuk Amazon S3 dan tindakan. AWS Glue

Saat mendaftarkan lokasi Amazon S3 dengan Lake Formation, Anda memiliki opsi untuk menerapkan izin Lake Formation untuk semua sumber daya di lokasi ini atau menggunakan mode akses hybrid. Mode akses hybrid hanya memberlakukan`CREATE_TABLE`,`CREATE_PARTITION`, `UPDATE_TABLE` izin secara default. Saat lokasi Amazon S3 berada dalam mode hibrida, Anda dapat mengaktifkan izin Lake Formation dengan memilih prinsipal untuk objek Katalog Data di bawah lokasi tersebut. Ini berarti izin Lake Formation dan izin IAM dapat mengontrol akses ke data tersebut. Ini berarti bahwa memilih prinsipal akan memerlukan izin Lake Formation dan izin IAM untuk mengakses data, sementara prinsipal yang tidak ikut serta akan terus mengakses data hanya dengan menggunakan izin IAM.

Dengan demikian, mode akses hybrid memberikan fleksibilitas untuk mengaktifkan Lake Formation secara selektif untuk katalog, database, dan tabel di Katalog Data Anda untuk kumpulan pengguna tertentu tanpa mengganggu akses untuk pengguna atau beban kerja lain yang ada.



Untuk pertimbangan dan batasan, lihat Pertimbangan dan batasan mode akses hibrida.

Istilah dan definisi

Berikut adalah definisi sumber daya Katalog Data berdasarkan cara Anda mengatur izin akses:

Sumber daya Lake Formation

Sumber daya yang terdaftar di Lake Formation. Pengguna memerlukan izin Lake Formation untuk mengakses sumber daya.

AWS Glue sumber daya

Sumber daya yang tidak terdaftar di Lake Formation. Pengguna hanya memerlukan izin IAM untuk mengakses sumber daya karena memiliki izin `IAMAllowedPrincipals` grup. Izin Lake Formation tidak diberlakukan.

Untuk informasi selengkapnya tentang izin `IAMAllowedPrincipals` grup, lihatIzin metadata.

Sumber daya hibrida

Sumber daya yang terdaftar dalam mode akses hybrid. Berdasarkan pengguna yang mengakses sumber daya, sumber daya secara dinamis beralih antara menjadi sumber daya Lake Formation atau sumber daya. AWS Glue

Kasus penggunaan mode akses hybrid umum

Anda dapat menggunakan mode akses hybrid untuk menyediakan akses dalam skenario berbagi data akun tunggal dan lintas akun:

Skenario akun tunggal

- Konversikan AWS Glue sumber daya menjadi sumber daya hibrida — Dalam skenario ini, saat ini Anda tidak menggunakan Lake Formation tetapi ingin mengadopsi izin Lake Formation untuk objek Katalog Data. Saat mendaftarkan lokasi Amazon S3 dalam mode akses hybrid, Anda dapat memberikan izin Lake Formation kepada pengguna yang memilih database dan tabel tertentu yang menunjuk ke lokasi tersebut.
- Mengonversi sumber daya Lake Formation menjadi sumber daya hibrida — Saat ini, Anda menggunakan izin Lake Formation untuk mengontrol akses database Katalog Data tetapi ingin memberikan akses ke prinsipal baru menggunakan izin IAM untuk Amazon S3 dan tanpa AWS Glue mengganggu izin Lake Formation yang ada.

  Saat Anda memperbarui pendaftaran lokasi data ke mode akses hibrid, prinsipal baru dapat mengakses database Katalog Data yang menunjuk lokasi Amazon S3 menggunakan kebijakan izin IAM tanpa mengganggu izin Lake Formation pengguna yang ada.

  Sebelum memperbarui pendaftaran lokasi data untuk mengaktifkan mode akses hibrida, Anda harus terlebih dahulu memilih prinsipal yang saat ini mengakses sumber daya dengan izin Lake Formation.
  Ini untuk mencegah potensi gangguan pada alur kerja saat ini.
  Anda juga perlu memberikan `Super` izin pada tabel dalam database ke `IAMAllowedPrincipal` grup.

Cross-account skenario berbagi data

- Bagikan AWS Glue sumber daya menggunakan mode akses hibrid — Dalam skenario ini, akun produsen memiliki tabel dalam database yang saat ini dibagikan dengan akun konsumen

menggunakan kebijakan izin IAM untuk Amazon AWS Glue S3 dan tindakan. Lokasi data database tidak terdaftar di Lake Formation.

Sebelum mendaftarkan lokasi data dalam mode akses hybrid, Anda perlu memperbarui pengaturan versi Cross account ke versi 4. Versi 4 menyediakan kebijakan AWS RAM izin baru yang diperlukan untuk berbagi lintas akun ketika `IAMAllowedPrincipal` grup memiliki `Super` izin pada sumber daya. Untuk sumber daya dengan izin `IAMAllowedPrincipal` grup, Anda dapat memberikan izin Lake Formation ke akun eksternal dan memilihnya untuk menggunakan izin Lake Formation. Administrator data lake di akun penerima dapat memberikan izin Lake Formation kepada kepala sekolah di akun dan memilihnya untuk menerapkan izin Lake Formation.

- Bagikan sumber daya Lake Formation menggunakan mode akses hybrid — Saat ini, akun produsen memiliki tabel dalam database yang dibagikan dengan akun konsumen yang memberlakukan izin Lake Formation. Lokasi data database terdaftar di Lake Formation.

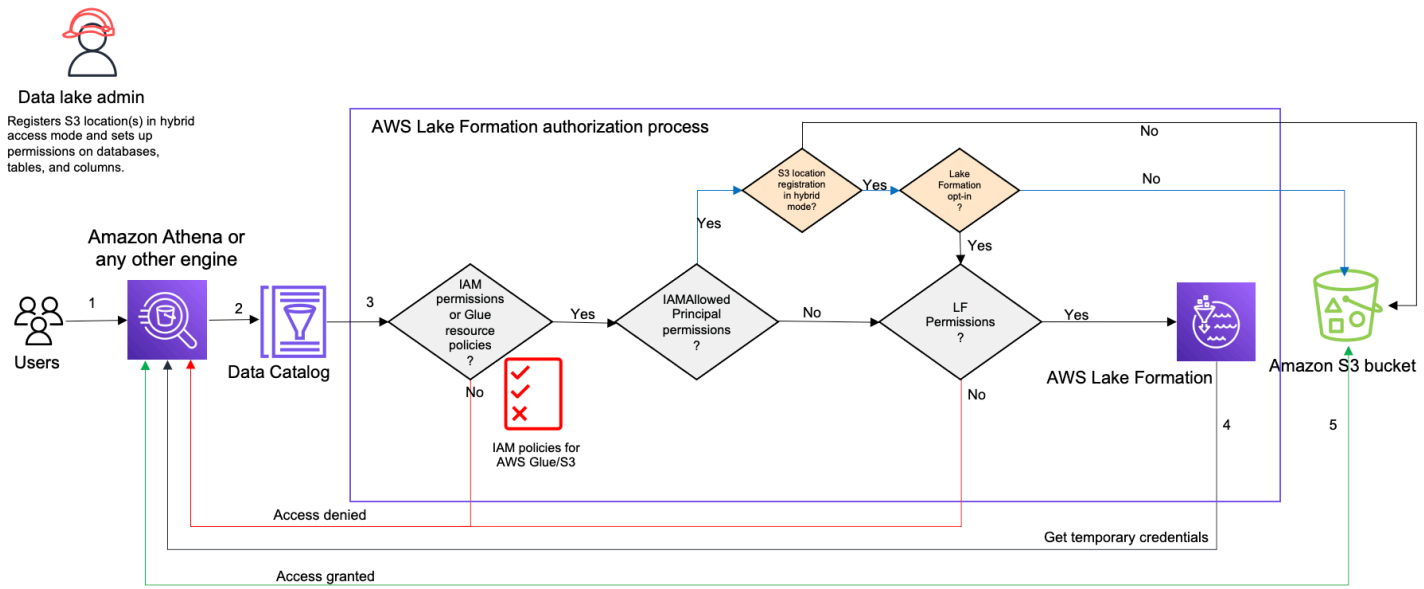
  Dalam hal ini, Anda dapat memperbarui pendaftaran lokasi Amazon S3 ke mode akses hibrid, dan membagikan data dari Amazon S3 dan metadata dari Katalog Data menggunakan kebijakan bucket Amazon S3 dan kebijakan sumber daya Katalog Data ke prinsipal di akun konsumen. Anda perlu memberikan kembali izin Lake Formation yang ada dan memilih prinsipal sebelum memperbarui pendaftaran lokasi Amazon S3. Juga, Anda perlu memberikan `Super` izin pada tabel dalam database ke `IAMAllowedPrincipals` grup.

Topik

- Cara kerja mode akses hybrid
- Menyiapkan mode akses hybrid - skenario umum
- Menghapus prinsip dan sumber daya dari mode akses hybrid
- Melihat prinsip dan sumber daya dalam mode akses hybrid
- Sumber daya tambahan

Cara kerja mode akses hybrid

Diagram berikut menunjukkan cara kerja otorisasi Lake Formation dalam mode akses hibrid saat Anda menanyakan sumber daya Katalog Data.



Sebelum mengakses data di data lake Anda, administrator data lake atau pengguna dengan izin administratif menyiapkan kebijakan pengguna tabel Katalog Data individual untuk mengizinkan atau menolak akses ke tabel di Katalog Data Anda. Kemudian, kepala sekolah yang memiliki izin untuk melakukan `RegisterResource` operasi mendaftarkan lokasi Amazon S3 dari tabel dengan Lake Formation dalam mode akses hybrid. Jika lokasi data tidak terdaftar di Lake Formation, administrator memberikan izin Lake Formation kepada pengguna tertentu pada database dan tabel Katalog Data dan memilihnya untuk menggunakan izin Lake Formation untuk database dan tabel tersebut dalam mode akses hibrid.

1. Mengirimkan kueri - Prinsipal mengirimkan kueri atau skrip ETL menggunakan layanan terintegrasi seperti Amazon Athena, Amazon EMR, atau AWS Glue Amazon Redshift Spectrum.
2. Permintaan data - Mesin analitik terintegrasi mengidentifikasi tabel yang diminta dan mengirimkan permintaan metadata ke Katalog Data (`GetTable`,). `GetDatabase`
3. Memeriksa izin - Katalog Data memverifikasi izin akses prinsipal kueri dengan Lake Formation.
   a. Jika tabel tidak memiliki izin `IAMAllowedPrincipals` grup yang dilampirkan, izin Lake Formation diberlakukan.
   b. Jika kepala sekolah telah memilih untuk menggunakan izin Lake Formation dalam mode akses hybrid, dan tabel memiliki izin `IAMAllowedPrincipals` grup yang dilampirkan, izin Lake Formation diberlakukan. Mesin kueri menerapkan filter yang diterimanya dari Lake Formation dan mengembalikan data ke pengguna.

c. Jika lokasi tabel tidak terdaftar di Lake Formation dan kepala sekolah belum memilih untuk menggunakan izin Lake Formation dalam mode akses hibrida, Katalog Data akan memeriksa apakah tabel memiliki izin `IAMAllowedPrincipals` grup yang dilampirkan padanya. Jika izin ini ada di atas meja, semua kepala sekolah di akun mendapat `Super` atau `All` izin di atas meja.

   Penjual kredenal Lake Formation tidak tersedia, bahkan ketika memilih, kecuali lokasi data terdaftar di Lake Formation.

4. Dapatkan kredensil — Katalog Data memeriksa dan memberi tahu mesin apakah lokasi tabel terdaftar di Lake Formation atau tidak. Jika data yang mendasarinya terdaftar di Lake Formation, mesin analitik meminta kredensi sementara Lake Formation untuk mengakses data di bucket Amazon S3.
5. Dapatkan data — Jika kepala sekolah berwenang untuk mengakses data tabel, Lake Formation menyediakan akses sementara ke mesin analitik terintegrasi. Menggunakan akses sementara, mesin analitik mengambil data dari Amazon S3, dan melakukan pemfilteran yang diperlukan seperti pemfilteran kolom, baris, atau sel. Ketika mesin selesai menjalankan pekerjaan, ia mengembalikan hasilnya kembali ke pengguna. Proses ini disebut credential vending. Untuk informasi lebih lanjut, lihatIntegrasi dengan Lake Formation.
6. Jika lokasi data tabel tidak terdaftar di Lake Formation, panggilan kedua dari mesin analitik dilakukan langsung ke Amazon S3. Kebijakan bucket Amazon S3 terkait dan kebijakan pengguna IAM dievaluasi untuk akses data. Setiap kali Anda menggunakan kebijakan IAM, pastikan bahwa Anda mengikuti praktik terbaik IAM. Untuk informasi selengkapnya, lihat Praktik terbaik keamanan di IAM di Panduan Pengguna IAM.

Menyiapkan mode akses hybrid - skenario umum

Seperti halnya izin Lake Formation, Anda biasanya memiliki dua jenis skenario di mana Anda dapat menggunakan mode akses hibrida untuk mengelola akses data: Menyediakan akses ke prinsipal dalam satu Akun AWS dan menyediakan akses ke eksternal atau prinsipal. Akun AWS

Bagian ini memberikan petunjuk untuk mengatur mode akses hybrid dalam skenario berikut:

Kelola izin dalam mode akses hybrid dalam satu Akun AWS

- Mengonversi AWS Glue sumber daya ke sumber daya hibrida— Saat ini Anda menyediakan akses ke tabel dalam database untuk semua prinsipal di akun Anda menggunakan izin IAM untuk

Amazon S3 dan tetapi AWS Glue ingin mengadopsi Lake Formation untuk mengelola izin secara bertahap.

- Mengonversi sumber daya Lake Formation menjadi sumber daya hibrida— Saat ini Anda menggunakan Lake Formation untuk mengelola akses tabel dalam database untuk semua kepala sekolah di akun Anda tetapi ingin menggunakan Lake Formation hanya untuk prinsipal tertentu. Anda ingin memberikan akses ke prinsipal baru dengan menggunakan izin IAM untuk dan Amazon AWS Glue S3 pada database dan tabel yang sama.

Mengelola izin dalam mode akses hybrid di seluruh Akun AWS detik

- Berbagi AWS Glue sumber daya menggunakan mode akses hybrid— Saat ini Anda tidak menggunakan Lake Formation untuk mengelola izin untuk tabel tetapi ingin menerapkan izin Lake Formation untuk menyediakan akses bagi kepala sekolah di akun lain.
- Berbagi sumber daya Lake Formation menggunakan mode akses hybrid— Anda menggunakan Lake Formation untuk mengelola akses untuk tabel tetapi ingin memberikan akses untuk prinsipal di akun lain dengan menggunakan izin IAM untuk dan Amazon AWS Glue S3 pada database dan tabel yang sama.

Menyiapkan mode akses hybrid - High-level langkah-langkah

1. Daftarkan lokasi data Amazon S3 dengan Lake Formation dengan memilih mode akses Hybrid.
2. Prinsipal harus memiliki `DATA_LOCATION` izin pada lokasi danau data untuk membuat tabel Katalog Data atau database yang mengarah ke lokasi tersebut.
3. Setel pengaturan Cross-account versi ke Versi 4.
4. Berikan izin berbutir halus kepada pengguna atau peran IAM tertentu pada database dan tabel. Pada saat yang sama, pastikan untuk mengatur `Super` atau `All` izin ke `IAMAllowedPrincipals` grup pada database dan semua atau tabel yang dipilih dalam database.
5. Pilih prinsip dan sumber daya. Prinsipal lain di akun dapat terus mengakses database dan tabel menggunakan kebijakan izin IAM untuk dan tindakan Amazon S3. AWS Glue
6. Secara opsional bersihkan kebijakan izin IAM untuk Amazon S3 untuk prinsipal yang memilih untuk menggunakan izin Lake Formation.

Prasyarat untuk mengatur mode akses hybrid

Berikut ini adalah prasyarat untuk mengatur mode akses hybrid:

> **Note**
>
> Kami menyarankan agar administrator Lake Formation mendaftarkan lokasi Amazon S3 dalam mode akses hybrid, dan memilih prinsip dan sumber daya.

1. Berikan izin lokasi data (`DATA_LOCATION_ACCESS`) untuk membuat sumber daya Katalog Data yang mengarah ke lokasi Amazon S3. Izin lokasi data mengontrol kemampuan untuk membuat katalog Katalog Data, database, dan tabel yang mengarah ke lokasi Amazon S3 tertentu.
2. Untuk berbagi sumber daya Katalog Data dengan akun lain dalam mode akses hybrid (tanpa menghapus izin `IAMAllowedPrincipals` grup dari sumber daya), Anda perlu memperbarui pengaturan versi Cross account ke Versi 4 atau lebih tinggi. Untuk memperbarui versi menggunakan konsol Lake Formation, pilih Versi 4 atau Versi 5 di bawah Pengaturan versi Cross account pada halaman pengaturan Katalog Data.

   Anda juga dapat menggunakan `put-data-lake-settings` AWS CLI perintah untuk mengatur `CROSS_ACCOUNT_VERSION` parameter ke versi 4 atau 5:

   ```
   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": "5"
       }
   }
   ```
3.

Untuk memberikan izin lintas akun dalam mode akses hybrid, pemberi harus memiliki izin IAM yang diperlukan untuk dan layanan. AWS Glue AWS RAM Kebijakan AWS terkelola `AWSLakeFormationCrossAccountManager` memberikan izin yang diperlukan.
Untuk mengaktifkan berbagi data lintas akun dalam mode akses hybrid, kami telah memperbarui kebijakan `AWSLakeFormationCrossAccountManager` terkelola dengan menambahkan dua izin IAM baru:

- domba jantan: ListResourceSharePermissions
- domba jantan: AssociateResourceSharePermission

> **Note**
>
> Jika Anda tidak menggunakan kebijakan AWS terkelola untuk peran pemberi, tambahkan kebijakan di atas ke kebijakan khusus Anda.

Lokasi bucket Amazon S3 dan akses pengguna

Saat membuat katalog, database, atau tabel di dalam AWS Glue Data Catalog, Anda dapat menentukan lokasi bucket Amazon S3 dari data yang mendasarinya dan mendaftarkannya ke Lake Formation. Tabel di bawah ini menjelaskan cara kerja izin AWS Glue dan pengguna Lake Formation (prinsipal) berdasarkan lokasi data Amazon S3 dari tabel atau database.

Lokasi Amazon S3 terdaftar di Lake Formation

Lokasi Amazon S3 dari database	AWS Glue pengguna	Pengguna Lake Formation
Terdaftar dengan Lake Formation (dalam mode akses hybrid atau dalam mode Lake Formation)	Memiliki read/write akses ke lokasi data Amazon S3 dengan mewarisi izin dari izin IAMAllowedPrincipals grup (akses super).	Mewarisi izin untuk membuat tabel dari izin CREATE TABLE yang diberikan.
Tidak ada lokasi Amazon S3 terkait	Memerlukan izin LOKASI DATA eksplisit untuk menjalankan pernyataan	Memerlukan izin LOKASI DATA eksplisit untuk menjalankan pernyataan

Lokasi Amazon S3 dari database	AWS Glue pengguna	Pengguna Lake Formation
	CREATE TABLE dan INSERT TABLE.	CREATE TABLE dan INSERT TABLE.

IsRegisteredWithLakeFormation properti meja

`IsRegisteredWithLakeFormation`Properti tabel menunjukkan apakah lokasi data tabel terdaftar dengan Lake Formation untuk pemohon. Jika mode izin lokasi terdaftar sebagai Lake Formation, maka `IsRegisteredWithLakeFormation` properti adalah `true` untuk semua pengguna yang mengakses lokasi data karena semua pengguna dianggap memilih untuk tabel tersebut. Jika lokasi terdaftar dalam mode akses hybrid, maka nilainya diatur `true` hanya untuk pengguna yang telah memilih tabel tersebut.

Cara kerja **`IsRegisteredWithLakeFormation`**

Modus izin	Users/Roles	**`IsRegisteredWithLakeFormation`**	Deskripsi
Lake Formation	Semua	True	Ketika lokasi terdaftar di Lake Formation , `IsRegisteredWithLakeFormation` properti akan disetel ke true untuk semua pengguna. Ini berarti bahwa izin yang ditentukan dalam Lake Formation berlaku untuk lokasi yang terdaftar. Credential vending akan dilakukan oleh Lake Formation.

Modus izin	Users/Roles	**IsRegisteredWithLakeFormation**	Deskripsi
Mode akses hibrid	Memilih	True	Untuk pengguna yang telah memilih untuk menggunakan Lake Formation untuk akses data dan tata kelola untuk tabel, `IsRegisteredWithLakeFormation` properti akan diatur `true` untuk tabel tersebut. Mereka tunduk pada kebijakan izin yang ditentukan dalam Lake Formation untuk lokasi terdaftar.

Modus izin	Users/Roles	**IsRegisteredWithLakeFormation**	Deskripsi
Mode akses hibrid	Tidak memilih	False	Untuk pengguna yang belum memilih untuk menggunakan izin Lake Formation, `IsRegisteredWithLakeFormation` properti disetel ke. `false` Mereka tidak tunduk pada kebijakan izin yang ditentukan dalam Lake Formation untuk lokasi terdaftar. Sebagai gantinya, pengguna akan mengikuti kebijakan izin Amazon S3.

Mengonversi AWS Glue sumber daya ke sumber daya hibrida

Ikuti langkah-langkah berikut untuk mendaftarkan lokasi Amazon S3 dalam mode akses hybrid dan mengaktifkan pengguna Lake Formation baru tanpa mengganggu akses data pengguna Katalog Data yang ada.

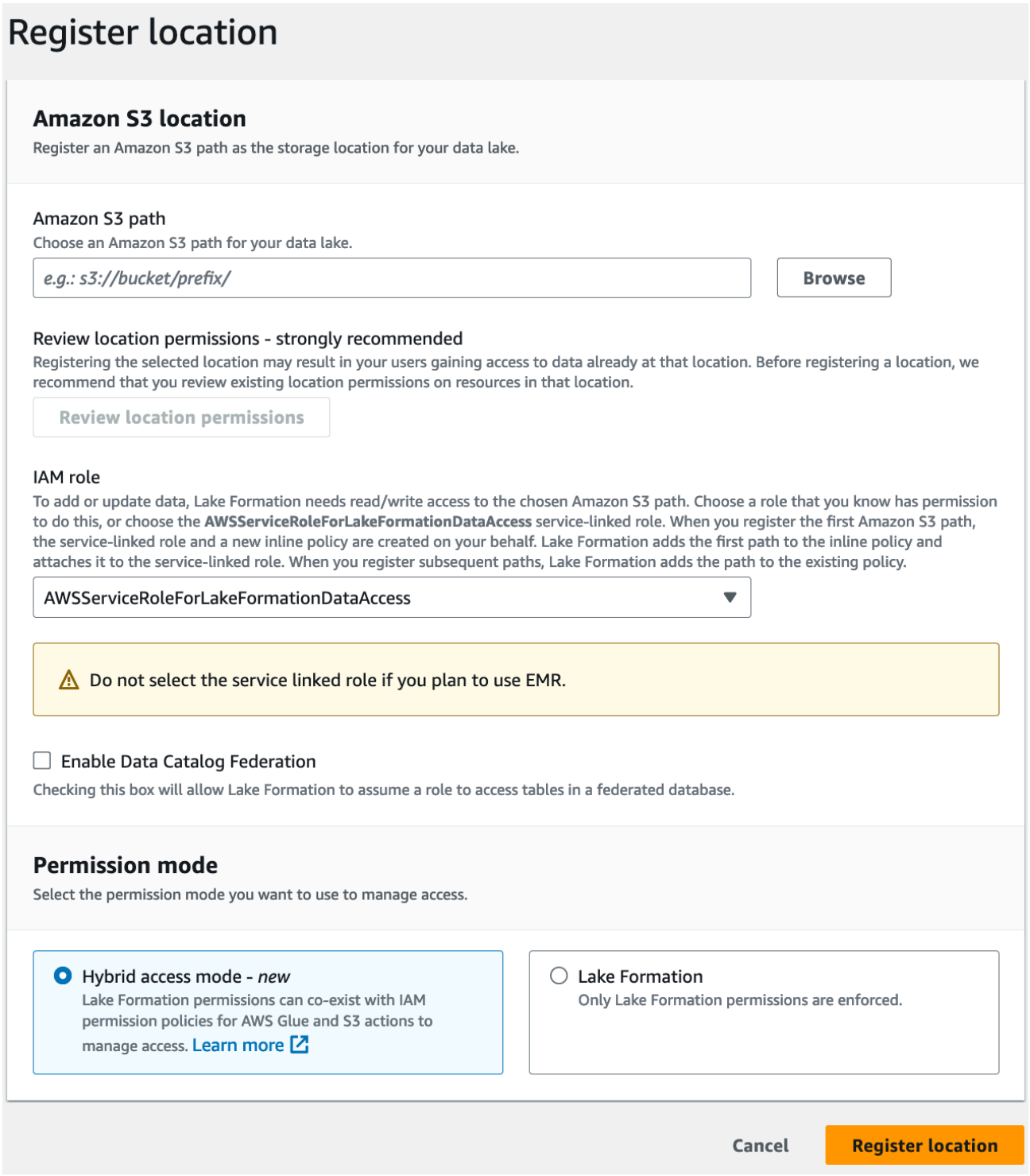
Deskripsi skenario - Lokasi data tidak terdaftar di Lake Formation, dan akses pengguna ke database dan tabel Katalog Data ditentukan oleh kebijakan izin IAM untuk Amazon S3 dan tindakan. AWS Glue
`IAMAllowedPrincipals`Grup secara default memiliki `Super` izin pada semua tabel dalam database.

Untuk mengaktifkan mode akses hibrida untuk lokasi data yang tidak terdaftar di Lake Formation

1. Daftarkan lokasi Amazon S3 yang mengaktifkan mode akses hybrid.

   Console

   1. Masuk ke konsol Lake Formation sebagai administrator danau data.
   2. Di panel navigasi, pilih Lokasi danau data di bawah Administrasi.
   3. Pilih Daftar lokasi.



4. Pada jendela Register location, pilih jalur Amazon S3 yang ingin Anda daftarkan dengan Lake Formation.
5. Untuk peran IAM, pilih peran `AWSServiceRoleForLakeFormationDataAccess` terkait layanan (default) atau IAM kustom peran yang memenuhi persyaratan di[Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi](#).

6. Pilih mode akses Hybrid untuk menerapkan kebijakan kontrol akses Lake Formation berbutir halus ke prinsipal keikutsertaan serta database dan tabel Katalog Data yang menunjuk ke lokasi terdaftar.

   Pilih Lake Formation untuk mengizinkan Lake Formation mengotorisasi permintaan akses ke lokasi yang terdaftar.

7. Pilih Daftar lokasi.

AWS CLI

Berikut ini adalah contoh untuk mendaftarkan lokasi data dengan Lake Formation dengan HybridAccessEnabled:true/false. Nilai default untuk `HybridAccessEnabled` parameter adalah false. Ganti jalur Amazon S3, nama peran, dan id AWS akun dengan nilai yang valid.

```
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. Berikan izin dan pilih prinsipal untuk menggunakan izin Lake Formation untuk sumber daya dalam mode akses hybrid

Sebelum Anda memilih prinsipal dan sumber daya dalam mode akses hibrida, verifikasi bahwa `Super` atau `All` izin untuk `IAMAllowedPrincipals` mengelompokkan ada di database dan tabel yang memiliki lokasi terdaftar dengan Lake Formation dalam mode akses hybrid.

> **Note**
>
> Anda tidak dapat memberikan izin `IAMAllowedPrincipals` grup di `All tables` dalam database. Anda harus memilih setiap tabel secara terpisah dari menu drop-down, dan memberikan izin. Juga, ketika Anda membuat tabel baru dalam database, Anda dapat memilih untuk menggunakan `Use only IAM access control for new tables in new databases` dalam Pengaturan Katalog Data. Opsi ini memberikan

`Super` izin ke `IAMAllowedPrincipals` grup secara otomatis saat Anda membuat tabel baru dalam database.

Console

1. Di konsol Lake Formation, di bawah Katalog Data, pilih Katalog, Database, atau Tabel.
2. Pilih katalog, database, atau tabel dari daftar, dan pilih Hibah dari menu Tindakan.
3. Pilih prinsipal untuk memberikan izin pada database, tabel, dan kolom menggunakan metode sumber daya bernama atau. LF-Tags

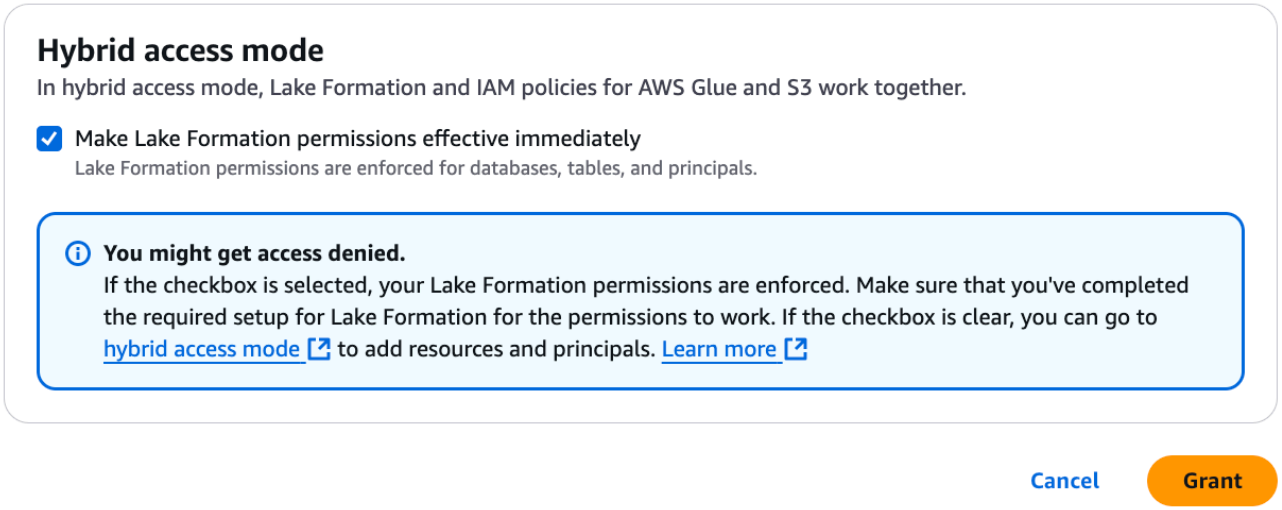
   Atau, pilih Izin data, pilih prinsipal untuk memberikan izin dari daftar, dan pilih Hibah.

   Untuk detail selengkapnya tentang pemberian izin data, lihat. Memberikan izin pada sumber daya Katalog Data

   > **Note**
   >
   > Jika Anda memberikan izin Create table utama, Anda juga perlu memberikan izin lokasi data (`DATA_LOCATION_ACCESS`) kepada prinsipal. Izin ini tidak diperlukan untuk memperbarui tabel.
   > Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

4. Saat Anda menggunakan metode sumber daya bernama untuk memberikan izin, opsi untuk memilih prinsipal dan sumber daya tersedia di bagian bawah halaman izin data Hibah.

   Pilih Jadikan izin Lake Formation segera efektif untuk mengaktifkan izin Lake Formation untuk kepala sekolah dan sumber daya.



5. PilihIzin.

   Ketika Anda memilih prinsipal A pada tabel A yang menunjuk ke lokasi data, ini memungkinkan prinsipal A untuk memiliki akses ke lokasi tabel ini menggunakan izin Lake Formation jika lokasi data terdaftar dalam mode hybrid.

AWS CLI

Berikut ini adalah contoh untuk memilih dalam prinsipal dan tabel dalam mode akses hybrid. Ganti nama peran, id AWS akun, nama database, dan nama tabel dengan nilai yang valid.

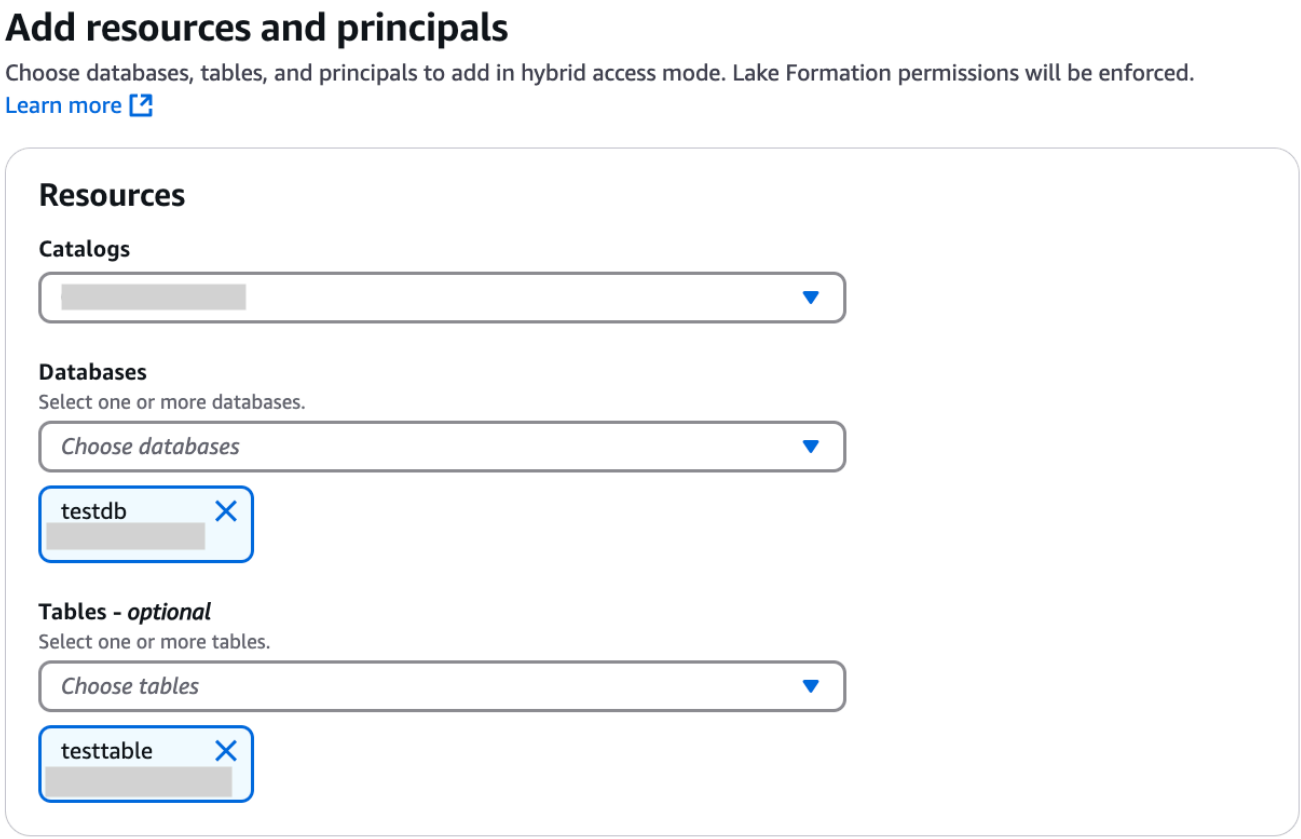
```
aws lakeformation create-lake-formation-opt-in --cli-input-json file://file path
json:
  {
        "Principal": {
            "DataLakePrincipalIdentifier":
 "arn:aws:iam::<123456789012>:role/<hybrid-access-role>"
        },
        "Resource": {
            "Table": {
                "CatalogId": "<123456789012>",
                "DatabaseName": "<hybrid_test>",
                "Name": "<hybrid_test_table>"
            }
        }
    }
```

a. (Optional) Jika Anda memilih LF-Tags untuk memberikan izin, Anda dapat memilih prinsipal untuk menggunakan izin Lake Formation dalam langkah terpisah. Anda dapat melakukan ini dengan memilih mode akses Hybrid di bawah Izin dari bilah navigasi kiri.

b. Pada bagian bawah halaman mode akses Hybrid, pilih Tambahkan untuk menambahkan sumber daya dan prinsipal ke mode akses hybrid.

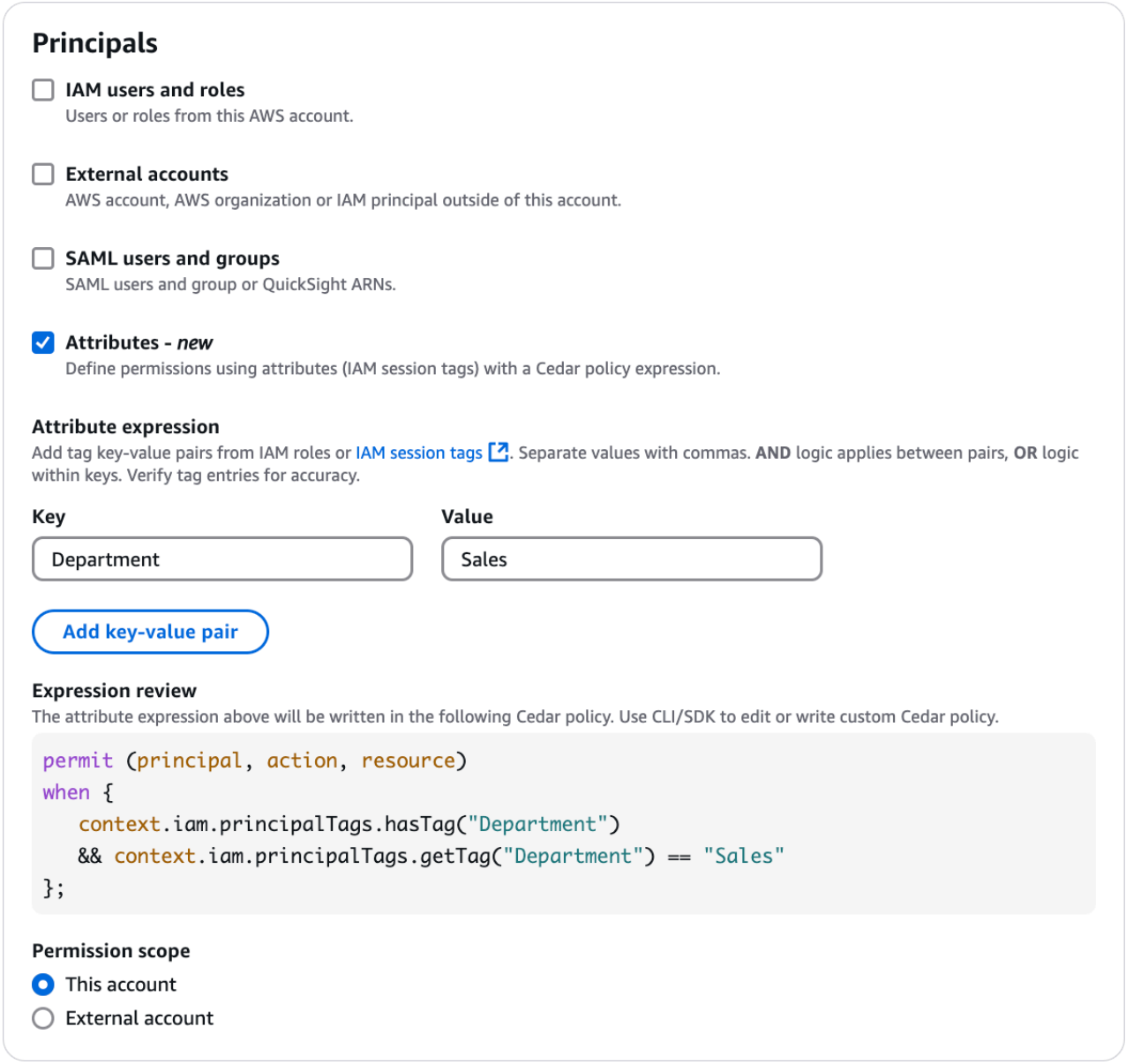
c. Pada halaman Tambah sumber daya dan prinsipal, pilih katalog, database, dan tabel yang terdaftar dalam mode akses hybrid.

   Anda dapat memilih `All tables` di bawah database untuk memberikan akses.



d. Pilih kepala sekolah yang ikut serta untuk menggunakan izin Lake Formation dalam mode akses hybrid.

   - Prinsipal — Anda dapat memilih pengguna dan peran IAM di akun yang sama atau di akun lain. Anda juga dapat memilih pengguna dan grup SAMB.
   - Atribut - Pilih atribut untuk memberikan izin berdasarkan atribut.



- Masukkan pasangan kunci-nilai untuk membuat hibah berdasarkan atribut. Tinjau ekspresi kebijakan Cedar di konsol. Untuk informasi lebih lanjut tentang Cedar, lihat Apa itu Cedar? | Referensi GuideLink Bahasa Kebijakan Cedar.
- Pilih Tambahkan.

  Semua IAM roles/users dengan atribut yang cocok diberikan akses.

e. Pilih Tambahkan.

Mengonversi sumber daya Lake Formation menjadi sumber daya hibrida

Jika Anda saat ini menggunakan izin Lake Formation untuk database dan tabel Katalog Data, Anda dapat mengedit properti pendaftaran lokasi untuk mengaktifkan mode akses hybrid. Ini memungkinkan Anda memberikan akses kepada prinsipal baru ke sumber daya yang sama menggunakan kebijakan izin IAM untuk Amazon S3 dan tindakan tanpa AWS Glue mengganggu izin Lake Formation yang ada.

Deskripsi skenario - Langkah-langkah berikut mengasumsikan bahwa Anda memiliki lokasi data yang terdaftar di Lake Formation, dan Anda telah menyiapkan izin untuk prinsipal pada database, tabel, atau kolom yang menunjuk ke lokasi tersebut. Jika lokasi terdaftar dengan peran yang ditautkan layanan, Anda tidak dapat memperbarui parameter lokasi dan mengaktifkan mode akses hibrid. `IAMAllowedPrincipals`Grup secara default memiliki izin Super pada database dan semua tabelnya.

> **Important**
>
> Jangan memperbarui pendaftaran lokasi ke mode akses hybrid tanpa memilih prinsipal yang mengakses data di lokasi ini.

Mengaktifkan mode akses hybrid untuk lokasi data yang terdaftar di Lake Formation

1.
> **Warning**
>
> Kami tidak menyarankan untuk mengonversi lokasi data terkelola Lake Formation ke mode akses hibrid untuk menghindari gangguan kebijakan izin pengguna atau beban kerja lain yang ada.

   Pilih kepala sekolah yang ada yang memiliki izin Lake Formation.

   1. Buat daftar dan tinjau izin yang Anda berikan kepada prinsipal di katalog, database, dan tabel. Untuk informasi selengkapnya, lihat Melihat izin database dan tabel di Lake Formation.
   2. Pilih mode akses Hybrid di bawah Izin dari bilah navigasi kiri, dan pilih Tambah.
   3. Pada halaman Tambah prinsip dan sumber daya, pilih katalog, database, dan tabel dari lokasi data Amazon S3 yang ingin Anda gunakan dalam mode akses hibrid. Pilih kepala sekolah yang sudah memiliki izin Lake Formation.

4. Pilih Tambah untuk memilih prinsipal untuk menggunakan izin Lake Formation dalam mode akses hybrid.

2. Perbarui bucket/prefix pendaftaran Amazon S3 dengan memilih opsi mode akses Hybrid.

   Console

   1. Masuk ke konsol Lake Formation sebagai administrator danau data.
   2. Di panel navigasi, di bawah Daftar dan Ingest, pilih Lokasi danau data.
   3. Pilih lokasi, dan pada menu Tindakan, pilih Edit.
   4. Pilih mode akses Hybrid.
   5. Pilih Simpan.
   6. Di bawah Katalog Data, pilih database atau tabel dan berikan `Super` atau `All` izin ke grup virtual yang disebut`IAMAllowedPrincipals`.
   7. Verifikasi bahwa akses pengguna Lake Formation yang ada tidak terganggu saat Anda memperbarui properti pendaftaran lokasi. Masuk ke konsol Athena sebagai kepala sekolah Lake Formation dan jalankan contoh kueri pada tabel yang menunjuk ke lokasi yang diperbarui.

      Demikian pula, verifikasi akses AWS Glue pengguna yang menggunakan kebijakan izin IAM untuk mengakses database dan tabel.

   AWS CLI

   Berikut ini adalah contoh untuk mendaftarkan lokasi data dengan Lake Formation dengan HybridAccessEnabled:true/false. Nilai default untuk `HybridAccessEnabled` parameter adalah false. Ganti jalur Amazon S3, nama peran, dan id AWS akun dengan nilai yang valid.

```
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
}
```

Berbagi AWS Glue sumber daya menggunakan mode akses hybrid

Bagikan data dengan orang lain Akun AWS atau prinsipal dalam Akun AWS menegakkan izin Lake Formation lainnya tanpa mengganggu akses berbasis IAM pengguna Katalog Data yang ada.

Deskripsi skenario - Akun produsen memiliki database Katalog Data yang memiliki akses yang dikontrol menggunakan kebijakan utama IAM untuk Amazon S3 AWS Glue dan tindakan. Lokasi data database tidak terdaftar di Lake Formation. `IAMAllowedPrincipals`Grup, secara default, memiliki `Super` izin pada database dan semua tabelnya.

Memberikan izin Lake Formation lintas akun dalam mode akses hybrid

1. Pengaturan akun produser

    1. Masuk ke konsol Lake Formation menggunakan peran yang memiliki izin `lakeformation:PutDataLakeSettings` IAM.
    2. Buka pengaturan Katalog Data, dan pilih `Version 4` pengaturan versi Cross account.

        Jika saat ini Anda menggunakan versi 1 atau 2, lihat Memperbarui pengaturan versi berbagi data lintas akun petunjuk tentang memperbarui ke versi 3.

        Tidak ada perubahan kebijakan izin yang diperlukan saat memutakhirkan dari versi 3 ke 4.
    3. Daftarkan lokasi Amazon S3 dari database atau tabel yang akan Anda bagikan dalam mode akses hybrid.
    4. Verifikasi bahwa `Super` izin ke `IAMAllowedPrincipals` grup ada di database dan tabel tempat Anda mendaftarkan lokasi data dalam mode akses hibrida pada langkah di atas.
    5. Berikan izin Lake Formation ke AWS organisasi, unit organisasi (OU), atau langsung dengan kepala IAM di akun lain.
    6. Jika Anda memberikan izin langsung ke prinsipal IAM, pilih prinsipal dari akun konsumen untuk menerapkan izin Lake Formation dalam mode akses hybrid dengan mengaktifkan opsi Jadikan izin Lake Formation segera berlaku.

        Jika Anda memberikan izin lintas akun ke AWS akun lain, saat Anda memilih akun, izin Lake Formation diberlakukan hanya untuk admin akun tersebut. Administrator danau data akun penerima perlu menurunkan izin dan memilih prinsipal di akun untuk menerapkan izin Lake Formation untuk sumber daya bersama yang berada dalam mode akses hibrid.

        Jika Anda memilih Sumber daya yang cocok dengan LF-Tags opsi untuk memberikan izin lintas akun, Anda harus terlebih dahulu menyelesaikan langkah pemberian izin. Anda dapat

memilih prinsip dan sumber daya ke mode akses hibrid sebagai langkah terpisah dengan memilih mode akses Hybrid di bawah Izin di bilah navigasi kiri konsol Lake Formation. Kemudian pilih Tambah untuk menambahkan sumber daya dan prinsip yang ingin Anda terapkan izin Lake Formation.

2. Pengaturan akun konsumen

    1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/sebagai administrator danau data.
    2. Pergi ke https://console.aws.amazon.com/ram/home, dan terima undangan berbagi sumber daya. Tab Dibagikan dengan saya di AWS RAM konsol menampilkan database dan tabel yang dibagikan dengan akun Anda.
    3. Buat tautan sumber daya ke and/or tabel database bersama di Lake Formation.
    4. Berikan `Describe` izin pada tautan sumber daya dan `Grant on target` izin (pada sumber daya bersama asli) ke prinsipal IAM di akun (konsumen) Anda.
    5. Berikan izin Lake Formation pada database atau tabel yang dibagikan dengan Anda kepada kepala sekolah di akun Anda. Pilih prinsip dan sumber daya untuk menerapkan izin Lake Formation dalam mode akses hybrid dengan mengaktifkan opsi Jadikan izin Lake Formation segera efektif.
    6. Uji izin Lake Formation kepala sekolah dengan menjalankan contoh kueri Athena. Uji akses AWS Glue pengguna Anda yang ada dengan kebijakan utama IAM untuk Amazon S3 AWS Glue dan tindakan.

        (Opsional) Hapus kebijakan bucket Amazon S3 untuk akses data dan kebijakan utama IAM untuk dan akses data Amazon AWS Glue S3 untuk prinsipal yang Anda konfigurasikan untuk menggunakan izin Lake Formation.

Berbagi sumber daya Lake Formation menggunakan mode akses hybrid

Izinkan pengguna Katalog Data baru di akun eksternal untuk mengakses database dan tabel Katalog Data menggunakan kebijakan berbasis IAM tanpa mengganggu izin berbagi lintas akun Lake Formation yang ada.

Deskripsi skenario - Akun produsen memiliki database dan tabel terkelola Lake Formation yang dibagikan dengan akun eksternal (konsumen) di tingkat akun atau tingkat utama IAM. Lokasi data database terdaftar di Lake Formation. `IAMAllowedPrincipals`Grup tidak memiliki `Super` izin pada database dan tabelnya.

Memberikan akses lintas akun ke pengguna Katalog Data baru melalui kebijakan berbasis IAM tanpa mengganggu izin Lake Formation yang ada

1. Pengaturan akun produser

    1. Masuk ke konsol Lake Formation menggunakan peran itu`lakeformation:PutDataLakeSettings`.
    2. Di bawah Pengaturan Katalog Data, pilih `Version 4` pengaturan versi Cross account.

        Jika saat ini Anda menggunakan versi 1 atau 2, lihat Memperbarui pengaturan versi berbagi data lintas akun petunjuk tentang memperbarui ke versi 3.

        Tidak ada perubahan kebijakan izin yang diperlukan untuk meningkatkan dari versi 3 ke 4.
    3. Buat daftar izin yang Anda berikan kepada prinsipal pada database dan tabel. Untuk informasi selengkapnya, lihat Melihat izin database dan tabel di Lake Formation.
    4. Berikan kembali izin lintas akun Lake Formation yang ada dengan memilih prinsip dan sumber daya.

        > **Note**
        >
        > Sebelum memperbarui pendaftaran lokasi data ke mode akses hibrid untuk memberikan izin lintas akun, Anda harus memberikan kembali setidaknya satu pembagian data lintas akun per akun. Langkah ini diperlukan untuk memperbarui izin AWS RAM terkelola yang dilampirkan pada pembagian AWS RAM sumber daya. Pada Juli 2023, Lake Formation telah memperbarui izin AWS RAM terkelola yang digunakan untuk berbagi database dan tabel:
        >
        > - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueAllTablesReadWriteForDatabase`(kebijakan pembagian tingkat basis data)
        > - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueTableReadWrite`(kebijakan berbagi tingkat tabel)
        >
        > Hibah izin lintas akun yang dibuat sebelum Juli 2023 tidak memiliki izin yang diperbarui ini. AWS RAM
        > Jika Anda telah memberikan izin lintas akun secara langsung ke kepala sekolah, Anda perlu memberikan kembali izin tersebut secara individual kepada prinsipal. Jika

Anda melewati langkah ini, kepala sekolah yang mengakses sumber daya bersama mungkin mendapatkan kesalahan kombinasi ilegal.

5. Kunjungi https://console.aws.amazon.com/ram/home.
6. Tab Dibagikan oleh saya di AWS RAM konsol menampilkan database dan nama tabel yang telah Anda bagikan dengan akun eksternal atau prinsipal.

   Pastikan bahwa izin yang dilampirkan ke sumber daya bersama memiliki ARN yang benar.
7. Verifikasi sumber daya dalam AWS RAM pembagian dalam `Associated` status. Jika status menunjukkan sebagai`Associating`, tunggu sampai mereka masuk ke `Associated` negara bagian. Jika statusnya menjadi`Failed`, hentikan dan hubungi tim layanan Lake Formation.
8. Pilih mode akses Hybrid di bawah Izin dari bilah navigasi kiri, dan pilih Tambah.
9. Halaman Add prinsipal dan sumber daya menunjukkan database, and/or tabel, dan prinsipal yang memiliki akses. Anda dapat membuat pembaruan yang diperlukan dengan menambahkan atau menghapus prinsip dan sumber daya.
10. Pilih prinsipal dengan izin Lake Formation untuk database dan tabel yang ingin Anda ubah ke mode akses hybrid. Pilih database dan tabel.
11. Pilih Tambah untuk memilih prinsipal untuk menerapkan izin Lake Formation dalam mode akses hybrid.
12. Berikan `Super` izin ke grup virtual `IAMAllowedPrincipals` pada database Anda dan tabel yang dipilih.
13. Edit pendaftaran Lake Formation lokasi Amazon S3 ke mode akses hybrid.
14. Berikan izin untuk AWS Glue pengguna di akun eksternal (konsumen) menggunakan kebijakan izin IAM untuk tindakan Amazon AWS Glue S3.

2. Pengaturan akun konsumen

   1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/sebagai administrator danau data.
   2. Pergi ke https://console.aws.amazon.com/ram/homedan terima undangan berbagi sumber daya. Tab Sumber Daya yang dibagikan dengan saya di AWS RAM halaman menampilkan database dan nama tabel yang dibagikan dengan akun Anda.

      Untuk AWS RAM pembagian, pastikan bahwa izin terlampir memiliki ARN yang benar dari undangan bersama AWS RAM . Periksa apakah sumber daya dalam AWS RAM pembagian dalam `Associated` status. Jika status menunjukkan sebagai`Associating`, tunggu sampai

mereka masuk ke `Associated` negara bagian. Jika statusnya menjadi`Failed`, hentikan dan hubungi tim layanan Lake Formation.

3. Buat tautan sumber daya ke and/or tabel database bersama di Lake Formation.
4. Berikan `Describe` izin pada tautan sumber daya dan `Grant on target` izin (pada sumber daya bersama asli) ke prinsipal IAM di akun (konsumen) Anda.
5. Selanjutnya, siapkan izin Lake Formation untuk kepala sekolah di akun Anda di database atau tabel bersama.

   Di bilah navigasi kiri, di bawah Izin, pilih mode akses Hybrid.
6. Pilih Tambah di bagian bawah halaman mode akses Hybrid untuk memilih prinsipal dan database atau tabel yang dibagikan dengan Anda dari akun produsen.
7. Berikan izin untuk AWS Glue pengguna di akun Anda menggunakan kebijakan izin IAM untuk tindakan Amazon AWS Glue S3.
8. Uji izin dan AWS Glue izin Lake Formation pengguna dengan menjalankan kueri sampel terpisah di atas meja menggunakan Athena

   (Opsional) Bersihkan kebijakan izin IAM untuk Amazon S3 untuk prinsipal yang berada dalam mode akses hybrid.

Menghapus prinsip dan sumber daya dari mode akses hybrid

Ikuti langkah-langkah ini untuk menghapus database, tabel, dan prinsipal dari mode akses hybrid.

Console

1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di bawah Izin, pilih mode akses Hybrid.
3. Pada halaman mode akses Hybrid, pilih kotak centang di sebelah database atau nama tabel dan pilih`Remove`.
4. Pesan peringatan meminta Anda untuk mengonfirmasi tindakan tersebut. Pilih Hapus.

   Lake Formation tidak lagi memberlakukan izin untuk sumber daya tersebut, dan akses ke sumber daya ini akan dikontrol menggunakan IAM dan izin. AWS Glue Hal ini dapat menyebabkan pengguna tidak lagi memiliki akses ke sumber daya ini jika mereka tidak memiliki izin IAM yang sesuai.

AWS CLI

Contoh berikut menunjukkan cara menghapus sumber daya dari mode akses hybrid.

```
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>"
          }
    }
}
```

Melihat prinsip dan sumber daya dalam mode akses hybrid

Ikuti langkah-langkah ini untuk melihat database, tabel, dan prinsipal dalam mode akses hybrid.

Console

1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di bawah Izin, pilih mode akses Hybrid.
3. Halaman mode akses Hybrid menunjukkan sumber daya dan prinsipal yang saat ini dalam mode akses hybrid..

AWS CLI

Contoh berikut menunjukkan bagaimana untuk daftar semua opt in prinsipal dan sumber daya yang berada dalam mode akses hybrid.

```
aws lakeformation list-lake-formation-opt-ins
```

Contoh berikut menunjukkan cara membuat daftar opt in untuk pasangan sumber daya utama tertentu.

```
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>"
          }
    }
}
```

Sumber daya tambahan

Dalam posting blog berikut, kami memandu Anda melalui instruksi untuk izin Lake Formation onboard dalam mode akses hybrid untuk pengguna yang dipilih sementara database sudah dapat diakses oleh pengguna lain melalui izin IAM dan Amazon S3. Kami akan meninjau instruksi untuk mengatur mode akses hybrid dalam AWS akun dan di antara dua akun.

- Memperkenalkan mode akses hybrid AWS Glue Data Catalog untuk mengamankan akses menggunakan kebijakan Lake Formation dan IAM dan Amazon S3.

Membuat objek di AWS Glue Data Catalog

AWS Lake Formation menggunakan AWS Glue Data Catalog (Katalog Data) untuk menyimpan metadata tentang data lake, sumber data, transformasi, dan target. Metadata adalah data tentang data yang mendasari dalam set data Anda. Setiap AWS akun memiliki satu Katalog Data per AWS Wilayah.

Metadata dalam Katalog Data diatur dalam hierarki data tiga tingkat yang terdiri dari katalog, database, dan tabel. Ini mengatur data dari berbagai sumber ke dalam wadah logis yang disebut katalog. Setiap katalog mewakili data dari sumber seperti gudang data Amazon Redshift, Amazon DynamoDB database, dan sumber data pihak ketiga seperti Snowflake, MySQL, dan lebih dari 30 sumber data eksternal, yang terintegrasi melalui konektor federasi. Anda juga dapat membuat katalog baru di Katalog Data untuk menyimpan data di Bucket Tabel S3 atau Redshift Managed Storage (RMS).

Tabel menyimpan informasi tentang data yang mendasarinya, termasuk informasi skema, informasi partisi, dan lokasi data. Database adalah kumpulan tabel. Katalog Data juga berisi tautan sumber daya, yang merupakan tautan ke katalog bersama, database, dan tabel di akun eksternal, dan digunakan untuk akses lintas akun ke data di danau data.

Katalog Data adalah objek katalog bersarang yang berisi katalog, database, dan tabel. Ini direferensikan oleh Akun AWS ID, dan merupakan katalog default di akun Wilayah AWS dan file. Katalog Data menggunakan hierarki tiga tingkat (catalog.database.table) untuk mengatur tabel.

- Katalog — Tingkat paling atas dari tiga tingkat hierarki metadata Katalog Data. Anda dapat menambahkan beberapa katalog dalam Katalog Data melalui federasi.
- Database — Tingkat kedua dari hierarki metadata yang terdiri dari tabel dan tampilan. Database juga disebut sebagai skema di banyak sistem data seperti Amazon Redshift dan Trino.
- Tabel dan tampilan — Tingkat ketiga hierarki data 3 tingkat Katalog Data.

Semua tabel Iceberg di Amazon S3 disimpan dalam Katalog Data default yang memiliki ID Katalog = ID. Akun AWS Anda dapat membuat katalog federasi dalam menyimpan definisi tabel di AWS Glue Data Catalog Amazon Redshift, penyimpanan Tabel Amazon S3, atau sumber data pihak ketiga lainnya melalui federasi.

Topik

- Membuat katalog
- Membuat basis data
- Membuat tabel
- Membangun AWS Glue Data Catalog pandangan

Membuat katalog

Katalog mewakili tingkat tertinggi atau paling atas dalam hierarki metadata tiga tingkat. AWS Glue Data Catalog Anda dapat menggunakan beberapa metode untuk membawa data ke dalam Katalog Data dan membuat katalog multi-level.

Untuk informasi selengkapnya tentang membuat katalog dari sumber data eksternal, lihat. Membawa data Anda ke AWS Glue Data Catalog

Untuk membuat katalog menggunakan konsol Lake Formation, Anda harus masuk sebagai administrator danau data atau pembuat katalog. Pencipta katalog adalah kepala sekolah yang telah diberikan `CREATE_CATALOG` izin Lake Formation. Anda dapat melihat daftar pembuat katalog di halaman peran dan tugas Administratif konsol Lake Formation. Untuk melihat daftar ini, Anda harus memiliki izin `lakeformation:ListPermissions` IAM dan masuk sebagai administrator danau data atau sebagai pembuat katalog dengan opsi hibah pada `CREATE_CATALOG` izin.

Membuat basis data

Tabel metadata dalam Katalog Data disimpan dalam database. Anda dapat membuat database sebanyak yang Anda butuhkan, dan Anda dapat memberikan izin Lake Formation yang berbeda di setiap database.

Database dapat memiliki properti lokasi opsional. Lokasi ini biasanya berada dalam lokasi Amazon Simple Storage Service (Amazon S3) yang terdaftar di Lake Formation. Saat Anda menentukan lokasi, kepala sekolah tidak memerlukan izin lokasi data untuk membuat tabel Katalog Data yang mengarah ke lokasi dalam lokasi database. Untuk informasi selengkapnya, lihat Underlying data access control.

Untuk membuat database menggunakan konsol Lake Formation, Anda harus masuk sebagai administrator data lake atau pembuat database. Pembuat database adalah kepala sekolah yang telah diberikan `CREATE_DATABASE` izin Lake Formation. Anda dapat melihat daftar pembuat basis data di halaman peran dan tugas Administratif konsol Lake Formation. Untuk melihat daftar ini, Anda harus memiliki izin `lakeformation:ListPermissions` IAM dan masuk sebagai administrator data lake atau sebagai pembuat basis data dengan opsi hibah pada `CREATE_DATABASE` izin.

Untuk membuat basis data

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator danau data atau pembuat basis data.

2. Di panel navigasi, di bawah Katalog data, pilih Database.
3. Pilih Buat basis data.
4. Dalam kotak dialog Buat database, masukkan nama database, lokasi opsional, dan deskripsi opsional.
5. Secara opsional pilih Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini.

   Untuk informasi tentang opsi ini, lihatthe section called “Mengubah pengaturan default untuk data lake”.
6. Pilih Buat basis data.

Membuat tabel

AWS Lake Formation tabel metadata berisi informasi tentang data di danau data, termasuk informasi skema, informasi partisi, dan lokasi data. Tabel ini disimpan dalam Katalog AWS Glue Data. Anda menggunakannya untuk mengakses data dasar di data lake dan mengelola data tersebut dengan izin Lake Formation. Tabel disimpan dalam database dalam Katalog Data.

Ada beberapa cara untuk membuat tabel Katalog Data:

- Jalankan crawler diAWS Glue. Lihat Mendefinisikan crawler di Panduan AWS Glue Pengembang.
- Buat dan jalankan alur kerja. Lihat the section called “Mengimpor data menggunakan alur kerja”.
- Buat tabel secara manual menggunakan konsol Lake Formation, AWS Glue API, atau AWS Command Line Interface (AWS CLI).
- Buat tabel menggunakan Amazon Athena.
- Buat tautan sumber daya ke tabel di akun eksternal. Lihat the section called “Membuat tautan sumber daya”.

Membuat tabel Apache Iceberg

AWS Lake Formation mendukung pembuatan tabel Apache Iceberg yang menggunakan format data Apache Parquet di AWS Glue Data Catalog dengan data yang berada di Amazon S3. Tabel dalam Katalog Data adalah definisi metadata yang mewakili data dalam penyimpanan data. Secara default, Lake Formation membuat tabel Iceberg v2. Untuk perbedaan antara tabel v1 dan v2, lihat Format perubahan versi dalam dokumentasi Apache Iceberg.

Apache Iceberg adalah format tabel terbuka untuk dataset analitik yang sangat besar. Iceberg memungkinkan perubahan mudah pada skema Anda, juga dikenal sebagai evolusi skema, yang berarti bahwa pengguna dapat menambahkan, mengganti nama, atau menghapus kolom dari tabel data tanpa mengganggu data yang mendasarinya. Iceberg juga menyediakan dukungan untuk pembuatan versi data, yang memungkinkan pengguna untuk melacak perubahan data dari waktu ke waktu. Ini memungkinkan fitur perjalanan waktu, yang memungkinkan pengguna untuk mengakses dan menanyakan versi historis data dan menganalisis perubahan data antara pembaruan dan penghapusan.

Anda dapat menggunakan konsol Lake Formation atau `CreateTable` operasi di AWS Glue API untuk membuat tabel Gunung Es di Katalog Data. Untuk informasi selengkapnya, lihat CreateTable tindakan (Python: create_table).

Saat Anda membuat tabel Gunung Es di Katalog Data, Anda harus menentukan format tabel dan jalur file metadata di Amazon S3 agar dapat melakukan pembacaan dan penulisan.

Anda dapat menggunakan Lake Formation untuk mengamankan tabel Gunung Es menggunakan izin kontrol akses berbutir halus saat Anda mendaftarkan lokasi data Amazon S3. AWS Lake Formation Untuk data sumber di Amazon S3 dan metadata yang tidak terdaftar di Lake Formation, akses ditentukan oleh kebijakan izin IAM untuk Amazon S3 dan tindakan. AWS Glue Untuk informasi selengkapnya, lihat Mengelola izin Lake Formation.

> **Note**
>
> Data Catalog tidak mendukung pembuatan partisi dan menambahkan properti tabel Iceberg.

Topik

- Prasyarat
- Membuat tabel Iceberg

Prasyarat

Untuk membuat tabel Gunung Es di Katalog Data, dan mengatur izin akses data Lake Formation, Anda harus melengkapi persyaratan berikut:

1. Izin diperlukan untuk membuat tabel Gunung Es tanpa data yang terdaftar di Lake Formation.

   Selain izin yang diperlukan untuk membuat tabel di Katalog Data, pembuat tabel memerlukan izin berikut:

   - `s3:PutObject`pada sumber daya arn:aws:s3::: {bucketName}
   - `s3:GetObject`pada sumber daya arn:aws:s3::: {bucketName}
   - `s3:DeleteObject`pada sumber daya arn:aws:s3::: {bucketName}
2. Izin yang diperlukan untuk membuat tabel Gunung Es dengan data yang terdaftar di Lake Formation:

   Untuk menggunakan Lake Formation untuk mengelola dan mengamankan data di danau data Anda, daftarkan lokasi Amazon S3 Anda yang memiliki data untuk tabel dengan Lake Formation. Ini agar Lake Formation dapat menjual kredensil ke layanan AWS analitis seperti Athena, Redshift Spectrum, dan Amazon EMR untuk mengakses data. Untuk informasi selengkapnya tentang mendaftarkan lokasi Amazon S3, lihat. Menambahkan lokasi Amazon S3 ke danau data Anda

   Kepala sekolah yang membaca dan menulis data dasar yang terdaftar di Lake Formation memerlukan izin berikut:

   - `lakeformation:GetDataAccess`
   - `DATA_LOCATION_ACCESS`

     Kepala sekolah yang memiliki izin lokasi data di lokasi juga memiliki izin lokasi di semua lokasi anak.

     Untuk informasi selengkapnya tentang izin lokasi data, lihatKontrol akses data yang mendasari.

Untuk mengaktifkan pemadatan, layanan perlu mengambil peran IAM yang memiliki izin untuk memperbarui tabel di Katalog Data. Untuk detailnya, lihat Prasyarat pengoptimalan tabel.

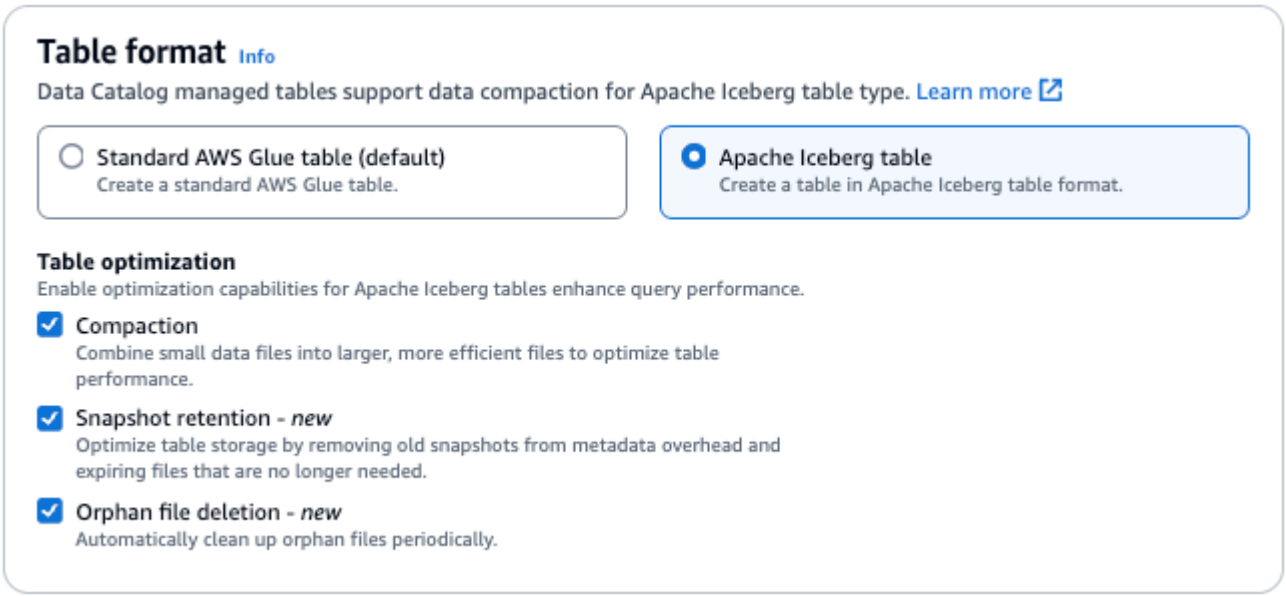
Membuat tabel Iceberg

Anda dapat membuat tabel Iceberg v1 dan v2 menggunakan konsol Lake Formation atau AWS Command Line Interface seperti yang didokumentasikan di halaman ini. Anda juga dapat membuat tabel Iceberg menggunakan AWS Glue konsol atau. Perayap AWS Glue Untuk informasi selengkapnya, lihat Katalog Data dan Crawler di Panduan AWS Glue Pengembang.

Untuk membuat tabel Iceberg

Console

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di bawah Katalog Data, pilih Tabel, dan gunakan tombol Buat tabel untuk menentukan atribut berikut:
   - Nama tabel: Masukkan nama untuk tabel. Jika Anda menggunakan Athena untuk mengakses tabel, gunakan tips penamaan ini di Panduan Pengguna Amazon Athena.
   - Database: Pilih database yang ada atau buat yang baru.
   - Deskripsi:Deskripsi tabel. Anda dapat menulis deskripsi untuk membantu Anda memahami isi tabel tersebut.
   - Format tabel: Untuk format Tabel, pilih Apache Iceberg.



   - Optimalisasi tabel
     - Pemadatan — File data digabungkan dan ditulis ulang menghapus data usang dan mengkonsolidasikan data yang terfragmentasi menjadi file yang lebih besar dan lebih efisien.
     - Retensi snapshot — Snapshot adalah versi stempel waktu dari tabel Iceberg. Konfigurasi retensi snapshot memungkinkan pelanggan untuk menerapkan berapa lama untuk menyimpan snapshot dan berapa banyak snapshot yang akan disimpan. Mengonfigurasi pengoptimal retensi snapshot dapat membantu mengelola overhead penyimpanan dengan

menghapus snapshot yang lebih lama dan tidak perlu serta file yang mendasarinya yang terkait.

- Penghapusan file yatim piatu — File yatim piatu adalah file yang tidak lagi direferensikan oleh metadata tabel Iceberg. File-file ini dapat terakumulasi dari waktu ke waktu, terutama setelah operasi seperti penghapusan tabel atau pekerjaan ETL yang gagal. Mengaktifkan penghapusan file yatim memungkinkan AWS Glue untuk secara berkala mengidentifikasi dan menghapus file-file yang tidak perlu ini, membebaskan penyimpanan.

Untuk informasi selengkapnya, lihat Mengoptimalkan tabel Gunung Es.

- Peran IAM: Untuk menjalankan pemadatan, layanan mengasumsikan peran IAM atas nama Anda. Anda dapat memilih peran IAM menggunakan drop-down. Pastikan peran memiliki izin yang diperlukan untuk mengaktifkan pemadatan.

  Untuk mempelajari lebih lanjut tentang izin yang diperlukan, lihat Prasyarat pengoptimalan tabel.

- Lokasi: Tentukan jalur ke folder di Amazon S3 yang menyimpan tabel metadata. Iceberg membutuhkan file metadata dan lokasi di Katalog Data untuk dapat melakukan pembacaan dan penulisan.

- Skema: Pilih Tambahkan kolom untuk menambahkan kolom dan tipe data kolom. Anda memiliki opsi untuk membuat tabel kosong dan memperbarui skema nanti. Katalog Data mendukung tipe data Hive. Untuk informasi selengkapnya, lihat Tipe data sarang.

  Iceberg memungkinkan Anda untuk mengembangkan skema dan partisi setelah Anda membuat tabel. Anda dapat menggunakan kueri Athena untuk memperbarui skema tabel dan kueri Spark untuk memperbarui partisi.

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/"
            }
        }'
```

Mengoptimalkan tabel Iceberg

Lake Formation mendukung beberapa opsi pengoptimalan tabel untuk meningkatkan manajemen dan kinerja tabel Apache Iceberg yang digunakan oleh mesin AWS analitik dan pekerjaan ETL. Pengoptimal ini menyediakan pemanfaatan penyimpanan yang efisien, peningkatan kinerja kueri, dan manajemen data yang efektif. Ada tiga jenis pengoptimal tabel yang tersedia di Lake Formation:

- Pemadatan — Pemadatan data memadatkan file data kecil untuk mengurangi penggunaan penyimpanan dan meningkatkan kinerja baca. File data digabungkan dan ditulis ulang untuk menghapus data usang dan mengkonsolidasikan data yang terfragmentasi menjadi file yang lebih besar dan lebih efisien. Pemadatan dapat dikonfigurasi untuk berjalan secara otomatis atau dipicu secara manual sesuai kebutuhan.
- Retensi snapshot — Snapshot adalah versi stempel waktu dari tabel Iceberg. Konfigurasi retensi snapshot memungkinkan pelanggan untuk menerapkan berapa lama untuk menyimpan snapshot dan berapa banyak snapshot yang akan disimpan. Mengonfigurasi pengoptimal retensi snapshot dapat membantu mengelola overhead penyimpanan dengan menghapus snapshot yang lebih lama dan tidak perlu serta file yang mendasarinya yang terkait.
- Penghapusan file yatim piatu — File yatim piatu adalah file yang tidak lagi direferensikan oleh metadata tabel Iceberg. File-file ini dapat terakumulasi dari waktu ke waktu, terutama setelah operasi seperti penghapusan tabel atau pekerjaan ETL yang gagal. Mengaktifkan penghapusan file yatim memungkinkan AWS Glue untuk secara berkala mengidentifikasi dan menghapus file-file yang tidak perlu ini, membebaskan penyimpanan.

Anda dapat mengaktifkan atau menonaktifkan pemadatan, retensi snapshot, dan pengoptimal penghapusan file yatim untuk tabel Iceberg individual di Katalog Data menggunakan konsol, atau operasi API. AWS Glue AWS CLI AWS Glue

Untuk informasi selengkapnya, lihat Mengoptimalkan tabel Gunung Es di Panduan Pengembang. AWS Glue

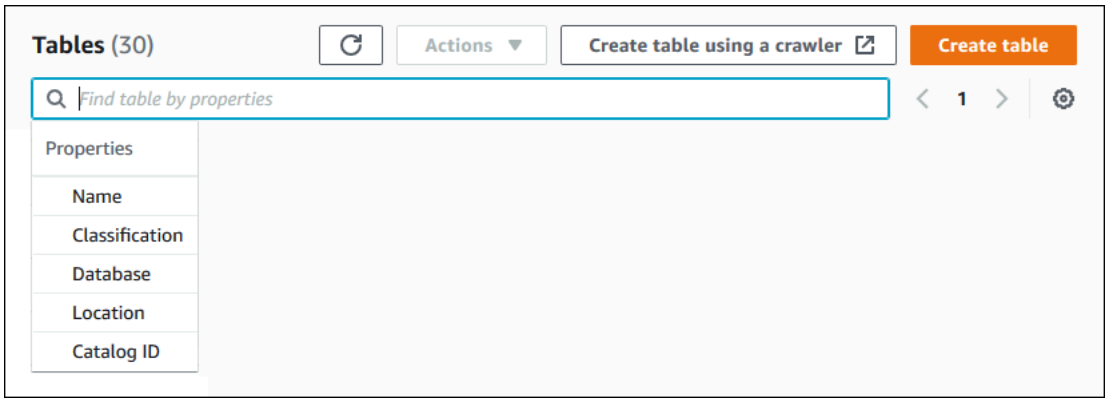
Mencari tabel

Anda dapat menggunakan AWS Lake Formation konsol untuk mencari tabel Katalog Data berdasarkan nama, lokasi, berisi database, dan lainnya. Hasil pencarian hanya menampilkan tabel tempat Anda memiliki izin Lake Formation.

Untuk mencari tabel (konsol)

1. Masuk ke Konsol Manajemen AWS dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Tabel.
3. Posisikan kursor di bidang pencarian. Bidang ini memiliki teks placeholder Temukan tabel berdasarkan properti.

   Menu Properties muncul, menunjukkan berbagai properti tabel untuk dicari.



4. Lakukan salah satu tindakan berikut:

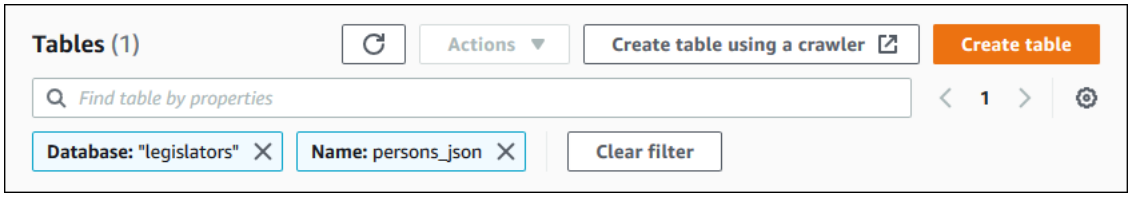
   - Cari dengan memuat database.

     1. Pilih Database dari menu Properties, lalu pilih database dari menu Database yang muncul atau ketik nama database dan tekan Enter.

        Tabel yang Anda memiliki izin dalam database terdaftar.

2. (Opsional) Untuk mempersempit daftar ke satu tabel dalam database, posisikan kursor di bidang pencarian lagi, pilih Nama dari Properti menu, dan pilih nama tabel dari menu Tabel yang muncul atau ketik nama tabel dan tekan Memasukkan.

   Tabel tunggal terdaftar, dan nama database dan nama tabel muncul sebagai ubin di bawah bidang pencarian.



   Untuk menyesuaikan filter, tutup salah satu ubin atau pilih Clear filter.

- Cari berdasarkan properti lain.

  1. Pilih properti pencarian dari menu Properties.

     Untuk mencari berdasarkan ID AWS akun, pilih ID Katalog dari menu Properti, masukkan ID AWS akun yang valid (misalnya, 111122223333), dan tekan Enter.

     Untuk mencari berdasarkan lokasi, pilih Lokasi dari menu Properti, dan pilih lokasi dari menu Lokasi yang muncul. Semua tabel di lokasi root lokasi yang dipilih (misalnya, Amazon S3) dikembalikan.

Mencari tabel menggunakan AWS CLI

- Contoh berikut menunjukkan bagaimana menjalankan serach paral. `--search-text`Parameter ini memungkinkan Anda untuk mencari tabel yang berisi teks yang ditentukan dalam metadata mereka. Dalam hal ini, ia mengembalikan semua tabel yang memiliki “pelanggan” dalam nama, deskripsi, atau bidang metadata lainnya.

```
aws glue search-tables
      --search-text "customer"
      --region Wilayah AWS
      --max-results 10
      --sort-criteria "FieldName=Name,Sort=ASC"
```

Berbagi tabel Katalog Data dan database di seluruh AWS Akun

Anda dapat membagikan sumber daya Katalog Data (database dan tabel) dengan AWS akun eksternal dengan memberikan izin Lake Formation pada sumber daya ke akun eksternal. Pengguna kemudian dapat menjalankan kueri dan pekerjaan yang bergabung dan menanyakan tabel di beberapa akun. Dengan beberapa batasan, saat Anda membagikan sumber daya Katalog Data dengan akun lain, prinsipal di akun tersebut dapat beroperasi pada sumber daya tersebut seolah-olah sumber daya tersebut berada di Katalog Data mereka.

Anda tidak berbagi sumber daya dengan prinsipal tertentu di AWS akun eksternal—Anda berbagi sumber daya dengan akun atau organisasi. AWS Saat berbagi sumber daya dengan AWS organisasi, Anda membagikan sumber daya dengan semua akun di semua tingkatan di organisasi tersebut. Administrator data lake di setiap akun eksternal kemudian harus memberikan izin pada sumber daya bersama kepada kepala sekolah di akun mereka.

Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation dan Memberikan izin pada sumber daya Katalog Data.

Lihat Juga:

- Mengakses dan melihat tabel dan database Katalog Data bersama
- Prasyarat

Membangun AWS Glue Data Catalog pandangan

Dalam AWS Glue Data Catalog, tampilan adalah tabel virtual di mana konten didefinisikan oleh kueri SQL yang mereferensikan satu atau lebih tabel. Anda dapat membuat tampilan Katalog Data yang mereferensikan hingga 10 tabel menggunakan editor SQL untuk Amazon Athena, Amazon Redshift, atau Apache Spark menggunakan EMR Tanpa Server atau versi 5.0. AWS Glue Tabel referensi yang mendasari untuk tampilan dapat menjadi milik database yang sama atau database yang berbeda dalam Katalog Data yang sama Akun AWS.

Anda dapat mereferensikan AWS Glue tabel dan tabel standar dalam format tabel terbuka (OTF) seperti Apache Hudi, Linux Foundation Delta Lake, dan Apache Iceberg, dengan data dasar yang disimpan di lokasi Amazon S3 yang terdaftar. AWS Lake Formation Selain itu, Anda dapat membuat tampilan dari tabel federasi dari datashares Amazon Redshift yang dibagikan dengan Lake Formation.

Membedakan tampilan Katalog Data dari jenis tampilan lainnya

Tampilan Katalog Data berbeda dari tampilan Apache Hive, Apache Spark, dan Amazon Athena. Tampilan Katalog Data adalah fitur asli dari AWS Glue Data Catalog, dan merupakan tampilan definer-made multi-dialek. Anda dapat membuat tampilan Katalog Data menggunakan salah satu layanan analitik yang didukung, seperti Athena atau Amazon Redshift Spectrum, dan mengakses tampilan yang sama menggunakan layanan analitik lain yang didukung. Di sisi lain, tampilan Apache Hive, Apache Spark, dan Athena dibuat secara independen di setiap layanan analitik, seperti Athena dan Amazon Redshift, dan hanya terlihat dan dapat diakses dalam layanan itu.

Apa itu pandangan yang lebih pasti?

Tampilan definer adalah tampilan SQL yang beroperasi berdasarkan izin prinsipal yang membuatnya. Peran definer memiliki izin yang diperlukan untuk mengakses tabel referensi, dan menjalankan pernyataan SQL yang mendefinisikan tampilan. Definer membuat tampilan dan membagikannya dengan pengguna lain melalui kontrol akses AWS Lake Formation berbutir halus.

Saat pengguna menanyakan tampilan definer, mesin kueri menggunakan izin peran definer untuk mengakses tabel referensi yang mendasarinya. Pendekatan ini memungkinkan pengguna untuk berinteraksi dengan tampilan tanpa memerlukan akses langsung ke tabel sumber, meningkatkan keamanan dan menyederhanakan manajemen akses data.

Untuk menyiapkan tampilan definer, peran IAM definer dapat berada dalam AWS akun yang sama dengan tabel dasar, atau di akun yang berbeda menggunakan peran penentu lintas akun. Untuk informasi selengkapnya tentang izin yang diperlukan untuk peran definer, lihat. Prasyarat untuk membuat tampilan

Kerangka kerja untuk tampilan multi-dialek

Katalog Data mendukung pembuatan tampilan menggunakan beberapa dialek bahasa kueri terstruktur (SQL). SQL adalah bahasa yang digunakan untuk menyimpan dan memproses informasi dalam database relasional dan setiap mesin AWS analitik menggunakan variasi SQL, atau dialek SQL sendiri.

Anda membuat tampilan Katalog Data dalam satu dialek SQL menggunakan salah satu mesin kueri analitik yang didukung. Selanjutnya, Anda dapat memperbarui tampilan menggunakan `ALTER VIEW` pernyataan dalam dialek SQL yang berbeda dalam mesin analitik lain yang didukung. Namun, setiap dialek harus mereferensikan kumpulan tabel, kolom, dan tipe data yang sama.

Anda dapat mengakses beberapa dialek yang tersedia untuk tampilan menggunakan `GetTable` API, AWS CLI dan AWS konsol. Dengan demikian, tampilan Katalog Data terlihat dan tersedia untuk kueri di berbagai mesin analitik yang didukung.

Dengan mendefinisikan skema tampilan umum dan objek metadata yang dapat Anda kueri dari beberapa mesin, tampilan Katalog Data memungkinkan Anda menggunakan tampilan seragam di seluruh data lake Anda.

Untuk detail selengkapnya tentang cara skema diselesaikan untuk setiap dialek, lihat tautan ke referensi API. Untuk detail selengkapnya tentang aturan pencocokan untuk berbagai jenis, lihat tautan ke bagian yang relevan di dokumen API.

Mengintegrasikan dengan izin Lake Formation

Anda dapat menggunakan AWS Lake Formation untuk memusatkan manajemen izin pada AWS Glue Data Catalog tampilan untuk pengguna. Anda dapat memberikan izin berbutir halus pada tampilan Katalog Data menggunakan metode sumber daya bernama atau LF-Tags, dan membagikannya di seluruh AWS organisasi Akun AWS, dan unit organisasi. Anda juga dapat berbagi dan mengakses tampilan Katalog Data Wilayah AWS menggunakan tautan sumber daya. Hal ini memungkinkan pengguna untuk menyediakan akses data tanpa menduplikasi sumber data, dan berbagi tabel yang mendasarinya.

Pernyataan `CREATE VIEW` DDL dari tampilan Katalog Data dapat mereferensikan AWS Glue tabel dan tabel standar dalam format tabel terbuka (OTF) seperti Hudi, Delta Lake, dan Iceberg dengan data dasar yang disimpan di lokasi Amazon S3 yang terdaftar di Lake Formation serta tabel federasi dari Amazon Redshift datashare yang dibagikan dengan Lake Formation. Tabel dapat berupa format file apa pun selama mesin yang digunakan untuk menanyakan tampilan mendukung format itu. Anda juga dapat mereferensikan fungsi bawaan dari mesin tempat ia dijalankan tetapi sumber daya khusus mesin lainnya mungkin tidak diizinkan. Untuk detail selengkapnya, lihat Katalog Data melihat pertimbangan dan batasan

Kasus penggunaan

Berikut ini adalah kasus penggunaan penting untuk tampilan Katalog Data:

- Buat dan kelola izin pada skema tampilan tunggal. Ini membantu Anda menghindari risiko izin yang tidak konsisten pada tampilan duplikat yang dibuat di beberapa mesin.
- Berikan izin kepada pengguna pada tampilan yang mereferensikan beberapa tabel tanpa memberikan izin langsung pada tabel referensi yang mendasarinya.

- Mencapai pemfilteran tingkat baris pada tabel menggunakan LF-Tags (di mana LF-Tags kaskade hanya sampai tingkat kolom) dengan menerapkan LF-Tags pada tampilan dan memberikan izin LF-Tags berbasis kepada pengguna.

Didukung AWS layanan analitik untuk tampilan

Layanan AWS analitik berikut mendukung pembuatan tampilan Katalog Data:

- Amazon Redshift
- Amazon Athena versi 3
- Apache Spark di EMR Tanpa Server
- Apache Spark pada versi 5.0 AWS Glue

Sumber daya tambahan

Anda dapat mempelajari lebih lanjut tentang Katalog Data dalam panduan ini, serta menggunakan sumber daya berikut:

Video berikut menunjukkan cara membuat tampilan dan menanyakannya dari Athena dan Amazon Redshift.

Topik

- Prasyarat untuk membuat tampilan
- Membuat tampilan Katalog Data menggunakan pernyataan DDL
- Membuat tampilan Katalog Data menggunakan AWS Glue API
- Memberikan izin pada tampilan Katalog Data
- Tampilan terwujud

Prasyarat untuk membuat tampilan

- Untuk membuat tampilan di Katalog Data, Anda harus mendaftarkan lokasi data Amazon S3 yang mendasari tabel referensi dengan Lake Formation. Untuk detail tentang mendaftarkan data dengan Lake Formation, lihatMenambahkan lokasi Amazon S3 ke danau data Anda.
- Hanya peran IAM yang dapat membuat tampilan Katalog Data. Identitas IAM lainnya tidak dapat membuat tampilan Katalog Data.

- Peran IAM yang mendefinisikan tampilan harus memiliki izin berikut:
  - `SELECT`Izin Lake Formation dengan `Grantable` opsi pada semua tabel referensi, semua kolom disertakan.
  - `CREATE_TABLE`Izin Lake Formation pada database target tempat tampilan sedang dibuat.
  - Kebijakan kepercayaan untuk Formasi Danau dan AWS Glue layanan untuk mengambil peran.

    JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "DataCatalogViewDefinerAssumeRole1",
            "Effect": "Allow",
            "Principal": {
               "Service": [
                    "glue.amazonaws.com",
                    "lakeformation.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

  - Iam: PassRole izin untuk AWS Glue dan 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 dan izin 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": "*"
        }
    ]
}
```

- Anda tidak dapat membuat tampilan dalam database yang memiliki `Super` atau `ALL` izin yang diberikan kepada `IAMAllowedPrincipals` grup. Anda dapat mencabut `Super` izin untuk `IAMAllowedPrincipals` grup pada database, melihatLangkah 4: Alihkan penyimpanan data Anda ke model izin Lake Formation, atau membuat database baru dengan Gunakan kontrol akses

IAM hanya untuk tabel baru dalam kotak database ini yang tidak dicentang di bawah Izin default untuk tabel yang baru dibuat.

Membuat tampilan Katalog Data menggunakan pernyataan DDL

Anda dapat membuat AWS Glue Data Catalog tampilan menggunakan editor SQL untuk Athena, Amazon Redshift, dan menggunakan API/. AWS Glue AWS CLI

Untuk membuat tampilan Katalog Data menggunakan editor SQL, pilih Athena atau Redshift Spectrum, dan buat tampilan menggunakan `CREATE VIEW` pernyataan Data Definition Language (DDL). Setelah membuat tampilan dalam dialek mesin pertama, Anda dapat menggunakan pernyataan `ALTER VIEW` DDL dari mesin kedua untuk menambahkan dialek tambahan.

Saat mendefinisikan pandangan, penting untuk mempertimbangkan hal-hal berikut:

- Mendefinisikan tampilan multi-dialek — Ketika Anda mendefinisikan tampilan dengan beberapa dialek, skema dialek yang berbeda harus cocok. Setiap dialek SQL akan memiliki spesifikasi sintaks yang sedikit berbeda. Sintaks kueri yang mendefinisikan tampilan Katalog Data harus diselesaikan ke daftar kolom yang sama persis, termasuk jenis dan nama, di semua dialek. Informasi ini disimpan `StorageDescriptor` dalam tampilan. Dialek juga harus mereferensikan objek tabel dasar yang sama dari Katalog Data.

  Untuk menambahkan dialek lain ke tampilan menggunakan DDL, Anda dapat menggunakan pernyataan. `ALTER VIEW` Jika `ALTER VIEW` pernyataan mencoba memperbarui definisi tampilan, seperti memodifikasi deskriptor penyimpanan atau tabel yang mendasari tampilan, pernyataan tersebut error dengan mengatakan “Input and existing storage descriptor mismatch”. Anda dapat menggunakan operasi cast SQL untuk memastikan bahwa jenis kolom tampilan cocok.
- Memperbarui tampilan - Untuk memperbarui tampilan, Anda dapat menggunakan `UpdateTable` API. Jika Anda memperbarui tampilan tanpa mencocokkan deskriptor penyimpanan atau tabel referensi, Anda dapat memberikan `FORCE` bendera (lihat dokumentasi SQL mesin untuk sintaks). Setelah pembaruan paksa, tampilan akan mengambil tabel paksa `StorageDescriptor` dan referensi. Setiap `ALTER VIEW` DDL lebih lanjut harus sesuai dengan nilai yang dimodifikasi. Tampilan yang telah diperbarui untuk memiliki dialek yang tidak kompatibel akan berada dalam status “Basi”. Status tampilan terlihat di konsol Lake Formation dan menggunakan `GetTable` operasi.
- Merujuk jenis kolom varchar sebagai string - Tidak mungkin untuk melemparkan jenis kolom varchar dari Redshift Spectrum ke string. Jika tampilan dibuat di Redshift Spectrum dengan tipe

kolom varchar dan dialek berikutnya mencoba mereferensikan bidang itu sebagai string, Katalog Data akan memperlakukannya sebagai string tanpa perlu bendera. `FORCE`

- Perawatan bidang tipe kompleks — Amazon Redshift memperlakukan semua tipe kompleks sebagai tipe SUPER sementara Athena menentukan tipe kompleks. Jika tampilan memiliki bidang `SUPER` tipe, dan mesin lain mereferensikan kolom tersebut sebagai tipe kompleks tertentu, seperti struct (`<street_address:struct<street_number:int, street_name:string, street_type:string>>`), Katalog Data mengasumsikan bahwa bidang tersebut menjadi tipe kompleks tertentu, dan menggunakannya dalam deskriptor penyimpanan, tanpa memerlukan tanda. `Force`

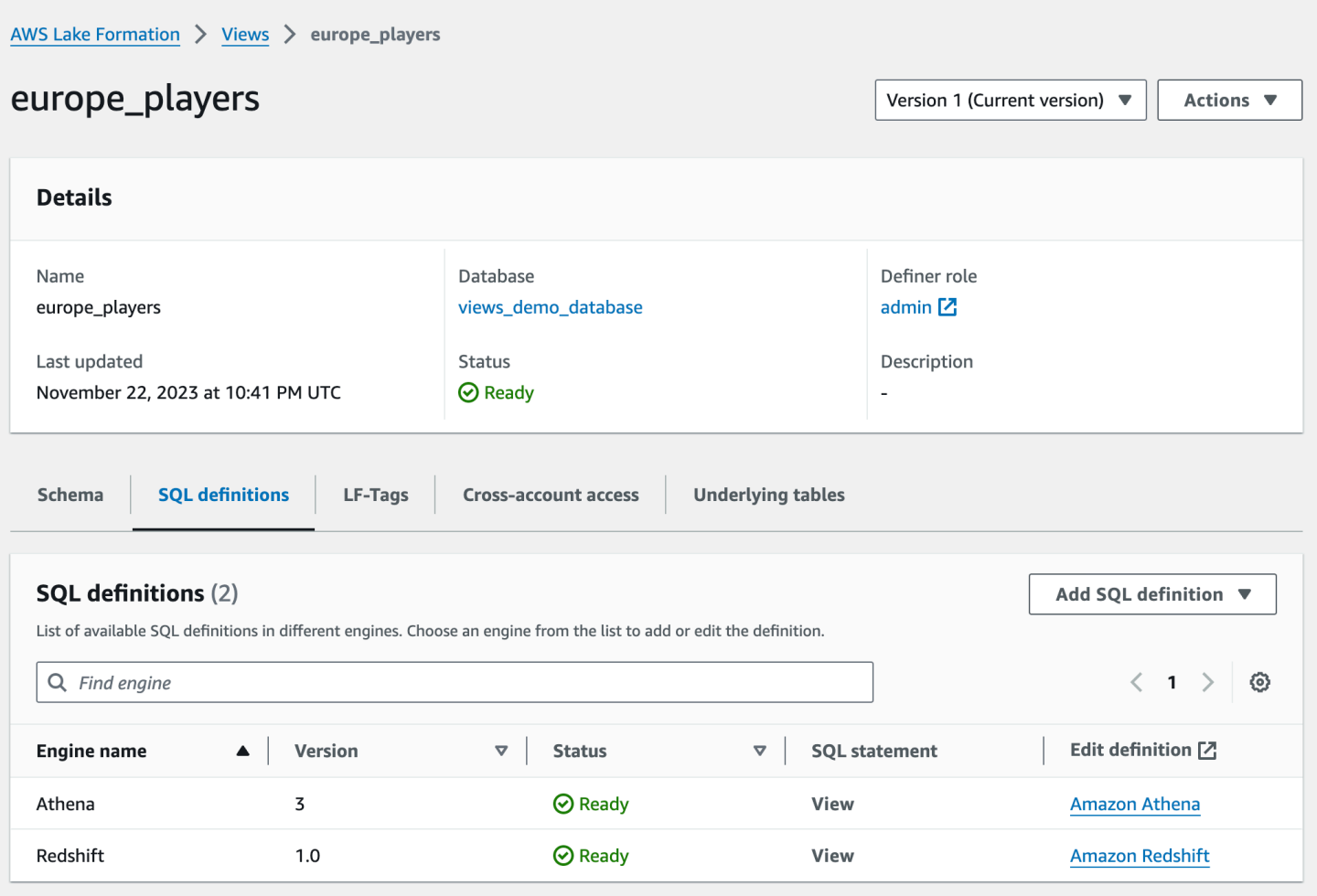
Untuk informasi selengkapnya tentang sintaks untuk membuat dan mengelola tampilan Katalog Data, lihat:

- Menggunakan AWS Glue Data Catalog tampilan di Panduan Pengguna Amazon Athena.
- Glue Data Catalog melihat sintaks kueri di Panduan Pengguna Amazon Athena.
- Membuat tampilan AWS Glue Data Catalog di Panduan Pengembang Database Amazon Redshift.

  Untuk informasi selengkapnya tentang perintah SQL yang terkait dengan tampilan di Katalog Data, lihat MEMBUAT TAMPILAN EKSTERNAL, MENGUBAH TAMPILAN EKSTERNAL, dan MENJATUHKAN TAMPILAN EKSTERNAL.

Setelah Anda membuat tampilan Katalog Data, detail tampilan tersedia di konsol Lake Formation.

1. Pilih Tampilan di bawah Katalog Data di konsol Lake Formation.
2. Daftar tampilan yang tersedia muncul di halaman tampilan.
3. Pilih tampilan dari daftar dan halaman detail menunjukkan atribut tampilan.



Skema

Pilih `Column` baris, dan pilih Edit LF-Tags untuk memperbarui nilai tag atau menetapkan baru LF-Tags.

Definisi SQL

Anda dapat melihat daftar definisi SQL yang tersedia. Pilih Tambahkan definisi SQL, dan pilih mesin kueri untuk menambahkan definisi SQL. Pilih mesin kueri (Athena atau Amazon Redshift) di bawah kolom untuk memperbarui `Edit definition` definisi SQL.

LF-Tags

Pilih Edit LF-Tags untuk mengedit nilai tag atau menetapkan tag baru. Anda dapat menggunakan LF-Tags untuk memberikan izin pada tampilan.

Cross-account akses

Anda dapat melihat daftar Akun AWS, organisasi, dan unit organisasi (OU) yang telah Anda bagikan tampilan Katalog Data.

Tabel yang mendasari

Tabel yang mendasari direferensikan dalam definisi SQL yang digunakan untuk membuat tampilan ditampilkan di bawah tab ini.

Membuat tampilan Katalog Data menggunakan AWS Glue API

Anda dapat menggunakan AWS Glue CreateTable, dan UpdateTableAPI untuk membuat dan memperbarui tampilan di Katalog Data. `UpdateTable`Operasi `CreateTable` dan memiliki `TableInput` struktur baru untuk`ViewDefinition`, sementara,`SearchTables`,`GetTable`, `GetTables``GetTableVersion`, `GetTableVersions` operasi menyediakan sintaks keluaran mereka untuk tampilan. `ViewDefinition` Selain itu, ada `Status` bidang baru dalam output `GetTable` API.

Dua AWS Glue koneksi baru tersedia untuk memvalidasi dialek SQL untuk setiap mesin kueri yang didukung, dan Amazon Amazon Athena Redshift.

`UpdateTable`API `CreateTable` dan bersifat asinkron saat digunakan dengan tampilan. Ketika API ini dipanggil dengan beberapa dialek SQL, panggilan divalidasi dengan setiap mesin untuk menentukan apakah dialek dapat dijalankan pada mesin itu, dan apakah skema tampilan yang dihasilkan dari masing-masing dialek cocok. AWS Glue Layanan ini menggunakan koneksi ini untuk melakukan panggilan internal ke mesin analitik. Panggilan ini mensimulasikan apa yang mesin lakukan untuk memvalidasi jika `CREATE VIEW` atau `ALTER VIEW` SQL DDL dijalankan pada mesin.

Jika SQL yang disediakan valid, dan skema cocok antara dialek tampilan, AWS Glue API secara atomik melakukan hasilnya. Atomicity memungkinkan tampilan dengan beberapa dialek dibuat atau diubah tanpa downtime.

Topik

- Membuat AWS Glue koneksi untuk memvalidasi status
- Memvalidasi status pembuatan tampilan
- Status dan operasi asinkron
- Lihat skenario kegagalan pembuatan selama operasi asinkron

Membuat AWS Glue koneksi untuk memvalidasi status

Untuk membuat atau memperbarui AWS Glue Data Catalog tampilan menggunakan `CreateTable` atau `UpdateTable` operasi, Anda harus membuat jenis AWS Glue koneksi baru untuk validasi, dan menyediakannya ke mesin analitik yang didukung. Koneksi ini diperlukan untuk menggunakan tampilan Katalog Data dengan Athena atau Amazon Redshift. Anda dapat membuat koneksi ini hanya menggunakan AWS CLI, AWS SDK, atau AWS Glue API. Anda tidak dapat menggunakan Konsol Manajemen AWS untuk membuat AWS Glue koneksi.

> **Note**
>
> Jika peran penentu tampilan dan pemanggilan peran `CreateTable` atau `UpdateTable` berbeda, maka keduanya memerlukan `glue:PassConnection` izin dalam pernyataan kebijakan IAM mereka.

Untuk informasi selengkapnya, lihat dokumentasi create-connection AWS CLI .

AWS CLI perintah untuk membuat koneksi

Berikut ini adalah AWS CLI perintah untuk membuat koneksi:

```
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 masukan JSON

Untuk 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"
        }
```

```
    }
}
```

Untuk 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"
        }
    }
}
```

Memvalidasi status pembuatan tampilan

Saat Anda menjalankan `UpdateTable` operasi `CreateTable` atau, `Status` bidang untuk output `GetTable` API menunjukkan detail status pembuatan tampilan. Untuk `create` permintaan di mana tabel belum ada, AWS Glue buat tabel kosong selama proses asinkron. Saat meneleponGetTable, Anda dapat meneruskan flag boolean opsionalIncludeStatusDetails, yang menunjukkan informasi diagnostik tentang permintaan tersebut. Dalam kasus kegagalan, bendera ini menunjukkan pesan kesalahan dengan status individual dari setiap dialek.

Kesalahan selama operasi pembuatan, baca, perbarui, dan hapus (CRUD) tampilan dapat terjadi baik selama pemrosesan di layanan AWS Glue/Lake Formation atau selama validasi SQL tampilan di Amazon Redshift atau Athena. Ketika kesalahan terjadi selama validasi di mesin, AWS Glue layanan memberikan pesan kesalahan bahwa mesin kembali.

Bidang status

Berikut ini adalah bidang status:

- Status: status generik, yang agnostik untuk berbagai jenis pekerjaan:
  - DIANTREKAN
  - IN_PROGRESS
  - BERHASIL
  - FAILED

- Tindakan - Menunjukkan tindakan mana yang dipanggil di atas meja, saat ini hanya `CREATE` atau `UPDATE` operasi yang tersedia.

  Membedakan antara `UPDATE` dan `CREATE` operasi penting saat bekerja dengan pandangan. Jenis operasi menentukan bagaimana Anda harus melanjutkan dengan query tabel.

  `UPDATE`Operasi menandakan bahwa tabel sudah ada di Katalog Data. Dalam hal ini, Anda dapat melanjutkan kueri tabel yang dibuat sebelumnya tanpa masalah. Di sisi lain, `CREATE` operasi menunjukkan bahwa tabel belum pernah berhasil dibuat sebelumnya. Jika tabel ditandai untuk`CREATE`, mencoba untuk query itu akan gagal karena tabel belum ada dalam sistem. Oleh karena itu, penting untuk mengidentifikasi jenis operasi (UPDATE atau CREATE) sebelum mencoba untuk query tabel.
- RequestedBy — ARN pengguna yang meminta perubahan asinkron.
- UpdatedBy ARN pengguna yang terakhir mengubah proses perubahan asinkron secara manual, seperti meminta pembatalan atau modifikasi.
- Kesalahan - Bidang ini hanya muncul ketika status GAGAL. Ini adalah pesan pengecualian tingkat induk. Mungkin ada kesalahan yang berbeda untuk setiap dialek.
  - ErrorCode Jenis pengecualian.
  - ErrorMessage — deskripsi singkat tentang pengecualian.
- RequestTime — string tanggal berformat ISO 8601 yang menunjukkan waktu perubahan dimulai.
- UpdateTime — string tanggal berformat ISO 8601 yang menunjukkan waktu status terakhir diperbarui.

Status dan operasi asinkron

Saat Anda menjalankan `glue:CreateTable` permintaan, pembuatan asinkron tampilan Katalog Data dimulai. Di bagian berikut, dokumen ini menjelaskan AWS Glue tampilan yang tersedia dalam `glue:GetTable` respons. `Status` Untuk singkatnya, bagian ini menghilangkan respons penuh.

```
{
    "Table": {
        ...
        "Status": {
            ...
            "Action": "CREATE",
            "State": "QUEUED",
        }
```

```
    }
}
```

Kedua atribut di atas mewakili informasi diagnostik penting yang menunjukkan keadaan operasi asinkron, serta tindakan yang dapat dilakukan pada tampilan ini. Di bawah ini adalah nilai yang mungkin yang dapat diambil oleh atribut ini.

1. `Status.Action`
   a. CREATE
   b. UPDATE
2. `Status.State`
   a. DIANTREKAN
   b. IN_PROGRESS
   c. BERHASIL
   d. FAILED

Penting juga untuk dicatat bahwa beberapa pembaruan pada tampilan Katalog Data tidak memerlukan operasi asinkron. Misalnya, seseorang mungkin ingin memperbarui `Description` atribut tabel. Karena ini tidak memerlukan operasi asinkron, metadata tabel yang dihasilkan tidak akan memilikinya`Status`, dan atributnya akan. `NULL`

```
{
    "Table": {
        ...,
        "Description": "I changed this attribute!"
    }
}
```

Selanjutnya, topik ini mengeksplorasi bagaimana informasi status di atas dapat memengaruhi operasi yang dapat dilakukan pada suatu AWS Glue tampilan.

Iem: CreateTable

Tidak ada perubahan untuk API ini jika dibandingkan dengan bagaimana `glue:CreateTable` fungsi untuk tabel Glue apa pun. `CreateTable`dapat dipanggil untuk nama tabel apa pun yang belum ada.

Iem: UpdateTable

Operasi ini tidak dapat dilakukan pada AWS Glue tampilan yang memiliki informasi status berikut:

1. Tindakan == CREATE dan State == ANTRIAN
2. Tindakan == BUAT dan Status == IN_PROGRESS
3. Tindakan == BUAT dan status == GAGAL
4. Tindakan == PEMBARUAN dan status == ANTRIAN
5. Tindakan == PEMBARUAN dan status == IN_PROGRESS

Untuk meringkas, Anda dapat memperbarui tampilan Katalog Data hanya jika memenuhi persyaratan berikut.

1. Ini telah berhasil dibuat untuk pertama kalinya.
   a. Tindakan == CREATE dan State == SUKSES
2. Ini telah mencapai status terminal setelah operasi pembaruan asinkron.
   a. Tindakan == PEMBARUAN dan Status == SUKSES
   b. Tindakan == PEMBARUAN dan Status == GAGAL
3. Ini memiliki atribut `NULL` status sebagai hasil dari pembaruan sinkron.

lem: DeleteTable

Tidak ada perubahan untuk operasi ini jika dibandingkan dengan bagaimana `glue:DeleteTable` fungsi untuk AWS Glue tabel apa pun. Anda dapat menghapus tampilan Katalog Data terlepas dari statusnya.

lem: GetTable

Tidak ada perubahan untuk operasi ini jika dibandingkan dengan bagaimana `glue:GetTable` fungsi untuk AWS Glue tabel apa pun. Namun, Anda tidak dapat menanyakan tampilan Katalog Data dari mesin analitik hingga berhasil dibuat untuk pertama kalinya. `Action == CREATE and State == SUCCESS`. Setelah berhasil membuat tampilan Katalog Data untuk pertama kalinya, Anda dapat melakukan kueri tampilan terlepas dari statusnya.

> **Note**
>
> Semua informasi di bagian ini berlaku untuk semua API baca tabel seperti`GetTable`,`GetTables`, dan`SearchTables`.

Lihat skenario kegagalan pembuatan selama operasi asinkron

Contoh berikut mewakili jenis kesalahan yang mungkin timbul dari `CreateTable` atau `UpdateTable` melihat panggilan API. Mereka tidak lengkap karena permukaan kesalahan kegagalan kueri SQL cukup besar.

Skenario 1: Kegagalan kueri Amazon Redshift

Kueri yang disediakan untuk Amazon Redshift menyertakan nama tabel yang salah eja tidak dapat ditemukan di Katalog Data selama validasi. Kesalahan yang dihasilkan ditampilkan di `Status` bidang dalam `GetTable` respons untuk tampilan.

`GetTable`permintaan:

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

GetTablerespon:

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

Skenario 2: Koneksi Amazon Redshift tidak valid

Koneksi Amazon Redshift dalam contoh berikut salah bentuk karena mengacu pada database Amazon Redshift yang tidak ada di titik akhir yang disediakan. cluster/serverless Amazon Redshift tidak dapat memvalidasi tampilan dan `Status` bidang dalam `GetTable` respons menunjukkan kesalahan (dari Amazon `"State": "FAILED"` Redshift.

`GetTable`permintaan:

```
{
```

```
    "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/GDCViewDdefiner",
            "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"
                }
            ]
        }
    }
}
```

`GetTable`respon:

```
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 \"devooo\" 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
\".\"table_1\"",
                        "UpdateTime": "2024-07-11T11:43:40-07:00",
                        "State": "SUCCESS"
                    },
                    {
                        "Dialect": "REDSHIFT",
                        "DialectVersion": "1.0",
                        "ViewValidationText": "SELECT * FROM \"gdc-view-playground-
external-schema\".\"table_1\";",
                        "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"
                        }
                    }
                ]
            }
        }
    }
}
```

Skenario 3: Kegagalan kueri Athena

SQL untuk Athena di sini tidak valid karena kueri salah mengeja nama database. Validasi kueri Athena menangkap ini dan kesalahan yang dihasilkan muncul melalui objek dalam `Status` panggilan. `GetTable`

`GetTable`permintaan:

```
{
    "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/GDCViewDdefiner",
            "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"
                }
            ]
        }
    }
}
```

`GetTable`respon:

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

Skenario 4: Deskriptor penyimpanan ketidakcocokan

SQL yang disediakan untuk dialek Athena memilih `col2` dan sementara SQL untuk Redshift hanya `col1` memilih. `col1` Hal ini menyebabkan kesalahan ketidakcocokan deskriptor penyimpanan.

`GetTable`permintaan:

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

`GetTable`respon:

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

Memberikan izin pada tampilan Katalog Data

Setelah membuat tampilan di AWS Glue Data Catalog, Anda dapat memberikan izin data lake pada tampilan ke kepala sekolah di seluruh Akun AWS, organisasi, dan unit organisasi. Anda dapat memberikan izin menggunakan LF-Tags atau metode sumber daya bernama. Untuk informasi selengkapnya tentang penandaan sumber daya, lihatKontrol akses berbasis tag Lake Formation. Untuk informasi selengkapnya tentang pemberian izin pada tampilan secara langsung, lihat. Memberikan izin pada tampilan menggunakan metode sumber daya bernama

Tampilan terwujud

Topik

- Membedakan tampilan terwujud dari jenis tampilan lain
- Kasus penggunaan
- Konsep utama
- Izin untuk tampilan terwujud
- Membuat dan mengelola tampilan yang terwujud

- Penyimpanan dan akses data
- Integrasi dengan AWS Lake Formation izin
- Pemantauan dan debugging
- Mengelola pekerjaan penyegaran
- Pemantauan dan pemecahan masalah
- Pertimbangan dan batasan

Dalam Katalog AWS Glue Data, tampilan terwujud adalah tabel terkelola yang menyimpan hasil yang telah dihitung sebelumnya dari kueri SQL dalam format Apache Iceberg. Tidak seperti tampilan Katalog Data standar yang mengeksekusi kueri setiap kali diakses, tampilan terwujud secara fisik menyimpan hasil kueri dan memperbaruinya saat tabel sumber yang mendasarinya berubah. Anda dapat membuat tampilan terwujud menggunakan Apache Spark versi 3.5.6+ di Amazon Athena, Amazon EMR, atau. AWS Glue

Tampilan terwujud mereferensikan tabel Apache Iceberg yang terdaftar di Katalog AWS Glue Data, dengan data yang telah dihitung sebelumnya disimpan sebagai tabel Apache Iceberg di bucket Amazon S3 Tables atau bucket tujuan umum Amazon S3, membuatnya dapat diakses dari beberapa mesin kueri termasuk Amazon Athena, Amazon Redshift, dan mesin pihak ketiga. Iceberg-compatible

Membedakan tampilan terwujud dari jenis tampilan lain

Tampilan terwujud berbeda dari tampilan Katalog AWS Glue Data, tampilan Apache Spark, dan tampilan Amazon Athena dengan cara mendasar. Sementara tampilan Katalog Data adalah tabel virtual yang mengeksekusi definisi kueri SQL setiap kali diakses, tampilan terwujud secara fisik menyimpan hasil kueri yang telah dihitung sebelumnya. Ini menghilangkan komputasi yang berlebihan dan secara signifikan meningkatkan kinerja kueri untuk transformasi kompleks yang sering diakses.

Tampilan terwujud juga berbeda dari jalur transformasi data tradisional yang dibangun dengan AWS Glue ETL atau pekerjaan Spark khusus. Alih-alih menulis kode khusus untuk menangani deteksi perubahan, pembaruan tambahan, dan orkestrasi alur kerja, Anda menentukan tampilan terwujud menggunakan sintaks SQL standar. Katalog AWS Glue Data secara otomatis memonitor tabel sumber, mendeteksi perubahan, dan menyegarkan tampilan terwujud menggunakan infrastruktur komputasi yang dikelola sepenuhnya.

Kasus penggunaan

Berikut ini adalah kasus penggunaan penting untuk tampilan terwujud:

- Mempercepat kueri analitik yang kompleks — Buat tampilan terwujud yang menghitung sebelumnya gabungan, agregasi, dan fungsi jendela yang mahal. Spark engine secara otomatis menulis ulang kueri berikutnya untuk menggunakan hasil yang telah dihitung sebelumnya, mengurangi latensi kueri dan biaya komputasi.
- Sederhanakan jalur transformasi data — Ganti pekerjaan ETL kompleks yang menangani deteksi perubahan, pembaruan tambahan, dan orkestrasi alur kerja dengan definisi tampilan terwujud sederhana. SQL-based Katalog AWS Glue Data mengelola semua kompleksitas operasional secara otomatis.
- Aktifkan analitik swalayan dengan akses data yang diatur — Buat tampilan terwujud yang dikuratori yang mengubah data mentah menjadi kumpulan data siap bisnis. Berikan pengguna akses ke tampilan yang terwujud tanpa mengekspos tabel sumber yang mendasarinya, menyederhanakan manajemen keamanan sekaligus memberdayakan analitik swalayan.
- Optimalkan rekayasa fitur untuk pembelajaran mesin — Tentukan tampilan terwujud yang menerapkan transformasi fitur untuk model ML. Kemampuan penyegaran otomatis memastikan penyimpanan fitur tetap terkini saat data sumber berkembang, sementara penyegaran tambahan meminimalkan biaya komputasi.
- Menerapkan berbagi data yang efisien — Buat tampilan terwujud yang menyaring dan mengubah data untuk konsumen tertentu. Bagikan tampilan terwujud di seluruh akun dan wilayah yang menggunakan AWS Lake Formation, menghilangkan kebutuhan akan duplikasi data sambil mempertahankan tata kelola terpusat.

Konsep utama

Penyegaran otomatis

Penyegaran otomatis adalah kemampuan yang terus memantau tabel sumber Anda dan memperbarui tampilan yang terwujud sesuai dengan jadwal yang Anda tentukan. Saat membuat tampilan terwujud, Anda dapat menentukan frekuensi penyegaran menggunakan penjadwalan berbasis waktu dengan interval sesering satu jam. Katalog AWS Glue Data menggunakan infrastruktur komputasi Spark terkelola untuk menjalankan operasi penyegaran di latar belakang, menangani semua aspek deteksi perubahan dan pembaruan tambahan secara transparan.

Saat data sumber berubah di antara interval penyegaran, tampilan yang terwujud menjadi basi sementara. Kueri yang secara langsung mengakses tampilan terwujud dapat mengembalikan hasil yang sudah ketinggalan zaman hingga penyegaran terjadwal berikutnya selesai. Untuk skenario yang membutuhkan akses langsung ke data terbaru, Anda dapat menjalankan penyegaran manual menggunakan perintah `REFRESH MATERIALIZED VIEW` SQL.

Penyegaran tambahan

Penyegaran tambahan adalah teknik pengoptimalan yang hanya memproses data yang telah berubah dalam tabel sumber sejak penyegaran terakhir, daripada menghitung ulang seluruh tampilan yang terwujud. Katalog AWS Glue Data memanfaatkan lapisan metadata Apache Iceberg untuk melacak perubahan tabel sumber secara efisien dan menentukan bagian mana dari tampilan terwujud yang memerlukan pembaruan.

Pendekatan ini secara signifikan mengurangi biaya komputasi dan durasi penyegaran dibandingkan dengan operasi penyegaran penuh, terutama untuk kumpulan data besar di mana hanya sebagian kecil perubahan data di antara siklus penyegaran. Mekanisme penyegaran inkremental beroperasi secara otomatis; Anda tidak perlu menulis logika khusus untuk mendeteksi atau memproses data yang diubah.

Penulisan ulang kueri otomatis

Penulisan ulang kueri otomatis adalah kemampuan pengoptimalan kueri yang tersedia di mesin Spark di Amazon Athena, Amazon EMR, dan. AWS Glue Saat Anda menjalankan kueri terhadap tabel dasar, pengoptimal Spark menganalisis rencana kueri Anda dan secara otomatis menentukan apakah tampilan terwujud yang tersedia dapat memenuhi kueri dengan lebih efisien. Jika tampilan terwujud yang sesuai ada, pengoptimal secara transparan menulis ulang kueri untuk menggunakan hasil yang telah dihitung sebelumnya alih-alih memproses tabel dasar.

Pengoptimalan ini terjadi tanpa memerlukan perubahan apa pun pada kode aplikasi atau pernyataan kueri Anda. Pengoptimal Spark memastikan bahwa penulisan ulang kueri otomatis hanya berlaku ketika tampilan terwujud saat ini dan dapat menghasilkan hasil yang akurat. Jika tampilan terwujud sudah basi atau tidak sepenuhnya sesuai dengan persyaratan kueri, pengoptimal akan mengeksekusi rencana kueri asli terhadap tabel dasar, memprioritaskan kebenaran daripada kinerja.

Lihat peran yang lebih pasti

Tampilan terwujud beroperasi berdasarkan izin peran IAM yang membuatnya, yang dikenal sebagai peran penentu tampilan. Peran definer harus memiliki akses baca ke semua tabel dasar yang direferensikan dalam definisi tampilan terwujud dan membuat izin tabel pada database target. Ketika Katalog AWS Glue Data menyegarkan tampilan terwujud, ia akan mengasumsikan peran definer untuk mengakses tabel sumber dan menulis hasil yang diperbarui.

Model keamanan ini memungkinkan Anda memberi pengguna akses ke tampilan terwujud tanpa memberi mereka izin langsung pada tabel sumber yang mendasarinya. Jika peran penentu tampilan

kehilangan akses ke tabel dasar apa pun, operasi penyegaran berikutnya akan gagal hingga izin dipulihkan.

Izin untuk tampilan terwujud

Untuk membuat dan mengelola tampilan terwujud, Anda harus mengonfigurasi AWS Lake Formation izin. Peran IAM yang menciptakan tampilan terwujud (peran definer) memerlukan izin khusus pada tabel sumber dan database target.

Izin yang diperlukan untuk peran definer

Peran definer harus memiliki izin Lake Formation berikut:

- Pada tabel sumber - PILIH atau SEMUA izin tanpa filter baris, kolom, atau sel
- Pada database target - izin CREATE_TABLE
- Pada Katalog AWS Glue Data — GetTable dan izin CreateTable API

Saat Anda membuat tampilan terwujud, ARN peran definer disimpan dalam definisi tampilan. Katalog AWS Glue Data mengasumsikan peran ini saat menjalankan operasi penyegaran otomatis. Jika peran definer kehilangan akses ke tabel sumber, operasi refresh akan gagal hingga izin dipulihkan.

Izin IAM untuk AWS Glue pekerjaan

Peran IAM AWS Glue pekerjaan Anda memerlukan izin berikut:

```
{
    "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": "*"
        }
    ]
}
```

Peran yang Anda gunakan untuk penyegaran otomatis Tampilan Materialisasi harus memiliki PassRole izin iam: pada peran tersebut.

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

Agar Glue secara otomatis menyegarkan tampilan terwujud untuk Anda, peran tersebut juga harus memiliki kebijakan kepercayaan berikut yang memungkinkan layanan untuk mengambil peran tersebut.

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "glue.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Jika Tampilan Materialisasi disimpan di Bucket Tabel S3, Anda juga perlu menambahkan izin berikut ke peran.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3tables:PutTableMaintenanceConfiguration"
      ],
      "Resource": "arn:aws:s3tables:*:123456789012:*"
    }
  ]
}
```

Memberikan akses ke tampilan yang terwujud

Untuk memberi pengguna lain akses ke kueri tampilan terwujud, gunakan AWS Lake Formation untuk memberikan izin SELECT pada tabel tampilan terwujud. Pengguna dapat menanyakan tampilan terwujud tanpa memerlukan akses langsung ke tabel sumber yang mendasarinya.

Untuk informasi lebih lanjut tentang mengonfigurasi izin Lake Formation, lihat Memberikan dan mencabut izin pada sumber daya Katalog Data di Panduan Pengembang. AWS Lake Formation

Membuat dan mengelola tampilan yang terwujud

Anda membuat tampilan terwujud menggunakan pernyataan `CREATE MATERIALIZED VIEW` SQL di mesin Spark. Definisi tampilan menentukan kueri SQL yang mendefinisikan logika transformasi, database target dan nama tabel, dan konfigurasi penyegaran opsional. Anda dapat menentukan transformasi kompleks termasuk agregasi, bergabung di beberapa tabel, filter, dan fungsi jendela.

```
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;
```

Untuk mengonfigurasi penyegaran otomatis, sertakan jadwal penyegaran dalam definisi tampilan Anda:

```
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;
```

Anda dapat menyegarkan tampilan terwujud secara manual kapan saja menggunakan `REFRESH MATERIALIZED VIEW` perintah:

```
REFRESH MATERIALIZED VIEW sales_summary;
```

Untuk mengubah jadwal penyegaran tampilan terwujud yang ada, gunakan `ALTER MATERIALIZED VIEW` pernyataan:

```
ALTER MATERIALIZED VIEW sales_summary
ADD SCHEDULE REFRESH EVERY 2 HOURS;
```

Pandangan terwujud bersarang

Anda dapat membuat tampilan terwujud yang mereferensikan tampilan terwujud lainnya sebagai tabel dasar, memungkinkan transformasi data multi-tahap. Saat Anda membuat tampilan terwujud bersarang, Katalog AWS Glue Data melacak dependensi dan secara otomatis menyebarkan pembaruan melalui hierarki tampilan terwujud. Saat tampilan terwujud dasar diperbarui, semua tampilan terwujud hilir yang bergantung padanya akan diperbarui sesuai dengan itu.

Kemampuan ini memungkinkan Anda untuk menguraikan transformasi kompleks menjadi tahapan logis, meningkatkan pemeliharaan dan memungkinkan penyegaran selektif lapisan transformasi berdasarkan persyaratan kesegaran data Anda.

Penyimpanan dan akses data

Tampilan terwujud menyimpan hasil yang telah dihitung sebelumnya sebagai tabel Apache Iceberg di bucket Tabel S3 atau bucket S3 tujuan umum dalam akun Anda. AWS Katalog AWS Glue Data

mengelola semua aspek pemeliharaan tabel Iceberg, termasuk pemadatan dan retensi snapshot, melalui kemampuan pengoptimalan otomatis S3 Tables.

Karena tampilan terwujud disimpan sebagai tabel Iceberg, Anda dapat membacanya langsung dari Iceberg-compatible mesin apa pun, termasuk Amazon Athena, Amazon Redshift, dan platform analitik pihak ketiga. Aksesibilitas multi-mesin ini memastikan data yang telah dihitung sebelumnya tetap dapat diakses di seluruh ekosistem analitik Anda tanpa duplikasi data atau konversi format.

Integrasi dengan AWS Lake Formation izin

Anda dapat menggunakannya AWS Lake Formation untuk mengelola izin berbutir halus pada tampilan terwujud. Pembuat tampilan secara otomatis menjadi pemilik tampilan terwujud dan dapat memberikan izin kepada pengguna atau peran lain menggunakan metode sumber AWS Lake Formation daya bernama atau. LF-Tags

Saat Anda memberikan `SELECT` izin kepada pengguna pada tampilan terwujud, mereka dapat menanyakan hasil yang telah dihitung sebelumnya tanpa memerlukan akses ke tabel sumber yang mendasarinya. Model keamanan ini menyederhanakan manajemen akses data dan memungkinkan Anda menerapkan prinsip hak istimewa paling sedikit, memberi pengguna akses hanya ke transformasi data spesifik yang mereka butuhkan.

Anda dapat membagikan tampilan terwujud di seluruh AWS akun, AWS organisasi, dan unit organisasi menggunakan AWS Lake Formation kemampuan berbagi lintas akun. Anda juga dapat mengakses tampilan terwujud di seluruh AWS Wilayah menggunakan tautan sumber daya, memungkinkan tata kelola data terpusat dengan akses data terdistribusi.

Pemantauan dan debugging

Katalog AWS Glue Data menerbitkan semua operasi penyegaran tampilan terwujud dan metrik terkait ke Amazon. CloudWatch Anda dapat memantau waktu mulai penyegaran, waktu akhir, durasi, volume data yang diproses, dan menyegarkan status melalui CloudWatch metrik. Ketika operasi penyegaran gagal, pesan kesalahan dan informasi diagnostik ditangkap di CloudWatch Log.

Anda dapat mengatur CloudWatch alarm untuk menerima notifikasi saat penyegaran pekerjaan melebihi durasi yang diharapkan atau gagal berulang kali. Katalog AWS Glue Data juga menerbitkan peristiwa perubahan untuk penyegaran yang berhasil dan gagal, memungkinkan Anda mengintegrasikan operasi tampilan terwujud ke dalam otomatisasi alur kerja yang lebih luas.

Untuk memeriksa status tampilan terwujud saat ini, gunakan perintah `DESCRIBE MATERIALIZED VIEW` SQL, yang menampilkan metadata termasuk status kebuntuan, stempel waktu penyegaran terakhir, dan konfigurasi jadwal penyegaran.

Mengelola pekerjaan penyegaran

Memulai penyegaran manual

Picu penyegaran langsung di luar interval yang dijadwalkan.

Izin yang Diperlukan: AWS Kredensil yang digunakan untuk melakukan panggilan API harus memiliki `glue:GetTable` izin untuk tampilan yang terwujud.

Untuk Katalog Tabel S3:

```
aws glue start-materialized-view-refresh-task-run \
    --catalog-id <ACCOUNT_ID>:s3tablescatalog/<CATALOG_NAME> \
    --database-name <DATABASE_NAME> \
    --table-name <MV_TABLE_NAME>
```

Untuk Katalog Root:

```
aws glue start-materialized-view-refresh-task-run \
    --catalog-id <ACCOUNT_ID> \
    --database-name <DATABASE_NAME> \
    --table-name <MV_TABLE_NAME>
```

Memeriksa status penyegaran

Dapatkan status pekerjaan penyegaran tertentu:

```
aws glue get-materialized-view-refresh-task-run \
    --catalog-id <CATALOG_ID> \
    --materialized-view-refresh-task-run-id <TASK_RUN_ID>
```

Daftar riwayat penyegaran

Lihat semua pekerjaan penyegaran untuk tampilan terwujud:

```
aws glue list-materialized-view-refresh-task-runs \
    --catalog-id <CATALOG_ID> \
    --database-name <DATABASE_NAME> \
    --table-name <MV_TABLE_NAME>
```

> **Note**
>
> Gunakan <ACCOUNT_ID>:s3tablescatalog/<CATALOG_NAME> untuk Tabel S3 atau <ACCOUNT_ID> untuk katalog root.

Menghentikan penyegaran yang sedang berjalan

Membatalkan pekerjaan penyegaran yang sedang berlangsung:

```
aws glue stop-materialized-view-refresh-task-run \
    --catalog-id <CATALOG_ID> \
    --database-name <DATABASE_NAME> \
    --table-name <MV_TABLE_NAME>
```

Pemantauan dan pemecahan masalah

Ada tiga cara untuk memantau pekerjaan penyegaran tampilan yang terwujud:

CloudWatch Metrik

Lihat metrik agregat untuk semua pekerjaan penyegaran tampilan terwujud Anda di: CloudWatch

Metrik yang Tersedia:

- AWS/Lem namespace dengan dimensi:
  - CatalogId: Pengidentifikasi katalog Anda
  - DatabaseName: Database yang berisi tampilan terwujud
  - TableName: Nama tampilan terwujud
  - TaskType: Setel ke "MaterializedViewRefresh”

Melihat di Konsol:

1. Arahkan ke CloudWatch Konsol → Metrik
2. AWS Pilih/Lem namespace
3. Filter berdasarkan dimensi: CatalogId, DatabaseName, TableName, TaskType
4. Lihat metrik untuk keberhasilan, kegagalan, dan durasi pekerjaan

Contoh Kueri CloudWatch Metrik:

```
{AWS/Glue,CatalogId,DatabaseName,TableName,TaskType} MaterializedViewRefresh
```

Menggunakan 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 Log

Lihat log eksekusi terperinci untuk menjalankan tugas penyegaran individu:

Grup Log: `/aws-glue/materialized-views/<task_run_id>`

Dimana `<task_run_id>` adalah UUID (misalnya, abc12345-def6-7890-ghij-klmnopqrstuv).

Melihat Log:

```
# 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>
```

Di CloudWatch Konsol:

1. Arahkan ke CloudWatch → Grup log
2. Cari /aws- -views/ glue/materialized
3. Pilih grup log dengan ID task run Anda
4. Lihat log eksekusi terperinci, kesalahan, dan output pekerjaan Spark

Notifikasi

Berlangganan acara untuk pemberitahuan waktu nyata tentang menyegarkan perubahan status pekerjaan:

Jenis Acara yang Tersedia:

- Glue Materialized View Refresh Task Dimulai
- Glue Materialized View Refresh Task Berhasil
- Gagal Glue Materialized View Refresh Task
- Glue Materialized View Gagal Auto-Refresh Pemanggilan

Membuat Aturan:

```
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>
```

Menambahkan Target (misalnya, Topik SNS):

```
aws events put-targets \
    --rule materialized-view-refresh-notifications \
    --targets "Id"="1","Arn"="arn:aws:sns:<REGION>:<ACCOUNT_ID>:<TOPIC_NAME>" \
    --region <REGION>
```

Melihat status penyegaran

Periksa status pekerjaan penyegaran tampilan terwujud Anda menggunakan AWS Glue API:

```
aws glue get-materialized-view-refresh-task-run \
    --catalog-id <CATALOG_ID> \
    --materialized-view-refresh-task-run-id <TASK_RUN_ID> \
    --region <REGION>
```

Atau daftar semua proses penyegaran terbaru:

```
aws glue list-materialized-view-refresh-task-runs \
    --catalog-id <CATALOG_ID> \
    --database-name <DATABASE_NAME> \
    --table-name <MV_TABLE_NAME> \
    --region <REGION>
```

Ini menunjukkan:

- Waktu penyegaran terakhir
- Segarkan status (BERHASIL, GAGAL, BERJALAN, BERHENTI)
- ID jalankan tugas
- Pesan galat (jika gagal)

Status Penyegaran Umum:

- RUNNING: Segarkan pekerjaan saat ini sedang dijalankan
- BERHASIL: Refresh berhasil diselesaikan
- GAGAL: Penyegaran mengalami kesalahan
- STOPTED: Refresh dibatalkan secara manual

Pemecahan Masalah Gagal Menyegarkan:

Jika penyegaran gagal, periksa:

1. Izin IAM: Pastikan peran definer memiliki akses ke semua tabel dasar dan lokasi tampilan yang terwujud

2. Ketersediaan Tabel Dasar: Verifikasi semua tabel yang direferensikan ada dan dapat diakses
3. Validitas Kueri: Konfirmasikan kueri SQL valid untuk dialek Spark SQL
4. Batas Sumber Daya: Periksa apakah Anda telah mencapai batas penyegaran bersamaan untuk akun Anda

Gunakan GetMaterializedViewRefreshTaskRun API untuk mengambil pesan kesalahan terperinci.

Pertimbangan dan batasan

- Tampilan terwujud hanya dapat mereferensikan tabel Apache Iceberg yang terdaftar di Katalog AWS Glue Data sebagai tabel dasar.
- Pembuatan tampilan dan penulisan ulang kueri otomatis hanya tersedia dari mesin Spark di Apache Spark versi 3.5.6 ke atas di Amazon Athena, Amazon EMR, dan (Versi 5.1). AWS Glue
- Tampilan terwujud pada akhirnya konsisten dengan tabel dasar. Selama jendela penyegaran, kueri yang secara langsung mengakses tampilan terwujud dapat mengembalikan data yang sudah ketinggalan zaman. Untuk akses langsung ke data saat ini, jalankan penyegaran manual.
- Interval penyegaran otomatis minimum adalah satu jam. Untuk kasus penggunaan yang membutuhkan pembaruan yang lebih sering, jalankan penyegaran manual secara terprogram menggunakan perintah. `REFRESH MATERIALIZED VIEW`
- Penulisan ulang kueri memprioritaskan kebenaran daripada kinerja. Jika tampilan terwujud basi atau tidak dapat memenuhi persyaratan kueri secara akurat, mesin Spark mengeksekusi kueri asli terhadap tabel dasar.

Mengimpor data menggunakan alur kerja di Lake Formation

Dengan AWS Lake Formation, Anda dapat mengimpor data menggunakan alur kerja. Alur kerja menentukan sumber data dan jadwal untuk mengimpor data ke danau data Anda. Ini adalah wadah untuk AWS Glue crawler, pekerjaan, dan pemicu yang digunakan untuk mengatur proses untuk memuat dan memperbarui data lake.

Topik

- Cetak biru dan alur kerja di Lake Formation
- Membuat alur kerja
- Menjalankan alur kerja

Cetak biru dan alur kerja di Lake Formation

Alur kerja merangkum aktivitas multi-job extract, transform, and load (ETL) yang kompleks. Alur kerja menghasilkan AWS Glue crawler, pekerjaan, dan pemicu untuk mengatur pemuatan dan pembaruan data. Lake Formation mengeksekusi dan melacak alur kerja sebagai satu entitas. Anda dapat mengonfigurasi alur kerja agar berjalan sesuai permintaan atau sesuai jadwal.

> **Note**
>
> Penulis parket Spark tidak mendukung karakter khusus dalam nama kolom. Ini adalah batasan teknis dari penulis itu sendiri, bukan masalah konfigurasi.

Alur kerja yang Anda buat di Lake Formation terlihat di AWS Glue konsol sebagai grafik asiklik terarah (DAG). Setiap node DAG adalah job, crawler, atau trigger. Untuk memantau kemajuan dan pemecahan masalah, Anda dapat melacak status setiap node dalam alur kerja.

Ketika alur kerja Lake Formation telah selesai, pengguna yang menjalankan alur kerja diberikan `SELECT` izin Lake Formation pada tabel Katalog Data yang dibuat oleh alur kerja.

Anda juga dapat membuat alur kerja diAWS Glue. Namun, karena Lake Formation memungkinkan Anda membuat alur kerja dari cetak biru, membuat alur kerja jauh lebih sederhana dan lebih otomatis di Lake Formation. Lake Formation menyediakan jenis cetak biru berikut:

- Snapshot database — Memuat atau memuat ulang data dari semua tabel ke dalam data lake dari sumber JDBC. Anda dapat mengecualikan beberapa data dari sumber berdasarkan pola pengecualian.
- Database tambahan — Memuat hanya data baru ke dalam data lake dari sumber JDBC, berdasarkan bookmark yang ditetapkan sebelumnya. Anda menentukan tabel individual dalam database sumber JDBC untuk disertakan. Untuk setiap tabel, Anda memilih kolom bookmark dan urutan sortir bookmark untuk melacak data yang sebelumnya telah dimuat. Pertama kali Anda menjalankan cetak biru database inkremental terhadap sekumpulan tabel, alur kerja memuat semua data dari tabel dan menetapkan bookmark untuk menjalankan cetak biru database inkremental berikutnya. Oleh karena itu, Anda dapat menggunakan cetak biru database tambahan alih-alih cetak biru snapshot database untuk memuat semua data, asalkan Anda menentukan setiap tabel di sumber data sebagai parameter.
- File log - memuat data massal dari sumber file log, termasuk AWS CloudTrail, log Elastic Load Balancing, dan log Application Load Balancer.

Gunakan tabel berikut untuk membantu memutuskan apakah akan menggunakan snapshot database atau cetak biru database inkremental.

Gunakan snapshot database saat...	Gunakan database inkremental saat...
• Evolusi skema fleksibel. (Kolom diberi nama ulang, kolom sebelumnya dihapus, dan kolom baru ditambahkan di tempatnya.) • Diperlukan konsistensi yang lengkap antara sumber dan tujuan.	• Evolusi skema bersifat inkremental. (Hanya ada penambahan kolom berturut-turut.) • Hanya baris baru yang ditambahkan; baris sebelumnya tidak diperbarui.

> **Note**
>
> Pengguna tidak dapat mengedit cetakan biru dan alur kerja yang dibuat oleh Lake Formation.

Membuat alur kerja

Sebelum memulai, pastikan bahwa Anda telah memberikan izin data yang diperlukan dan izin lokasi data untuk peran tersebut. `LakeFormationWorkflowRole` Ini agar alur kerja dapat membuat tabel metadata di Katalog Data dan menulis data ke lokasi target di Amazon S3. Untuk informasi selengkapnya, lihat (Opsional) Membuat peran IAM untuk alur kerja dan Ikhtisar izin Lake Formation.

> **Note**
>
> Lake Formation menggunakan`GetTemplateInstance`,`GetTemplateInstances`, dan `InstantiateTemplate` operasi untuk membuat alur kerja dari cetak biru. Operasi ini tidak tersedia untuk umum, dan hanya digunakan secara internal untuk membuat sumber daya atas nama Anda. Anda menerima CloudTrail acara untuk membuat alur kerja.

Untuk membuat alur kerja dari cetak biru

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna yang memiliki izin insinyur data. Untuk informasi selengkapnya, lihat Referensi personas Lake Formation dan izin IAM.

2. Di panel navigasi, pilih Blueprints, lalu pilih Use blueprint.
3. Pada halaman Gunakan cetak biru, pilih ubin untuk memilih jenis cetak biru.
4. Di bawah Impor sumber, tentukan sumber data.

   Jika Anda mengimpor dari sumber JDBC, tentukan yang berikut ini:

   - Koneksi database —Pilih koneksi dari daftar. Buat koneksi tambahan menggunakan AWS Glue konsol. Nama pengguna dan kata sandi JDBC dalam koneksi menentukan objek database yang dapat diakses oleh alur kerja.
   - Jalur data sumber — Masukkan*<database>*/*<schema>*/*<table>*atau*<database>*/*<table>*, tergantung pada produk database. Basis Data Oracle dan MySQL tidak mendukung skema dalam path. Anda dapat mengganti persen (%) karakter untuk *<schema>* atau*<table>*. Misalnya, untuk database Oracle dengan pengenal sistem (SID) dari`orcl`, masukkan `orcl/%` untuk mengimpor semua tabel yang pengguna nama dalam koneksi memiliki akses ke.

     > **Important**
     >
     > Bidang ini peka huruf besar/kecil. Alur kerja akan gagal jika ada ketidakcocokan kasus untuk salah satu komponen.

     Jika Anda menentukan database MySQL AWS Glue , ETL menggunakan driver Mysql5 JDBC secara default, sehingga MySQL8 tidak didukung secara native. Anda dapat mengedit skrip pekerjaan ETL untuk menggunakan `customJdbcDriverS3Path` parameter seperti yang dijelaskan dalam JDBC ConnectionType Values di Panduan Pengembang untuk menggunakan driver JDBC AWS Glue berbeda yang mendukung MySQL8.

   Jika Anda mengimpor dari file log, pastikan bahwa peran yang Anda tentukan untuk alur kerja (“peran alur kerja”) memiliki izin IAM yang diperlukan untuk mengakses sumber data. Misalnya, untuk mengimpor AWS CloudTrail log, pengguna harus memiliki `cloudtrail:LookupEvents` izin `cloudtrail:DescribeTrails` dan untuk melihat daftar CloudTrail log saat membuat alur kerja, dan peran alur kerja harus memiliki izin pada lokasi di CloudTrail Amazon S3.
5. Lakukan salah satu tindakan berikut:

   - Untuk jenis cetak biru snapshot Database, identifikasi subset data yang akan diimpor secara opsional dengan menentukan satu atau beberapa pola pengecualian. Pola pengecualian ini

adalah Unix-style `glob` pola. Mereka disimpan sebagai properti dari tabel yang dibuat oleh alur kerja.

Untuk detail tentang pola pengecualian yang tersedia, lihat Menyertakan dan Mengecualikan Pola di Panduan AWS Glue Pengembang.

- Untuk jenis cetak biru database Incremental, tentukan bidang berikut. Tambahkan baris untuk setiap tabel untuk diimpor.

  Nama tabel

  Tabel untuk mengimpor. Harus semua huruf kecil.

  Tombol bookmark

  Comma-delimited daftar nama kolom yang menentukan tombol bookmark. Jika kosong, kunci utama digunakan untuk menentukan data baru. Kasus untuk setiap kolom harus cocok dengan kasus seperti yang didefinisikan dalam sumber data.

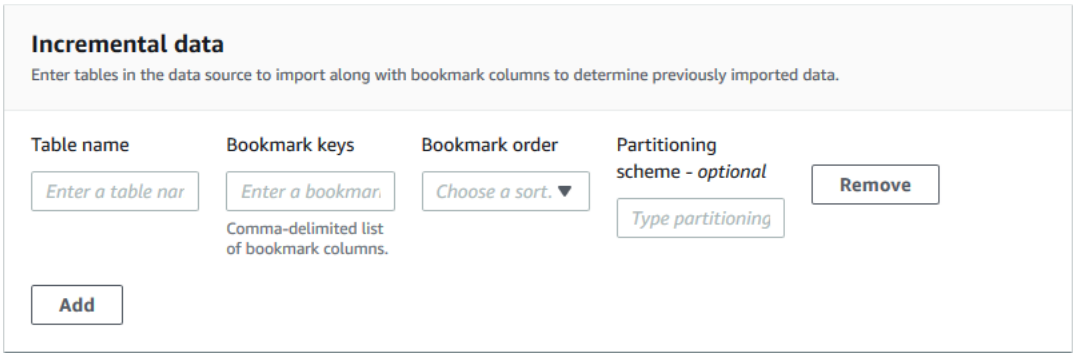
  > **Note**
  >
  > Kunci primer memenuhi syarat sebagai kunci bookmark default hanya jika secara berurutan meningkat atau menurun (tanpa celah). Jika Anda ingin menggunakan kunci primer sebagai kunci bookmark dan memiliki celah, Anda harus memberi nama kolom kunci utama sebagai kunci bookmark.

  Urutan bookmark

  Saat Anda memilih Ascending, baris dengan nilai lebih besar dari nilai yang ditandai diidentifikasi sebagai baris baru. Saat Anda memilih Menurun, baris dengan nilai kurang dari nilai yang ditandai diidentifikasi sebagai baris baru.

  Skema pembagian

  (Opsional) Daftar kolom kunci partisi, dibatasi oleh garis miring (/). Contoh: `year/month/day`.



Untuk informasi selengkapnya, lihat Melacak Data yang Diproses Menggunakan Bookmark Job di Panduan AWS Glue Pengembang.

6. Di bawah Impor target, tentukan database target, target lokasi Amazon S3, dan format data.

   Pastikan bahwa peran alur kerja memiliki izin Lake Formation yang diperlukan pada database dan lokasi target Amazon S3.

   > **Note**
   >
   > Saat ini, cetak biru tidak mendukung enkripsi data pada target.

7. Pilih frekuensi impor.

   Anda dapat menentukan `cron` ekspresi dengan opsi Kustom.

8. Di bawah opsi Impor:

   a. Masukkan nama alur kerja.

   b. Untuk peran, pilih peran`LakeFormationWorkflowRole`, yang Anda buat(Opsional) Membuat peran IAM untuk alur kerja.

   c. Secara opsional menentukan awalan tabel. Awalan ditambahkan ke nama tabel Katalog Data yang dibuat alur kerja.

9. Pilih Buat, dan tunggu konsol melaporkan bahwa alur kerja berhasil dibuat.

   > **Tip**
   >
   > Apakah Anda mendapatkan pesan kesalahan berikut?
   > `User: arn:aws:iam::<account-id>:user/<username> is not authorized to perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/<rolename>...`

Jika demikian, periksa apakah Anda mengganti *<account-id>* dengan nomor AWS akun yang valid di semua kebijakan.

Lihat juga:

- Cetak biru dan alur kerja di Lake Formation

Menjalankan alur kerja

Anda dapat menjalankan alur kerja menggunakan konsol Lake Formation, AWS Glue konsol, atau AWS Glue Command Line Interface (AWS CLI), atau API.

Untuk menjalankan alur kerja (konsol Lake Formation)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pengguna yang memiliki izin insinyur data. Untuk informasi selengkapnya, lihat Referensi personas Lake Formation dan izin IAM.
2. Di panel navigasi, pilih Cetak biru.
3. Pada halaman Blueprints, pilih alur kerja. Kemudian pada menu Tindakan, pilih Mulai.
4. Saat alur kerja berjalan, lihat kemajuannya di kolom Status Last run. Pilih tombol refresh sesekali.

   Status berubah dari RUNNING, ke Discovering, ke Importing, ke COMPLETED.

   Saat alur kerja selesai:

   - Katalog Data memiliki tabel metadata baru.
   - Data Anda tertelan ke danau data.

   Jika alur kerja gagal, lakukan hal berikut:

   a. Pilih alur kerja. Pilih Tindakan, lalu pilih Lihat grafik.

      Alur kerja terbuka di AWS Glue konsol.

   b. Pastikan bahwa alur kerja sudah dipilih, dan pilih tab Riwayat.

c. Di bawah Riwayat, pilih proses terbaru dan pilih Lihat detail jalankan.

d. Pilih job atau crawler yang gagal dalam grafik dinamis (runtime), dan tinjau pesan galatnya. Node yang gagal berwarna merah atau kuning.

Lihat juga:

- Cetak biru dan alur kerja di Lake Formation

Membawa data Anda ke AWS Glue Data Catalog

Anda dapat membuat katalog gabungan di AWS Glue Data Catalog (Katalog Data) dan menyatukan data di seluruh danau data Amazon S3 dan gudang data Amazon Redshift. Anda juga dapat mengintegrasikan data dari database operasional Anda seperti Amazon DynamoDB, dan sumber data pihak ketiga seperti PostgreSQL, BigQuery Google, MySQL, dan lainnya. Katalog Data menyediakan repositori metadata terpusat yang membuat pengelolaan dan penemuan data di seluruh sistem yang berbeda menjadi lebih mudah.

Katalog Data terintegrasi dengan lebih dari 30 sumber data eksternal melalui konektor federasi. Dengan integrasi ini, Anda dapat melakukan kueri data dari sumber eksternal ini tanpa harus membangun pipeline data untuk memasukkan AWS data terlebih dahulu.

Setelah membuat katalog data eksternal, Anda dapat menggunakan AWS Lake Formation untuk mengelola izin akses data secara terpusat di Katalog Data. Administrator data lake dapat memberikan izin akses berbutir halus ke prinsipal IAM lainnya (pengguna atau peran) dalam akun yang sama atau di seluruh akun menggunakan kontrol akses berbasis tag (LF-tag) dan metode sumber daya bernama.

Dengan menggunakan LF-tag, administrator data dapat secara logis mengatur sumber daya berdasarkan atribut seperti domain dan tingkat sensitivitas, menyederhanakan manajemen izin sambil memastikan kontrol akses yang konsisten di seluruh layanan analitik dan pembelajaran mesin termasuk Athena, Amazon EMR, atau Redshift Spectrum. AWS Glue

Katalog Data menyediakan metode berikut untuk mengelola data dan izin pada kumpulan data eksternal dan metastor eksternal:

- Integrasikan bucket Tabel Amazon S3 dengan Katalog Data — Anda dapat menerbitkan dan membuat katalog Tabel Amazon S3 sebagai objek Katalog Data dan mendaftarkan katalog sebagai lokasi data Lake Formation dari konsol Lake Formation atau menggunakan operasi API. AWS Glue
- Bawa data di gudang data Amazon Redshift ke dalam AWS Glue Data Catalog — Daftarkan namespace Amazon Redshift yang ada atau cluster dengan Katalog Data, dan buat katalog federasi multi-level di Katalog Data.

  Anda dapat mengakses data Anda menggunakan mesin kueri apa pun yang kompatibel dengan spesifikasi OpenAPI katalog Apache Iceberg REST, seperti Amazon EMR Tanpa Server, dan Amazon Athena.

- Gabungkan Iceberg REST katalog jarak jauh ke Katalog Data — Gabungkan katalog jarak jauh ke Iceberg REST Katalog Data dan akses tabel jarak jauh yang disimpan Iceberg di Amazon S3 dengan aman menggunakan mesin analitik. AWS
- Federasi ke dalam Katalog Data dari sumber data eksternal — Hubungkan Katalog Data ke sumber data eksternal menggunakan AWS Glue koneksi, dan buat katalog federasi untuk mengelola izin akses secara terpusat pada kumpulan data menggunakan Lake Formation. Tidak diperlukan migrasi metadata ke dalam Katalog Data.
- Buat katalog untuk mengelola tabel Amazon Redshift di Katalog Data — Anda mungkin tidak memiliki klaster produsen Amazon Redshift atau datashare Amazon Redshift yang tersedia saat ini, tetapi ingin membuat dan mengelola tabel Amazon Redshift menggunakan Katalog Data. Anda dapat memulai dengan membuat katalog AWS Glue terkelola menggunakan operasi `glue:CreateCatalog` API atau AWS Lake Formation konsol dengan menyetel jenis katalog sebagai `Managed` dan `Catalog source` sebagai Redshift.
- Publikasikan datashares Amazon Redshift dengan Katalog Data — Publikasikan datashares Amazon Redshift ke Katalog Data, dan gunakan Lake Formation untuk mengelola akses data data secara terpusat dan membatasi akses pengguna.

  Anda dapat menanyakan data Anda menggunakan Amazon Redshift Spectrum.
- Connect Data Catalog ke metastor Hive eksternal — Hubungkan Katalog Data ke metastor eksternal untuk mengelola izin akses pada kumpulan data di Amazon S3 menggunakan Lake Formation. Tidak diperlukan migrasi metadata ke dalam Katalog Data.
- Integrasikan Lake Formation dengan AWS Data Exchange — Lake Formation mendukung lisensi akses ke data Anda melalui AWS Data Exchange. Jika Anda ingin melisensikan data Lake Formation Anda, lihat Apa yang ada AWS Data Exchange di Panduan AWS Data Exchange Pengguna.

Topik

- Federasi katalog ke katalog Gunung Es jarak jauh
- Membawa data Amazon Redshift ke dalam AWS Glue Data Catalog
- Menggabungkan ke sumber data eksternal di AWS Glue Data Catalog
- Integrasi Tabel Amazon S3 dengan AWS Glue Data Catalog and AWS Lake Formation
- Mengubah kontrol akses untuk integrasi Tabel S3
- Membuat katalog terkelola Amazon Redshift di AWS Glue Data Catalog
- Mengelola izin untuk data dalam data Amazon Redshift

- Mengelola izin pada kumpulan data yang menggunakan metastor eksternal

Federasi katalog ke katalog Gunung Es jarak jauh

Federasi katalog AWS Glue menyediakan akses langsung dan aman ke tabel Iceberg, disimpan di Amazon S3 dan dikatalogkan dalam katalog jarak jauh, menggunakan mesin analitik. AWS Federasi katalog menyinkronkan metadata di seluruh Katalog Data dan katalog jarak jauh saat Anda mengakses tabel jarak jauh. Ini didukung oleh berbagai mesin analitik, termasuk Amazon Redshift, Amazon EMR, Amazon Athena,, mesin pihak ketiga seperti Apache Spark AWS Glue, dan banyak lagi.

Federasi katalog menggunakan AWS Glue Data Catalog untuk berkomunikasi dengan sistem katalog jarak jauh untuk menemukan tabel dan Lake Formation untuk mengotorisasi akses ke data tabel di Amazon S3. Saat Anda menanyakan tabel federasi, Katalog Data menemukan informasi tabel terbaru di katalog jarak jauh pada waktu kueri, mendapatkan lokasi Amazon S3 tabel, skema saat ini, dan informasi partisi. Mesin analitik Anda (Amazon Athena, Amazon Redshift, Amazon EMR) kemudian menggunakan informasi ini untuk mengakses file data Iceberg langsung dari Amazon S3. Lake Formation mengelola akses ke tabel dengan menjual kredenal cakupan ke data tabel yang disimpan di Amazon S3, memungkinkan engine menerapkan izin berbutir halus ke tabel federasi.

Fitur Federasi Katalog

Diatur menggunakan Lake Formation

Katalog Gunung Es Federasi di Katalog Data adalah sumber daya terdaftar Lake Formation, memungkinkan Anda memberikan izin baris, kolom, tingkat sel berbutir halus ke tabel Gunung Es di katalog Gunung Es federasi menggunakan hibah Lake Formation. Katalog Iceberg Federasi dan objek terkait dapat dibagikan dengan aman di seluruh akun. AWS Katalog Federated Iceberg juga bekerja dengan kontrol Tag-based akses Lake Formation yang memungkinkan Anda menskalakan tata kelola menggunakan tag.

Konfigurasi jaringan

Federasi katalog mendukung koneksi langsung ke sumber katalog jarak jauh menggunakan konektivitas HTTPS standar. Ini juga mendukung konektivitas melalui Amazon VPC ketika Anda ingin mempertahankan isolasi jaringan dan konektivitas menggunakan dukungan proxy ketika Anda menginginkan komunikasi yang aman melalui firewall organisasi.

Batasan

- Batas ukuran metadata tabel gunung es — Untuk katalog Iceberg federasi, AWS Glue Data Catalog mendukung ukuran metadata maksimum 5 MB per panggilan REST API. AWS Glue Data Catalog menolak permintaan ke tabel dengan metadata melebihi batas ini.

Topik

- Federasi untuk Snowflake Iceberg Katalog
- Federasi ke Katalog Unity Databricks

Federasi untuk Snowflake Iceberg Katalog

AWS Glue Data Catalog bergabung untuk Snowflake menggunakan kredenSIAL OAuth2 dari prinsipal layanan. Snowflake Anda dapat menggunakan federasi ini untuk terhubung ke Snowflake Horizon dan Snowflake Polaris katalog. Mekanisme otentikasi ini memungkinkan Katalog Data untuk mengakses metadata berbagai objek (seperti katalog, database, dan tabel) dalam katalog Anda, berdasarkan hak istimewa yang terkait dengan prinsipal layanan. Untuk memastikan akses ke objek yang tepat, penting untuk memberikan prinsipal layanan Snowflake dengan izin yang diperlukan untuk membaca metadata objek ini.

Prasyarat

Sebelum Anda membuat katalog federasi di Katalog Data yang diatur oleh Lake Formation, pastikan Anda memiliki izin berikut:

Prinsipal IAM Anda (pengguna atau peran) harus memiliki izin berikut:

- Izin Formasi Danau —`lakeformation:RegisterResource`, `lakeformation:DescribeResource`
- AWS Glue izin —`glue:CreateConnection`,`glue:CreateCatalog`,`glue:GetConnection`, `glue:PassConnection`
- Izin Manajer Rahasia `-secretsmanager:CreateSecret`, `secretsmanager:GetSecretValue`
- Izin IAM —`iam:CreateRole`,, `iam:AttachRolePolicy` `iam:PassRole`

Anda harus menjadi administrator danau data Lake Formation atau memiliki `CREATE_CATALOG` izin di Katalog Data

Buat Katalog Federasi

Menggunakan Konsol

1. Masuk ke konsol dan buka konsol Formasi Danau di https://console.aws.amazon.com/lakeformation/.
2. Pilih Wil AWS ayah yang disukai di pemilih Wilayah.
3. Di panel navigasi kiri, pilih Katalog.
4. Pilih Buat Katalog untuk membuka Alur Kerja Buat Katalog.
5. Di langkah Pilih sumber data, pilih Snowflake dari opsi yang tersedia.
6. Pada langkah Setel detail katalog, Anda memberikan tiga informasi - detail katalog, detail koneksi, dan detail pendaftaran.
7. Dalam wadah detail katalog, berikan nama unik untuk katalog AWS Glue federasi Anda dan masukkan nama Snowflake katalog yang ada.
8. Dalam wadah rincian koneksi, Anda dapat memilih dari koneksi yang ada yang Anda akses atau menyediakan konfigurasi untuk membuat konektor baru.
9. Konfigurasi koneksi baru meliputi:

   - Nama Koneksi — Nama unik dari objek AWS Glue koneksi.
   - URL Instance — URL titik akhir Snowflake akun Anda yang ada.
   - Otentikasi — Tentukan konfigurasi otentikasi AWS Glue yang digunakan untuk menyambung ke server katalog jarak jauh. AWS Glue mendukung otentikasi OAuth2 dan Kustom.
   - URL Token – Tentukan URL penyedia identitas katalog jarak jauh.
   - ID Klien OAuth2 — (Opsional) Tentukan ID Klien dari kredensi OAuth2 yang terkait dengan katalog jarak jauh Anda. Jika Anda tidak memberikan ID klien, permintaan token akan diautentikasi hanya menggunakan rahasia klien yang disimpan di AWS Secrets Manager dalamnya.
   - Rahasia — Simpan dan gunakan rahasia klien OAuth2 menggunakan AWS Secrets Manager atau masukkan nilai rahasia di kotak teks. Ketika Anda memasukkan rahasia secara manual di konsol AWS Glue , buat rahasia atas nama Anda.
   - Cakupan URL Token – Tentukan cakupan OAuth untuk autentikasi.

- Filter Casing Katalog - Pilih apakah akan membawa objek huruf kecil atau huruf besar dari katalog jarak jauh Anda ke Katalog Data.

10. Buat peran IAM yang dapat digunakan AWS Glue oleh kepala sekolah layanan Lake Formation untuk mengakses lokasi rahasia di dalam AWS Secrets Manager dan Amazon S3 dari tabel Iceberg jarak jauh. Pilih peran IAM di dropdown pendaftaran. Lihat langkah 2 dan 3 di bagian CLI berikut untuk detail kebijakan IAM.
11. Pilih Uji Koneksi untuk menguji apakah properti koneksi dan akses peran IAM dikonfigurasi dengan benar.
12. Pilih Berikutnya untuk meninjau pengaturan Anda.
13. Pilih Buat Katalog di halaman ulasan.

Menggunakan CLI

1. Buat AWS Secrets Manager rahasia

   AWS Glue Konektor mendukung dua jenis otentikasi - OAuth2 dan Custom. Saat menggunakan opsi OAuth2, gunakan AWS Secrets Manager untuk menyimpan rahasia klien dari prinsipal Snowflake layanan. Anda nantinya akan menggunakan rahasia ini saat membuat Kon AWS Glue eksi. Untuk otentikasi Kustom, gunakan AWS Secrets Manager untuk menyimpan dan mengambil token akses.

   Dalam contoh berikut, ganti*your-snowflake-secret*,*client_secret*, *region* dengan informasi Anda sendiri.

   ```
   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_SECRETadalah kata kunci cadangan yang AWS Glue digunakan untuk merujuk pada nilai rahasia klien dalam rahasia.

Gunakan kata kunci yang sama saat Anda membuat rahasia di konsol Lake Formation juga.

2. Buat peran IAM yang memberikan akses objek AWS Glue koneksi ke rahasia yang dibuat pada langkah sebelumnya

   Ob AWS Glue jek koneksi memerlukan akses ke AWS Secrets Manager rahasia saat Anda menggunakan AWS Secrets Manager untuk menyimpan, mengambil, dan menyegarkan token rahasia OAuth Anda. Ob AWS Glue jek koneksi juga memerlukan akses untuk membuat, menjelaskan, dan menggunakan antarmuka jaringan Amazon VPC saat Anda menggunakan titik akhir Amazon VPC untuk membatasi konektivitas ke akun AndaSnowflake.

   Buat kebijakan IAM dan lampirkan ke peran IAM. Tambahkan AWS Glue prinsipal layanan ke kebijakan kepercayaan.

   Dalam contoh berikut, ganti*your-secrets-manager-ARN*,*your-vpc-id*, dan *your-subnet-id1* dengan informasi Anda sendiri.

   Example Kebijakan 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 Kebijakan Kepercayaan

```
{
    "Version": "2012-10-17",
    "Statement": [{
        "Effect": "Allow",
        "Principal": {
            "Service": "glue.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    }]
}
```

3. Buat kebijakan IAM yang memberikan akses baca Lake Formation ke lokasi Amazon S3 katalog

   Sebagai pemilik katalog katalog federasi di Katalog Data, Anda menggunakan Lake Formation untuk memberikan akses tabel kasar, akses berbutir halus - tingkat kolom, tingkat baris, dan tingkat sel -, dan akses berbasis tag ke tim data Anda. Lake Formation menggunakan peran IAM yang memberinya akses ke lokasi Amazon S3 yang mendasari tabel Iceberg jarak jauh Anda. Akses ini memungkinkan Lake Formation untuk menjual kredenSIAL akses cakupan ke mesin analitik yang menanyakan tabel jarak jauh.

   Buat kebijakan IAM dan lampirkan ke peran IAM. Tambahkan prinsip layanan Formasi Danau ke kebijakan kepercayaan peran.

   Dalam contoh berikut, ganti *amzn-s3-demo-bucketN* dan *your-kms-key* dengan informasi Anda sendiri.

   Example Kebijakan 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 Kebijakan Kepercayaan

```
{
    "Version": "2012-10-17",
    "Statement": [{
        "Sid": "",
        "Effect": "Allow",
        "Principal": {
            "Service": "lakeformation.amazonaws.com"
```

```
        },
        "Action": "sts:AssumeRole"
    }]
}
```

> **Note**
>
> Saat Anda menggunakan konsol Lake Formation untuk membuat katalog federasi, konsol menggunakan peran IAM tunggal dengan kedua kebijakan yang dilampirkan untuk menyelesaikan penyiapan.

4. Buat objek AWS Glue koneksi

AWS Glue konektor mendukung metode otentikasi OAuth2 dan Kustom. Objek koneksi Katalog Data mendukung `CATALOG_CASING_FILTER` konfigurasi yang memungkinkan Anda membawa objek huruf kecil atau huruf besar dalam katalog jarak jauh ke Katalog Data.

Contoh berikut menggunakan konfigurasi otentikasi OAuth2 untuk membuat koneksi AWS Glue . Ganti *highlighted sections* dengan informasi Anda.

```
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"
}
}'
```

> **Note**
>
> Bid `0Auth2ClientApplication` ang ini opsional. Jika Anda menghilangkannya, permintaan token akan diautentikasi hanya menggunakan rahasia klien yang disimpan AWS Secrets Manager, tanpa mengirim ID klien.

5. Daftar AWS Glue kan koneksi sebagai sumber daya Formasi Danau

   Menggunakan objek AWS Glue koneksi (dibuat pada Langkah 4) dan peran IAM (dibuat pada Langkah 3), Anda sekarang dapat mendaftarkan objek AWS Glue koneksi sebagai sumber daya yang dikelola Lake Formation.

   Ganti *your-glue-connector-arn* dan *your-IAM-role-ARN-having-LF-access* dengan informasi Anda.

   ```
   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. Membuat katalog federasi di Katalog Data

   Setelah membuat objek AWS Glue koneksi dan mendaftarkannya dengan Lake Formation, Anda dapat membuat katalog federasi di Katalog Data:

   Berikan katalog federasi nama unik di*your-federated-catalog-name*, referensi katalog Snowflake di at*catalog-name-in-Snowflake*, dan masukkan nama koneksi yang dibuat sebelumnya di*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": []
}'
```

Pertimbangan saat berintegrasi dengan Snowflake

- Saat Anda memasukkan sumber daya (seperti database dan tabel)Snowflake, Lake Formation tidak secara otomatis mencabut izin yang diberikan pada sumber daya federasi tersebut. Untuk menghapus izin akses, Anda harus secara eksplisit mencabut izin yang sebelumnya diberikan pada sumber daya federasi menggunakan Formasi Danau.
- Saat Anda memasang katalog jarak jauh dengan `CATALOG_CASING_FILTER='UPPERCASE_ONLY'` konfigurasi, database dan tabel dengan pengidentifikasi huruf besar digabungkan tetapi objek dengan pengidentifikasi huruf kecil tidak.
- Anda dapat menggunakan kembali AWS Glue koneksi yang sama untuk membuat beberapa katalog federasi. Menghapus katalog tidak akan menghapus koneksi terkait. Untuk menghapus koneksi, gunakan `aws glue delete-connection` perintah AWS CLI dan pastikan semua katalog terkait dihapus terlebih dahulu.
- Ruang nama bersarang dalam Polaris katalog tidak didukung. Artinya, federasi katalog dapat mengakses Iceberg tabel jarak jauh yang mengikuti not `catalog.database.table` asi 3 bagian.

Federasi ke Katalog Unity Databricks

AWS Glue Data Catalog federasi untuk Databricks menggunakan kredensi OAuth2 dari kepala layanan. Databricks Mekanisme otentikasi ini memungkinkan AWS Glue Data Catalog untuk mengakses metadata berbagai objek (seperti katalog, database, dan tabel) diDatabricks Unity Catalog, berdasarkan hak istimewa yang terkait dengan prinsip layanan. Untuk memastikan akses ke objek yang tepat, penting untuk memberikan prinsipal layanan dengan izin yang diperlukan Databricks untuk membaca metadata objek ini.

Selanjutnya, federasi katalog memungkinkan penemuan dan kueri Iceberg tabel di Katalog Unity Databricks Anda. Untuk membaca tabel delta, pastikan Iceberg metadata tersedia untuk tabel ini menggunakan. Uniform Ikuti Databricks tutorial dan dokumentasi untuk membuat prinsip layanan dan hak istimewa terkait di ruang Databricks kerja Anda.

Prasyarat

Sebelum Anda membuat katalog federasi di Katalog Data yang diatur oleh Lake Formation, pastikan Anda memiliki izin berikut:

Prinsipal IAM Anda (pengguna atau peran) harus memiliki izin berikut:

- Izin Lake Formation —`lakeformation:RegisterResource`, `lakeformation:DescribeResource`
- AWS Glue izin —`glue:CreateConnection`,,`glue:CreateCatalog`, `glue:GetConnection` `glue:PassConnection`
- Izin Secrets Manager —`secretsmanager:CreateSecret`, `secretsmanager:GetSecretValue`
- Izin IAM —`iam:CreateRole`,, `iam:AttachRolePolicy` `iam:PassRole`

Anda harus menjadi administrator danau data Lake Formation atau memiliki `CREATE_CATALOG` izin pada Katalog Data

Buat Katalog Federasi

Menggunakan Konsol

1. Masuk ke konsol dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Pilih AWS Region yang disukai di pemilih Region.
3. Di panel navigasi kiri, pilih Katalog.
4. Pilih Create Catalog untuk membuka Create Catalog Workflow.
5. Di Pilih langkah sumber data, pilih Databricks dari opsi yang tersedia.
6. Pada langkah Atur detail katalog, Anda memberikan tiga informasi - detail katalog, detail koneksi, dan detail pendaftaran.
7. Dalam wadah detail katalog, berikan nama unik ke katalog AWS Glue federasi Anda dan masukkan nama Databricks katalog yang ada.
8. Dalam kontainer detail koneksi, Anda dapat memilih dari koneksi yang ada yang dapat Anda akses atau menyediakan konfigurasi untuk membuat konektor baru.
9. Konfigurasi koneksi baru meliputi:

- Nama Koneksi — Nama unik dari objek AWS Glue koneksi.
- URL Workspace — URL endpoint dari ruang kerja Anda yang adaDatabricks.
- Otentikasi — Tentukan konfigurasi otentikasi yang AWS Glue digunakan untuk terhubung ke server katalog jarak jauh. AWS Glue mendukung otentikasi OAuth2 dan Kustom.
- URL Token – Tentukan URL penyedia identitas katalog jarak jauh.
- ID Klien OAuth2 - Tentukan ID Klien dari kredensi OAuth2 yang terkait dengan katalog jarak jauh Anda.
- Rahasia — Simpan dan gunakan rahasia klien OAuth2 menggunakan AWS Secrets Manager atau masukkan nilai rahasia di kotak teks. Saat Anda memasukkan rahasia secara manual di konsol, AWS Glue buat rahasia atas nama Anda.
- Cakupan URL Token – Tentukan cakupan OAuth untuk autentikasi.

10. Buat peran IAM yang dapat digunakan oleh kepala layanan Lake Formation untuk mengakses rahasia di AWS Glue dan lokasi Amazon S3 dari tabel jarak jauh masing-masing. Iceberg Pilih peran IAM di dropdown pendaftaran. Lihat langkah 2 dan 3 di bagian CLI berikut untuk detail kebijakan IAM.
11. Pilih Uji Koneksi untuk menguji apakah properti koneksi dan akses peran IAM dikonfigurasi dengan benar. Fungsionalitas koneksi uji tidak tersedia saat menghubungkan Databricks menggunakan Amazon VPC.
12. Pilih Berikutnya untuk meninjau pengaturan Anda.
13. Pilih Buat Katalog di halaman ulasan.

Menggunakan CLI

1. Buat AWS Secrets Manager rahasia

   AWS Glue Konektor mendukung dua jenis otentikasi - OAuth2 dan Custom. Saat menggunakan opsi OAuth2, gunakan AWS Secrets Manager untuk menyimpan rahasia klien dari kepala Databricks layanan. Anda nantinya akan menggunakan rahasia ini saat membuat AWS Glue Koneksi. Untuk otentikasi kustom, gunakan AWS Secrets Manager untuk menyimpan dan mengambil token akses.

   Dalam contoh berikut, ganti*<databricks-secret>*,*<client_secret>*, *<region>* dengan informasi Anda sendiri.

   ```
   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_SECRETadalah kata kunci yang dicadangkan yang AWS Glue digunakan untuk merujuk ke nilai rahasia klien dalam rahasia. Gunakan kata kunci yang sama saat Anda membuat rahasia di konsol Lake Formation juga.

2. Buat peran IAM yang memberikan akses objek AWS Glue koneksi ke rahasia yang dibuat pada langkah sebelumnya

   Objek AWS Glue koneksi memerlukan akses ke AWS Secrets Manager rahasia saat Anda gunakan AWS Secrets Manager untuk menyimpan, mengambil, dan menyegarkan token rahasia OAuth Anda. Objek AWS Glue koneksi juga memerlukan akses untuk membuat, mendeskripsikan, dan menggunakan antarmuka jaringan Amazon VPC saat Anda menggunakan titik akhir VPC Amazon untuk membatasi konektivitas ke ruang kerja Anda. Databricks

   Buat kebijakan IAM dan lampirkan ke peran IAM. Tambahkan prinsip AWS Glue layanan ke kebijakan kepercayaan.

   Dalam contoh berikut, ganti*<your-secrets-manager-ARN>*,*<your-vpc-id>*, dan *<your-subnet-id1>* dengan informasi Anda sendiri.

   Example Kebijakan 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 Kebijakan Kepercayaan

```
{
    "Version": "2012-10-17",
    "Statement": [{
        "Effect": "Allow",
        "Principal": {
            "Service": "glue.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    }]
}
```

3. Buat kebijakan IAM yang memberikan akses baca Lake Formation ke lokasi Amazon S3 katalog

   Sebagai pemilik katalog katalog federasi di Katalog Data, Anda menggunakan Lake Formation untuk memberikan akses tabel berbutir kasar, akses berbutir halus - tingkat kolom, tingkat baris, dan tingkat sel - akses, dan akses berbasis tag ke tim data Anda. Lake Formation menggunakan peran IAM yang memberikan akses ke lokasi Amazon S3 yang mendasari tabel Iceberg jarak

jauh Anda. Akses ini memungkinkan Lake Formation untuk menjual kredenal akses cakupan ke mesin analitik yang menanyakan tabel jarak jauh.

Buat kebijakan IAM dan lampirkan ke peran IAM. Tambahkan kepala layanan Lake Formation ke kebijakan kepercayaan peran IAM.

Dalam contoh berikut, ganti *<your-s3-bucket-N>* dan *<your-kms-key>* dengan informasi Anda sendiri.

Example Kebijakan 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 Kebijakan Kepercayaan

```
{
    "Version": "2012-10-17",
    "Statement": [{
        "Sid": "",
        "Effect": "Allow",
        "Principal": {
            "Service": "lakeformation.amazonaws.com"
        },
        "Action": "sts:AssumeRole"
    }]
}
```

> **Note**
>
> Saat Anda menggunakan konsol Lake Formation untuk membuat katalog federasi, konsol menggunakan peran IAM tunggal dengan kedua kebijakan yang dilampirkan untuk menyelesaikan penyiapan.

4. Buat objek AWS Glue koneksi

Data Catalog mendukung ConnectionType`DATABRICKSICEBERGRESTCATALOG`: untuk menghubungkan Data Catalog ke. Databricks AWS Glue Konektor ini mendukung metode otentikasi OAuth2 dan Custom.

Contoh berikut menggunakan konfigurasi otentikasi OAuth2 untuk membuat koneksi. AWS Glue Ganti *`highlighed sections`* dengan informasi Anda.

```
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. Daftarkan AWS Glue koneksi sebagai sumber daya Lake Formation

   Menggunakan objek AWS Glue koneksi (dibuat di Langkah 4) dan peran IAM (dibuat pada Langkah 3), Anda sekarang dapat mendaftarkan objek AWS Glue koneksi sebagai sumber daya yang dikelola Lake Formation.

   Ganti *<your-glue-connector-arn>* dan *<your-IAM-role-ARN-having-LF-access>* dengan informasi Anda.

   ```
   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. Buat katalog federasi di Katalog Data

   Setelah membuat objek AWS Glue koneksi dan mendaftarkannya dengan Lake Formation, Anda dapat membuat katalog federasi di Katalog Data:

   Berikan katalog federasi nama unik di*<your-federated-catalog-name>*, referensi katalog Databricks di at*"<catalog-name-in-Databricks>*, dan masukkan nama koneksi yang dibuat sebelumnya di*"<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": []
}'
```

Pertimbangan saat berintegrasi dengan Databricks

- Saat Anda menghapus sumber daya (seperti database dan tabel) diDatabricks, Lake Formation tidak secara otomatis mencabut izin yang diberikan pada sumber daya federasi tersebut. Untuk menghapus izin akses, Anda harus secara eksplisit mencabut izin yang sebelumnya diberikan pada sumber daya federasi menggunakan Lake Formation.
- Anda dapat menanyakan tabel Iceberg yang disimpan di Amazon S3 menggunakan integrasi ini. Saat menggunakan format tabel atau penyimpanan objek lainnya, Anda dapat menggabungkan metadata dalam katalog jarak jauh ke AWS Glue dan mencantumkan database dan tabelnya tetapi operasi kueri seperti `SELECT ColumnFoo from TableBar` akan gagal selama kueri dengan kesalahan 'Gagal membaca tabel. Apache Iceberg Lokasi penyimpanan objek tidak didukung. '
- Anda dapat menggunakan kembali AWS Glue koneksi yang sama untuk membuat beberapa katalog federasi. Menghapus katalog tidak akan menghapus objek koneksi terkait. Untuk menghapus objek koneksi, gunakan `aws glue delete-connection` perintah AWS CLI dan pastikan semua katalog terkait dihapus terlebih dahulu.

Membawa data Amazon Redshift ke dalam AWS Glue Data Catalog

Anda dapat mengelola data analitik di gudang data Amazon Redshift di AWS Glue Data Catalog (Katalog Data), dan menyatukan data lake Amazon S3 dan gudang data Amazon Redshift. Amazon Redshift adalah layanan gudang data skala petabyte yang dikelola sepenuhnya di Cloud. AWS Gudang data Amazon Redshift adalah kumpulan sumber daya komputasi yang disebut node, yang diatur ke dalam grup yang disebut cluster. Setiap klaster menjalankan mesin Amazon Redshift dan berisi satu atau lebih database.

Di Amazon Redshift, Anda dapat membuat klaster yang disediakan Amazon Redshift dan ruang nama tanpa server, dan mendaftarkannya ke Katalog Data. Dengan melakukan ini, Anda dapat menyatukan data di Amazon Redshift managed storage (RMS) dan Amazon S3 bucket, dan mengakses data dari mesin analitik yang kompatibel dengan Apache Iceberg.

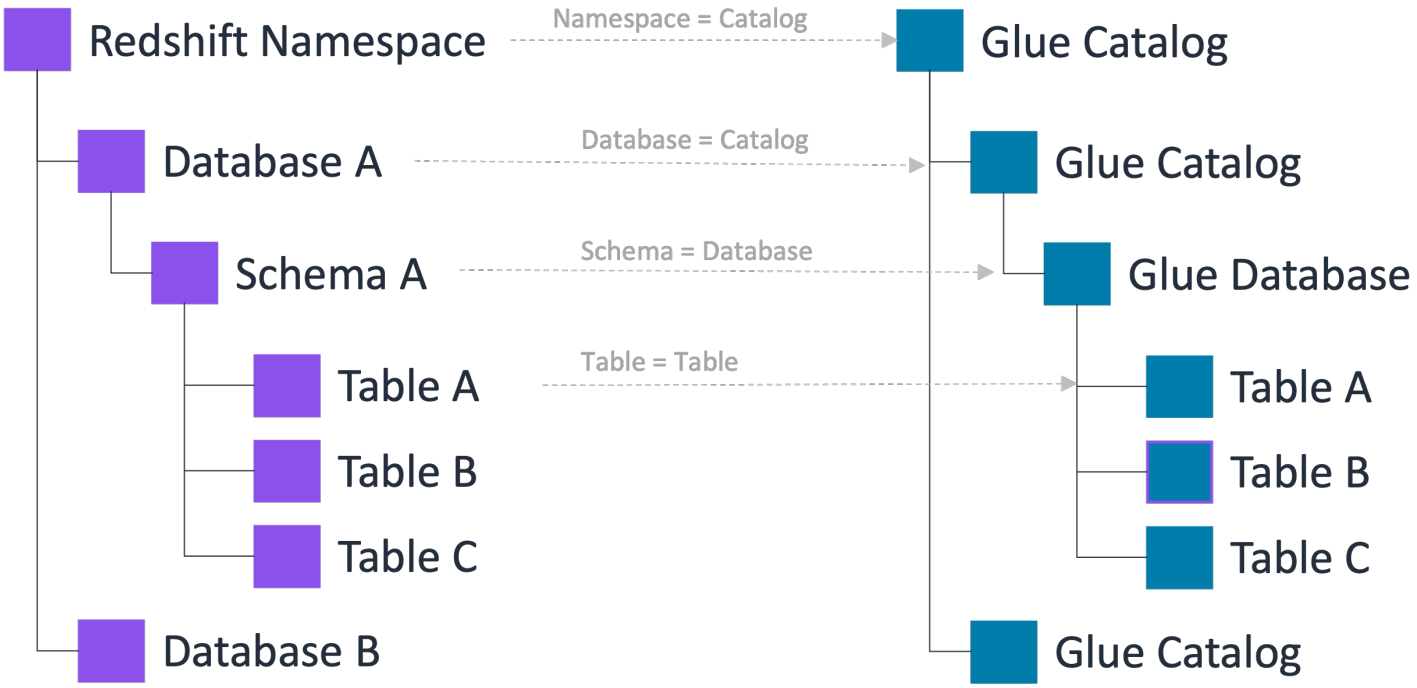
Dengan mendaftarkan ruang nama dan cluster, Anda dapat memberikan akses ke data tanpa perlu menyalin atau memindahkannya. Untuk informasi selengkapnya tentang mendaftarkan cluster dan ruang nama di Amazon Redshift, lihat Mendaftarkan klaster dan ruang nama Amazon Redshift ke. AWS Glue Data Catalog

Di Amazon Redshift, Anda dapat melakukan berbagi data melalui datashares atau dengan mendaftarkan namespace dan cluster dengan Data Catalog. Dengan datashares, yang beroperasi pada tingkat objek database individual, Anda harus mengaktifkan berbagi untuk setiap tabel atau tampilan. Sebaliknya, penerbitan namespace berfungsi pada tingkat cluster atau namespace. Saat Anda mendaftarkan cluster atau namespace dengan Katalog Data, semua database dan tabel di dalamnya akan dibagikan secara otomatis, tanpa Anda harus mengonfigurasi berbagi untuk objek individual.

Di Katalog Data, Anda dapat membuat katalog federasi untuk setiap namespace atau cluster. Katalog disebut sebagai katalog federasi ketika menunjuk ke entitas di luar Katalog Data. Tabel dan tampilan di namespace Amazon Redshift dicantumkan sebagai tabel individual dalam Katalog Data. Anda dapat berbagi database dan tabel dalam katalog federasi dengan prinsipal IAM dan pengguna SAMB yang dipilih dalam akun yang sama, atau di akun lain dengan Lake Formation. Anda juga dapat menyertakan ekspresi filter baris dan kolom untuk membatasi akses ke data tertentu. Untuk informasi selengkapnya, lihat Pemfilteran data dan keamanan tingkat sel di Lake Formation.

Katalog Data mendukung hierarki metadata tiga tingkat yang terdiri dari katalog, database, dan tabel (dan tampilan). Saat Anda mendaftarkan namespace dengan Katalog Data, hierarki data Amazon Redshift dipetakan ke hierarki 3 tingkat Katalog Data sebagai berikut:

- Namespace Amazon Redshift menjadi katalog multi-level di Katalog Data.
- Basis data Amazon Redshift terkait terdaftar sebagai katalog di Katalog Data.
- Skema Amazon Redshift menjadi database dalam Katalog Data.
- Tabel Amazon Redshift menjadi tabel di Katalog Data.



Dengan hierarki metadata tiga tingkat ini, Anda dapat mengakses tabel Amazon Redshift dengan menggunakan notasi 3 bagian - ".database.table" di Katalog Data. catalog1/catalog2 Selain itu, tim data dapat mempertahankan organisasi yang sama yang digunakan Amazon Redshift untuk mengatur tabel dalam akun Katalog Data.

Di Lake Formation, Anda dapat mengelola data dengan aman dari Amazon Redshift menggunakan kontrol akses berbutir halus untuk sumber daya Katalog Data. Dengan integrasi ini, Anda dapat mengelola, mengamankan, dan menanyakan data analitik dari satu katalog dengan mekanisme kontrol akses umum.

Untuk batasan, lihat Keterbatasan untuk membawa data gudang data Amazon Redshift ke dalam AWS Glue Data Catalog.

Topik

- Manfaat utama
- Peran dan tanggung jawab
- Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog
- Membuat katalog federasi Amazon Redshift
- Melihat objek katalog
- Memperbarui katalog federasi

- Mengakses katalog federasi bersama
- Menghapus katalog federasi
- Menanyakan katalog federasi
- Sumber daya tambahan

Manfaat utama

Mendaftarkan cluster dan ruang nama Amazon Redshift dengan dan menyatukan data di seluruh danau data Amazon S3 AWS Glue Data Catalog dan gudang data Amazon Redshift, menawarkan manfaat berikut:

- Pengalaman kueri seragam — Kueri data dan data terkelola Amazon Redshift Anda di bucket Amazon S3 menggunakan mesin kueri apa pun yang kompatibel dengan Apache Iceberg, seperti Amazon EMR Tanpa Server dan Amazon Athena tanpa harus memindahkan atau menyalin data.
- Akses data yang konsisten di seluruh layanan — Anda tidak perlu memperbarui nama database dan tabel di jalur data saat mengakses sumber data gabungan yang sama dari layanan AWS analitik yang berbeda, karena sumber data terdaftar di Katalog Data.
- Fine-grained kontrol akses — Anda dapat menerapkan izin Lake Formation untuk mengelola akses ke sumber data federasi menggunakan izin kontrol akses berbutir halus.

Peran dan tanggung jawab

Peran	Tanggung jawab
Administrator klaster produsen Amazon Redshift	Mendaftarkan cluster atau namespace dengan Katalog Data.
Administrator danau data Lake Formation	Menerima undangan cluster atau namespace , membuat katalog federasi, dan memberika n akses pada katalog federasi ke prinsipal lainnya.
Lake Formation hanya membaca administrator	Menemukan katalog federasi, menanyakan tabel Amazon Redshift di katalog federasi.

Peran transfer data	Amazon Redshift mengasumsikan atas nama Anda untuk mentransfer data ke dan dari bucket Amazon S3.

Berikut ini adalah langkah-langkah tingkat tinggi untuk memberi pengguna akses ke namespace Amazon Redshift:

1. Di Amazon Redshift, administrator cluster produser mendaftarkan klaster atau namespace dengan Katalog Data.
2. Administrator data lake menerima undangan namespace dari administrator klaster produsen Amazon Redshift, dan membuat katalog federasi di Katalog Data.

   Setelah menyelesaikan langkah ini, Anda dapat mengelola katalog namespace Amazon Redshift dalam Katalog Data.
3. Berikan izin kepada pengguna di katalog, database, dan tabel. Anda dapat membagikan seluruh katalog namespace atau subset tabel dengan pengguna di akun yang sama atau akun lain.

Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog

1. Buat administrator data lake - Buat peran IAM yang diizinkan untuk menerima undangan namespace, dan buat AWS Glue Data Catalog objek (katalog, database,), tables/views dan berikan izin Lake Formation ke pengguna lain.

   Untuk petunjuk langkah demi langkah tentang membuat administrator danau data, lihatBuat administrator danau data.
2. Perbarui izin administrator danau data.

   Selain izin administrator data lake, administrator data lake memerlukan izin berikut untuk menerima undangan namespace Amazon Redshift di Lake Formation, membuat atau memperbarui sumber daya Katalog Data, dan mengaktifkan akses 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. Jika peran IAM yang digunakan untuk membuat katalog federasi bukan administrator data lake, Anda harus memberikan izin kepada peran tersebut. `Create catalog`

   Untuk membuat pembuat katalog

   a. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   b. Pilih peran dan tugas administratif di bawah Administrasi.

   c. PilihIzin.

   d. Pada layar Hibah izin, pilih pengguna atau peran IAM.

   e. Pilih Buat izin katalog.

   f. Secara opsional, Anda juga dapat memberikan izin Buat katalog yang dapat diberikan. Izin yang dapat diberikan memungkinkan pembuat katalog untuk memberikan `Create catalog` izin kepada prinsipal lain.

   g. PilihIzin.

   AWS CLI contoh untuk memberikan izin untuk membuat katalog federasi.

```
aws lakeformation grant-permissions \
--cli-input-json \
'{
    "Principal": {
     "DataLakePrincipalIdentifier":"arn:aws:iam::123456789012:role/Admin"
    },
    "Resource": {
        "Catalog": {
        }
    },
    "Permissions": [
        "CREATE_CATALOG",
        "DESCRIBE"
    ]
}'
```

4. Buat peran administrator hanya baca untuk menemukan katalog federasi Amazon Redshift di Katalog Data dari Amazon Redshift Query Editor v2.

   Untuk menanyakan tabel Amazon Redshift di katalog federasi dari Amazon Redshift Query Editor v2, pastikan bahwa kebijakan peran administrator Baca saja berisi ARN untuk peran terkait layanan 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. Buat peran transfer data yang dapat diasumsikan Amazon Redshift atas nama Anda untuk mentransfer data ke dan dari bucket Amazon S3.

   Saat Anda mengaktifkan akses data lake untuk mesin kueri yang kompatibel dengan Apache Iceberg seperti Athena, Amazon EMR di Amazon EC2 untuk mengakses sumber daya Amazon Redshift di Katalog Data, Anda perlu membuat peran IAM dengan izin yang diperlukan untuk melakukan transfer data ke dan dari 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. Tambahkan kebijakan kepercayaan berikut ke peran transfer data untuk AWS Glue dan layanan Amazon Redshift untuk mengambil peran mentransfer data ke dan dari bucket Amazon S3.

JSON

```
      {
    "Version":"2012-10-17",
    "Statement": [{
        "Effect": "Allow",
        "Principal": {
        "Service": [
        "redshift.amazonaws.com",
        "glue.amazonaws.com"
        ]
    },
        "Action": "sts:AssumeRole"
    }]
}
```

7. Tambahkan kebijakan kunci berikut ke AWS KMS kunci jika Anda menggunakan kunci yang dikelola pelanggan untuk mengenkripsi data di Amazon cluster/namespace Redshift. Ganti nomor akun dengan nomor AWS akun yang valid, dan tentukan nama peran transfer data. Secara default, data di cluster Amazon Redshift dienkripsi menggunakan kunci KMS. Lake Formation menyediakan opsi untuk membuat kunci KMS kustom Anda untuk enkripsi. Jika Anda menggunakan kunci yang dikelola pelanggan, Anda harus menambahkan kebijakan kunci tertentu ke kunci tersebut.

   Untuk informasi selengkapnya tentang mengelola izin kunci terkelola pelanggan, lihat Kunci yang dikelola pelanggan.

   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"
                }
            }
        }
    ]
}
```

Membuat katalog federasi Amazon Redshift

Topik ini menjelaskan langkah-langkah yang perlu Anda ikuti untuk menerima undangan cluster atau namespace, membuat katalog multi-level gabungan, dan memberikan izin kepada prinsipal lainnya. Anda dapat menyelesaikan tugas-tugas ini menggunakan konsol Lake Formation, AWS Command Line Interface (AWS CLI), atau APIs/SDKs. Contoh dalam topik ini menunjukkan produsen cluster/namespace, Katalog Data, dan konsumen data dalam akun yang sama.

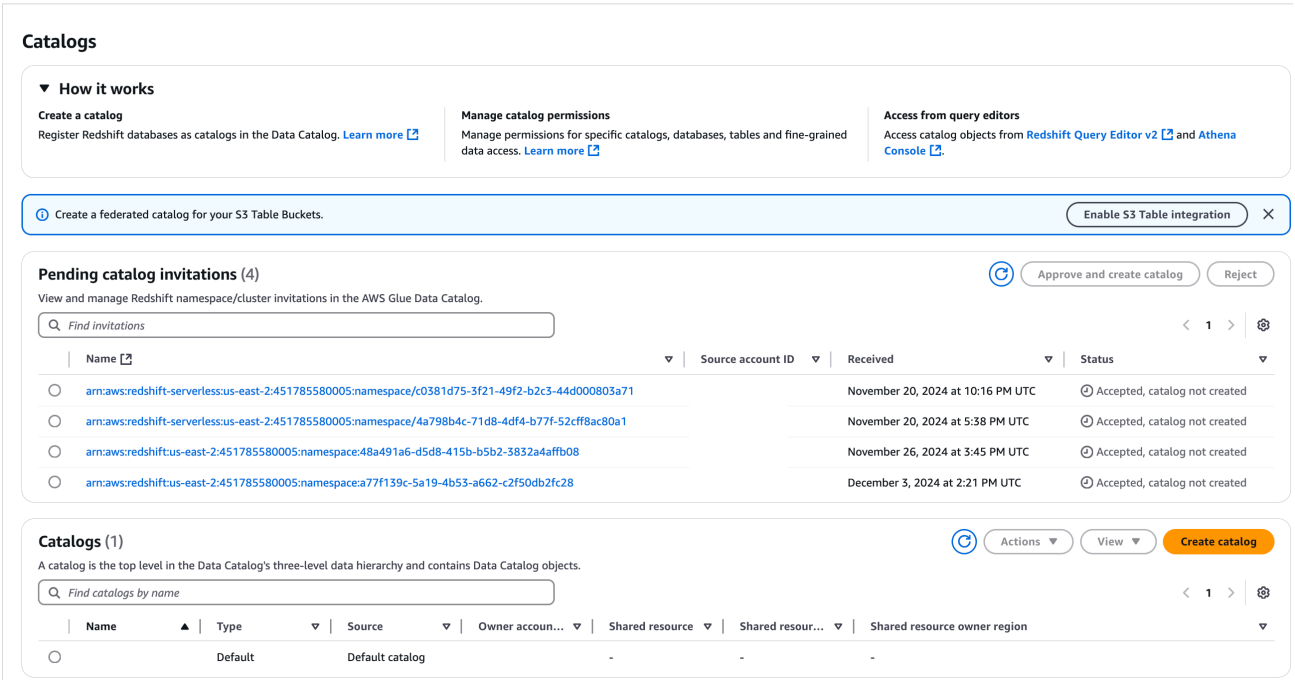
Untuk mempelajari lebih lanjut tentang kemampuan lintas akun Lake Formation, lihatCross-account berbagi data di Lake Formation.

Untuk mengelola namespace Amazon Redshift di Katalog Data

1. Tinjau undangan namespace dan terima.

   Console

   1. Masuk ke konsol Lake Formation sebagai administrator danau data di https://console.aws.amazon.com/lakeformation/. Arahkan ke halaman Katalog di bawah Katalog Data.
   2. Tinjau undangan namespace yang diizinkan untuk diakses. Kolom Status menunjukkan status partisipasi Anda saat ini untuk namespace. Status Tidak diterima menunjukkan bahwa Anda telah ditambahkan ke namespace, tetapi Anda belum menerimanya atau menolak undangan.



   3. Untuk menanggapi namespace atau undangan cluster, pilih nama undangan dan pilih Tinjau undangan. Dalam Menerima atau menolak undangan, tinjau detail undangan. Pilih Terima untuk menerima undangan atau Tolak untuk menolak undangan. Anda tidak mendapatkan akses ke namespace jika Anda menolak undangan.

AWS CLI

Contoh berikut menunjukkan cara melihat, menerima, dan mendaftarkan undangan. Ganti Akun AWS ID dengan ID yang valid Akun AWS . Ganti `data-share-arn` dengan Amazon Resource Name (ARN) aktual yang mereferensikan namespace.

1. Lihat undangan yang tertunda.

```
aws redshift describe-data-shares \
 --data-share-arn 'arn:aws:redshift:us-
east-1:123456789012:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
ds_internal_namespace' \
```

2. Terima undangan.

```
 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. Daftarkan cluster atau namespace di akun Lake Formation. Gunakan operasi RegisterResourceAPI untuk mendaftarkan datashare di Lake Formation. `DataShareArn`adalah parameter input untuk`ResourceArn`.

> **Note**
>
> Ini adalah langkah wajib.

```
aws lakeformation register-resource \
 --resource-arn 'arn:aws:redshift:us-
east-1:123456789012:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
ds_internal_namespace'
```

2. Buat katalog federasi.

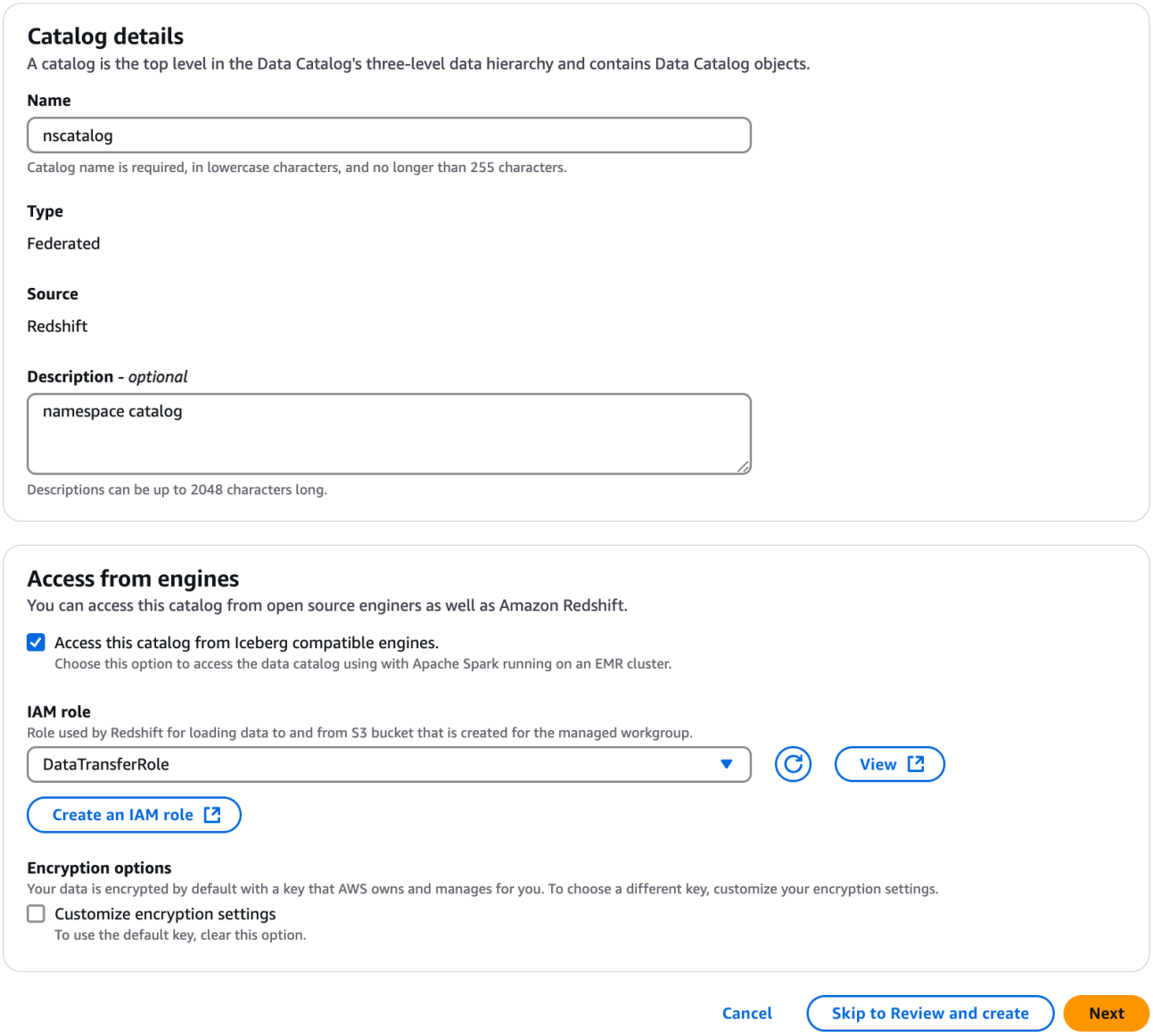
Setelah menerima undangan, Anda perlu membuat katalog federasi di Katalog Data yang memetakan objek di namespace Amazon Redshift ke Katalog Data. Anda harus menjadi

administrator data lake atau pengguna atau peran yang memerlukan izin untuk membuat katalog.

Console

1. Setelah menerima Undangan namespace, halaman Setel detail katalog ditampilkan.
2. Pada halaman Tetapkan detail katalog, masukkan nama unik untuk katalog. Gunakan huruf kecil untuk nama katalog. Nama katalog harus kurang dari atau sama dengan 255 karakter. Anda menggunakan pengenal ini untuk memetakan namespace secara internal dalam hierarki metadata (catalogid.db). Name.schema.table
3. Masukkan deskripsi untuk katalog. Deskripsi harus kurang dari atau sama dengan 2048 karakter.
4. Selanjutnya, pilih kotak centang Akses katalog ini dari mesin yang kompatibel dengan Iceberg untuk memungkinkan mengakses sumber daya Amazon Redshift menggunakan mesin analitik yang kompatibel dengan Apache Iceberg seperti Athena dan Apache Spark di Amazon EMR.

   Anda tidak perlu mengaktifkan akses danau data untuk mengakses katalog gabungan menggunakan Amazon Redshift.



5. Untuk mengaktifkan mesin kueri ini membaca dan menulis ke ruang nama Amazon Redshift, AWS Glue buat klaster Amazon Redshift terkelola dengan sumber daya komputasi dan penyimpanan yang diperlukan untuk melakukan operasi baca dan tulis tanpa memengaruhi beban kerja gudang data Amazon Redshift.

   Anda juga perlu memberikan peran IAM dengan izin yang diperlukan untuk mentransfer data ke dan dari bucket Amazon S3.

6. Secara default, data di klaster Amazon Redshift dienkripsi menggunakan kunci terkelola. AWS Lake Formation menyediakan opsi untuk membuat kunci KMS kustom Anda untuk enkripsi. Jika Anda menggunakan kunci yang dikelola pelanggan, Anda harus menambahkan kebijakan kunci tertentu ke kunci tersebut.

Pilih pengaturan Sesuaikan enkripsi jika Anda menggunakan kunci yang dikelola pelanggan untuk mengenkripsi data di Amazon cluster/namespace Redshift. Untuk menggunakan kunci kustom, Anda harus menambahkan kebijakan kunci kustom terkelola tambahan ke kunci KMS Anda. Untuk informasi selengkapnya, lihat Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog.

AWS CLI

Gunakan kode contoh berikut untuk membuat katalog dengan data Amazon Redshift yang dipublikasikan ke Katalog Data menggunakan. AWS CLI

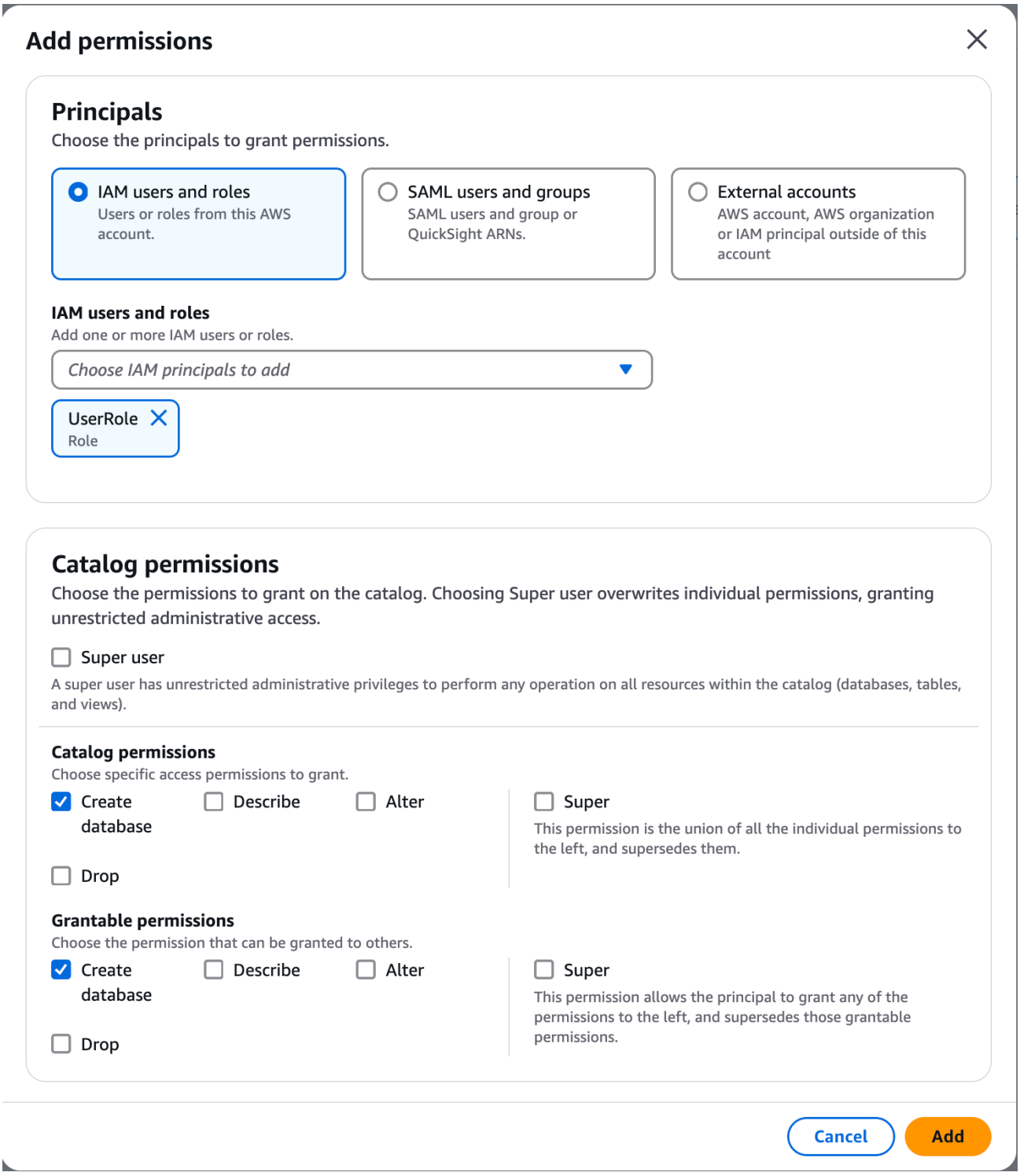
```
aws glue create-catalog
--cli-input-json \
'{
    "Name": "nscatalog",
    "CatalogInput": {
        "Description": "Redshift federated catalog",
        "CreateDatabaseDefaultPermissions" : [],
        "CreateTableDefaultPermissions": [],
        "FederatedCatalog": {
            "Identifier": "arn:aws:redshift:us-
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. Berikan izin kepada pengguna di akun Anda atau di akun eksternal.

Konsol Manajemen AWS

1. Pilih Berikutnya untuk memberikan izin kepada pengguna lain di katalog, database, dan tabel bersama.

2. Pada layar Tambahkan izin, pilih prinsip dan jenis izin yang akan diberikan.



a. Di bagian Principal, pilih jenis principal lalu tentukan principal yang akan diberi izin.

- Peran dan pengguna IAM – Pilih satu atau beberapa pengguna atau peran dari daftar peran dan pengguna IAM.

- Pengguna dan grup SAMB — Untuk SALL dan Amazon Quick pengguna dan grup, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Amazon Cepat. Tekan Enter setelah setiap ARN.

  Untuk informasi tentang cara membuat ARN, lihat AWS CLI memberikan dan mencabut AWS CLI perintah.
- Akun eksternal — Untuk AWS, AWS organisasi, atau IAM Principal masukkan satu atau beberapa ID AWS akun, ID organisasi, ID unit organisasi, atau ARN yang valid untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID. ID organisasi terdiri dari “o-” diikuti oleh 10-32 huruf kecil atau digit. ID unit organisasi dimulai dengan “ou-” diikuti oleh 4-32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

b. Di bagian Izin, pilih izin dan izin yang dapat diberikan.

   Di Izin katalog, pilih satu atau beberapa izin untuk diberikan. Di bawah Izin yang dapat diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.

   Pilih Super user untuk memberikan pengguna izin tak terbatas ke sumber daya (database, tabel, tampilan) dalam katalog.

3. Pilih Tambahkan.

AWS CLI

Gunakan contoh berikut untuk memberikan izin katalog, database, dan tabel menggunakan: AWS CLI

- Contoh berikut menunjukkan cara memberikan izin pada katalog federasi.

```
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": [
             ]
    }'
```

- Gunakan contoh berikut untuk memberikan izin pada 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"
              ]
          }'
```

- Contoh berikut menunjukkan cara memberikan izin pada tabel di 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. Pilih Berikutnya untuk meninjau detail katalog dan membuat katalog gabungan. Katalog federasi yang baru dibuat dan objek katalog muncul di halaman Katalog.

   Katalog federasi Amazon Redshift direferensikan dengan. `catalogID = 123456789012:Redshift-federated catalog id`

Melihat objek katalog

Setelah Anda membuat katalog federasi, Anda dapat melihat objek dalam katalog menggunakan konsol Lake Formation atau AWS CLI.

Konsol Manajemen AWS

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Pilih Katalog di bawah Katalog Data.
3. Pilih katalog gabungan dari daftar di halaman Katalog.
4. Halaman ringkasan katalog menunjukkan objek katalog (database dan tabel) tempat Anda memiliki izin. Tab Izin menunjukkan prinsipal IAM yang telah diberikan izin pada objek ini.

AWS CLI

- AWS CLI Contoh berikut menunjukkan cara meminta katalog tingkat atas.

```
aws glue get-catalog \
--catalog-id 123456789012:nscatalog
```

Respons

```
{
    "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": []
    }
}
```

- Contoh berikut menunjukkan cara meminta semua katalog di akun.

```
aws glue get-catalogs \
  --recursive
```

- Contoh permintaan berikut menunjukkan cara mendapatkan katalog tingkat database Amazon Redshift.

```
aws glue get-catlog \
 --catalog-id 123456789012:namespace catalog name/redshift database name
```

- Contoh permintaan berikut menunjukkan cara mendapatkan database di katalog tingkat database Amazon Redshift.

```
aws glue get-databases \
--catalog-id 123456789012:namespace catalog name/redshift database name
```

- Contoh permintaan berikut menunjukkan cara mendapatkan tabel Amazon Redshift di katalog.

```
aws glue get-table \
  --catalog-id 123456789012:parent catalog name/redshift database \
  --database-name redshift schema name \
  --name table name
```

- Berikut contoh menunjukkan bagaimana untuk mendapatkan semua tabel dari database Amazon Redshift.

```
aws glue get-tables \
 --catalog-id 123456789012:namespace catalog name/redshift database name \
 --database-name RS schema name
```

Memperbarui katalog federasi

Anda dapat memperbarui katalog federasi Amazon Redshift di Katalog Data menggunakan konsol Lake Formation, AWS CLI atau operasi API. UpdateCatalog

Konsol Manajemen AWS

Ikuti langkah-langkah ini untuk memperbarui katalog federasi Anda menggunakan konsol Lake Formation.

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi kiri, pilih Katalog di bawah Katalog Data.
3. Pada halaman Katalog, pilih katalog federasi Amazon Redshift yang ingin Anda perbarui.
4. Di bawah Tindakan, pilih Edit.

5. Pada layar Set Catalog details, di bawah bagian Access from engine, pilih Access this catalog from Iceberg compatible engine. Memeriksa opsi ini akan memungkinkan akses data lake untuk mesin kueri yang kompatibel dengan Apache Iceberg.
6. Selanjutnya, buat peran IAM baru atau pilih peran IAM yang ada yang memiliki kebijakan yang memberikan izin untuk melakukan transfer data ke dan dari bucket Amazon S3.

   Untuk informasi selengkapnya tentang izin, lihatPrasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog.
7. Secara default, data Anda di klaster Amazon Redshift dienkripsi menggunakan file. Kunci yang dikelola AWS Jika Anda memilih untuk mengenkripsi data menggunakan kunci yang dikelola pelanggan, buat kunci KMS atau pilih kunci yang ada yang memiliki izin yang ditentukan di bagian tersebut. Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog
8. Pilih Simpan.

   Setelah berhasil diselesaikan, halaman detail Katalog menunjukkan nama grup kerja terkelola dengan status sebagai “Sukses”.

AWS CLI

Berikut ini adalah contoh input `update-catalog` CLI dengan akses danau data dinonaktifkan dengan menetapkan nilai `DataLakeAacess` parameter sebagai. `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
        }
```

```
        }
    }
}'
```

Mengakses katalog federasi bersama

AWS Lake Formation Kemampuan lintas akun memungkinkan pengguna untuk berbagi data lake terdistribusi dengan aman di beberapa Akun AWS s, AWS organisasi, atau langsung dengan prinsipal IAM di akun lain yang menyediakan akses halus ke metadata dan data yang mendasarinya.

Lake Formation menggunakan layanan AWS Resource Access Manager (AWS RAM) untuk memfasilitasi berbagi sumber daya. Saat Anda membagikan sumber daya katalog dengan akun lain, AWS RAM kirimkan undangan ke akun penerima hibah untuk menerima atau menolak hibah sumber daya.

Layanan analitik terintegrasi seperti Amazon Athena dan Redshift Spectrum memerlukan tautan sumber daya untuk dapat menyertakan sumber daya bersama dalam kueri. Prinsipal perlu membuat tautan sumber daya ke sumber daya bersama dari yang lain. AWS Glue Data Catalog Akun AWS Untuk informasi selengkapnya tentang tautan sumber daya, lihat Cara kerja tautan sumber daya di Lake Formation.

Kontainer tautan Katalog adalah objek Katalog Data, yang mereferensikan katalog tingkat basis data federasi lokal atau lintas akun dari akun lain. AWS Anda juga dapat membuat tautan database dan tautan tabel dalam wadah tautan katalog. Saat membuat tautan database atau tautan tabel, Anda harus menentukan sumber daya target yang berada di bawah katalog tingkat database Amazon Redshift target yang sama (database Amazon Redshift).

Untuk membuat wadah tautan katalog, Anda memerlukan Lake Formation `CREATE_CATALOG` atau `glue:CreateCatalog` izin.

Anda harus memiliki pengaturan versi Cross account versi 4 atau lebih tinggi untuk berbagi database atau tabel dalam katalog federasi di seluruh Akun AWS s.

Membuat wadah tautan katalog ke katalog federasi lintas akun

Anda dapat membuat wadah tautan katalog yang mengarah ke katalog federasi tingkat database Redshift di AWS Wilayah mana pun dengan menggunakan AWS Lake Formation konsol, API, AWS Glue `CreateCatalog` atau (). AWS Command Line Interface AWS CLI

Untuk membuat wadah tautan katalog ke katalog bersama (konsol)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai kepala sekolah yang memiliki `CREATE_CATALOG` izin Lake Formation.
2. Di panel navigasi, pilih Katalog, lalu pilih Buat katalog.
3. Pada halaman Setel detail katalog, berikan informasi berikut:

   Nama

   Masukkan nama yang mematuhi aturan yang sama dengan nama katalog. Namanya bisa sama dengan katalog bersama target.

   Tipe

   Pilih wadah tautan Katalog sebagai jenis katalog.

   Sumber

   Pilih `Redshift`.

   Katalog Target Redshift

   Pilih katalog federasi tingkat database Redshift atau pilih katalog lokal (milik) dari daftar.

   Daftar ini berisi semua katalog yang dibagikan dengan akun Anda. Perhatikan ID akun pemilik katalog terdaftar dengan setiap katalog. Jika Anda tidak melihat katalog yang Anda tahu telah dibagikan dengan akun Anda, periksa hal berikut:

   - Jika Anda bukan administrator data lake, periksa apakah administrator danau data memberi Anda izin Lake Formation di katalog.
   - Jika Anda adalah administrator data lake, dan akun Anda tidak berada dalam AWS organisasi yang sama dengan akun pemberi, pastikan bahwa Anda telah menerima undangan berbagi sumber daya AWS Resource Access Manager (AWS RAM) untuk katalog. Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

   > **Note**
   >
   > Saat membuat wadah tautan katalog melalui konsol, tarik-turun Katalog Redshift Target mungkin ditampilkan **`No matches`** saat mencoba memilih katalog Redshift lintas akun. Terlepas dari tampilan ini, Anda dapat secara manual memasukkan ARN target dari katalog tingkat database federasi Amazon Redshift (database Amazon

Redshift) di bidang input, dan formulir akan tetap berfungsi dengan benar. Sebagai contoh: `arn:aws:glue:us-east-1:123456789012:catalog/federated-catalog-redshift/dev`.
Perilaku ini terjadi karena konsol hanya dapat mencari calon potensial dalam akun yang saat ini masuk. Dropdown dimaksudkan sebagai fitur pelengkapan otomatis, tetapi Anda masih dapat memasukkan ARN secara manual untuk akses lintas akun.

4. Untuk mengaktifkan mesin kueri Apache Iceberg membaca dan menulis ke ruang nama Amazon Redshift, buat klaster Amazon Redshift AWS Glue terkelola dengan sumber daya komputasi dan penyimpanan yang diperlukan untuk melakukan operasi baca dan tulis tanpa memengaruhi beban kerja gudang data Amazon Redshift. Anda perlu memberikan peran IAM dengan izin yang diperlukan untuk mentransfer data ke dan dari bucket Amazon S3.
5. Pilih Berikutnya.
6. (Opsional) Pilih Tambahkan izin untuk memberikan izin kepada prinsipal lain.

   Namun, pemberian izin pada wadah tautan katalog tidak memberikan izin pada katalog target (tertaut). Anda harus memberikan izin pada katalog target secara terpisah agar tautan katalog terlihat di Athena.
7. Selanjutnya, tinjau detail wadah tautan katalog dan pilih Buat katalog.

   Anda kemudian dapat melihat nama penampung tautan di bawah halaman Katalog.

   Sekarang, Anda dapat membuat tautan database dan tautan tabel di wadah tautan katalog untuk mengaktifkan akses dari mesin kueri.

Buat contoh CLI wadah tautan katalog

- Dalam contoh berikut, `TargetRedshiftCatalog` objek menentukan arn katalog tingkat database federasi Amazon Redshift (database Amazon Redshift). `DataLakeAccess`Harus diaktifkan saat Anda membuat wadah tautan katalog.

```
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"
              }
            }
        }
    }'
```

Membuat tautan sumber daya di bawah wadah tautan katalog

Anda dapat membuat tautan sumber daya ke database dan tautan tabel di bawah wadah tautan katalog. Saat membuat tautan sumber daya database atau tautan sumber daya tabel, Anda harus menentukan sumber daya target yang berada di bawah katalog tingkat database Amazon Redshift target yang sama (database Amazon Redshift) yang ditunjuk oleh penampung tautan.

Anda dapat membuat tautan sumber daya ke database Amazon Redshift bersama atau tabel menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface ()AWS CLI.

- Untuk petunjuk mendetail, lihat Membuat tautan sumber daya ke database Katalog Data bersama.

  Berikut ini adalah AWS CLI contoh untuk membuat link sumber daya database di bawah wadah link katalog.

```
aws glue create-database \
  --cli-input-json \
    '{
        "CatalogId": "111122223333:linkcontainer",
        "DatabaseInput": {
            "Name": "dblink",
             "TargetDatabase": {
               "CatalogId": "123456789012:nscatalog/dev",
                "DatabaseName": "schema1"
             }
        }
    }'
```

- Untuk membuat tautan sumber daya tabel di bawah wadah tautan katalog, Anda harus terlebih dahulu membuat AWS Glue database di lokal AWS Glue Data Catalog untuk memuat tautan sumber daya tabel.

  Untuk informasi selengkapnya tentang membuat tautan sumber daya ke tabel bersama, lihatMembuat tautan sumber daya ke tabel Katalog Data bersama.

  - Buat database untuk memuat contoh tautan sumber daya tabel

```
aws glue create-database \
  --cli-input-json \
      '{
          "CatalogId": "111122223333:linkcontainer",
          "DatabaseInput": {
              "Name": "db1",
              "Description": "creating parent database for table link"
          }
      }'
```

  - Buat contoh tautan sumber daya tabel

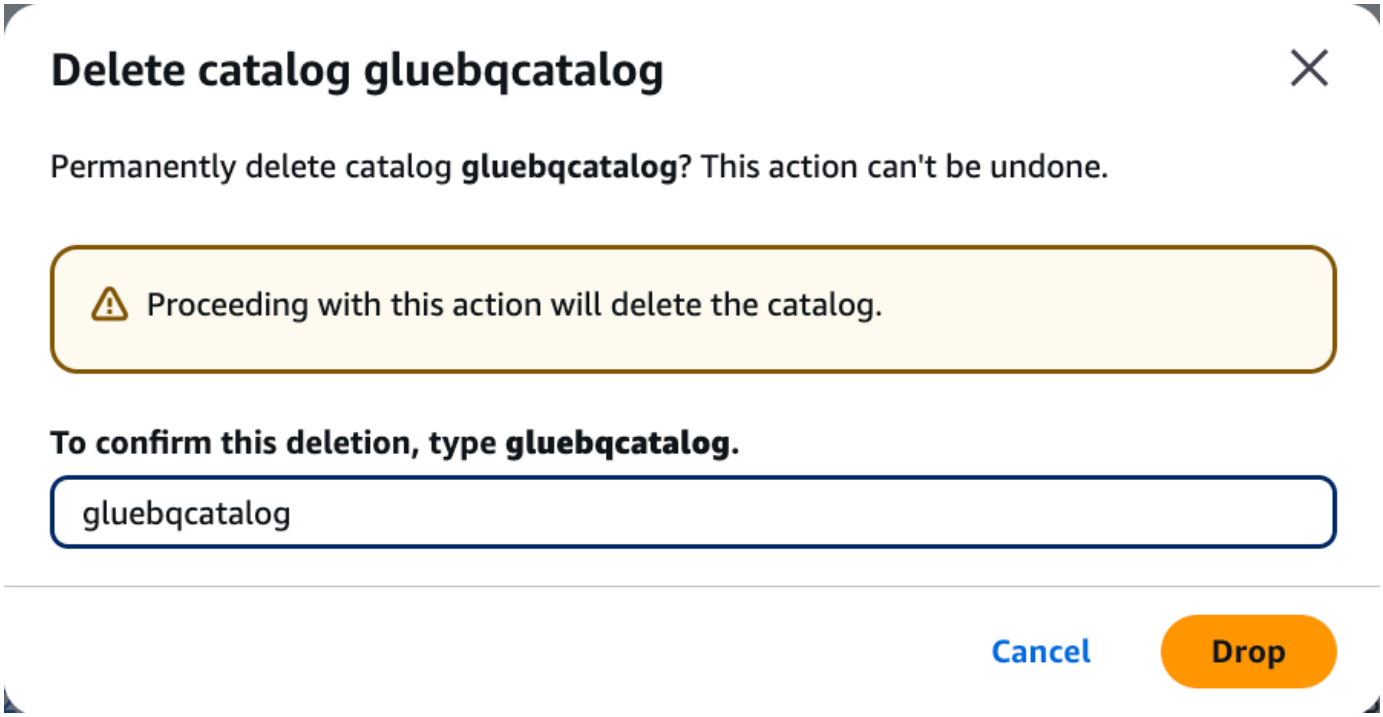
```
aws glue create-table \
  --cli-input-json \
    '{
        "CatalogId": "111122223333:linkcontainer",
         "DatabaseName": "db1",
        "TableInput": {
            "Name": "tablelink",
            "TargetTable": {
                "CatalogId": "123456789012:nscatalog/dev",
               "DatabaseName": "schema1",
                "Name": "table1"
             }
        }
    }'
```

Menghapus katalog federasi

Anda dapat menghapus katalog federasi yang Anda buat dalam AWS Glue Data Catalog menggunakan `glue:DeleteCatalog` operasi atau konsol. AWS Lake Formation

Untuk menghapus katalog federasi (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Katalog di bawah Katalog Data.
3. Pilih katalog yang ingin Anda hapus dari daftar katalog.
4. Pilih Hapus dari Tindakan.
5. Pilih Jatuhkan untuk mengonfirmasi dan katalog federasi akan dihapus dari Katalog Data.



Untuk menghapus katalog federasi (CLI)

- ```
  aws glue delete-catalog
   --catalog-id 123456789012:catalog name
  ```
  ```

Menanyakan katalog federasi

Setelah Anda memberikan izin kepada prinsipal lain, mereka dapat masuk dan mulai menanyakan tabel di katalog federasi dengan masuk ke alat SQL menggunakan Amazon Redshift, Amazon EMR, dan ETL. Amazon Athena AWS Glue

Untuk informasi selengkapnya tentang menghubungkan ke titik akhir ekstensi Apache Iceberg Rest atau aplikasi Spark mandiri, lihat Mengakses bagian dalam Panduan Pengembang. AWS Glue Data Catalog AWS Glue Data Catalog AWS Glue

Anda dapat menggunakan kueri bahasa definisi data (DDL) untuk membuat dan mengelola tabel dalam database menggunakan Apache Spark di Amazon EMR. Untuk membuat dan menghapus tabel di database Amazon Redshift, prinsipal harus memiliki izin Lake Formation`Create table`. `Drop`

Untuk informasi selengkapnya tentang pemberian izin Katalog Data, lihat. Memberikan izin pada sumber daya Katalog Data

Untuk informasi selengkapnya tentang menanyakan sumber daya katalog Amazon Athena, lihat Menanyakan AWS Glue Data Catalog dari Amazon Athena dalam Panduan Pengguna Amazon Athena.

Sumber daya tambahan

Anda dapat menggunakan Amazon SageMaker Lakehouse untuk mencapai akses terpadu ke data di gudang data dan data lake. Melalui SageMaker Lakehouse, Anda dapat menggunakan analitik pilihan, pembelajaran mesin, dan mesin intelijen bisnis melalui API REST Apache Iceberg yang terbuka untuk membantu memastikan akses yang aman ke data dengan kontrol akses yang konsisten dan halus.

- SageMaker Lokakarya Amazon
- Sederhanakan akses data untuk perusahaan Anda menggunakan Amazon SageMaker Lakehouse

Menggabungkan ke sumber data eksternal di AWS Glue Data Catalog

Anda dapat menghubungkan AWS Glue Data Catalog (Katalog Data) ke gudang data seperti Amazon Redshift, Snowflake, database cloud seperti Amazon RDS, Oracle Amazon DynamoDB, dan layanan

streaming seperti Amazon MSK, dan sistem lokal seperti Teradata menggunakan koneksi. AWS Glue Koneksi ini disimpan dalam AWS Glue Data Catalog dan terdaftar dengan AWS Lake Formation, memungkinkan Anda untuk membuat katalog federasi untuk setiap sumber data yang tersedia.

Katalog federasi adalah wadah tingkat atas yang menunjuk ke database dalam sistem data eksternal. Ini memungkinkan Anda untuk menanyakan data langsung dari sistem data eksternal tanpa proses ekstrak, transformasi, dan beban (ETL).

Untuk informasi selengkapnya tentang AWS Glue koneksi, lihat Menyambungkan ke data di Panduan AWS Glue Pengembang.

Administrator data lake dapat membuat katalog federasi menggunakan Amazon SageMaker Lakehouse atau. Amazon Athena

Administrator data lake kemudian dapat memberikan izin halus pada objek dalam katalog menggunakan Lake Formation, mengontrol akses di berbagai tingkatan seperti katalog, database, tabel, kolom, baris, atau sel. Analis data dapat menemukan dan menanyakan sumber data yang dikatalogkan menggunakan Athena, dengan Lake Formation menegakkan kebijakan akses yang ditentukan. Analis dapat menggabungkan data di berbagai sumber dalam satu kueri tanpa perlu terhubung ke setiap sumber secara individual.

Topik

- Alur kerja
- Prasyarat untuk menghubungkan Katalog Data ke sumber data eksternal
- Membuat katalog federasi menggunakan AWS Glue koneksi
- Melihat objek katalog
- Menghapus katalog federasi
- Menanyakan katalog federasi
- Sumber daya tambahan

Alur kerja

Administrator data lake atau pengguna dengan izin yang diperlukan menyelesaikan langkah-langkah berikut untuk menghubungkan AWS Glue Data Catalog ke sumber data eksternal.

1. Membuat AWS Glue koneksi ke sumber data. Saat Anda mendaftarkan koneksi, peran IAM yang digunakan dalam mendaftarkan koneksi harus memiliki akses ke fungsi Lambda dan lokasi bucket tumpahan Amazon S3.
2. Mendaftarkan koneksi dengan Lake Formation.
3. Membuat katalog federasi dalam Katalog Data menggunakan AWS Glue koneksi untuk terhubung ke sumber data yang tersedia. Database, tabel, dan tampilan secara otomatis dikatalogkan di Katalog Data, dan terdaftar di Lake Formation.
4. Memberikan akses ke katalog, database, dan tabel tertentu kepada analis data menggunakan izin Lake Formation. Fine-grained Kebijakan kontrol akses dapat didefinisikan di seluruh danau data, gudang, dan sumber OLTP menggunakan Lake Formation, memungkinkan filter keamanan tingkat baris dan tingkat kolom.

   Analis data kemudian dapat mengakses semua data melalui Katalog Data menggunakan kueri SQL di Athena, tanpa memerlukan koneksi terpisah atau kredensi sumber data. Analis dapat menjalankan kueri SQL federasi yang memindai data dari berbagai sumber, menggabungkan data di tempat tanpa jalur data yang kompleks.

Prasyarat untuk menghubungkan Katalog Data ke sumber data eksternal

Untuk menghubungkan AWS Glue Data Catalog ke sumber data eksternal, mendaftarkan koneksi dengan Lake Formation, dan mengatur katalog federasi, Anda harus melengkapi persyaratan berikut:

> **Note**
>
> Kami menyarankan agar administrator danau data Lake Formation membuat AWS Glue koneksi untuk terhubung ke sumber data eksternal, dan membuat katalog federasi.

1. Buat IAM role.

   - Buat peran yang memiliki izin yang diperlukan untuk menyebarkan sumber daya (fungsi Lambda, keranjang tumpahan Amazon S3, peran IAM, dan AWS Glue koneksi) yang diperlukan untuk membuat sambungan ke sumber data eksternal.
   - Buat peran yang memiliki izin minimum yang diperlukan untuk mengakses properti AWS Glue koneksi (fungsi Lambda dan bucket tumpahan Amazon S3). Ini adalah peran yang akan Anda sertakan saat mendaftarkan koneksi dengan Lake Formation.

Untuk menggunakan Lake Formation untuk mengelola dan mengamankan data di danau data Anda, Anda harus mendaftarkan AWS Glue koneksi dengan Lake Formation. Dengan demikian, Lake Formation dapat menjual kredensil ke Amazon Athena untuk menanyakan sumber data federasi.

Peran harus memiliki `Select` atau `Describe` izin pada bucket Amazon S3 dan fungsi 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": "*"
    }
  ]
```

```
}
```

- Tambahkan kebijakan kepercayaan berikut ke peran IAM yang digunakan dalam mendaftarkan koneksi:

  JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "lakeformation.amazonaws.com"
              ]
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

- Administrator data lake yang mendaftarkan koneksi harus memiliki `iam:PassRole` izin pada peran tersebut.

  Berikut ini adalah kebijakan inline yang memberikan izin ini. Ganti *<account-id>* dengan nomor AWS akun yang valid, dan ganti *<role-name>* dengan nama peran.

  JSON

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

```
            ]
        }
    ]
}
```

- Untuk membuat katalog federasi di Katalog Data, pastikan peran IAM yang Anda gunakan adalah administrator danau data Lake Formation dengan memeriksa pengaturan data lake (). `aws lakeformation get-data-lake-settings`

  Jika Anda bukan administrator danau data, Anda memerlukan `CREATE_CATALOG` izin Lake Formation untuk membuat katalog. Contoh berikut menunjukkan cara memberikan izin yang diperlukan untuk membuat katalog.

```
aws lakeformation grant-permissions \
--cli-input-json \
        '{
            "Principal": {
             "DataLakePrincipalIdentifier":"arn:aws:iam::123456789012:role/non-admin"
            },
            "Resource": {
                "Catalog": {
                }
            },
            "Permissions": [
                "CREATE_CATALOG",
                "DESCRIBE"
            ]
        }'
```

  Selain `CREATE_CATALOG` izin, Anda perlu `DATA_LOCATION_ACCESS` koneksi terdaftar untuk membuat katalog federasi. Administrator data lake memiliki izin lokasi data implisit dan tidak memerlukan hibah ini. Untuk informasi selengkapnya, lihat Izin Formasi Danau Implisit.

  Contoh berikut menunjukkan cara memberikan `DATA_LOCATION_ACCESS` pada koneksi terdaftar.

```
aws lakeformation grant-permissions \
  --cli-input-json \
    '{
```

```
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/non-admin-
role"
      },
      "Resource": {
        "DataLocation": {
          "CatalogId": "123456789012",
          "ResourceArn": "arn:aws:glue:us-east-1:123456789012:connection/
connection-name"
        }
      },
      "Permissions": ["DATA_LOCATION_ACCESS"]
    }'
```

2. Tambahkan kebijakan kunci berikut ke AWS KMS kunci jika Anda menggunakan kunci yang dikelola pelanggan untuk mengenkripsi data di sumber data. Ganti nomor akun dengan nomor AWS akun yang valid, dan tentukan nama peran. Secara default, data dienkripsi menggunakan kunci KMS. Lake Formation menyediakan opsi untuk membuat kunci KMS kustom Anda untuk enkripsi. Jika Anda menggunakan kunci yang dikelola pelanggan, Anda harus menambahkan kebijakan kunci tertentu ke kunci tersebut.

   Untuk informasi selengkapnya tentang mengelola izin kunci terkelola pelanggan, lihat Kunci terkelola pelanggan.

   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"
        }
    ]
```

```
}
```

Membuat katalog federasi menggunakan AWS Glue koneksi

Untuk menghubungkan AWS Glue Data Catalog ke sumber data eksternal, Anda perlu menggunakan AWS Glue koneksi yang memungkinkan komunikasi dengan sumber data eksternal. Anda dapat membuat AWS Glue koneksi menggunakan AWS Glue konsol, Buat API koneksi, dan konsol Amazon SageMaker Lakehouse.

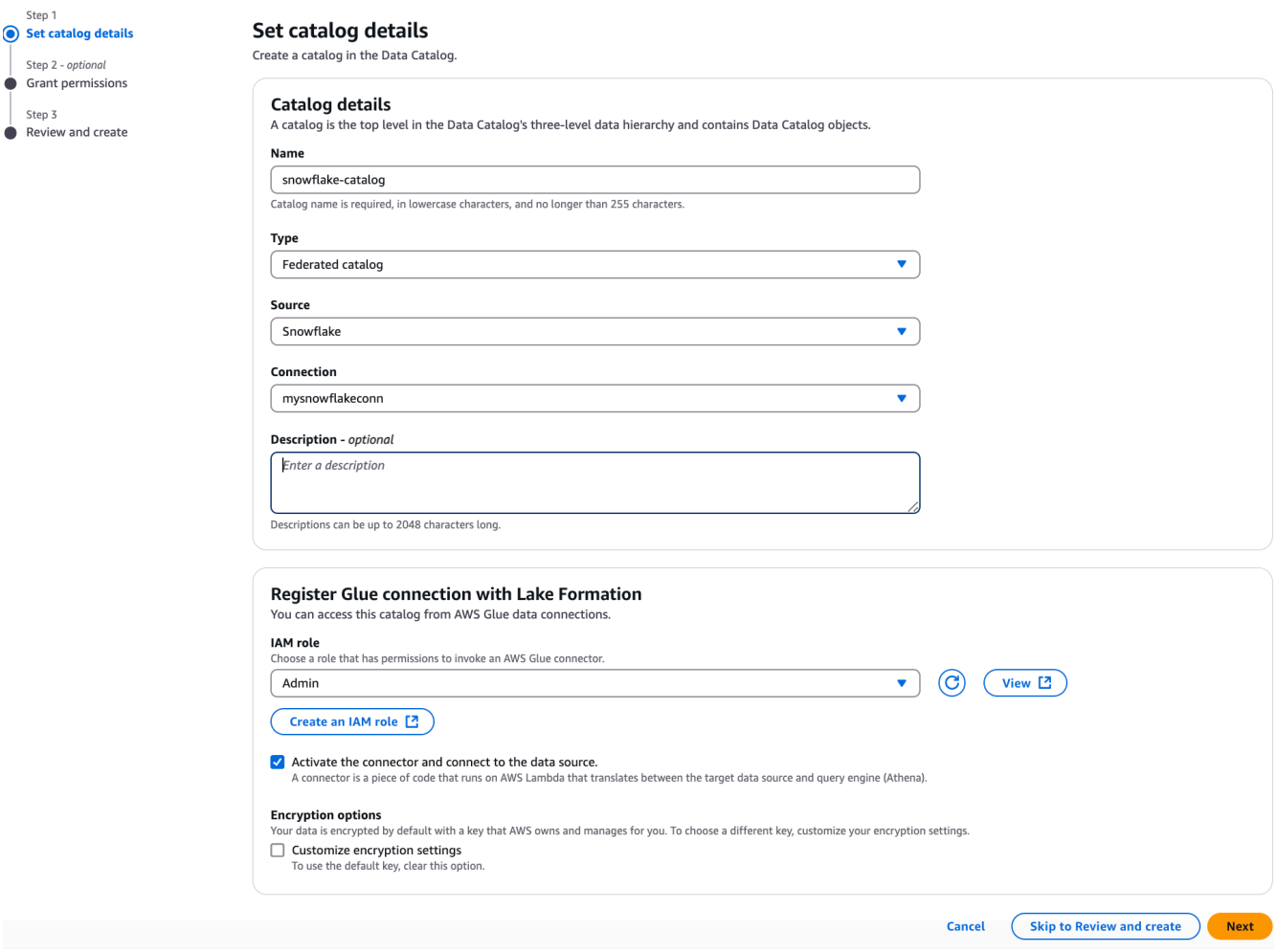
Untuk petunjuk langkah demi langkah untuk membuat AWS Glue sambungan, lihat Menyambungkan ke data di Panduan AWS Glue Pengembang atau Membuat koneksi di Amazon SageMaker Lakehouse.

Saat pengguna menjalankan kueri pada tabel federasi, Lake Formation menjual kredensil yang memanggil AWS Lambda fungsi yang ditentukan dalam AWS Glue koneksi untuk mengambil objek metadata dari sumber data.

Konsol Manajemen AWS

Untuk membuat katalog federasi dari sumber data eksternal dan menyiapkan izin (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Katalog di bawah Katalog Data.
3. Pilih opsi Buat katalog.
4. Pada halaman Set Catalog details, masukkan informasi berikut:



- Nama — Nama unik untuk katalog federasi Anda. Nama tidak dapat diubah, dan harus dalam huruf kecil. Nama dapat terdiri dari maksimum 255 karakter. akun.
- Jenis - Pilih katalog federasi sebagai jenis katalog.
- Sumber — Pilih sumber data dari dropdown. Sumber data yang telah Anda buat koneksinya akan ditampilkan. Untuk informasi selengkapnya tentang membuat AWS Glue sambungan ke sumber data eksternal, lihat Membuat koneksi untuk konektor di Panduan AWS Glue Pengembang atau Membuat koneksi di Amazon SageMaker Lakehouse.
- Koneksi — Pilih AWS Glue koneksi yang ada ke sumber data.
- Deskripsi — Masukkan deskripsi untuk katalog yang dibuat dari sumber data.

5. Pilih peran IAM untuk Lake Formation untuk berasumsi untuk menjual kredensil untuk mesin kueri untuk mengakses data dari sumber data. Peran ini harus memiliki izin yang diperlukan untuk mengakses AWS Glue koneksi dan memanggil fungsi Lambda untuk mengakses data dari sumber data eksternal.

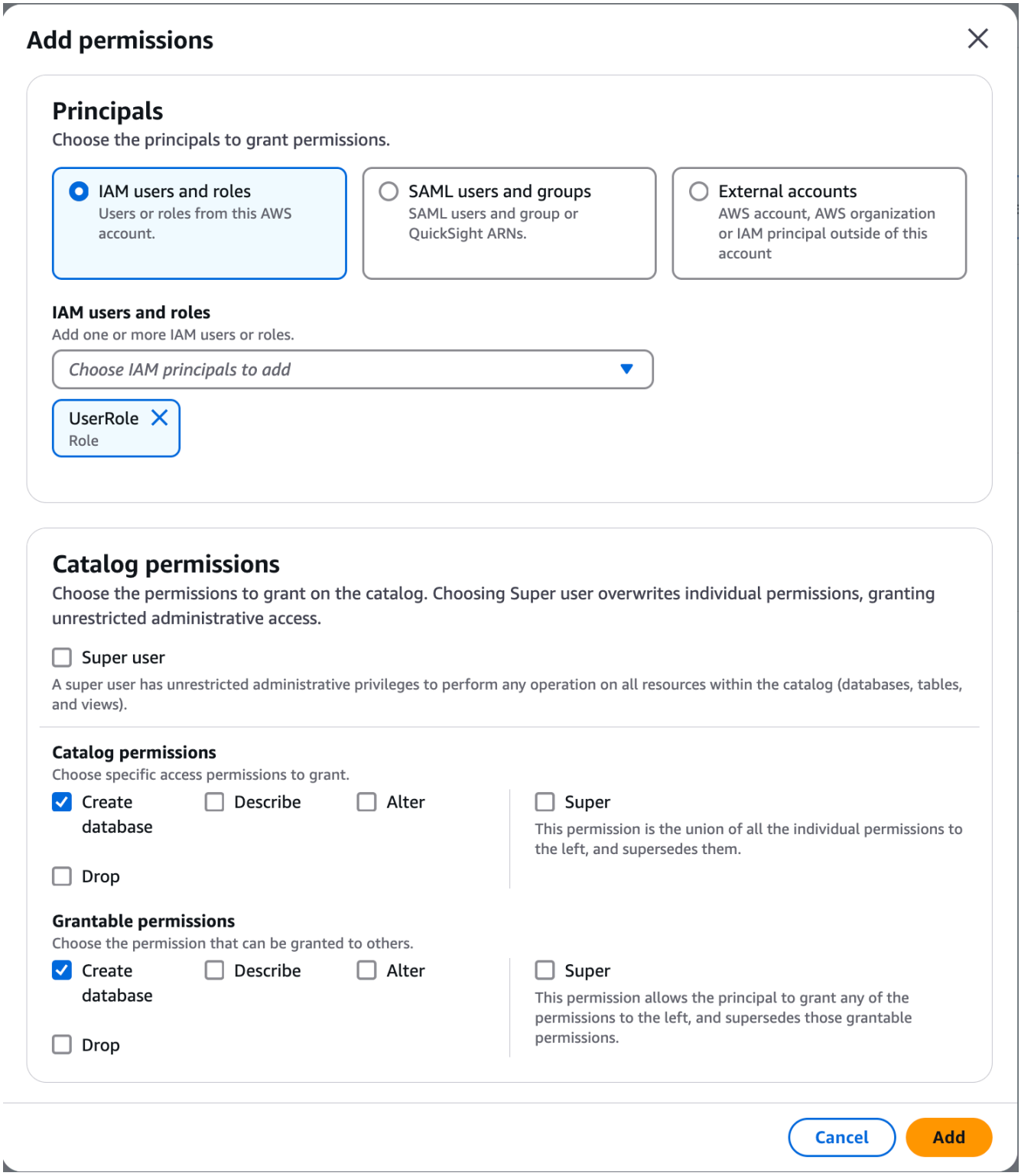
Anda juga dapat membuat peran baru di konsol IAM.

Lihat Prasyarat untuk menghubungkan Katalog Data ke sumber data eksternal bagian untuk izin yang diperlukan.

6. Pilih opsi Aktifkan konektor untuk terhubung ke sumber data untuk mengaktifkan Athena menjalankan kueri federasi.

   Untuk daftar konektor yang didukung, lihat Daftarkan koneksi Anda di Panduan Pengguna Amazon Athena.

7. Opsi enkripsi - Pilih opsi Sesuaikan pengaturan enkripsi jika Anda ingin menggunakan kunci khusus untuk mengenkripsi katalog. Untuk menggunakan kunci kustom, Anda harus menambahkan kebijakan kunci kustom terkelola tambahan ke kunci KMS Anda.
8. Pilih Berikutnya untuk memberikan izin kepada prinsipal lain.
9. Pada halaman Hibah izin, pilih Tambahkan izin.
10. Pada layar Tambahkan izin, pilih prinsip dan jenis izin yang akan diberikan.



- Di bagian Principal, pilih jenis principal lalu tentukan principal yang akan diberi izin.
  - Peran dan pengguna IAM – Pilih satu atau beberapa pengguna atau peran dari daftar peran dan pengguna IAM.

- Pengguna dan grup SAFL — Untuk SALL dan Amazon Quick pengguna dan grup, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAFL, atau ARN untuk pengguna atau grup Amazon Quick. Tekan Enter setelah setiap ARN.

- Di bagian Izin, pilih izin dan izin yang dapat diberikan.

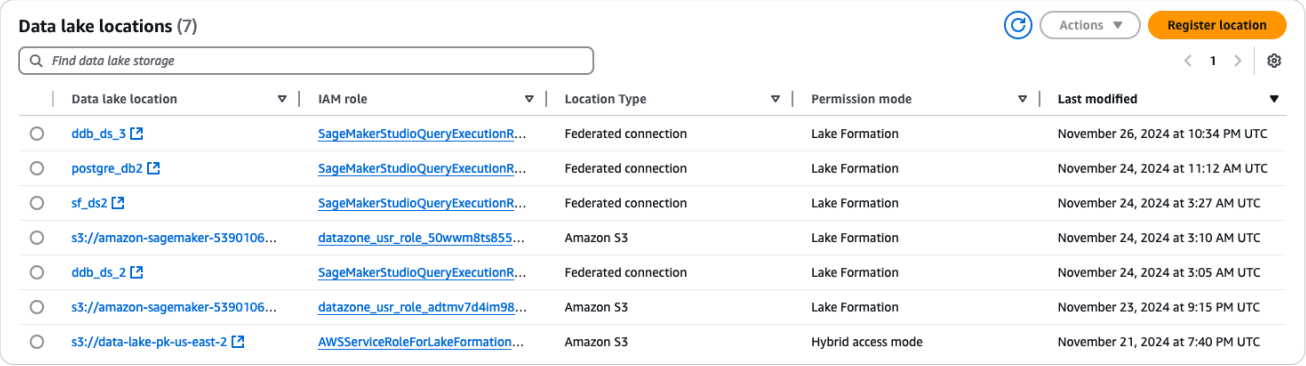
  Di Izin katalog, pilih satu atau beberapa izin untuk diberikan.

  Pilih Pengguna super untuk memberikan izin administratif yang tidak dibatasi pada semua sumber daya di dalam katalog.

  Di bawah Izin yang dapat diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.

11. Pilih Berikutnya untuk meninjau informasi dan membuat katalog. Daftar Katalog menunjukkan katalog federasi baru.

    Daftar lokasi Data menunjukkan koneksi federasi yang baru terdaftar.



AWS CLI

Untuk membuat katalog federasi dari sumber data eksternal dan menyiapkan izin

1. Contoh berikut menunjukkan cara membuat AWS Glue koneksi.

```
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. Contoh berikut menunjukkan cara mendaftarkan AWS Glue koneksi dengan 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. Jika Anda bukan administrator data lake, berikan `DATA_LOCATION_ACCESS` pada koneksi terdaftar ke kepala sekolah yang membuat katalog federasi. Administrator data lake memiliki izin lokasi data implisit dan tidak memerlukan hibah ini. Untuk informasi selengkapnya, lihat Izin Formasi Danau Implisit.

   Contoh berikut menunjukkan cara memberikan `DATA_LOCATION_ACCESS` pada koneksi terdaftar.

```
aws lakeformation grant-permissions \
  --cli-input-json \
    '{
      "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:iam::123456789012:role/non-
admin-role"
      },
      "Resource": {
        "DataLocation": {
          "CatalogId": "123456789012",
          "ResourceArn": "arn:aws:glue:us-east-1:123456789012:connection/
connection-name"
        }
      },
      "Permissions": ["DATA_LOCATION_ACCESS"]
    }'
```

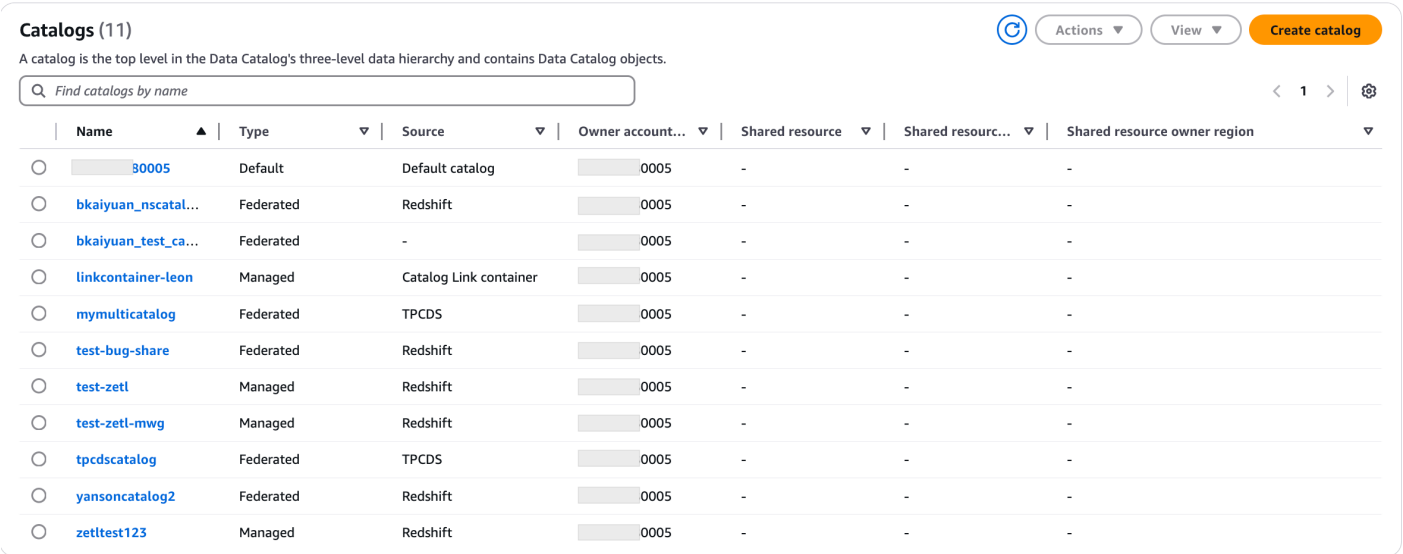
4. Contoh berikut menunjukkan cara membuat katalog federasi.

```
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"}
         }
       }'
```

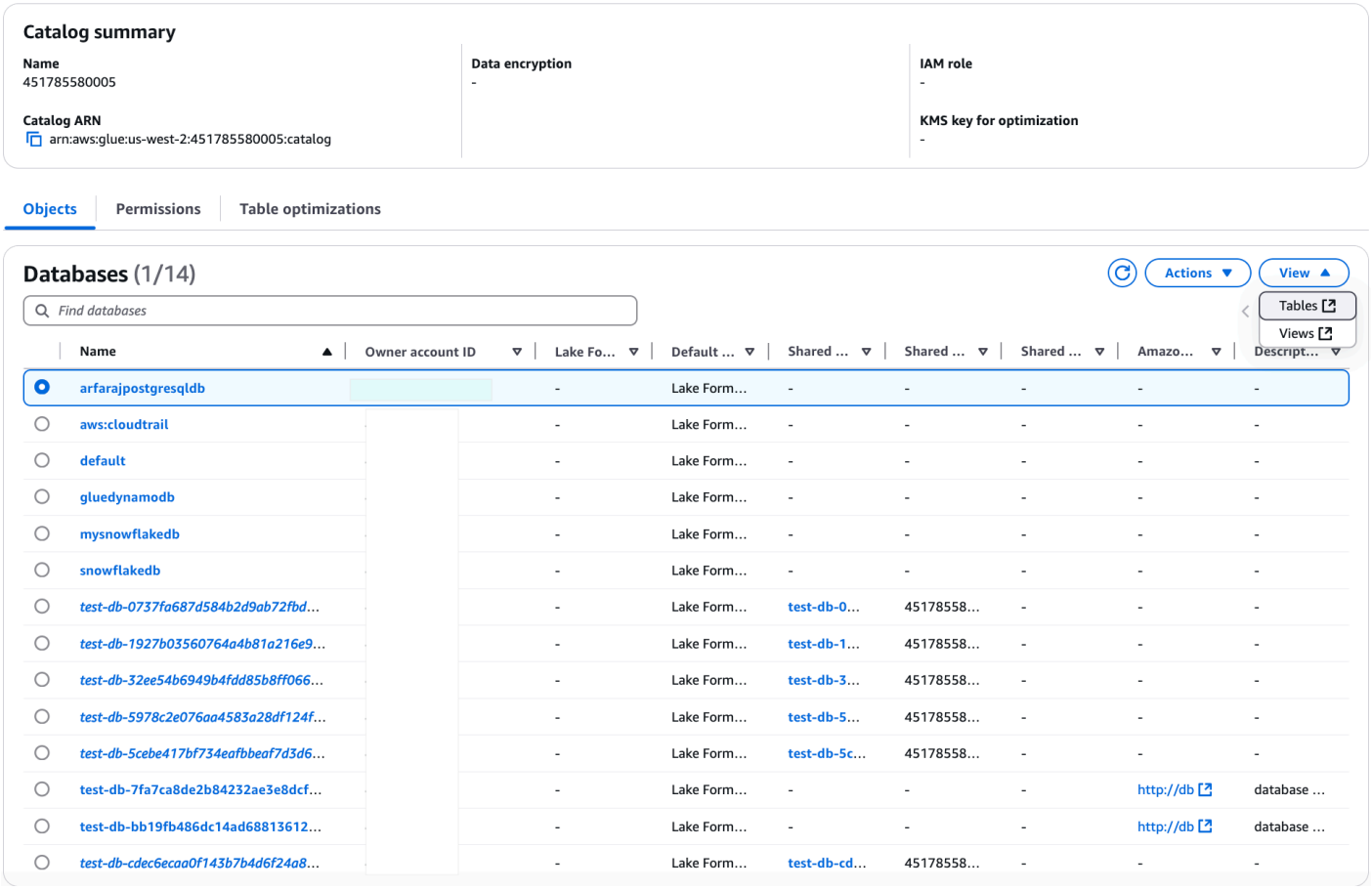
Melihat objek katalog

Untuk setiap sumber data yang tersedia, AWS Glue buat katalog yang sesuai di file AWS Glue Data Catalog. Setelah Anda membuat katalog, Anda dapat melihat database dan tabel di katalog menggunakan konsol Lake Formation atau AWS CLI. Untuk

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Pilih Katalog di bawah Katalog Data. Halaman katalog menunjukkan katalog tempat Anda memiliki izin.



3. Pilih katalog dari daftar untuk melihat database dan tabel yang terdapat dalam katalog. Daftar ini berisi database di akun dan tautan sumber daya Anda, yang merupakan tautan ke database dan tabel bersama di akun eksternal, dan digunakan untuk akses lintas akun ke data di data lake.



4. Pilih opsi Tabel di bawah Lihat untuk melihat dan mengelola tabel dalam database.

AWS CLI contoh untuk melihat katalog dan database

Contoh berikut menunjukkan cara melihat katalog menggunakan AWS CLI

```
aws glue get-catalog \
--catalog-id 123456789012:dynamodbcatalog
```

Contoh berikut menunjukkan cara meminta semua katalog di akun.

```
aws glue get-catalogs \
 --recursive
```

Contoh permintaan berikut menunjukkan bagaimana untuk mendapatkan database dalam katalog.

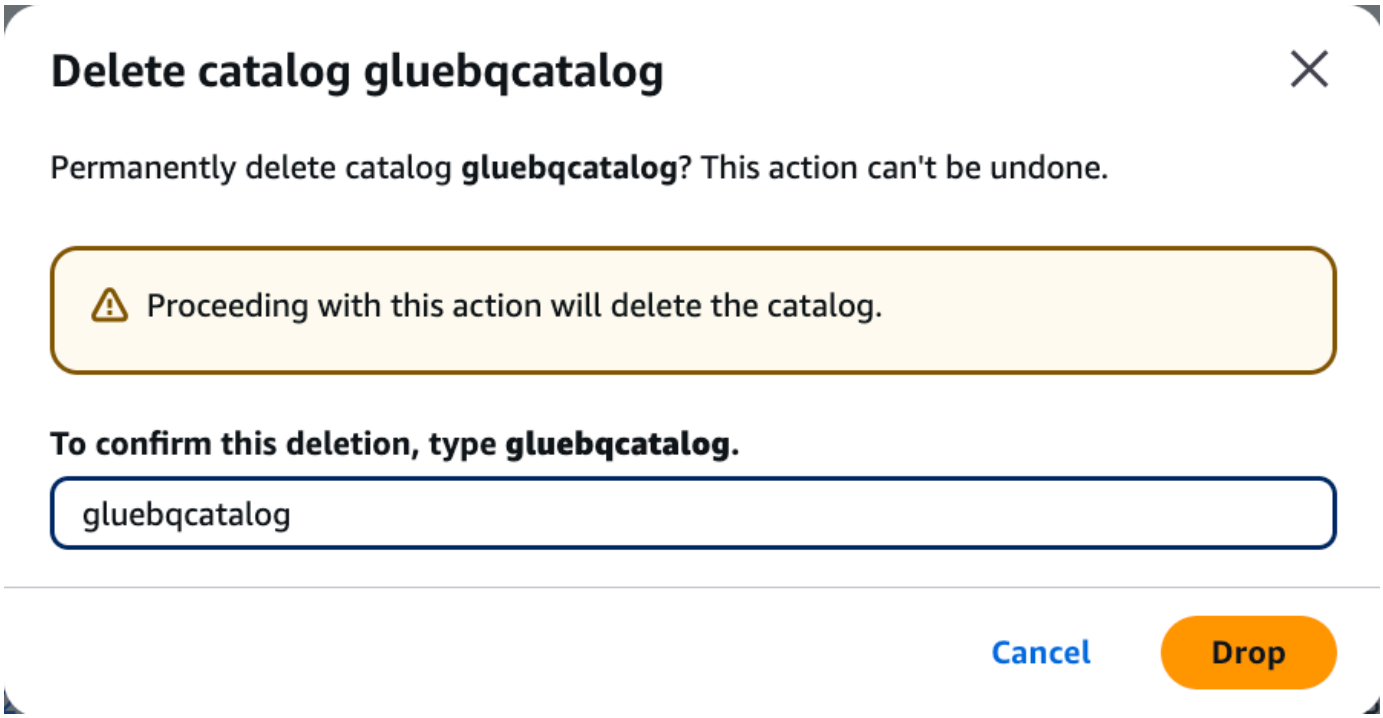
```
aws glue get-database \
--catalog-id 123456789012:dynamodbcatalog
--database-name database name
```

Menghapus katalog federasi

Anda dapat menghapus katalog federasi yang Anda buat dalam AWS Glue Data Catalog menggunakan `glue:DeleteCatalog` operasi atau konsol. AWS Lake Formation

Untuk menghapus katalog federasi (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Katalog di bawah Katalog Data.
3. Pilih katalog yang ingin Anda hapus dari daftar katalog.
4. Pilih Hapus dari Tindakan.
5. Pilih Jatuhkan untuk mengonfirmasi dan katalog federasi akan dihapus dari Katalog Data.



Untuk menghapus katalog federasi (CLI)

-
```
aws glue delete-catalog
  --catalog-id 123456789012:catalog name
```

Menanyakan katalog federasi

Setelah Anda memberikan izin kepada prinsipal lain, mereka dapat masuk dan mulai menanyakan tabel di katalog federasi menggunakan Athena.

Untuk membuat dan menghapus tabel di database federasi, kepala sekolah harus memiliki `Drop` izin Lake Formation`Create table`.

Untuk informasi selengkapnya tentang pemberian izin Katalog Data, lihat. Memberikan izin pada sumber daya Katalog Data

Untuk informasi selengkapnya tentang menanyakan Katalog Data dari Amazon Athena, lihat Menanyakan AWS Glue Data Catalog dari Amazon Athena dalam Panduan Pengguna Amazon Athena.

Sumber daya tambahan

Dalam posting blog ini, kami menunjukkan bagaimana analis data sekarang dapat dengan aman mengakses dan menanyakan data yang disimpan di luar danau data S3, termasuk gudang Amazon DynamoDB dan basis data Amazon Redshift, semuanya melalui satu pengalaman terpadu. Administrator sekarang dapat menerapkan kontrol akses pada berbagai tingkat perincian untuk memastikan data sensitif tetap terlindungi sambil memperluas akses data. Hal ini memungkinkan organisasi untuk mempercepat inisiatif data sambil menjaga keamanan dan kepatuhan, yang mengarah pada pengambilan keputusan yang lebih cepat dan berbasis data.

- Katalog dan atur kueri federasi Amazon Athena dengan Amazon Lakehouse SageMaker

Integrasi Tabel Amazon S3 dengan AWS Glue Data Catalog and AWS Lake Formation

Tabel Amazon S3 menyediakan penyimpanan S3 yang secara khusus dioptimalkan untuk beban kerja analitik, meningkatkan kinerja kueri sekaligus mengurangi biaya. Data dalam Tabel S3 disimpan dalam tipe bucket baru: bucket tabel, yang menyimpan tabel sebagai subresource. Tabel S3 memiliki dukungan bawaan untuk standar Apache Iceberg, yang memungkinkan Anda untuk dengan mudah menanyakan data tabel di bucket tabel Amazon S3 menggunakan mesin kueri populer seperti Apache Spark.

Anda dapat mengintegrasikan Tabel Amazon S3 dengan AWS Glue Data Catalog menggunakan kontrol akses IAM atau dengan hibah IAM dan Lake Formation:

- Kontrol akses IAM: Menggunakan kebijakan IAM untuk mengontrol akses ke Tabel S3 dan Katalog Data. Dalam pendekatan kontrol akses ini, Anda memerlukan izin IAM pada sumber daya Tabel S3 dan objek Katalog Data untuk mengakses sumber daya.
- Kontrol akses Lake Formation: Menggunakan AWS Lake Formation hibah selain izin AWS Glue IAM untuk mengontrol akses ke Tabel S3 melalui Katalog Data. Dalam mode ini, prinsipal memerlukan izin IAM untuk berinteraksi dengan Katalog Data, dan hibah Lake Formation menentukan sumber daya katalog (database, tabel, kolom, baris) yang dapat diakses prinsipal. Mode ini mendukung kontrol akses berbutir kasar (hibah tingkat database dan tingkat tabel) dan kontrol akses berbutir halus (keamanan tingkat kolom dan tingkat baris). Ketika peran terdaftar dikonfigurasi dan penjual kredensi diaktifkan, izin IAM Tabel S3 tidak diperlukan untuk kepala sekolah, karena Lake Formation menjual kredensil atas nama kepala sekolah menggunakan peran terdaftar. Kontrol akses Lake Formation juga mendukung penjual kredenal untuk mesin analitik pihak ketiga.

Bagian ini memberikan panduan untuk mengonfigurasi integrasi dengan AWS Lake Formation skenario berikut:

- Skenario A: Anda mengintegrasikan Tabel S3 dan Katalog Data menggunakan kontrol akses IAM dan sekarang berencana untuk menggunakannya. AWS Lake Formation Lihat Mengubah kontrol akses untuk integrasi Tabel S3 untuk mempelajari selengkapnya.
- Skenario B: Anda berencana untuk mengintegrasikan Tabel S3 dan Katalog Data menggunakan AWS Lake Formation dan tidak mengintegrasikannya di akun dan Wilayah Anda hari ini. Mulailah dengan Prasyarat untuk mengintegrasikan katalog tabel Amazon S3 dengan Katalog Data dan Lake Formation bagian dan ikutiMengaktifkan integrasi Tabel Amazon S3.
- Skenario C: Anda mengintegrasikan Tabel S3 dan Katalog Data menggunakan AWS Lake Formation dan sekarang berencana untuk menggunakan IAM. Lihat Mengubah kontrol akses untuk integrasi Tabel S3 untuk mempelajari selengkapnya.

Pastikan Anda mengikuti langkah-langkah dalam Mengintegrasikan Tabel S3 dengan layanan AWS analitik sehingga Anda memiliki izin yang sesuai untuk mengakses dan sumber daya tabel Anda, AWS Glue Data Catalog dan untuk bekerja dengan AWS layanan analitik.

Topik

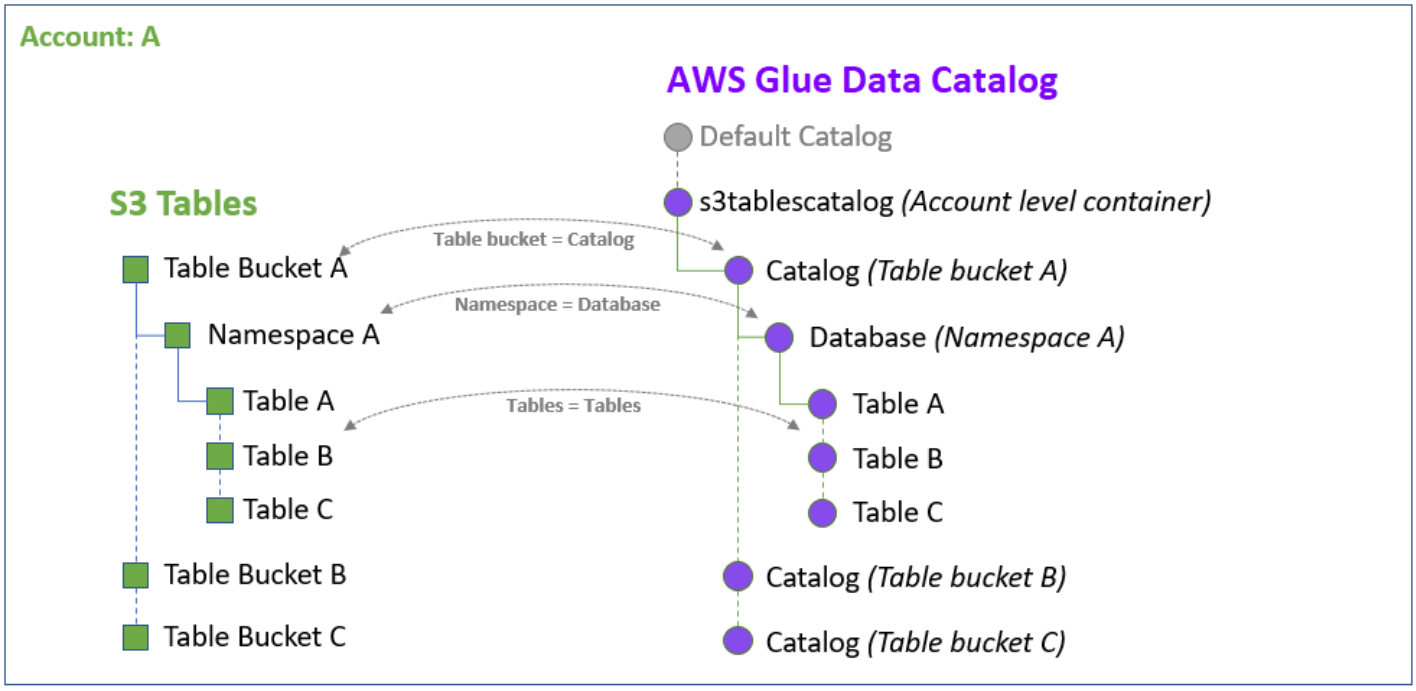
- Cara kerja integrasi Katalog Data dan Lake Formation

- Prasyarat untuk mengintegrasikan katalog tabel Amazon S3 dengan Katalog Data dan Lake Formation
- Mengaktifkan integrasi Tabel Amazon S3
- Membuat database dan tabel di katalog tabel S3
- Mendaftarkan ember meja Amazon S3 di ember lain AWS akun
- Memberikan izin

Cara kerja integrasi Katalog Data dan Lake Formation

Saat Anda mengintegrasikan katalog tabel S3 dengan Katalog Data dan Lake Formation, AWS Glue layanan akan membuat katalog gabungan tunggal yang disebut `s3tablescatalog` dalam Katalog Data default akun Anda khusus untuk Anda. Wilayah AWS Integrasi memetakan semua sumber daya bucket tabel Amazon S3 di akun Anda dan Wilayah AWS di bawah katalog federasi dengan cara berikut:

- Bucket tabel Amazon S3 menjadi katalog multi-level di Katalog Data.
- Namespace Amazon S3 terkait terdaftar sebagai database di Katalog Data.
- Tabel Amazon S3 di keranjang tabel menjadi tabel di Katalog Data.



Setelah terintegrasi dengan Lake Formation, Anda dapat membuat tabel Apache Iceberg di katalog bucket tabel, dan mengaksesnya melalui mesin AWS analitik terintegrasi seperti, Amazon Athena Amazon EMR serta mesin analitik pihak ketiga.

Saat Anda juga mengaktifkan Lake Formation dengan integrasi, ini memungkinkan kontrol akses berbutir halus. AWS Lake Formation Pendekatan keamanan ini berarti bahwa, selain izin AWS Identity and Access Management (IAM), Anda harus memberikan kepala IAM Anda dengan izin Lake Formation di tabel Anda sebelum Anda dapat bekerja dengannya.

Ada dua jenis izin utama di AWS Lake Formation:

- Izin akses metadata mengontrol kemampuan untuk membuat, membaca, memperbarui, dan menghapus database dan tabel metadata di Katalog Data.
- Izin akses data yang mendasari mengontrol kemampuan untuk membaca dan menulis data ke lokasi Amazon S3 yang mendasari yang ditunjuk oleh sumber daya Katalog Data.

Lake Formation menggunakan kombinasi model izinnya sendiri dan model izin IAM untuk mengontrol akses ke sumber daya Katalog Data dan data yang mendasarinya:

- Agar permintaan mengakses sumber daya Katalog Data atau data dasar agar berhasil, permintaan harus lulus pemeriksaan izin oleh IAM dan Lake Formation.
- Izin IAM mengontrol akses ke Lake Formation dan AWS Glue API serta sumber daya, sedangkan izin Lake Formation mengontrol akses ke sumber daya Katalog Data, lokasi Amazon S3, dan data yang mendasarinya.

Izin Lake Formation hanya berlaku di Wilayah di mana mereka diberikan, dan kepala sekolah harus diberi wewenang oleh administrator danau data atau kepala sekolah lain dengan izin yang diperlukan untuk diberikan izin Lake Formation.

Prasyarat untuk mengintegrasikan katalog tabel Amazon S3 dengan Katalog Data dan Lake Formation

Berikut ini adalah prasyarat untuk mengaktifkan integrasi tabel Amazon S3 dengan dan. AWS Glue Data Catalog AWS Lake Formation

1. Saat Anda mengaktifkan integrasi tabel Amazon S3, Lake Formation secara otomatis mendaftarkan lokasi tabel S3. Untuk mendaftarkan lokasi

bucket tabel dengan Lake Formation, Anda memerlukan IAM role/user dengan`lakeformation:RegisterResource`,`lakeformation:RegisterResourceWithPrivile` dan `lakeformation:CreateCatalog` izin. Ketika pengguna non-administrator dengan izin ini mendaftarkan lokasi katalog, Lake Formation secara otomatis memberi mereka izin untuk lokasi tersebut yang memungkinkan kepala pemanggil `DATA_LOCATION_ACCESS` izin untuk melakukan semua operasi Lake Formation yang didukung di lokasi data terdaftar.

2. Saat Anda mengaktifkan integrasi tabel S3, Anda harus memilih peran IAM untuk Lake Formation untuk menjual kredensil agar memungkinkan akses data. Buat peran IAM untuk akses data Lake Formation ke bucket tabel S3 Anda. Peran IAM yang digunakan saat mendaftarkan bucket tabel dengan Lake Formation memerlukan izin berikut:

   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/*"
            ]
        }
    ]
}
```

Untuk informasi selengkapnya, lihat Persyaratan untuk peran yang digunakan untuk mendaftarkan lokasi.

3. Tambahkan kebijakan kepercayaan berikut ke peran IAM untuk memungkinkan layanan Lake Formation mengambil peran dan menjual kredensi sementara ke mesin analitik terintegrasi.

```
{
  "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
  ]
}
```

> **Note**
>
> Proses integrasi layanan AWS analitik telah diperbarui. Jika Anda telah menyiapkan integrasi dengan rilis pratinjau, Anda dapat terus menggunakan integrasi Anda saat ini. Namun, proses integrasi yang diperbarui memberikan peningkatan kinerja. Untuk memperbarui integrasi:

1. Pertama, hapus katalog tabel S3 Anda yang ada di Lake Formation. Untuk menghapus katalog, pilih **`s3tablescatalog`** katalog dari daftar katalog, dan pilih Hapus dari Tindakan.
2. Selanjutnya, deregister lokasi data untuk file. `s3tablescatalog`
   a. Di konsol Lake Formation, di bawah bagian Administrasi, pilih Lokasi Data.
   b. Pilih lokasi, dan dari menu Tindakan, pilih Hapus.
   c. Saat diminta konfirmasi, pilih Hapus.

   Untuk petunjuk terperinci tentang membatalkan pendaftaran lokasi data, lihat bagian. Membatalkan pendaftaran lokasi Amazon S3
3. Kemudian, ikuti langkah-langkah integrasi yang diperbarui di Mengaktifkan integrasi Tabel Amazon S3 bagian ini.

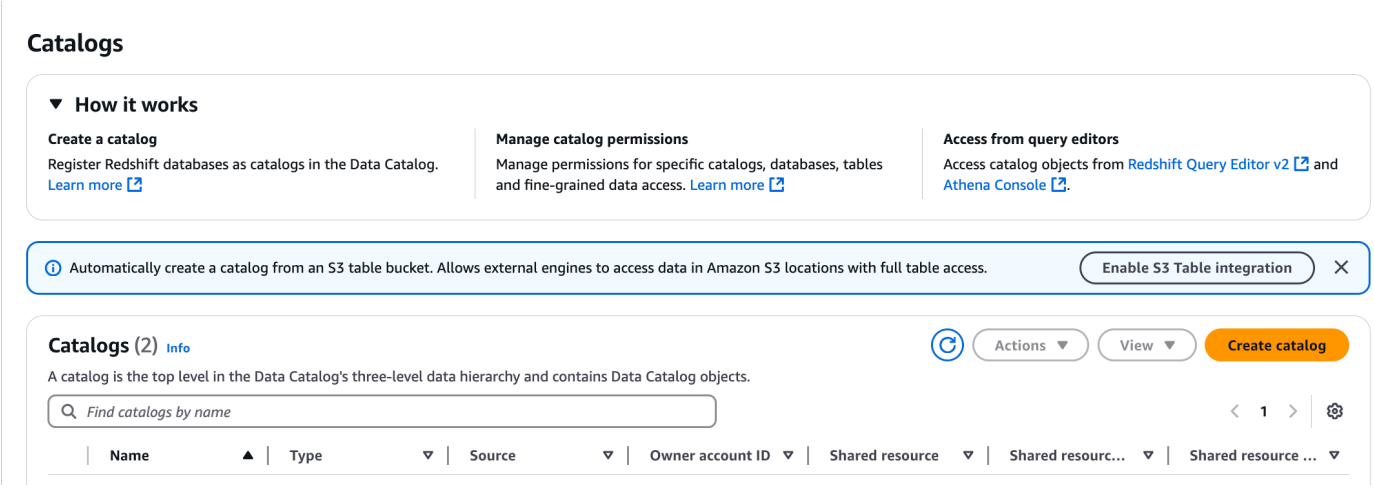
Mengaktifkan integrasi Tabel Amazon S3

Anda dapat membuat bucket tabel Amazon S3 menggunakan konsol Amazon S3, dan mengintegrasikannya dengan layanan analitik. AWS Untuk informasi selengkapnya, lihat Menggunakan Tabel Amazon S3 dengan layanan AWS analitik.

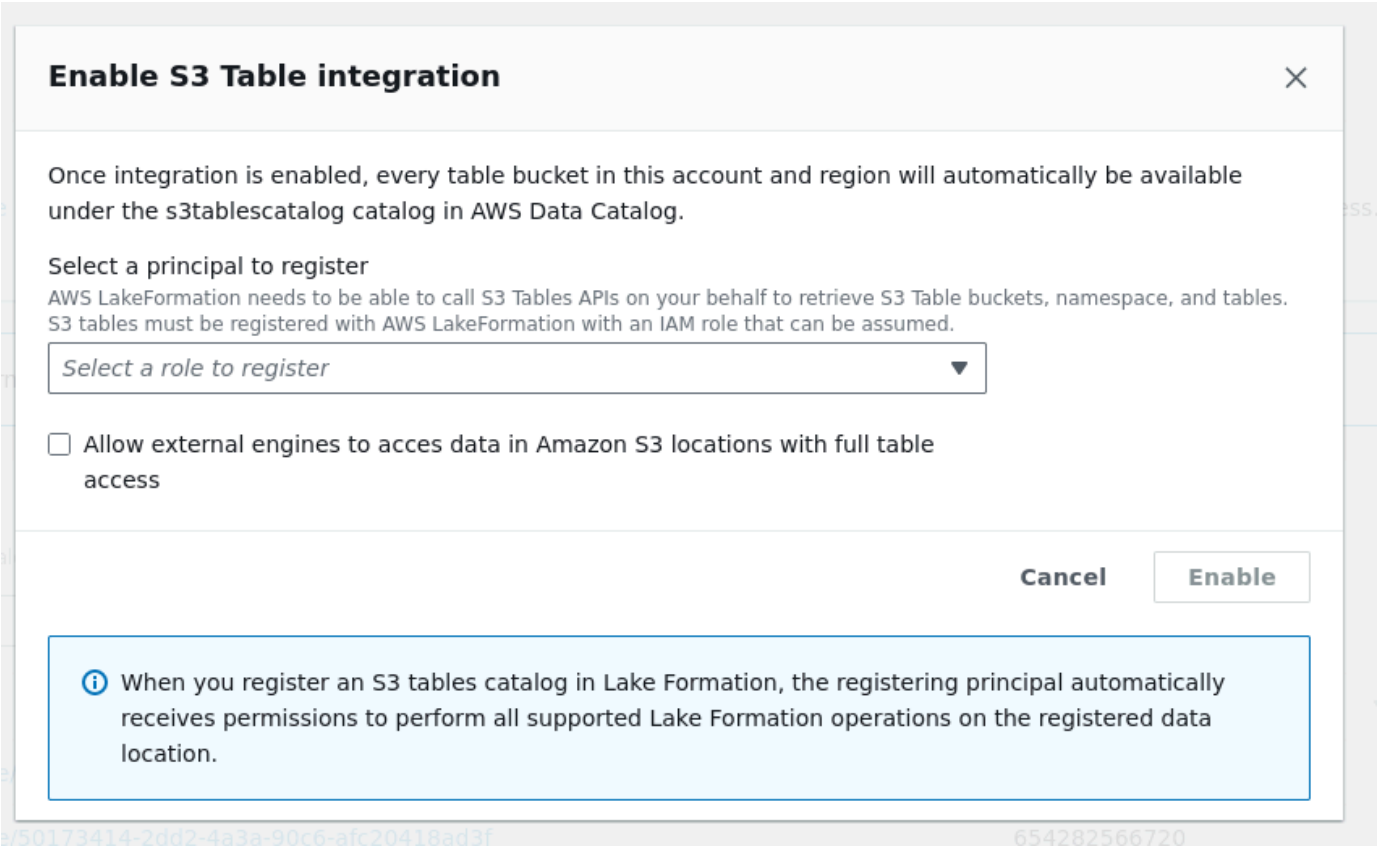
Di AWS Lake Formation, Anda dapat mengaktifkan integrasi Tabel Amazon S3 dengan AWS Glue Data Catalog dan AWS Lake Formation menggunakan konsol atau penggunaan Lake Formation. AWS CLI

Untuk mengintegrasikan Tabel Amazon S3 dengan Katalog Data dan Lake Formation (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Katalog di bawah Katalog Data.
3. Pilih Aktifkan integrasi Tabel S3 pada halaman Katalog.



4. Pilih peran IAM dengan izin yang diperlukan untuk Lake Formation untuk mengasumsikan untuk menjual kredensil ke mesin kueri analitik. Untuk izin yang diperlukan untuk peran mengakses data, lihat step3-permissions di bagian prasyarat.

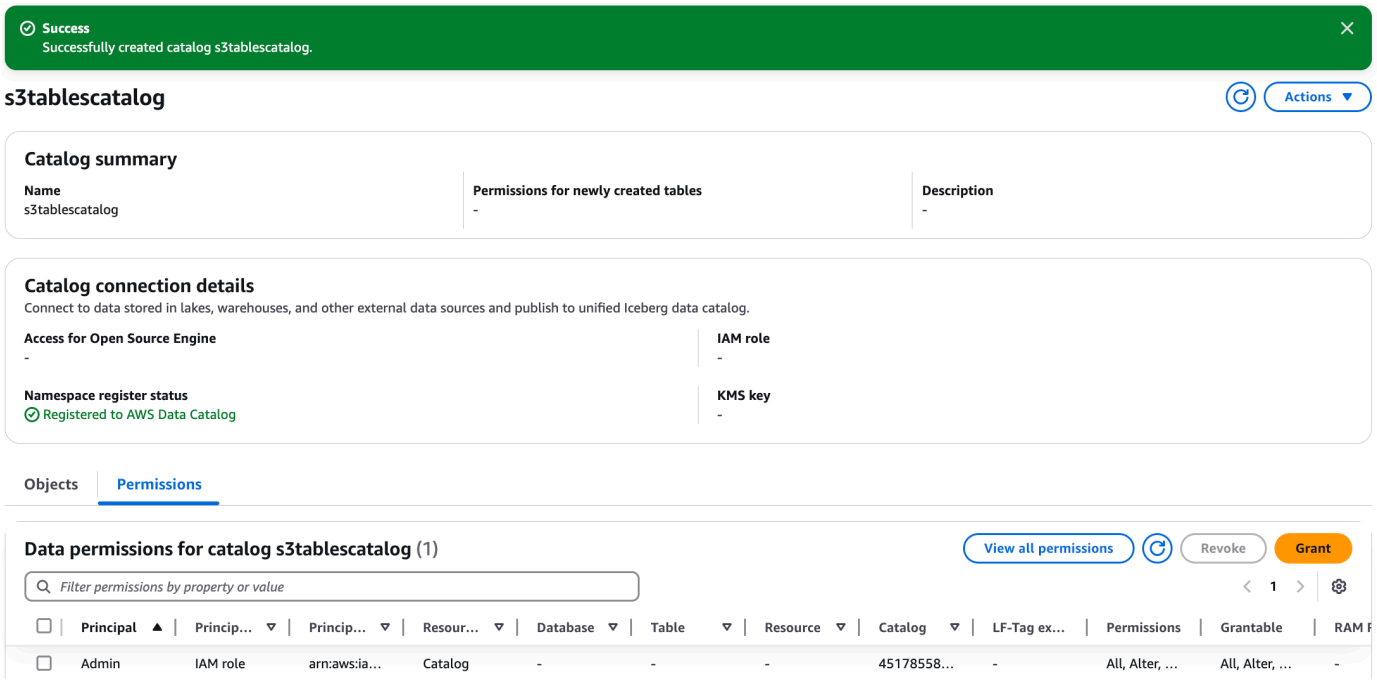


5. Pilih Izinkan mesin eksternal mengakses data di lokasi Amazon S3 dengan opsi akses tabel penuh. Saat Anda mengaktifkan akses tabel penuh untuk mesin pihak ketiga, Lake Formation mengembalikan kredensil ke mesin pihak ketiga secara langsung tanpa melakukan validasi tag

sesi IAM. Ini berarti Anda tidak dapat menerapkan kontrol akses berbutir halus Lake Formation ke tabel yang sedang diakses.

6. Pilih Aktifkan. Katalog baru untuk Tabel S3 ditambahkan ke daftar katalog. Saat Anda mengaktifkan integrasi katalog tabel S3, layanan akan mendaftarkan lokasi data bucket tabel S3 dengan Lake Formation.

7. Pilih katalog untuk melihat objek katalog dan memberikan izin kepada prinsipal lain.



Untuk membuat katalog multi-level, lihat bagian Membuat keranjang tabel di Panduan Pengguna Layanan Penyimpanan Sederhana Amazon.

Untuk mengintegrasikan tabel Amazon S3 dengan Katalog Data dan Lake Formation (CLI)

Mengikuti bagian prasyarat, buat peran layanan IAM yang memungkinkan Lake Formation mengakses sumber daya tabel Anda.

1. Buat file bernama `Role-Trust-Policy.json` yang berisi kebijakan kepercayaan berikut:

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Sid": "LakeFormationDataAccessPolicy",
        "Effect": "Allow",
```

```
        "Principal": {
          "Service": "lakeformation.amazonaws.com"
        },
        "Action": [
            "sts:AssumeRole",
            "sts:SetContext",
            "sts:SetSourceIdentity"
        ],
        "Condition": {
          "StringEquals": {
            "aws:SourceAccount": "111122223333"
          }
        }
      }
    ]
}
```

2. Buat peran layanan IAM dengan menggunakan perintah berikut:

```
aws iam create-role \
  --role-name S3TablesRoleForLakeFormation \
  --assume-role-policy-document file://Role-Trust-Policy.json
```

3. Buat file bernama `LF-GluePolicy.json` yang berisi kebijakan berikut:

```
{
    "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-1:111122223333:bucket/*"
            ]
        }
    ]
}
```

4. Lampirkan kebijakan ke peran dengan menggunakan perintah berikut:

```
aws iam put-role-policy \
  --role-name S3TablesRoleForLakeFormation \
  --policy-name LakeFormationDataAccessPermissionsForS3TableBucket \
  --policy-document file://LF-GluePolicy.json
```

5. Buat file bernama `input.json` yang berisi berikut ini:

```
{
    "ResourceArn": "arn:aws:s3tables:us-east-1:111122223333:bucket/*",
    "WithFederation": true,
    "RoleArn": "arn:aws:iam::111122223333:role/S3TablesRoleForLakeFormation"
}
```

6. Daftarkan bucket tabel dengan Lake Formation dengan menggunakan perintah berikut:

```
aws lakeformation register-resource \
  --region us-east-1 \
  --with-privileged-access \
```

```
  --cli-input-json file://input.json
```

7. Buat file bernama `catalog.json` yang berisi katalog berikut:

```
{
   "Name": "s3tablescatalog",
   "CatalogInput": {
      "FederatedCatalog": {
          "Identifier": "arn:aws:s3tables:us-east-1:111122223333:bucket/*",
          "ConnectionName": "aws:s3tables"
       },
       "CreateDatabaseDefaultPermissions": [],
       "CreateTableDefaultPermissions": [],
       "AllowFullTableExternalDataAccess": "True"
   }
}
```

8. Buat `s3tablescatalog` katalog dengan menggunakan perintah berikut. Membuat katalog ini mengisi dengan objek yang sesuai AWS Glue Data Catalog dengan bucket tabel, ruang nama, dan tabel.

```
aws glue create-catalog \
  --region us-east-1 \
  --cli-input-json file://catalog.json
```

9. Verifikasi bahwa `s3tablescatalog` katalog telah ditambahkan AWS Glue dengan menggunakan perintah berikut:

```
aws glue get-catalog --catalog-id s3tablescatalog
```

Membuat database dan tabel di katalog tabel S3

Anda dapat membuat database untuk mengatur tabel Apache Iceberg Anda, dan tabel untuk menentukan skema dan lokasi data Anda dalam katalog tabel S3.

Buat database (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator danau data atau pembuat database.
2. Di panel navigasi, pilih Database di bawah Katalog Data.

3. Pilih Buat basis data.
4. Pada halaman Buat database, pilih opsi Database, dan masukkan detail berikut:
   - Nama — Nama unik untuk database
   - Katalog data - Pilih katalog tabel S3. Database akan berada di katalog ini.
   - Deskripsi - (Opsional) Tambahkan deskripsi dan lokasi.
   - Kontrol akses IAM untuk tabel baru — Secara opsional pilih Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini. Untuk informasi tentang opsi ini, lihat bagian Mengubah pengaturan default untuk data lake Anda.
   - Pilih Buat database. Anda dapat melihat database yang dibuat di bawah katalog tabel S3.

Buat database menggunakan AWS CLI

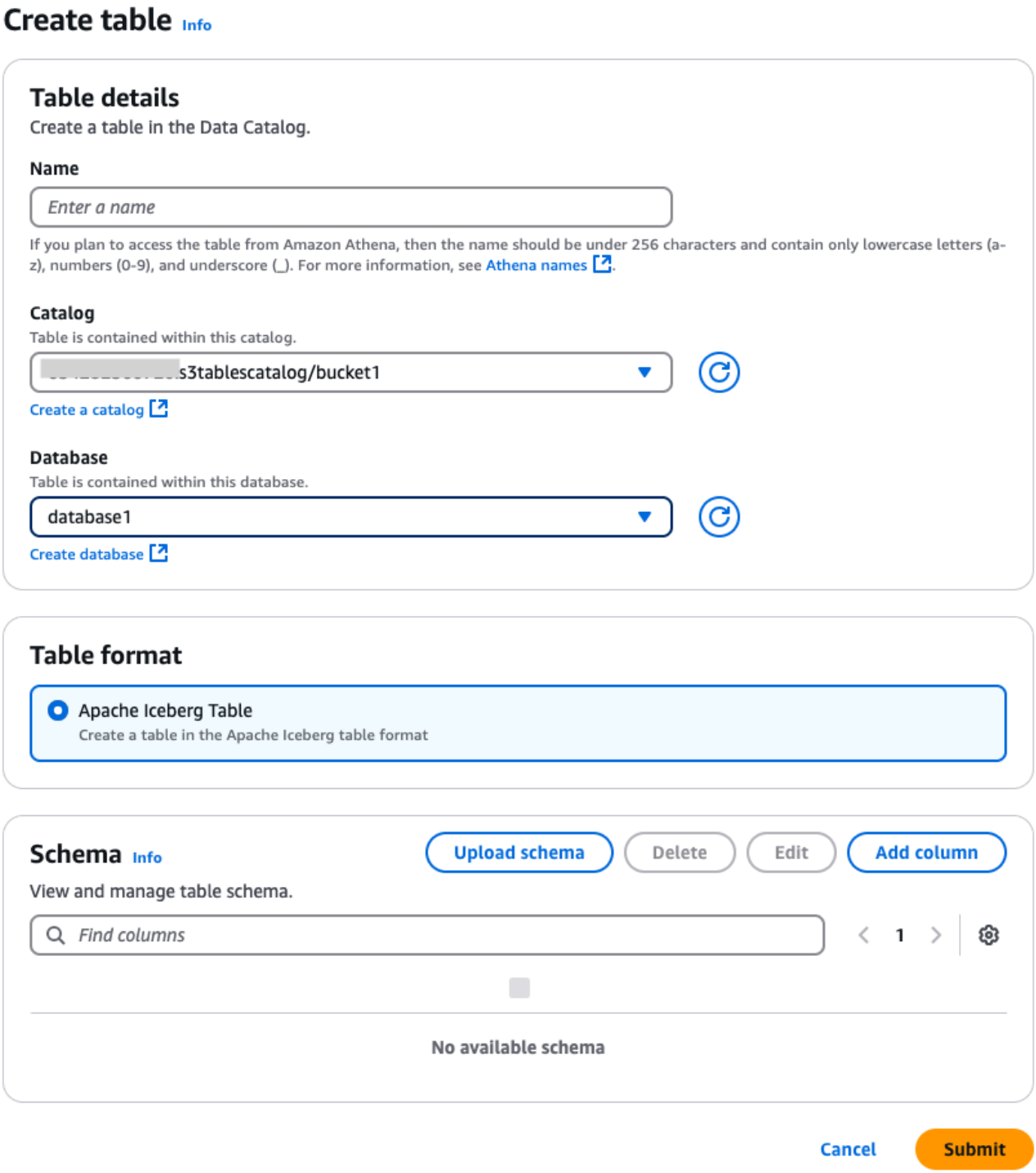
Perintah CLI berikut menunjukkan cara membuat database di katalog tabel S3.

```
aws glue create-database
--region us-east-1 \
--catalog-id "123456789012:s3tablescatalog/test" \
--database-input \
 '{ "Name": "testglueclidbcreation" }'
```

Buat tabel (Konsol Manajemen AWS)

Anda dapat membuat tabel metadata Apache Iceberg di katalog tabel S3 menggunakan konsol Lake Formation atau API. AWS Glue `CreateTable`

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator danau data atau pengguna dengan `CreateTable` izin.
2. Di panel navigasi, pilih Tabel di bawah Katalog Data.
3. Pilih Buat tabel.
4. Pada halaman Buat tabel, masukkan detail tabel:



- Nama — Masukkan nama unik untuk tabel.
- Katalog — Pilih katalog tabel S3 sebagai katalog.
- Database — Pilih database di bawah katalog tabel S3.

- Deskripsi — Masukkan deskripsi untuk tabel.
- Skema - Pilih Tambahkan kolom untuk menambahkan kolom dan tipe data kolom. Anda memiliki opsi untuk membuat tabel kosong dan memperbarui skema nanti. Iceberg memungkinkan Anda untuk mengembangkan skema dan partisi setelah Anda membuat tabel. Anda dapat menggunakan kueri Athena untuk memperbarui skema tabel dan kueri Spark untuk memperbarui partisi.

5. Pilih Kirim.

Buat tabel (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"}} ] } }'
```

Mendaftarkan ember meja Amazon S3 di ember lain AWS akun

Anda dapat mendaftarkan bucket tabel Amazon S3 individual (format ARN:`arn:aws:s3tables:us-east-1:`*`account-id`*`:bucket/`*`bucket-name`*) dari satu akun dengan AWS Lake Formation di akun lain. Misalnya, Anda dapat mendaftarkan keranjang meja dari akun A di Lake Formation akun B.

Prasyarat

Sebelum memulai registrasi bucket lintas akun:

- Buat keranjang meja di akun A.
- Buat peran IAM di akun B dengan izin yang sesuai untuk pendaftaran bucket.

  Untuk informasi selengkapnya tentang izin yang diperlukan untuk mendaftarkan keranjang meja dengan Lake Formation, lihatPrasyarat untuk mengintegrasikan katalog tabel Amazon S3 dengan Katalog Data dan Lake Formation.
- Daftarkan keranjang tabel di akun tempat Anda akan membuat katalog tabel S3.

- Untuk akses lintas akun, peran yang ditentukan saat mendaftarkan bucket tabel harus berupa peran dalam akun dengan izin yang sesuai untuk mengakses bucket lintas akun. Peran tersebut membutuhkan tindakan IAM Tabel S3 yang diperlukan untuk mengakses sumber daya bucket.

Cross-account pendaftaran ember meja

Dalam prosedur berikut, akun A adalah akun pemilik sumber daya, dan akun B adalah tempat keranjang tabel akan didaftarkan untuk mengelola izin akses.

1. Masuk ke akun A. Konsol Manajemen AWS

   Buka konsol Amazon S3 di. https://console.aws.amazon.com/s3/
2. Buat ember meja. Untuk informasi selengkapnya, lihat Membuat bucket tabel di Panduan Pengguna Amazon S3.
3. Daftarkan ember meja di akun B.

   Gunakan AWS CLI untuk mendaftarkan ember meja dari akun A dengan Lake Formation di akun 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
   ```

   Ganti akun-A-id, single-bucket-name, account-B-id, dan role-name dengan nilai spesifik Anda.
4. Selanjutnya, buat katalog untuk bucket tabel di akun B.

   Buat katalog menggunakan 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": []
   }
}'
```

Ganti catalog-name, account-A-id, dan bucket-name dengan nilai spesifik Anda.

Contoh CLI berikut menunjukkan cara melihat detail katalog.

```
aws glue get-catalog \
  --catalog-id account-id:catalog-name \
  --region us-east-1
```

5. Selanjutnya, buat database dan tabel di katalog yang baru dibuat di akun B.

   Buat database.

```
aws glue create-database \
  --region us-east-1 \
  --catalog-id "account-B-id:catalog-name" \
  --database-input \
'{
  "Name": "database-name"
}'
```

   Buat tabel.

```
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"}}
          ]
```

```
        }
}'
```

Ganti nama-database, akun-, nama-katalogB-id, dan nama tabel dengan nilai spesifik Anda.

6. Contoh berikut menunjukkan cara melihat objek dalam katalog.

   Lihat detail basis data.

```
aws glue get-database \
  --name database-name \
  --catalog-id account-B-id:catalog-name \
  --region us-east-1
```

   Lihat detail tabel.

```
aws glue get-table \
  --name table-name \
  --database-name database-name \
  --catalog-id account-B-id:catalog-name \
  --region us-east-1
```

   Ganti nama-database, akun-, nama-katalogB-id, dan nama tabel dengan nilai spesifik Anda.

Memberikan izin

Setelah mengintegrasikan tabel S3 AWS Lake Formation, Anda dapat memberikan izin pada katalog tabel S3 dan objek katalog (bucket tabel, database, tabel) ke peran dan pengguna IAM lainnya di akun Anda. Izin Lake Formation memungkinkan Anda menentukan kontrol akses pada tabel, kolom, dan granularitas tingkat baris untuk pengguna mesin analitik terintegrasi seperti Amazon Redshift Spectrum dan Athena.

Anda dapat memberikan izin dengan menggunakan metode sumber daya bernama atau metode kontrol akses (LF-TBAC) berbasis tag Lake Formation. Sebelum memberikan izin menggunakan LF-Tags dan LF-Tag ekspresi, Anda harus mendefinisikannya dan mengaitkannya ke objek Katalog Data.

Untuk informasi selengkapnya, lihat Mengelola LF-Tags kontrol akses metadata.

Anda dapat berbagi database dan tabel dengan AWS akun eksternal dengan memberikan izin Lake Formation ke akun eksternal. Pengguna kemudian dapat menjalankan kueri dan pekerjaan yang bergabung dan menanyakan tabel di beberapa akun. Saat Anda membagikan sumber daya katalog dengan akun lain, prinsipal di akun tersebut dapat beroperasi pada sumber daya tersebut seolah-olah sumber daya tersebut berada di Katalog Data mereka.

Saat Anda berbagi database dan tabel dengan akun eksternal, izin pengguna Super tidak tersedia.

Untuk petunjuk terperinci tentang pemberian izin, lihat bagian. Mengelola izin Lake Formation

AWS CLI contoh untuk memberikan izin pada Tabel 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"
    ]
}'
```

Berikut ini adalah parameter yang harus disertakan dalam perintah:

- DataLakePrincipalIdentifier — Pengguna IAM, peran, atau grup ARN untuk memberikan izin
- CatalogId — 12 digit ID AWS akun yang memiliki Katalog Data
- DatabaseName — Nama ruang nama keranjang meja Amazon S3
- Nama - Nama meja ember Amazon S3
- Izin — Izin untuk diberikan. Pilihan meliputi: SELECT, INSERT, DELETE, DESCRIPTION, ALTER, DROP, ALLL, dan SUPER

Mengakses tabel Amazon S3 bersama

Setelah Anda memberikan izin lintas akun pada database atau tabel di katalog tabel S3, untuk mengakses sumber daya, Anda perlu membuat tautan sumber daya ke database dan tabel bersama.

1. Di akun tujuan (akun yang menerima sumber daya bersama), buat tautan sumber daya basis data. Untuk petunjuk mendetail, lihat Membuat tautan sumber daya ke database Katalog Data bersama.

   Contoh CLI untuk membuat tautan sumber daya 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. Berikan izin lintas akun di atas meja.

   Contoh CLI untuk hibah izin lintas akun

   ```
   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. Berikan `DESCRIBE` izin Formasi Danau pada tautan sumber daya.

   Contoh CLI untuk memberikan izin deskripsi pada tautan sumber daya.

```
aws lakeformation grant-permissions \
    --principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/
S3TablesTestExecRole
    --resource Database='{CatalogId=111122223333;, Name=s3table_resourcelink}' \
    --permissions DESCRIBE
```

Mengubah kontrol akses untuk integrasi Tabel S3

Setelah mengintegrasikan Tabel Amazon S3 dengan AWS Glue Data Catalog, Anda dapat mengubah cara akses ke sumber daya katalog dikontrol. Bagian ini cara mengubah kontrol akses tergantung pada model kontrol akses Anda saat ini dan yang diinginkan. Mengaktifkan Lake Formation memungkinkan Anda untuk menggunakan izin berbutir halus seperti keamanan tingkat kolom dan tingkat baris melalui hibah Lake Formation, dan memungkinkan Lake Formation untuk menjual kredensi sementara atas nama kepala sekolah melalui peran terdaftar. Mengubah kontrol akses dari AWS Lake Formation IAM mengembalikan kontrol akses ke kebijakan IAM standar, yang mungkin sesuai jika beban kerja Anda tidak memerlukan akses halus dan Anda lebih memilih untuk mengelola izin sepenuhnya melalui IAM. Kedua jalur migrasi melibatkan memperbarui default Katalog Data, menyesuaikan pendaftaran sumber daya dengan Lake Formation, dan mengoordinasikan hibah izin untuk menghindari gangguan akses selama transisi.

Topik

- Aktifkan Lake Formation dengan integrasi Tabel S3 dengan Katalog Data
- Ubah kontrol akses dari AWS Lake Formation para IAM

Aktifkan Lake Formation dengan integrasi Tabel S3 dengan Katalog Data

Bagian ini menjelaskan alur kerja untuk memigrasikan kontrol akses dari hak istimewa IAM ke IAM dengan AWS Lake Formation hibah untuk Tabel Amazon S3 yang terintegrasi dengan. AWS Glue Data Catalog

> **Important**
>
> Mengaktifkan kontrol AWS Lake Formation akses akan mencabut semua akses yang ada ke sumber daya IAM-based Tabel S3 Anda. Setelah menyelesaikan Langkah 1, pengguna dan peran yang sebelumnya mengakses data melalui izin IAM akan segera kehilangan akses. Anda harus memberikan izin Lake Formation di Langkah 2 sebelum pengguna dapat melakukan kueri data lagi. Rencanakan migrasi ini selama jendela pemeliharaan dan koordinasikan dengan tim data Anda.

Prasyarat

Untuk read/write akses ke Tabel S3, selain izin Lake Formation, kepala sekolah juga memerlukan izin IAM. `lakeformation:GetDataAccess` Dengan izin ini, Lake Formation memberikan permintaan kredensil sementara untuk mengakses data.

Penggunaan AWS CLI

1. Langkah 1: Daftarkan bucket dengan Lake Formation menggunakan peran IAM

   Daftarkan sumber daya Tabel S3 dengan Lake Formation.

   > **Note**
   >
   > Jika Anda memiliki peran yang ada, pastikan akses hybrid salah.

   ```
   aws lakeformation register-resource \
     --resource-arn "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/*" \
     --role-arn "arn:aws:iam::AWSAccountID:role/service-role/LFAccessRole" \
     --with-federation
   ```

2. Langkah 2: Perbarui AWS Glue katalog untuk mengaktifkan kontrol akses Lake Formation

Perbarui katalog dengan kosong `CreateDatabaseDefaultPermissions` dan `CreateTableDefaultPermissions` (atur ke`[]`). Ini memastikan bahwa database dan tabel dalam katalog dikelola menggunakan hibah Lake Formation, bukan izin IAM-based default.

```
aws glue update-catalog \
  --catalog-id "s3tablescatalog" \
  --catalog-input '{
    "FederatedCatalog": {
        "Identifier": "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/*",
        "ConnectionName": "aws:s3tables"
    },
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": [],
    "AllowFullTableExternalDataAccess": "True"
  }'
```

3. Langkah 3: Berikan izin Lake Formation ke tim data Anda

   Berikan izin Lake Formation ke kepala sekolah (peran, pengguna, atau grup) yang memerlukan akses. Misalnya, untuk memberikan akses baca tabel lengkap ke peran:

```
aws lakeformation grant-permissions \
  --principal
 DataLakePrincipalIdentifier=arn:aws:iam::AWSAccountID:role/DataTeamRole \
  --resource '{
    "Table": {
        "CatalogId": "AWSAccountID",
        "DatabaseName": "s3tablescatalog/table-bucket-name/namespace",
        "TableWildcard": {}
    }
  }' \
  --permissions "SELECT" "DESCRIBE"
```

   Ulangi untuk setiap kombinasi prinsip dan sumber daya sesuai kebutuhan.

Ubah kontrol akses dari AWS Lake Formation para IAM

Bagian ini menjelaskan alur kerja untuk mengubah kontrol akses dari AWS Lake Formation hibah ke hak istimewa IAM untuk Tabel Amazon S3 yang terintegrasi dengan. AWS Glue Data Catalog

Important

Mengubah kontrol akses dari AWS Lake Formation hibah ke IAM akan mencabut semua Formation-based akses Danau yang ada ke sumber daya Tabel S3 Anda. Setelah menyelesaikan Langkah 2, pengguna dan peran yang sebelumnya mengakses data melalui hibah Lake Formation akan segera kehilangan akses. Anda harus memberikan akses IAM pada Langkah 1 sebelum memperbarui katalog. Rencanakan migrasi ini selama jendela pemeliharaan dan koordinasikan dengan tim data Anda.

Important

Fine-grained kontrol akses, seperti akses tingkat kolom dan filter sel data, dengan objek Katalog Data tersedia saat hanya menggunakan AWS Lake Formation . Sebelum melanjutkan untuk memigrasikan kontrol akses dari AWS Lake Formation IAM, audit hibah Lake Formation Anda yang ada menggunakan `aws lakeformation list-permissions` dan tentukan apakah kebijakan IAM yang setara dapat menyediakan akses yang dibutuhkan pengguna Anda. Setiap kepala sekolah yang mengandalkan hibah Lake Formation berbutir halus akan memerlukan akses IAM tingkat tabel penuh setelah memigrasikan kontrol akses.

Prasyarat

Sebelum Anda mulai, pastikan hal-hal berikut:

- Anda telah mengidentifikasi semua hibah Lake Formation yang saat ini berlaku untuk sumber daya yang dimigrasikan. Jalankan `aws lakeformation list-permissions --resource-type TABLE` untuk memeriksanya.
- Anda telah menyiapkan kebijakan IAM yang menyediakan akses setara untuk semua prinsipal yang terpengaruh.
- Peran IAM yang terdaftar di Lake Formation masih memiliki `lakeformation:GetDataAccess` (diperlukan selama periode transisi hibrida).

Penggunaan AWS CLI

1. Langkah 1: Berikan izin IAM kepada kepala sekolah

Lampirkan kebijakan IAM ke pengguna atau peran yang memerlukan akses. Kebijakan harus menyertakan izin AWS Glue metadata dan izin data Tabel S3.

> **Note**
>
> Kebijakan contoh berikut hanya menyediakan akses baca.

```
aws iam put-user-policy \
  --user-name GlueIAMAccessUser \
  --policy-name S3TablesIAMAccessPolicy \
  --policy-document '{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "GlueMetadataAccess",
            "Effect": "Allow",
            "Action": [
                "glue:GetCatalog",
                "glue:GetDatabase",
                "glue:GetTable"
            ],
            "Resource": [
                "arn:aws:glue:us-east-1:AWSAccountID:catalog/s3tablescatalog",
                "arn:aws:glue:us-east-1:AWSAccountID:database/
s3tablescatalog/table-bucket-name/namespace",
                "arn:aws:glue:us-east-1:AWSAccountID:table/s3tablescatalog/table-
bucket-name/namespace/*"
            ]
        },
        {
            "Sid": "S3TablesDataAccess",
            "Effect": "Allow",
            "Action": [
                "s3tables:GetTableBucket",
                "s3tables:GetTable",
                "s3tables:GetTableMetadataLocation",
                "s3tables:GetTableData"
            ],
            "Resource": [
                "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/table-bucket-name",
```

```
                "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/table-bucket-name/
table/*"
            ]
        }
    ]
  }'
```

Verifikasi bahwa semua pengguna dan peran yang terpengaruh dapat mengakses tabel yang diharapkan menggunakan kredensil IAM mereka sebelum melanjutkan.

2. Langkah 2: Perbarui katalog untuk mengembalikan izin default IAM

   Perbarui katalog sehingga `CreateDatabaseDefaultPermissions` dan `CreateTableDefaultPermissions` berikan `ALL` ke`IAM_ALLOWED_PRINCIPALS`. Setel `OverwriteChildResourcePermissionsWithDefault` `Accept` agar perubahan menyebar ke semua sumber daya anak yang ada, bukan hanya yang baru dibuat.

   ```
   aws glue update-catalog \
     --catalog-id "s3tablescatalog" \
     --catalog-input '{
       "FederatedCatalog": {
           "Identifier": "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/*",
           "ConnectionName": "aws:s3tables"
       },
       "CreateDatabaseDefaultPermissions": [{
           "Principal": {"DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"},
           "Permissions": ["ALL"]
       }],
       "CreateTableDefaultPermissions": [{
           "Principal": {"DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"},
           "Permissions": ["ALL"]
       }],
       "OverwriteChildResourcePermissionsWithDefault": "Accept"
     }'
   ```

3. Langkah 3: Membatalkan pendaftaran sumber daya dari Lake Formation

   Setelah Anda mengonfirmasi bahwa semua akses bekerja melalui kebijakan IAM dan tidak ada kepala sekolah yang bergantung pada hibah Lake Formation, Anda dapat membatalkan pendaftaran sumber daya dari Lake Formation untuk menyelesaikan migrasi.

   ```
   aws lakeformation deregister-resource \
   ```

```
--resource-arn "arn:aws:s3tables:us-east-1:AWSAccountID:bucket/*"
```

> **Note**
>
> Setelah membatalkan pendaftaran sumber daya, hapus `lakeformation:GetDataAccess` dari kepala sekolah IAM yang tidak lagi membutuhkannya.

Tidak ada `revoke-permissions` langkah yang diperlukan.

Membuat katalog terkelola Amazon Redshift di AWS Glue Data Catalog

Anda mungkin tidak memiliki kluster produsen Amazon Redshift atau datashare Amazon Redshift yang tersedia saat ini, tetapi ingin membuat dan mengelola tabel Amazon Redshift menggunakan. AWS Glue Data Catalog Anda dapat memulai dengan membuat katalog AWS Glue terkelola menggunakan `glue:CreateCatalog` API atau AWS Lake Formation konsol dengan menyetel jenis katalog sebagai `Managed` dan `Catalog source` sebagai Redshift. Langkah ini melakukan hal berikut:

- Membuat katalog di Katalog Data
- Mendaftarkan katalog sebagai lokasi data Lake Formation
- membuat grup kerja tanpa server terkelola Amazon Redshift
- Menautkan grup kerja tanpa server Amazon Redshift dan Katalog Data menggunakan objek datashare

Untuk membuat katalog terkelola dan menyiapkan izin (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, pilih Katalog di bawah Katalog Data.
3. Pilih opsi Buat katalog.
4. Pada halaman Setel detail Katalog, masukkan informasi berikut:
   - Nama — Nama unik untuk katalog terkelola Anda. Nama tidak dapat diubah, dan harus dalam huruf kecil. Nama dapat terdiri dari maksimum 255 karakter. akun.

- Jenis - Pilih `Managed catalog` sebagai jenis katalog.
- Penyimpanan — Pilih `Redshift` untuk penyimpanan.
- Deskripsi — Masukkan deskripsi untuk katalog yang dibuat dari sumber data.

5. Anda dapat menggunakan aplikasi Apache Spark yang berjalan di Amazon EMR di Amazon EC2 untuk mengakses database Amazon Redshift di. AWS Glue Data Catalog

   Untuk mengaktifkan Apache Spark membaca dan menulis ke penyimpanan terkelola Amazon Redshift AWS Glue , buat klaster Amazon Redshift terkelola dengan sumber daya komputasi dan penyimpanan yang diperlukan untuk melakukan operasi baca dan tulis tanpa memengaruhi beban kerja gudang data Amazon Redshift. Anda juga perlu memberikan peran IAM dengan izin yang diperlukan untuk mentransfer data ke dan dari bucket Amazon S3. Untuk izin yang diperlukan untuk peran transfer data, lihat langkah 5 di Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog bagian.

6. Secara default, data di klaster Amazon Redshift dienkripsi menggunakan kunci terkelola. AWS Lake Formation menyediakan opsi untuk membuat kunci KMS kustom Anda untuk enkripsi. Jika Anda menggunakan kunci yang dikelola pelanggan, Anda harus menambahkan kebijakan kunci tertentu ke kunci tersebut.

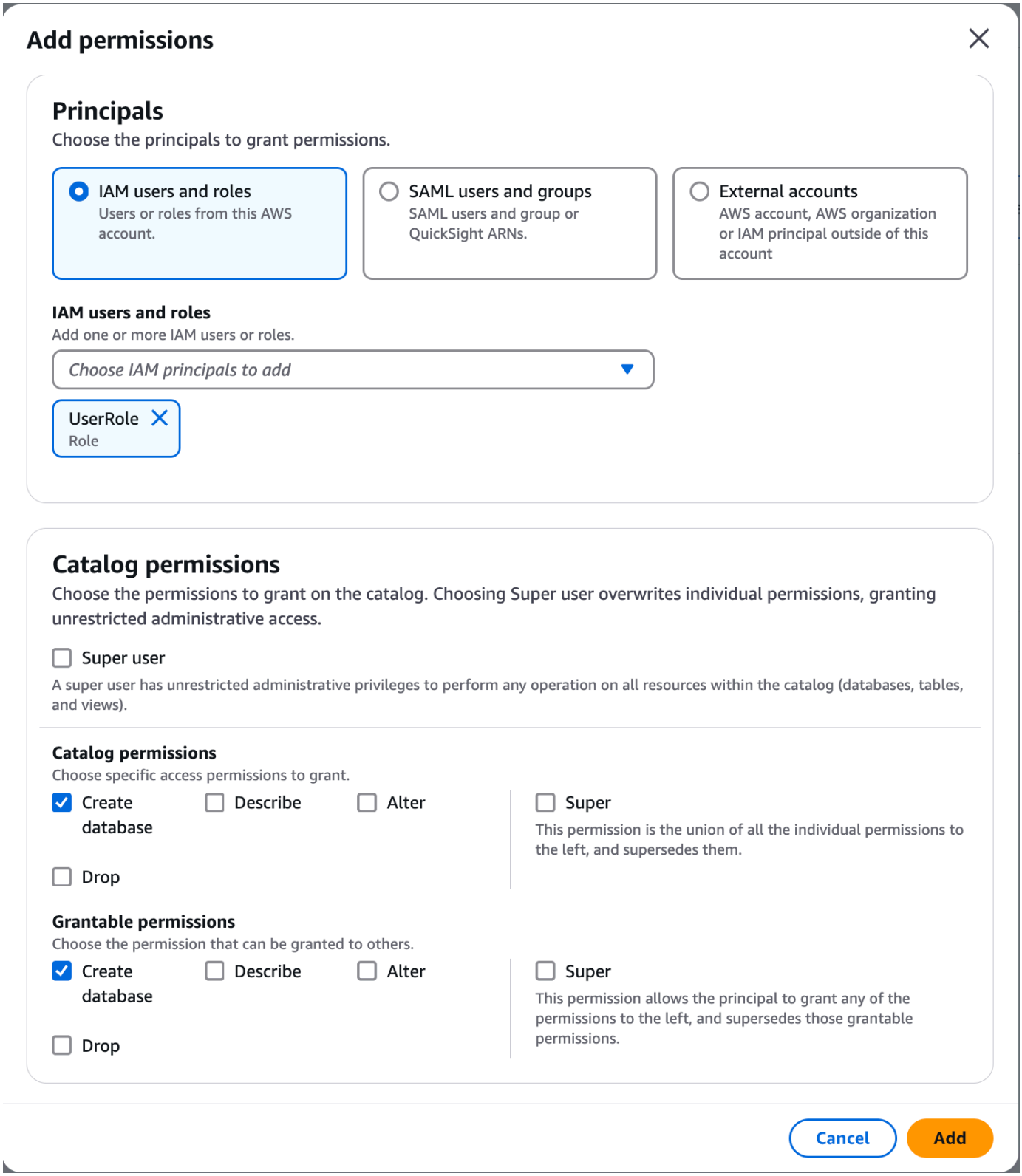
7. Pilih pengaturan Sesuaikan enkripsi jika Anda menggunakan kunci yang dikelola pelanggan untuk mengenkripsi data di penyimpanan terkelola Amazon Redshift. Untuk menggunakan kunci kustom, Anda harus menambahkan kebijakan kunci kustom terkelola tambahan ke kunci KMS Anda. Untuk informasi selengkapnya, lihat Prasyarat untuk mengelola ruang nama Amazon Redshift di AWS Glue Data Catalog.

8. Opsi enkripsi - Pilih opsi Sesuaikan pengaturan enkripsi jika Anda ingin menggunakan kunci khusus untuk mengenkripsi katalog. Untuk menggunakan kunci kustom, Anda harus menambahkan kebijakan kunci kustom terkelola tambahan ke kunci KMS Anda.

9. Pilih Berikutnya untuk memberikan izin kepada prinsipal lain.

10. Pada halaman Hibah izin, pilih Tambahkan izin.

11. Pada layar Tambahkan izin, pilih prinsip dan jenis izin yang akan diberikan.



- Di bagian Principal, pilih jenis principal lalu tentukan principal yang akan diberi izin.

- Peran dan pengguna IAM – Pilih satu atau beberapa pengguna atau peran dari daftar peran dan pengguna IAM.
- Pengguna dan grup SAMB — Untuk SALL dan Amazon Quick pengguna dan grup, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Amazon Cepat. Tekan Enter setelah setiap ARN.

  Untuk informasi tentang cara membuat ARN, lihat AWS CLI memberikan dan mencabut AWS CLI perintah.

- Di bagian Izin, pilih izin dan izin yang dapat diberikan.

  Di Izin katalog, pilih satu atau beberapa izin untuk diberikan.

  Pilih Pengguna super untuk memberikan izin administratif yang tidak dibatasi pada semua sumber daya di dalam katalog.

  Di bawah Izin yang dapat diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.

12. Pilih Berikutnya untuk meninjau informasi dan membuat katalog. Daftar Katalog menunjukkan katalog terkelola baru.

Untuk membuat katalog federasi (CLI)

- Contoh berikut menunjukkan cara membuat katalog federasi.

```
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 mendapatkan respons katalog

```
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"
         }
       }
    }
```

Mengelola izin untuk data dalam data Amazon Redshift

Dengan AWS Lake Formation, Anda dapat mengelola data dengan aman di datashare dari Amazon Redshift. Amazon Redshift adalah layanan gudang data skala petabyte yang dikelola sepenuhnya di Cloud. AWS Menggunakan kemampuan berbagi data, Amazon Redshift membantu Anda berbagi data. Akun AWS Untuk informasi selengkapnya tentang berbagi data Amazon Redshift, lihat Ikhtisar berbagi data di Amazon Redshift.

Di Amazon Redshift, administrator cluster produser membuat datashare, dan membagikannya dengan administrator data lake. Untuk step-by-step petunjuk tentang cara membuat administrator danau data, lihatBuat administrator danau data.

Setelah Anda (administrator data lake) menerima datashare, Anda harus membuat AWS Glue Data Catalog database untuk datashare tertentu. Ini agar Anda dapat mengontrol akses ke sana menggunakan izin Lake Formation. Lake Formation memetakan setiap datashare ke database Katalog Data yang sesuai. Ini muncul sebagai database federasi dalam Katalog Data.

Database disebut sebagai database federasi ketika menunjuk ke entitas di luar Katalog Data. Tabel dan tampilan dalam data Amazon Redshift dicantumkan sebagai tabel individual dalam Katalog Data. Anda dapat berbagi database federasi dengan prinsipal IAM terpilih dan pengguna SAFL dalam akun yang sama atau di akun lain dengan Lake Formation. Anda juga dapat menyertakan ekspresi filter baris dan kolom untuk membatasi akses ke data tertentu. Untuk informasi selengkapnya, lihat Pemfilteran data dan keamanan tingkat sel di Lake Formation.

Untuk memberi pengguna akses ke datashare Amazon Redshift, Anda harus melakukan hal berikut:

1. Perbarui pengaturan Katalog Data untuk mengaktifkan izin Lake Formation.
2. Terima undangan datashare dari administrator kluster produsen Amazon Redshift dan daftarkan datashare di Lake Formation.

   Setelah menyelesaikan langkah ini, Anda dapat mengelola datashare dalam Katalog Data Lake Formation.
3. Buat database federasi dan tentukan izin pada database itu.
4. Berikan izin kepada pengguna pada database dan tabel. Anda dapat berbagi seluruh database atau subset tabel dengan pengguna di akun yang sama atau akun lain.

Untuk batasan, lihat Batasan berbagi data Amazon Redshift.

Topik

- Prasyarat untuk menyiapkan izin di datashares Amazon Redshift
- Menyiapkan izin untuk datashares Amazon Redshift
- Menanyakan database federasi

Prasyarat untuk menyiapkan izin di datashares Amazon Redshift

Perbarui pengaturan Katalog Data default

Untuk mengaktifkan izin Lake Formation untuk sumber daya Katalog Data, sebaiknya Anda menonaktifkan pengaturan Katalog Data default di Lake Formation. Untuk informasi selengkapnya, lihat Ubah model izin default atau gunakan mode akses hybrid.

Perbarui izin

Selain izin administrator data lake (`AWSLakeFormationDataAdmin`), izin berikut juga diperlukan untuk menerima datashare Amazon Redshift di Lake Formation:

- `glue:PassConnection on aws:redshift`
- `redshift:AssociateDataShareConsumer`
- `redshift:DescribeDataSharesForConsumer`
- `redshift:DescribeDataShares`

Pengguna IAM administrator data lake memiliki izin berikut secara implisit.

- data_location_access
- create_database
- LakeFomation:RegisterResource

Menyiapkan izin untuk datashares Amazon Redshift

Topik ini menjelaskan langkah-langkah yang perlu Anda ikuti untuk menerima undangan datashare, membuat database federasi, dan memberikan izin. Anda dapat menggunakan konsol Lake Formation atau AWS Command Line Interface (AWS CLI). Contoh dalam topik ini menunjukkan cluster produsen, Katalog Data, dan konsumen data di akun yang sama.

Untuk mempelajari lebih lanjut tentang kemampuan lintas akun Lake Formation, lihatCross-account berbagi data di Lake Formation.

Untuk mengatur izin untuk datashare

1. Tinjau undangan datashare dan terima.

Console

1. Masuk ke konsol Lake Formation sebagai administrator danau data di https://console.aws.amazon.com/lakeformation/. Arahkan ke halaman berbagi data.
2. Tinjau datashares yang diizinkan untuk Anda akses. Kolom Status menunjukkan status partisipasi Anda saat ini untuk datashare. Status Tertunda menunjukkan bahwa Anda telah ditambahkan ke datashare, tetapi Anda belum menerimanya atau menolak undangan tersebut.
3. Untuk menanggapi undangan datashare, pilih nama datashare dan pilih Tinjau undangan. Di Menerima atau menolak datashare, tinjau detail undangan. Pilih Terima untuk menerima undangan atau Tolak untuk menolak undangan. Anda tidak mendapatkan akses ke datashare jika Anda menolak undangan.

AWS CLI

Contoh berikut menunjukkan cara melihat, menerima, dan mendaftarkan undangan. Ganti Akun AWS ID dengan ID yang valid Akun AWS . Ganti `data-share-arn` dengan Amazon Resource Name (ARN) aktual yang mereferensikan datashare.

1. Lihat undangan yang tertunda.

```
aws redshift describe-data-shares \
 --data-share-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds' \
```

2. Terima 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. Daftarkan datashare di akun Lake Formation. Gunakan operasi RegisterResourceAPI untuk mendaftarkan datashare di Lake Formation. `DataShareArn`adalah parameter input untuk`ResourceArn`.

> **Note**
>
> Ini adalah langkah wajib.

```
aws lakeformation register-resource \
 --resource-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds'
```

2. Buat database.

   Setelah menerima undangan datashare, Anda perlu membuat database yang mengarah ke database Amazon Redshift yang terkait dengan datashare. Anda harus menjadi administrator data lake untuk membuat database.

   Console

   1. Pilih datashare dari panel Undangan dan pilih Set rincian database.
   2. Di Set rincian database, masukkan nama unik dan identifier untuk datashare. Anda menggunakan pengenal ini untuk memetakan datashare secara internal dalam hierarki metadata (DbName.schema.table).
   3. Pilih Berikutnya untuk memberikan izin kepada pengguna lain pada database dan tabel bersama.

   AWS CLI

   Gunakan kode contoh berikut untuk membuat database yang menunjuk ke database Amazon Redshift yang dibagikan dengan Lake Formation menggunakan file. 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. Berikan izin.

Setelah membuat database, Anda dapat memberikan izin kepada pengguna di akun Anda atau eksternal Akun AWS dan organisasi. Anda tidak akan dapat memberikan izin menulis data (menyisipkan, menghapus) dan izin metadata (ubah, jatuhkan, buat) pada database federasi yang dipetakan ke datashare Amazon Redshift. Untuk informasi selengkapnya tentang pemberian izin, lihat. Mengelola izin Lake Formation

> **Note**
>
> Sebagai administrator data lake, Anda hanya dapat melihat tabel di database federasi. Untuk melakukan tindakan lain, Anda perlu memberi diri Anda lebih banyak izin pada tabel tersebut.

Console

1. Pada layar Hibah izin, pilih pengguna untuk memberikan izin.
2. PilihIzin.

AWS CLI

Gunakan contoh berikut untuk memberikan izin database dan tabel menggunakan: 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"
                   ]
 }
```

Menanyakan database federasi

Setelah Anda memberikan izin, pengguna dapat masuk dan mulai menanyakan database federasi menggunakan Amazon Redshift. Pengguna sekarang dapat menggunakan nama database lokal untuk mereferensikan datashare Amazon Redshift dalam kueri SQL. Di Amazon Redshift, tabel

pelanggan dalam skema publik yang dibagikan melalui datashare akan memiliki tabel yang sesuai yang dibuat seperti `public.customer` dalam Katalog Data.

1. Sebelum melakukan kueri pada database federasi menggunakan Amazon Redshift, administrator klaster membuat database dari database Data Catalog menggunakan perintah berikut:

```
CREATE DATABASE sharedcustomerdb FROM ARN
 'arn:aws:glue:<region>:111122223333:database/tahoedb' WITH DATA CATALOG SCHEMA
 tahoedb
```

2. Admin klaster memberikan izin penggunaan pada database.

```
GRANT USAGE ON DATABASE sharedcustomerdb TO IAM:user;
```

3. Anda (pengguna federasi) sekarang dapat masuk ke alat SQL untuk menanyakan tabel.

```
Select * from sharedcustomerdb.public.customer limit 10;
```

Untuk informasi selengkapnya, lihat Menanyakan AWS Glue Data Catalog di Panduan Manajemen Pergeseran Merah Amazon.

Mengelola izin pada kumpulan data yang menggunakan metastor eksternal

Dengan federasi AWS Glue Data Catalog metadata (federasi Katalog Data), Anda dapat menghubungkan Katalog Data ke metastor eksternal yang menyimpan metadata untuk data Amazon S3 Anda, dan mengelola izin akses data dengan aman menggunakan. AWS Lake Formation Anda tidak perlu memigrasikan metadata dari metastore eksternal ke Katalog Data.

Katalog Data menyediakan repositori metadata terpusat yang membuat pengelolaan dan penemuan data di seluruh sistem yang berbeda menjadi lebih mudah. Saat organisasi mengelola data di Katalog Data, Anda dapat menggunakannya AWS Lake Formation untuk mengontrol akses ke kumpulan data di Amazon S3.

> **Note**
>
> Saat ini, kami hanya mendukung federasi metastore Apache Hive (versi 3 ke atas).

Untuk mengatur federasi Katalog Data, kami menyediakan aplikasi AWS Serverless Application Model (AWS SAM) yang disebut GlueDataCatalogFederation-HiveMetastoredi AWS Serverless Application Repository.

Implementasi referensi disediakan GitHub sebagai proyek open source di AWS Glue Data Catalog Federation - Hive Metastore.

AWS SAM Aplikasi membuat dan menyebarkan sumber daya berikut yang diperlukan untuk menghubungkan Katalog Data ke metastore Hive:

- AWS Lambda Fungsi — Menyelenggarakan implementasi layanan federasi yang berkomunikasi antara Katalog Data dan metastore Hive. AWS Glue memanggil fungsi Lambda ini untuk mengambil objek metadata dari metastore Hive.
- Amazon API Gateway— Titik akhir koneksi untuk metastore Hive Anda yang bertindak sebagai proxy untuk merutekan semua pemanggilan ke fungsi Lambda.
- Peran IAM — Peran dengan izin yang diperlukan untuk membuat koneksi antara Katalog Data dan metastore Hive.
- AWS Glue koneksi — Amazon API Gateway Jenis AWS Glue koneksi yang menyimpan Amazon API Gateway titik akhir dan peran IAM untuk memanggilnya.

Saat Anda melakukan kueri tabel, AWS Glue layanan membuat panggilan runtime ke metastore Hive dan mengambil metadata. Fungsi Lambda bertindak sebagai penerjemah antara metastore Hive dan Katalog Data.

Setelah membuat koneksi, untuk menyinkronkan metadata di metastore Hive dengan Katalog Data, Anda perlu membuat database federasi di Katalog Data menggunakan detail koneksi metastore Hive, dan memetakan database ini ke database Hive. Database disebut sebagai database federasi ketika menunjuk ke entitas di luar Katalog Data.

Anda dapat menerapkan izin Lake Formation menggunakan kontrol akses berbasis tag dan metode sumber daya bernama pada database federasi, dan membagikannya di beberapa unit Akun AWS AWS Organizations, dan organisasi (OU). Anda juga dapat berbagi database federasi secara langsung dengan kepala sekolah IAM dari akun lain.

Anda dapat menentukan izin berbutir halus pada tingkat kolom, tingkat baris, dan tingkat sel menggunakan filter data Lake Formation pada tabel Hive eksternal. Anda dapat menggunakan Amazon Athena, Amazon Redshift, atau Amazon EMR untuk menanyakan tabel Hive eksternal yang dikelola Lake Formation.

Untuk informasi selengkapnya tentang berbagi data lintas akun dan pemfilteran data, lihat:

- Cross-account berbagi data di Lake Formation
- Pemfilteran data dan keamanan tingkat sel di Lake Formation

Katalog Data metadata federasi langkah-langkah tingkat tinggi

1. Anda membuat pengguna IAM dan peran yang memiliki izin yang sesuai untuk menyebarkan AWS SAM aplikasi dan membuat database federasi.
2. Anda mendaftarkan lokasi data Amazon S3 dengan Lake Formation dengan memilih `Enable Data Catalog federation` opsi untuk kumpulan data yang menggunakan metastore Hive eksternal.
3. Anda mengonfigurasi pengaturan AWS SAM aplikasi (nama AWS Glue koneksi, URL ke metastore Hive, dan parameter fungsi Lambda) dan menyebarkan aplikasi. AWS SAM
4. AWS SAM Aplikasi ini menyebarkan sumber daya yang diperlukan untuk menghubungkan metastore Hive eksternal dengan Katalog Data.
5. Untuk menerapkan izin Lake Formation pada database dan tabel Hive, Anda membuat database di Katalog Data menggunakan detail koneksi metastore Hive, dan memetakan database ini ke database Hive.
6. Berikan izin pada database federasi kepada kepala sekolah di akun Anda atau di akun lain.

> **Note**
>
> Anda dapat menghubungkan Katalog Data ke mestastore Hive eksternal, membuat database federasi, dan menjalankan kueri dan skrip ETL pada database dan tabel Hive tanpa menerapkan izin Lake Formation. Untuk data sumber di Amazon S3 yang tidak terdaftar di Lake Formation, akses ditentukan oleh kebijakan izin IAM untuk Amazon S3 dan tindakan. AWS Glue

Untuk batasan, lihat Metadata sarang menyimpan pertimbangan dan batasan berbagi data.

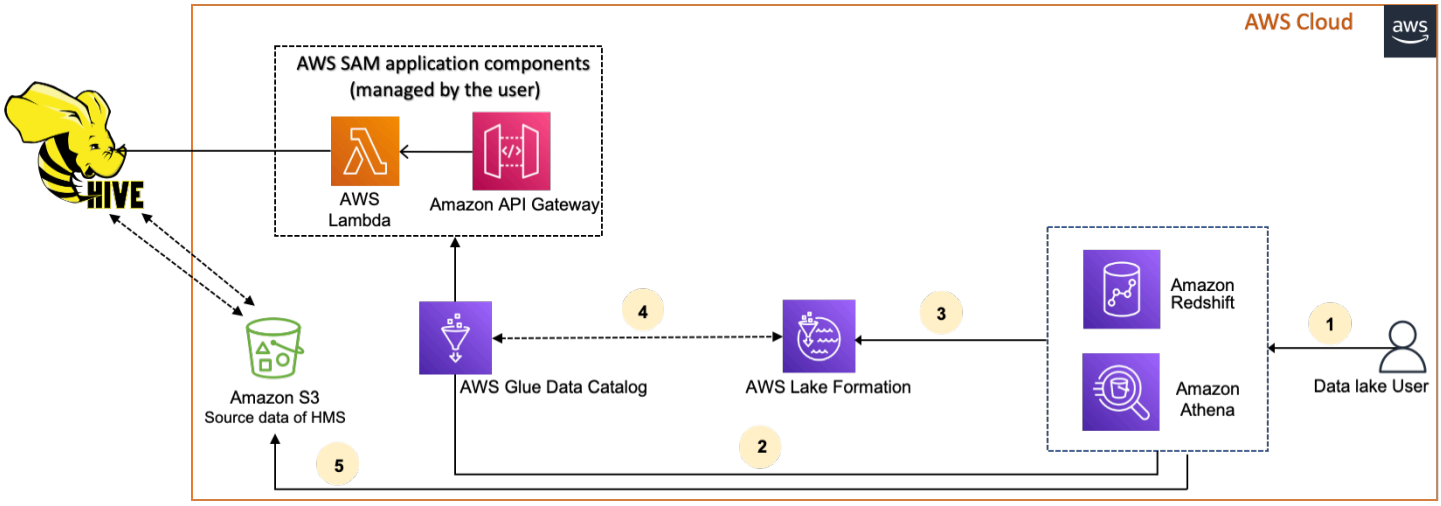
Topik

- Alur kerja
- Prasyarat untuk menghubungkan Katalog Data ke metastore Hive

- Menghubungkan Katalog Data ke metastore Hive eksternal
- Sumber daya tambahan

Alur kerja

Diagram berikut menunjukkan alur kerja untuk menghubungkan AWS Glue Data Catalog ke metastore Hive eksternal.



1. Seorang kepala sekolah mengirimkan kueri menggunakan layanan terintegrasi seperti Athena atau Redshift Spectrum.
2. Layanan terintegrasi membuat panggilan ke Katalog Data untuk metadata, yang pada gilirannya memanggil titik akhir metastore Hive yang tersedia di belakang Amazon API Gateway, dan menerima respons terhadap permintaan metadata.
3. Layanan terintegrasi mengirimkan permintaan ke Lake Formation untuk memverifikasi informasi tabel dan kredensi untuk mengakses tabel.
4. Lake Formation mengotorisasi permintaan dan menjual kredensil sementara ke aplikasi terintegrasi, yang memungkinkan akses data.
5. Menggunakan kredensi sementara yang diterima dari Lake Formation, layanan terintegrasi membaca data dari Amazon S3, dan membagikan hasilnya kepada kepala sekolah.

Prasyarat untuk menghubungkan Katalog Data ke metastore Hive

Untuk menghubungkan AWS Glue Data Catalog ke metastore Apache Hive eksternal dan mengatur izin akses data, Anda harus melengkapi persyaratan berikut:

> **Note**
>
> Kami menyarankan agar administrator Lake Formation menyebarkan AWS SAM aplikasi, dan hanya pengguna istimewa yang menggunakan koneksi metastore Hive untuk membuat database federasi yang sesuai.

1. Buat IAM role.

   Untuk menyebarkan AWS SAM application

   - Buat peran yang memiliki izin yang diperlukan untuk menyebarkan sumber daya (fungsi Lambda, peran IAM Amazon API Gateway, dan AWS Glue koneksi) yang diperlukan untuk membuat koneksi ke metastore Hive.

   Untuk membuat database federasi

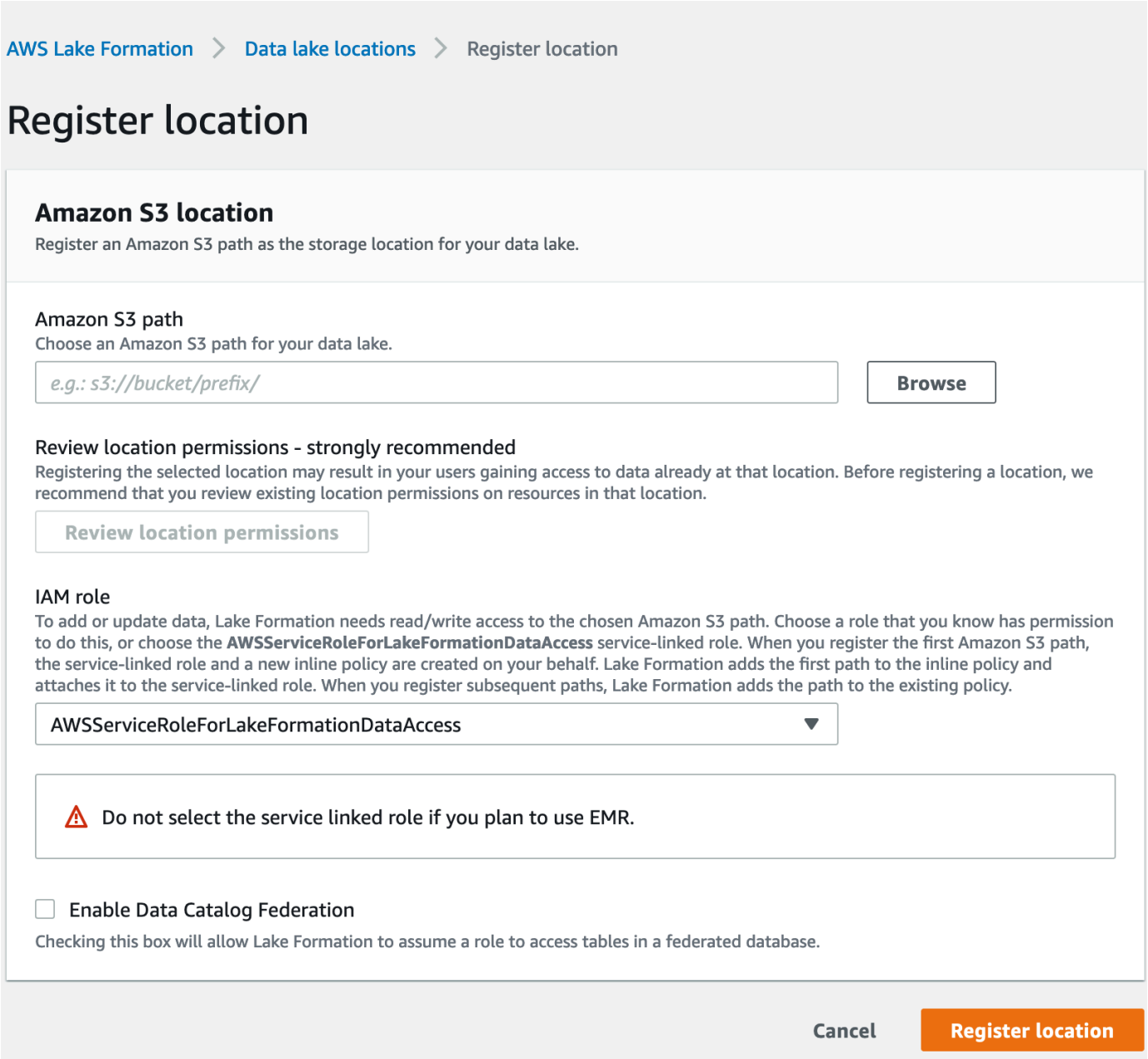
   Izin berikut diperlukan pada sumber daya:

   - `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. Daftarkan lokasi Amazon S3 dengan Lake Formation.

   Untuk menggunakan Lake Formation untuk mengelola dan mengamankan data di danau data Anda, Anda harus mendaftarkan lokasi Amazon S3 yang memiliki data untuk tabel di metastore Hive dengan Lake Formation. Dengan demikian, Lake Formation dapat menjual kredensil ke layanan AWS analitis seperti Athena, Redshift Spectrum, dan Amazon EMR.

   Untuk informasi selengkapnya tentang mendaftarkan lokasi Amazon S3, lihat. Menambahkan lokasi Amazon S3 ke danau data Anda

   Saat Anda mendaftarkan lokasi Amazon S3, pilih kotak centang Aktifkan Federasi Katalog Data untuk mengizinkan Lake Formation mengambil peran untuk mengakses tabel dalam database federasi.



Untuk informasi lebih lanjut tentang mendaftarkan lokasi data dengan Lake Formation, lihatKonfigurasikan lokasi Amazon S3 untuk danau data Anda.

3. Gunakan versi EMR Amazon yang benar.

   Untuk menggunakan Amazon EMR dengan database metastore Hive federasi, Anda harus memiliki Hive versi 3.x atau lebih tinggi dan Amazon EMR versi 6.x atau lebih tinggi.

Menghubungkan Katalog Data ke metastore Hive eksternal

Untuk menghubungkan AWS Glue Data Catalog ke metastore Hive, Anda perlu menyebarkan aplikasi yang disebut. AWS SAM GlueDataCatalogFederation-HiveMetastore Ini menciptakan sumber daya yang diperlukan untuk menghubungkan metastore Hive eksternal dengan Katalog Data. Anda dapat mengakses AWS SAM aplikasi di AWS Serverless Application Repository.

AWS SAM Aplikasi ini membuat koneksi untuk metastore Hive di belakang Amazon API Gateway menggunakan fungsi Lambda. AWS SAM Aplikasi ini menggunakan pengenal sumber daya seragam (URI) sebagai masukan dari pengguna dan menghubungkan metastore Hive eksternal ke Katalog Data. Saat pengguna menjalankan kueri pada tabel Hive, Katalog Data memanggil titik akhir API Gateway. Titik akhir memanggil fungsi Lambda untuk mengambil metadata tabel Hive.

Untuk menghubungkan Katalog Data ke metastore Hive dan mengatur izin

1. Menyebarkan AWS SAM aplikasi.
    1. Masuk ke Konsol Manajemen AWS dan buka AWS Serverless Application Repository.
    2. Di panel navigasi, pilihAplikasi yang tersedia.
    3. Pilih aplikasi Publik.
    4. Pilih opsiMenampilkan aplikasi yang membuat IAM role khusus atau kebijakan sumber daya.
    5. Di kotak pencarian, masukkan nama GlueDataCatalogFederation-HiveMetastore.
    6. Pilih GlueDataCatalogFederation-HiveMetastoreaplikasinya.
    7. Di bawah Pengaturan Aplikasi, masukkan pengaturan minimum yang diperlukan berikut untuk fungsi Lambda Anda:
        - Nama aplikasi - Nama untuk AWS SAM aplikasi Anda.
        - GlueConnectionName- Nama untuk koneksi.
        - HiveMetastoreURIs- URI host metastore Hive Anda.
        - LambdaMemory- Jumlah memori Lambda dalam MB dari 128-10240. Defaultnya adalah 1.024.
        - LambdaTimeout- Runtime pemanggilan Lambda maksimum dalam hitungan detik. Bawaannya adalah 30.
        - VPCSecurityGroupIdsdan VPCSubnetIds- Informasi untuk VPC tempat metastore Hive ada.
    8. Pilih Saya mengakui bahwa aplikasi ini membuat peran IAM khusus dan kebijakan sumber daya. Untuk informasi selengkapnya, pilih tautan Info.

9. Di bagian Pengaturan aplikasi, pilih Deploy. Saat penerapan selesai, fungsi Lambda muncul di bagian Sumber Daya di konsol Lambda.

Aplikasi ini digunakan untuk Lambda. Namanya dilengkapi dengan serverlessrepo- untuk menunjukkan bahwa aplikasi tersebut digunakan dari file. AWS Serverless Application Repository Memilih aplikasi akan membawa Anda ke halaman Sumber Daya tempat masing-masing sumber daya aplikasi yang digunakan terdaftar. Sumber daya termasuk fungsi Lambda yang memungkinkan komunikasi antara Katalog Data dan metastore Hive, AWS Glue koneksi, dan sumber daya lain yang diperlukan untuk federasi database.

2. Buat database federasi di Katalog Data.

Setelah membuat koneksi ke metastore Hive, Anda dapat membuat database federasi di Katalog Data yang mengarah ke database metastore Hive eksternal. Anda perlu membuat database yang sesuai di Katalog Data untuk setiap database metastore Hive yang Anda sambungkan ke Katalog Data.

Lake Formation console

1. Pada halaman Berbagi data, pilih tab Basis data bersama, lalu pilih Buat database.
2. Untuk nama Koneksi, pilih nama koneksi metastore Hive Anda dari menu tarik-turun.
3. Masukkan nama database unik dan pengidentifikasi sumber federasi untuk database. Ini adalah nama yang Anda gunakan dalam pernyataan SQL Anda ketika Anda menanyakan tabel. Nama dapat terdiri dari maksimum 255 karakter dan harus unik dalam akun Anda.
4. Pilih Buat basis data.

AWS CLI

```
aws glue create-database \
'{
 "CatalogId": "<111122223333>",
  "database-input": {
    "Name":"<fed_glue_db>",
    "FederatedDatabase":{
        "Identifier":"<hive_db_on_emr>",
        "ConnectionName":"<hms_connection>"
     }
   }
 }'
```

3. Lihat tabel dalam database federasi.

   Setelah membuat database federasi, Anda dapat melihat daftar tabel di metastore Hive Anda menggunakan konsol Lake Formation atau. AWS CLI

   Lake Formation console

   1. Pilih nama database dari tab Shared database.
   2. Pada halaman Database, pilih Lihat tabel.

   AWS CLI

   Contoh berikut menunjukkan bagaimana untuk mengambil definisi koneksi, nama database, dan beberapa atau semua tabel dalam database. Ganti ID Katalog Data dengan Akun AWS ID valid yang Anda gunakan untuk membuat database. Ganti `hms_connection` dengan nama koneksi.

   ```
   aws glue get-connection \
   --name <hms_connection>  \
   --catalog-id 111122223333
   ```

   ```
   aws glue get-database \
   --name <fed_glu_db> \
   --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. Berikan izin.

   Setelah membuat database, Anda dapat memberikan izin kepada pengguna dan peran IAM lainnya di akun Anda atau ke eksternal Akun AWS dan organisasi. Anda tidak akan dapat memberikan izin menulis data (menyisipkan, menghapus) dan izin metadata (mengubah, menjatuhkan, membuat) pada database federasi. Untuk informasi selengkapnya tentang pemberian izin, lihat. Mengelola izin Lake Formation

5. Kueri database federasi.

   Setelah Anda memberikan izin, pengguna dapat masuk dan mulai menanyakan database federasi menggunakan Athena dan Amazon Redshift. Pengguna sekarang dapat menggunakan nama database lokal untuk referensi database Hive dalam query SQL.

   Contoh sintaks Amazon Athena kueri

   Ganti `fed_glue_db` dengan nama database lokal yang Anda buat sebelumnya.

   ```
   Select * from fed_glue_db.customers limit 10;
   ```

Sumber daya tambahan

Posting blog berikut berisi instruksi terperinci untuk mengatur izin Lake Formation pada database dan tabel metastore Hive, dan menanyakannya menggunakan Athena. Kami juga mengilustrasikan kasus penggunaan berbagi lintas akun, di mana kepala Lake Formation di akun produsen A berbagi database dan tabel Hive federasi yang digunakan LF-Tag ke akun konsumen B.

- Kueri metastore Apache Hive Anda dengan izin AWS Lake Formation

Mengelola izin Lake Formation

Lake Formation menyediakan kontrol akses pusat untuk data di danau data Anda. Anda dapat menentukan aturan berbasis kebijakan keamanan untuk pengguna dan aplikasi Anda berdasarkan peran dalam Lake Formation, dan integrasi dengan AWS Identity and Access Management mengautentikasi pengguna dan peran tersebut. Setelah aturan ditetapkan, Lake Formation memberlakukan kontrol akses Anda pada tabel, kolom, dan granularitas tingkat ro untuk pengguna Amazon Redshift Spectrum dan Amazon Athena.

Topik

- Memberikan izin lokasi data
- Memberikan izin pada sumber daya Katalog Data
- Skenario contoh izin
- Pemfilteran data dan keamanan tingkat sel di Lake Formation
- Melihat izin database dan tabel di Lake Formation
- Mencabut izin menggunakan konsol Lake Formation
- Cross-account berbagi data di Lake Formation
- Mengakses dan melihat tabel dan database Katalog Data bersama
- Membuat tautan sumber daya
- Mengakses tabel di seluruh Wilayah

Memberikan izin lokasi data

Izin lokasi data AWS Lake Formation memungkinkan prinsipal untuk membuat dan mengubah sumber daya Katalog Data yang mengarah ke lokasi Amazon S3 terdaftar yang ditentukan. Izin lokasi data berfungsi selain izin data Lake Formation untuk mengamankan informasi di danau data Anda.

Lake Formation tidak menggunakan layanan AWS Resource Access Manager (AWS RAM) untuk pemberian izin lokasi data, jadi Anda tidak perlu menerima undangan berbagi sumber daya untuk izin lokasi data.

Anda dapat memberikan izin lokasi data menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

> **Note**
>
> Agar hibah berhasil, Anda harus terlebih dahulu mendaftarkan lokasi data dengan Lake Formation.

> **Lihat Juga:**
>
> - Underlying data access control

Topik

- Memberikan izin lokasi data (akun yang sama)
- Memberikan izin lokasi data (akun eksternal)
- Memberikan izin pada lokasi data yang dibagikan dengan akun Anda

Memberikan izin lokasi data (akun yang sama)

Ikuti langkah-langkah ini untuk memberikan izin lokasi data kepada kepala sekolah di akun Anda. AWS Anda dapat memberikan izin menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

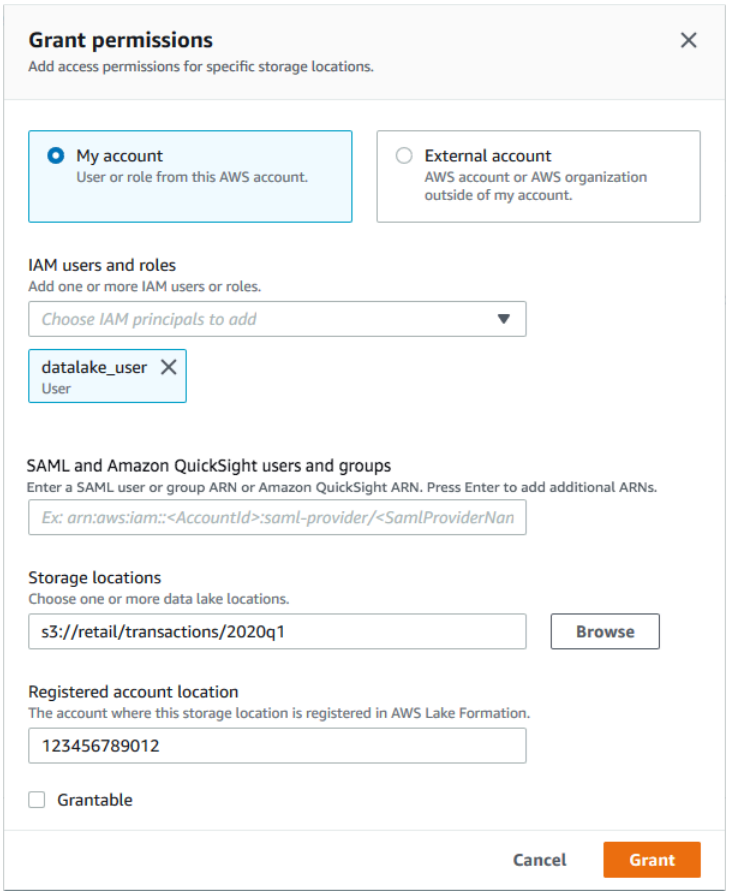
Konsol Manajemen AWS

Untuk memberikan izin lokasi data (akun yang sama)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai kepala sekolah yang memiliki izin pemberian pada lokasi data yang diinginkan.
2. Di panel navigasi, di bawah Izin, pilih Lokasi data.
3. PilihIzin.
4. Di kotak dialog Hibah izin, pastikan bahwa ubin Akun saya dipilih. Kemudian berikan informasi berikut:
   - Untuk pengguna dan peran IAM, pilih satu atau beberapa prinsipal.

- Untuk pengguna dan grup SAFL dan Amazon Quick, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAFL atau ARN untuk pengguna atau grup Amazon Quick.

  Masukkan satu ARN pada satu waktu, dan tekan Enter setelah setiap ARN. Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah
- Untuk lokasi Penyimpanan, pilih Browse, dan pilih lokasi penyimpanan Amazon Simple Storage Service (Amazon S3). Lokasi harus terdaftar di Lake Formation. Pilih Browse lagi untuk menambahkan lokasi lain. Anda juga dapat mengetik lokasi, tetapi pastikan Anda mendahului lokasi dengan. `s3://`
- Untuk lokasi akun terdaftar, masukkan ID AWS akun tempat lokasi terdaftar. Ini default ke ID akun Anda. Dalam skenario lintas akun, administrator data lake di akun penerima dapat menentukan akun pemilik di sini saat memberikan izin lokasi data ke kepala sekolah lain di akun penerima.
- (Opsional) Untuk mengaktifkan prinsipal yang dipilih untuk memberikan izin lokasi data pada lokasi yang dipilih, pilih Dapat Diberikan.



5. PilihIzin.

AWS CLI

Untuk memberikan izin lokasi data (akun yang sama)

- Jalankan `grant-permissions` perintah, dan berikan `DATA_LOCATION_ACCESS` kepada prinsipal, tentukan jalur Amazon S3 sebagai sumber daya.

  Example

  Contoh berikut memberikan izin lokasi data `s3://retail` ke pengguna. `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

Contoh berikut memberikan izin lokasi data `s3://retail` ke `ALLIAMPrincipals` grup.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "111122223333"}}'
```

Lihat Juga:

- Referensi izin Lake Formation

Memberikan izin lokasi data (akun eksternal)

Ikuti langkah-langkah ini untuk memberikan izin lokasi data ke AWS akun atau organisasi eksternal.

Anda dapat memberikan izin menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

Sebelum Anda mulai

Pastikan semua prasyarat akses lintas akun terpenuhi. Untuk informasi selengkapnya, lihat Prasyarat.

Konsol Manajemen AWS

Untuk memberikan izin lokasi data (akun eksternal, konsol)

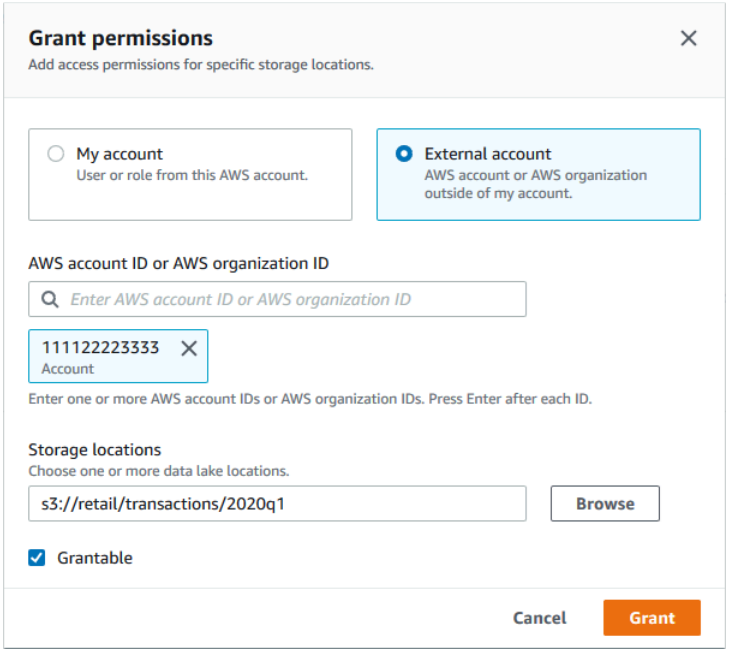
1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, pilih Lokasi data, lalu pilih Hibah.
3. Dalam kotak dialog Hibah izin, pilih ubin akun eksternal.
4. Saat diminta, berikan informasi berikut:
   - Untuk ID AWS akun atau ID AWS organisasi, masukkan nomor AWS akun, ID organisasi, atau ID unit organisasi yang valid.

Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10 hingga 32 huruf kecil atau digit.

ID unit organisasi terdiri dari “ou-” diikuti oleh 4 hingga 32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh “-” kedua (tanda hubung) dan 8 hingga 32 huruf kecil atau digit tambahan.

- Di bawah Lokasi penyimpanan, pilih Jelajahi, dan pilih lokasi penyimpanan Amazon Simple Storage Service (Amazon S3). Lokasi harus terdaftar di Lake Formation.



5. Pilih Grantable.
6. PilihIzin.

AWS CLI

Untuk memberikan izin lokasi data (akun eksternal, AWS CLI)

- Untuk memberikan izin ke AWS akun eksternal, masukkan perintah yang mirip dengan yang berikut ini.

```
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"}}'
```

Perintah ini memberikan opsi hibah ke akun 1111-2222-3333 di lokasi Amazon S3, yang dimiliki oleh akun `s3://retail/transactions/2020q1` 1234-5678-9012. `DATA_LOCATION_ACCESS`

Untuk memberikan izin ke organisasi, masukkan perintah yang mirip dengan yang berikut ini.

```
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"}}'
```

Perintah ini memberikan opsi hibah kepada organisasi `o-abcdefghijkl` di `s3://retail/transactions/2020q1` lokasi Amazon S3, yang dimiliki oleh akun 1234-5678-9012. `DATA_LOCATION_ACCESS`

Untuk memberikan izin kepada prinsipal di AWS akun eksternal, masukkan perintah yang mirip dengan yang berikut ini.

```
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"}}'
```

Perintah ini memberikan `DATA_LOCATION_ACCESS` kepada kepala sekolah di akun 1111-2222-3333 di lokasi Amazon S3, yang dimiliki oleh akun `s3://retail/transactions/2020q1` 1234-5678-9012.

Example

Contoh berikut memberikan izin lokasi data `s3://retail` untuk `ALLIAMPrincipals` mengelompokkan di akun eksternal.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "123456789012"}}'
```

Lihat Juga:

- Referensi izin Lake Formation

Memberikan izin pada lokasi data yang dibagikan dengan akun Anda

Setelah sumber daya Katalog Data dibagikan dengan AWS akun Anda, sebagai administrator data lake, Anda dapat memberikan izin pada sumber daya ke prinsipal lain di akun Anda. Jika `ALTER` izin diberikan pada tabel bersama, dan tabel menunjuk ke lokasi Amazon S3 terdaftar, Anda juga harus memberikan izin lokasi data di lokasi tersebut. Demikian juga, jika `ALTER` izin `CREATE_TABLE` atau diberikan pada database bersama dan database memiliki properti lokasi yang menunjuk ke lokasi terdaftar, Anda juga harus memberikan izin lokasi data di lokasi tersebut.

Untuk memberikan izin lokasi data pada lokasi bersama kepada prinsipal di akun Anda, akun Anda harus telah diberikan `DATA_LOCATION_ACCESS` izin di lokasi tersebut dengan opsi hibah. Ketika Anda kemudian memberikan `DATA_LOCATION_ACCESS` kepada prinsipal lain di akun Anda, Anda harus menyertakan ID Katalog Data (ID AWS akun) dari akun pemilik. Akun pemilik adalah akun yang mendaftarkan lokasi.

Anda dapat menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI untuk memberikan izin lokasi data.

Untuk memberikan izin pada lokasi data yang dibagikan dengan akun Anda (konsol)

- Ikuti langkah-langkah di Memberikan izin lokasi data (akun yang sama).

  Untuk lokasi Penyimpanan, Anda harus mengetik lokasi. Untuk lokasi akun terdaftar, masukkan ID AWS akun akun pemilik.

Untuk memberikan izin pada lokasi data yang dibagikan dengan akun Anda (AWS CLI)

- Masukkan salah satu perintah berikut untuk memberikan izin kepada pengguna atau peran.

```
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>"}}'
```

Memberikan izin pada sumber daya Katalog Data

Anda dapat memberikan izin Data kepada prinsipal AWS Lake Formation sehingga prinsipal dapat membuat dan mengelola sumber daya Katalog Data, serta dapat mengakses data yang mendasarinya. Anda dapat memberikan izin Data lake pada katalog, database, tabel, dan tampilan. Saat Anda memberikan izin pada tabel, Anda dapat membatasi akses ke kolom atau baris tabel tertentu untuk kontrol akses yang lebih halus.

Anda dapat memberikan izin pada setiap katalog, database, tabel, dan tampilan, atau dengan satu operasi hibah, Anda dapat memberikan izin pada semua database, tabel, dan tampilan dalam katalog atau database. Jika Anda memberikan izin pada semua tabel dalam database kepada prinsipal IAM, Anda secara implisit memberikan izin pada database. `DESCRIBE` Database kemudian muncul di halaman Database di konsol, dan dikembalikan oleh operasi `GetDatabases` API. Prinsip yang sama berlaku di tingkat katalog - ketika Anda menerima izin untuk database dalam katalog, Anda juga mendapatkan `DESCRIBE` izin untuk katalog itu.

> **Important**
>
> Izin implisit hanya berlaku ketika memberikan `DESCRIBE` izin kepada prinsipal IAM dalam akun yang sama. AWS Untuk sumber daya lintas akun, Anda harus secara eksplisit memberikan izin. `DESCRIBE` Pemberian `DESCRIBE` izin otomatis tidak berlaku saat menggunakan kontrol akses berbasis atribut (ABAC). Saat memberikan izin pada semua tabel dalam database menggunakan atribut, Lake Formation tidak secara implisit memberikan `DESCRIBE` izin ke database.

Anda dapat memberikan izin dengan menggunakan metode sumber daya bernama atau metode kontrol akses (LF-TBAC) berbasis tag Lake Formation.

Anda dapat memberikan izin kepada kepala sekolah yang sama Akun AWS atau ke akun atau organisasi eksternal. Ketika Anda memberikan kepada akun atau organisasi eksternal, Anda membagikan objek Katalog Data yang Anda miliki dengan akun atau organisasi tersebut. Prinsipal di akun atau organisasi tersebut kemudian dapat mengakses objek Katalog Data yang Anda miliki dan data yang mendasarinya.

> **Note**
>
> Saat ini, LF-TBAC metode ini mendukung pemberian izin lintas akun kepada kepala sekolah IAM, organisasi Akun AWS, dan unit organisasi (OU).

Saat Anda memberikan izin ke akun atau organisasi eksternal, Anda harus menyertakan opsi hibah. Hanya administrator data lake di akun eksternal yang dapat mengakses objek bersama hingga administrator memberikan izin pada objek bersama ke prinsipal lain di akun eksternal.

Anda dapat memberikan izin Katalog Data dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

> **Note**
>
> Saat Anda menghapus objek Katalog Data, semua izin yang terkait dengan objek menjadi tidak valid. Membuat ulang sumber daya yang sama dengan nama yang sama, tidak akan memulihkan izin Lake Formation. Pengguna harus mengatur izin baru lagi.

> **Lihat juga:**
>
> - Berbagi tabel Katalog Data dan database di seluruh AWS Akun
> - Kontrol akses metadata
> - Referensi izin Lake Formation

Izin IAM diperlukan untuk memberikan atau mencabut izin Lake Formation

Semua kepala sekolah, termasuk administrator data lake, memerlukan izin AWS Identity and Access Management (IAM) berikut untuk memberikan atau mencabut izin Katalog AWS Lake Formation Data atau izin lokasi data dengan Lake Formation API atau: AWS CLI

- `lakeformation:GrantPermissions`
- `lakeformation:BatchGrantPermissions`
- `lakeformation:RevokePermissions`
- `lakeformation:BatchRevokePermissions`
- `glue:GetTable`,`glue:GetDatabase`, atau `glue:GetCatalog` untuk tabel, database, atau katalog yang Anda berikan izin menggunakan metode sumber daya bernama.

> **Note**
>
> Administrator danau data memiliki izin Lake Formation implisit untuk memberikan dan mencabut izin Lake Formation. Tetapi mereka masih membutuhkan izin IAM pada hibah Lake Formation dan mencabut operasi API.
> Peran IAM dengan kebijakan `AWSLakeFormationDataAdmin` AWS terkelola tidak dapat menambahkan administrator data lake baru karena kebijakan ini berisi penolakan eksplisit untuk operasi Lake Formation API,. `PutDataLakeSetting`

Kebijakan IAM berikut direkomendasikan untuk prinsipal yang bukan administrator data lake dan yang ingin memberikan atau mencabut izin menggunakan konsol 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": "*"
        }
    ]
}
```

Semua izin `glue:` dan `iam:` izin dalam kebijakan ini tersedia dalam kebijakan AWS `AWSGlueConsoleFullAccess` terkelola.

Untuk memberikan izin dengan menggunakan kontrol akses berbasis tag Lake Formation (LF-TBAC), kepala sekolah memerlukan izin IAM tambahan. Untuk informasi selengkapnya, lihat Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation dan Referensi personas Lake Formation dan izin IAM.

Cross-account izin

Pengguna yang ingin memberikan izin Lake Formation lintas akun dengan menggunakan metode sumber daya bernama juga harus memiliki izin dalam kebijakan `AWSLakeFormationCrossAccountManager` AWS terkelola.

Administrator data lake memerlukan izin yang sama untuk memberikan izin lintas akun, ditambah izin AWS Resource Access Manager (AWS RAM) untuk mengaktifkan pemberian izin kepada organisasi. Untuk informasi selengkapnya, lihat Izin administrator danau data.

Pengguna administratif

Kepala sekolah dengan izin administratif—misalnya, dengan kebijakan `AdministratorAccess` AWS terkelola—memiliki izin untuk memberikan izin Lake Formation dan membuat administrator data

lake. Untuk menolak akses pengguna atau peran ke operasi administrator Lake Formation, lampirkan atau tambahkan `Deny` pernyataan kebijakannya untuk operasi API administrator.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Action": [
                "lakeformation:GetDataLakeSettings",
                "lakeformation:PutDataLakeSettings"
            ],
            "Effect": "Deny",
            "Resource": [
                "*"
            ]
        }
    ]
}
```

> **Important**
>
> Untuk mencegah pengguna menambahkan diri mereka sebagai administrator dengan skrip ekstrak, transformasi, dan muat (ETL), pastikan bahwa semua pengguna dan peran non-administrator ditolak akses ke operasi API ini. Kebijakan `AWSLakeFormationDataAdmin` AWS terkelola berisi penolakan eksplisit untuk operasi Lake Formation API, `PutDataLakeSetting` yang mencegah pengguna menambahkan administrator data lake baru.

Memberikan izin data menggunakan metode sumber daya bernama

Metode sumber daya Katalog Data bernama adalah cara memberikan izin ke AWS Glue Data Catalog objek, seperti katalog, database, tabel, kolom, dan tampilan, menggunakan pendekatan terpusat. Ini memungkinkan Anda untuk menentukan kebijakan berbasis sumber daya yang mengontrol akses ke sumber daya tertentu dalam data lake Anda.

Saat Anda menggunakan metode sumber daya bernama untuk memberikan izin, Anda dapat menentukan jenis sumber daya dan izin yang ingin Anda berikan atau cabut untuk sumber daya tersebut. Anda juga dapat mencabut izin nanti jika diperlukan, sehingga menghapus izin dari sumber daya terkait.

Anda dapat memberikan izin dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Topik

- Memberikan izin katalog menggunakan metode sumber daya bernama
- Memberikan izin database menggunakan metode sumber daya bernama
- Memberikan izin tabel menggunakan metode sumber daya bernama
- Memberikan izin pada tampilan menggunakan metode sumber daya bernama

Memberikan izin katalog menggunakan metode sumber daya bernama

Langkah-langkah berikut menjelaskan cara memberikan izin katalog dengan menggunakan metode sumber daya bernama.

Console

Gunakan halaman izin Hibah di konsol Lake Formation. Halaman ini dibagi menjadi beberapa bagian berikut:

- Jenis utama - Anda dapat memberikan izin kepada prinsipal tertentu atau menggunakan tag atribut.
  - Prinsipal — Pengguna IAM, peran, pengguna dan grup IAM Identity Center, pengguna dan grup SAMB, AWS akun, organisasi, atau unit organisasi untuk memberikan izin.

    Prinsipal berdasarkan atribut — Tambahkan pasangan nilai kunci tag dari tag sesi IamRoles atau IAM. Prinsipal dengan atribut yang cocok menerima akses ke sumber daya yang ditentukan.
  - LF-Tags atau sumber daya katalog — Katalog, database, tabel, tampilan, atau tautan sumber daya untuk memberikan izin pada.
  - Izin — Izin Lake Formation untuk diberikan.

> **Note**
>
> Untuk memberikan izin pada tautan sumber daya database, lihatMemberikan izin tautan sumber daya.

1. Buka halaman izin Hibah.

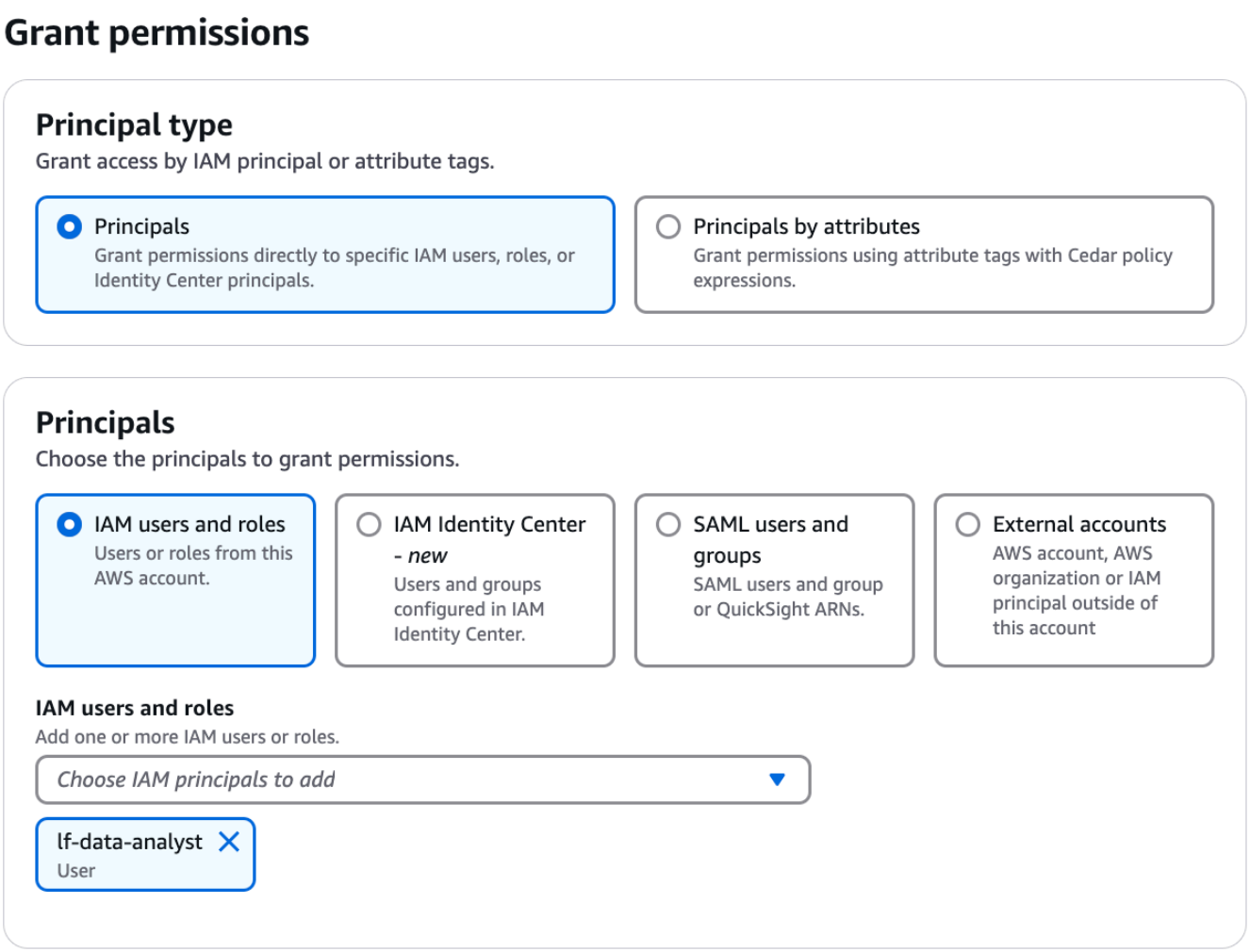
   Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake, pembuat katalog, atau pengguna IAM yang memiliki izin Grantable di katalog.

   Lakukan salah satu tindakan berikut:

   - Di panel navigasi, di bawah Izin, pilih Izin data. Kemudian pilih Grant.
   - Di panel navigasi, pilih Katalog di bawah Katalog Data. Kemudian, pada halaman Katalog, pilih katalog, dan dari menu Tindakan, di bawah Izin, pilih Hibah.

   > **Note**
   >
   > Anda dapat memberikan izin pada katalog melalui tautan sumber dayanya. Untuk melakukannya, pada halaman Katalog, pilih wadah tautan katalog, dan pada menu Tindakan, pilih Hibah sesuai target. Untuk informasi selengkapnya, lihat Cara kerja tautan sumber daya di Lake Formation.

2. Selanjutnya, di bagian Jenis utama, pilih prinsipal atau tentukan atribut yang dilampirkan pada prinsipal.



Tentukan kepala sekolah

Pengguna dan peran IAM

Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

Pusat Identitas IAM

Pilih satu atau beberapa pengguna atau grup dari daftar Pengguna dan grup. Pilih Tambah untuk menambahkan lebih banyak pengguna atau grup.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAMB dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Amazon Quick. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah

> **Note**
>
> Integrasi Lake Formation dengan Quick hanya didukung untuk Quick Enterprise Edition.

Akun eksternal

Untuk Akun AWS, AWS organisasi, atau IAM Principal masukkan satu atau beberapa ID AWS akun, ID organisasi, ID unit organisasi, atau ARN yang valid untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10-32 huruf kecil atau digit.

ID unit organisasi dimulai dengan “ou-” diikuti oleh 4-32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

Prinsipal berdasarkan atribut

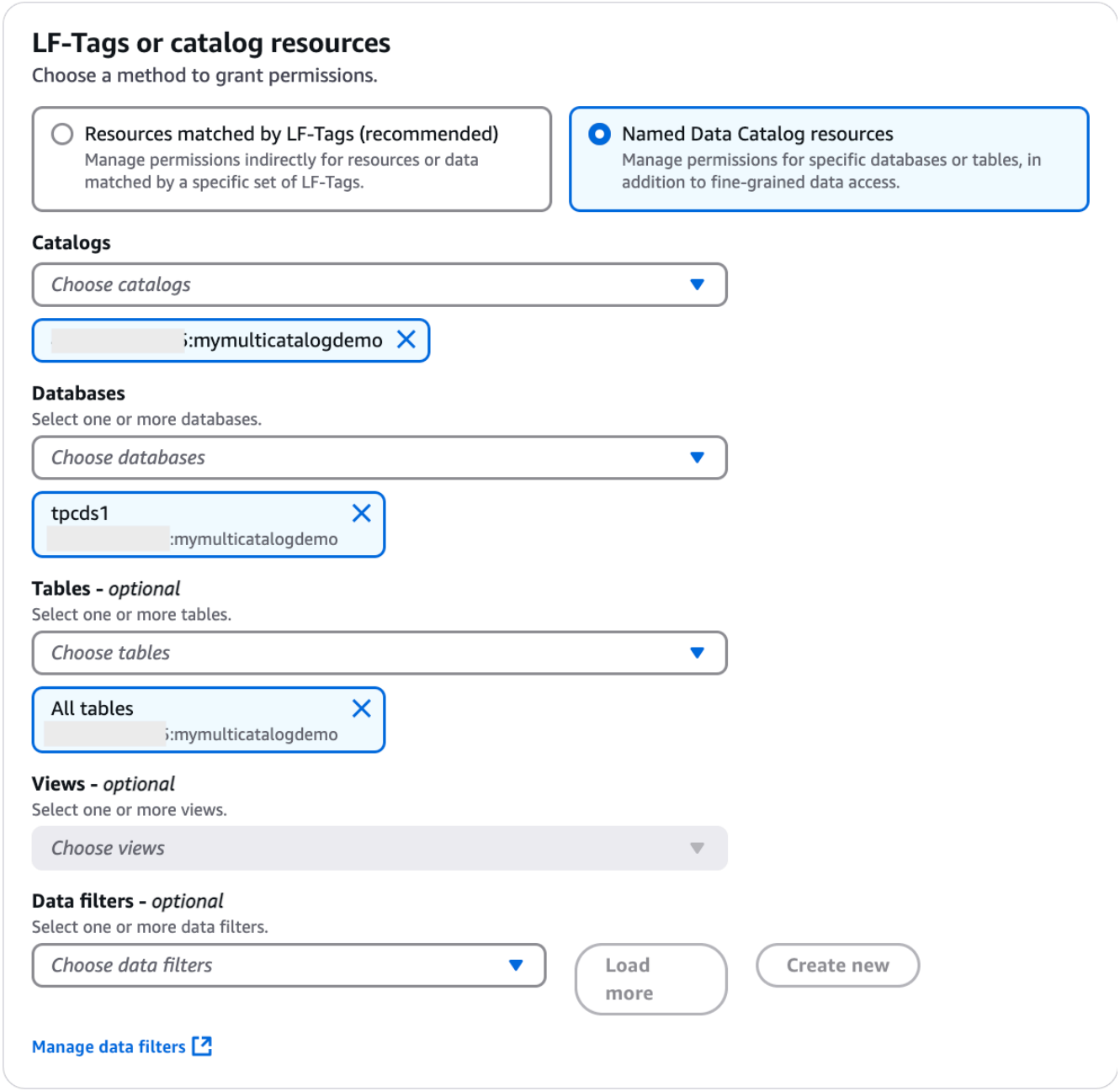
Atribut

Tambahkan pasangan nilai kunci tag IAM dari peran IAM.

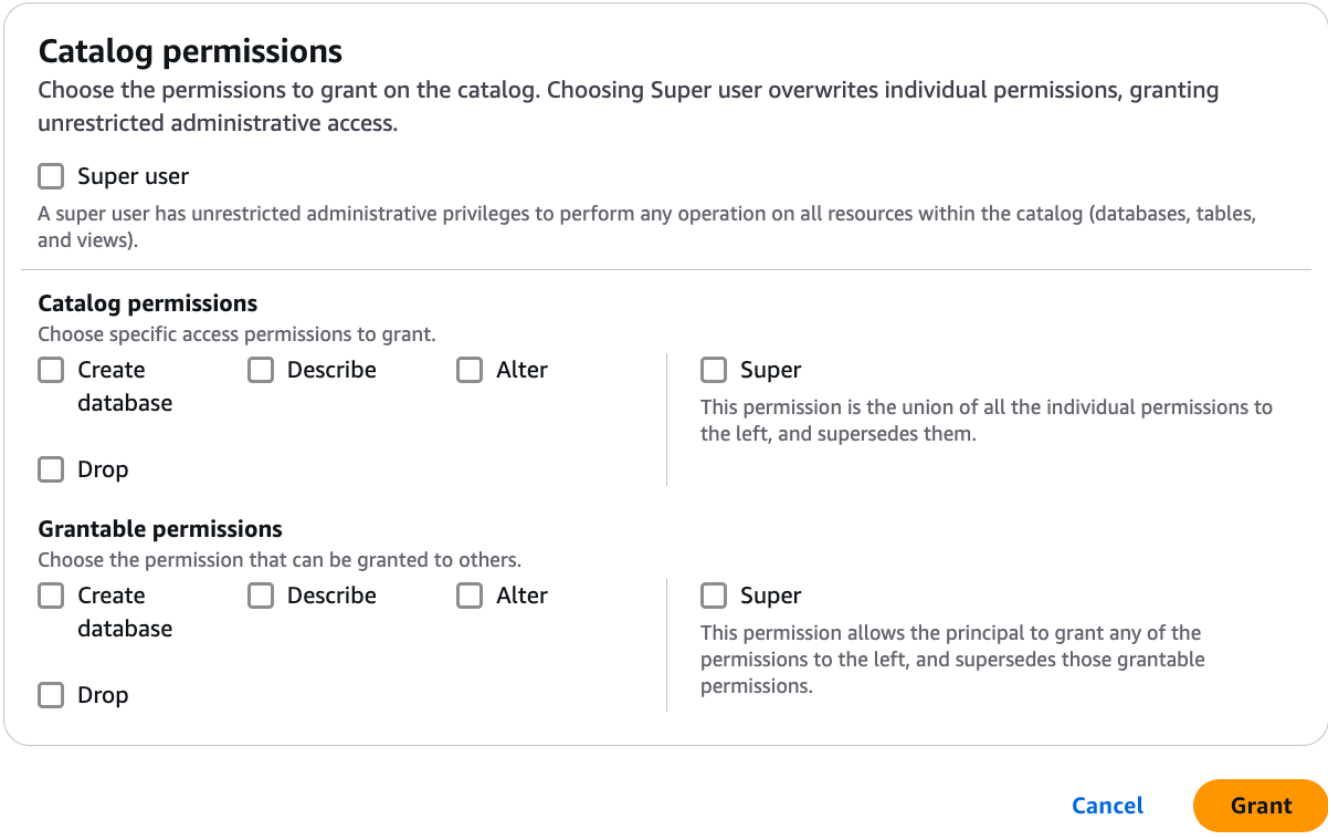
Ruang lingkup izin

Tentukan apakah Anda memberikan izin kepada kepala sekolah dengan atribut yang cocok di akun yang sama atau di akun lain.

3. Di bagian sumber daya LF-Tags atau katalog, pilih Sumber daya katalog data bernama.



4. Pilih satu atau beberapa katalog dari daftar Katalog. Anda juga dapat memilih satu atau beberapa Database, Tabel, Filter and/or data.
5. Di bagian Izin katalog, pilih izin dan izin yang dapat diberikan. Di Izin katalog, pilih satu atau beberapa izin untuk diberikan.



Pilih Super user untuk memberikan hak administratif tak terbatas untuk melakukan operasi apa pun pada semua sumber daya dalam katalog (database, tabel, dan tampilan).

> **Note**
>
> Setelah memberikan `Create database` atau `Alter` pada katalog yang memiliki properti lokasi yang menunjuk ke lokasi terdaftar, pastikan juga memberikan izin lokasi data pada lokasi tersebut kepada prinsipal. Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

6. (Opsional) Di bawah Izin yang Dapat Diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.
7. PilihIzin.

   Halaman izin data menunjukkan rincian izin. Jika Anda menggunakan opsi atribut Prinsip berdasarkan untuk memberikan izin, Anda dapat melihat pemberian izin dalam daftar. `ALLPrincipals`

AWS CLI

Untuk memberikan izin katalog menggunakan AWS CLI, lihat. Membuat katalog federasi Amazon Redshift

Memberikan izin database menggunakan metode sumber daya bernama

Langkah-langkah berikut menjelaskan cara memberikan izin database dengan menggunakan metode sumber daya bernama.

Console

Gunakan halaman izin Hibah di konsol Lake Formation. Halaman ini dibagi menjadi beberapa bagian berikut:

- Jenis utama - Bagian Prinsipal mencakup pengguna IAM, peran, pengguna dan grup Pusat Identitas IAM, pengguna dan grup SAMB, AWS akun, organisasi, atau unit organisasi untuk memberikan izin. Di bagian Prinsip menurut atribut, Anda dapat menentukan kunci dan nilai untuk atribut yang dilampirkan pada peran IAM.
- LF-Tags atau sumber daya katalog — Database, tabel, tampilan, atau tautan sumber daya untuk memberikan izin pada.
- Izin — Izin Lake Formation untuk diberikan.

> **Note**
>
> Untuk memberikan izin pada tautan sumber daya database, lihatMemberikan izin tautan sumber daya.

1. Buka halaman izin Hibah.

   Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake, pembuat database, atau pengguna IAM yang memiliki izin Grantable pada database.

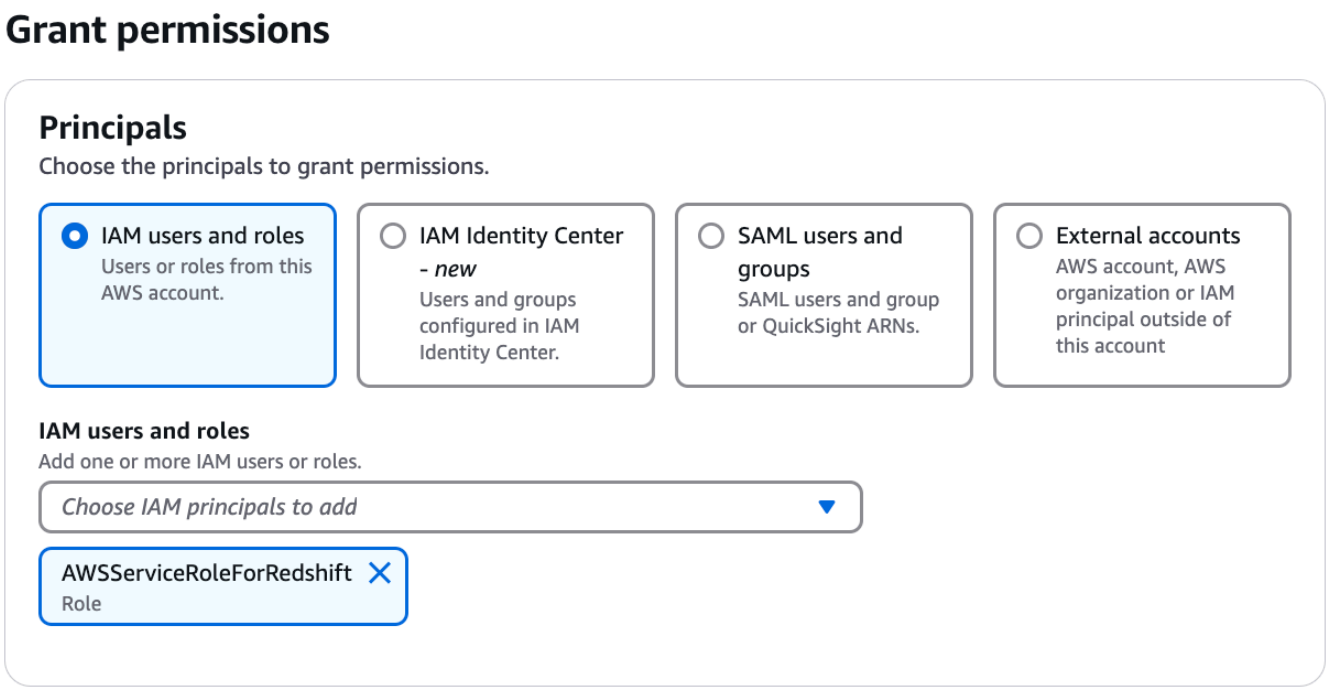
   Lakukan salah satu tindakan berikut:

   - Di panel navigasi, di bawah Izin, pilih Izin data. Kemudian pilih Grant.

- Di panel navigasi, pilih Database di bawah Katalog Data. Kemudian, pada halaman Database, pilih database, dan dari menu Tindakan, di bawah Izin, pilih Hibah.

> **Note**
>
> Anda dapat memberikan izin pada database melalui tautan sumber dayanya. Untuk melakukannya, pada halaman Database, pilih tautan sumber daya, dan pada menu Tindakan, pilih Hibah sesuai target. Untuk informasi selengkapnya, lihat Cara kerja tautan sumber daya di Lake Formation.

2. Di bagian Jenis utama, tentukan prinsipal atau berikan izin kepada kepala sekolah menggunakan atribut.



Pengguna dan peran IAM

Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

Pusat Identitas IAM

Pilih satu atau beberapa pengguna atau grup dari daftar Pengguna dan grup. Pilih Tambah untuk menambahkan lebih banyak pengguna atau grup.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAMB dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Amazon Quick. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah

> **Note**
>
> Integrasi Lake Formation dengan Quick hanya didukung untuk Quick Enterprise Edition.

Akun eksternal

Untuk Akun AWS, AWS organisasi, atau IAM Principal masukkan satu atau beberapa ID AWS akun, ID organisasi, ID unit organisasi, atau ARN yang valid untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10-32 huruf kecil atau digit.

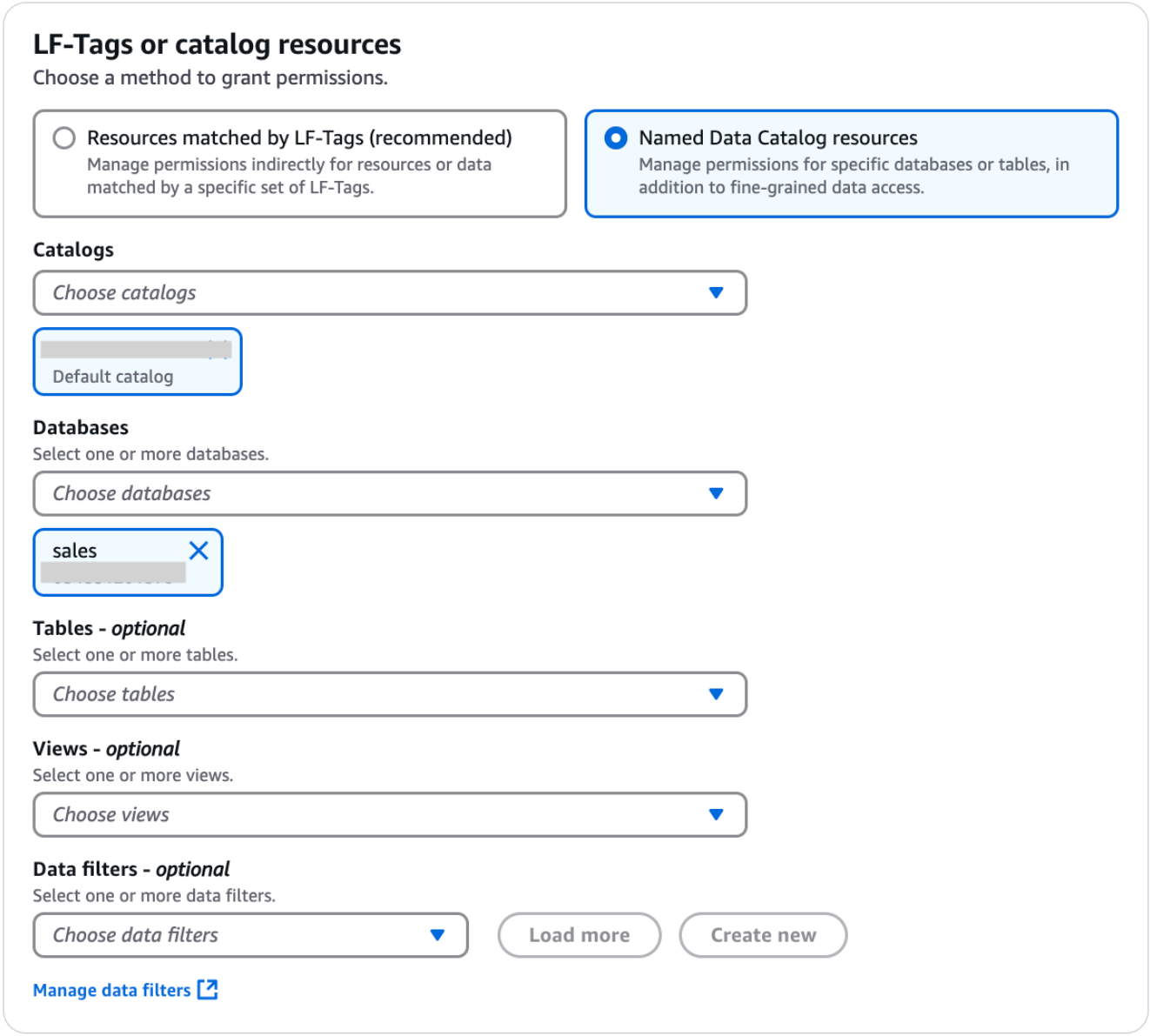
ID unit organisasi dimulai dengan “ou-” diikuti oleh 4-32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

Prinsipal berdasarkan atribut

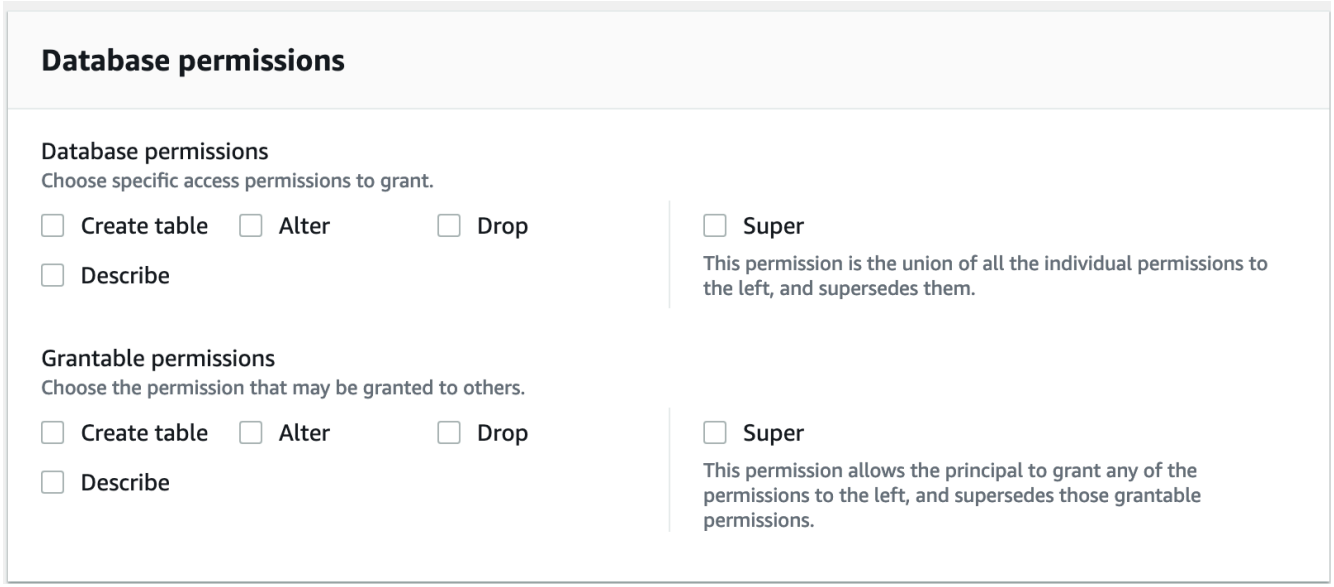
Tentukan kunci atribut dan nilai (s). Jika Anda memilih lebih dari satu nilai, Anda membuat ekspresi atribut dengan operator OR. Ini berarti bahwa jika salah satu nilai tag atribut yang ditetapkan ke peran IAM atau kecocokan pengguna, role/user memperoleh izin akses pada sumber daya.

Pilih cakupan izin dengan menentukan apakah Anda memberikan izin kepada kepala sekolah dengan atribut yang cocok di akun yang sama atau di akun lain.

3. Di bagian sumber daya LF-Tags atau katalog, pilih Sumber daya katalog data bernama.



4. Pilih satu atau beberapa database dari daftar Database. Anda juga dapat memilih satu atau beberapa filter and/or Data Tabel.
5. Di bagian Izin, pilih izin dan izin yang dapat diberikan. Di bawah Izin database, pilih satu atau beberapa izin untuk diberikan.



> **Note**
>
> Setelah memberikan `Create Table` atau `Alter` pada database yang memiliki properti lokasi yang menunjuk ke lokasi terdaftar, pastikan juga untuk memberikan izin lokasi data pada lokasi kepada prinsipal. Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

6. (Opsional) Di bawah Izin yang Dapat Diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.
7. PilihIzin.

AWS CLI

Anda dapat memberikan izin database dengan menggunakan metode sumber daya bernama dan AWS Command Line Interface (AWS CLI).

Untuk memberikan izin database menggunakan AWS CLI

- Jalankan `grant-permissions` perintah, dan tentukan database atau Katalog Data sebagai sumber daya, tergantung pada izin yang diberikan.

  Dalam contoh berikut, ganti *<account-id>* dengan ID AWS akun yang valid.

Example— Hibah untuk membuat database

Contoh ini diberikan `CREATE_DATABASE` kepada penggunadatalake_user1. Karena sumber daya di mana izin ini diberikan adalah Katalog Data, perintah menentukan `CatalogResource` struktur kosong sebagai `resource` parameter.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

Example— Hibah untuk membuat tabel dalam database yang ditunjuk

Contoh berikutnya memberikan `CREATE_TABLE` pada database `retail` kepada penggunadatalake_user1.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Example— Hibah ke eksternal AWS akun dengan opsi Hibah

Contoh berikutnya memberikan opsi hibah `CREATE_TABLE` pada database ke akun eksternal `retail` 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333 --permissions "CREATE_TABLE"
 --permissions-with-grant-option "CREATE_TABLE" --resource '{ "Database":
 {"Name":"retail"}}'
```

Example— Hibah untuk organisasi

Contoh berikutnya memberikan opsi hibah pada database `issues` ke organisasio-abcdefghijkl. `ALTER`

```
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- Hibah kepada AlliamPrincipals di akun yang sama

Contoh berikutnya memberikan `CREATE_TABLE` izin pada database `retail` ke semua kepala sekolah di akun yang sama. Opsi ini memungkinkan setiap prinsipal dalam akun untuk membuat tabel dalam database dan membuat tautan sumber daya tabel yang memungkinkan mesin kueri terintegrasi untuk mengakses database dan tabel bersama. Opsi ini sangat berguna ketika kepala sekolah menerima hibah lintas akun, dan tidak memiliki izin untuk membuat tautan sumber daya. Dalam skenario ini, administrator data lake dapat membuat database placeholder dan memberikan `CREATE_TABLE` izin ke `ALLIAMPrincipal` grup, memungkinkan setiap prinsipal IAM di akun untuk membuat tautan sumber daya di database placeholder.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"temp","CatalogId":"111122223333"}}'
```

Example- Hibah kepada AlliamPrincipals di akun eksternal

Contoh berikutnya memberikan `CREATE_TABLE` pada database `retail` ke semua kepala sekolah di akun eksternal. Opsi ini memungkinkan setiap prinsipal dalam akun untuk membuat tabel dalam database.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"retail","CatalogId":"123456789012"}}'
```

> **Note**
>
> Setelah memberikan `CREATE_TABLE` atau `ALTER` pada database yang memiliki properti lokasi yang menunjuk ke lokasi terdaftar, pastikan juga untuk memberikan izin lokasi data pada lokasi kepada prinsipal. Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

Lihat juga

- Referensi izin Lake Formation
- Memberikan izin pada database atau tabel yang dibagikan dengan akun Anda
- Mengakses dan melihat tabel dan database Katalog Data bersama

Memberikan izin tabel menggunakan metode sumber daya bernama

Anda dapat menggunakan konsol Lake Formation atau AWS CLI untuk memberikan izin Lake Formation pada tabel Katalog Data. Anda dapat memberikan izin pada tabel individual, atau dengan operasi hibah tunggal, Anda dapat memberikan izin pada semua tabel dalam database.

Jika Anda memberikan izin pada semua tabel dalam database, Anda secara implisit memberikan `DESCRIBE` izin pada database. Database kemudian muncul di halaman Database di konsol, dan dikembalikan oleh operasi `GetDatabases` API. Pemberian `DESCRIBE` izin otomatis ini tidak berlaku saat menggunakan kontrol akses berbasis atribut (ABAC). Saat memberikan izin pada semua tabel dalam database menggunakan atribut, Lake Formation tidak secara implisit memberikan `DESCRIBE` izin ke database.

Saat Anda memilih `SELECT` sebagai izin untuk diberikan, Anda memiliki opsi untuk menerapkan filter kolom, filter baris, atau filter sel.

Console

Langkah-langkah berikut menjelaskan cara memberikan izin tabel dengan menggunakan metode sumber daya bernama dan halaman izin danau data Grant di konsol Lake Formation. Halaman ini dibagi menjadi beberapa bagian ini:

- Jenis Prinsipal — Pengguna, peran, AWS akun, organisasi, atau unit organisasi untuk memberikan izin. Anda juga dapat memberikan izin kepada prinsipal dengan atribut yang cocok.
- LF-Tags atau sumber daya katalog — Database, tabel, atau tautan sumber daya untuk memberikan izin pada.
- Izin — Izin Lake Formation untuk diberikan.

> **Note**
>
> Untuk memberikan izin pada tautan sumber daya tabel, lihatMemberikan izin tautan sumber daya.

1. Buka halaman izin Hibah.

   Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake, pembuat tabel, atau pengguna yang telah diberikan izin di atas tabel dengan opsi hibah.

   Lakukan salah satu tindakan berikut:

   - Di panel navigasi, pilih Izin data di bawah Izin. Kemudian pilih Grant.
   - Di panel navigasi, pilih Tabel. Kemudian, pada halaman Tabel, pilih tabel, dan pada menu Tindakan, di bawah Izin, pilih Hibah.

   > **Note**
   >
   > Anda dapat memberikan izin pada tabel melalui tautan sumber dayanya. Untuk melakukannya, pada halaman Tabel, pilih tautan sumber daya, dan pada menu Tindakan, pilih Hibah sesuai target. Untuk informasi selengkapnya, lihat Cara kerja tautan sumber daya di Lake Formation.

2. Selanjutnya, di bagian Jenis utama, tentukan prinsipal atau prinsipal dengan attruasi yang cocok untuk memberikan izin.

   Pengguna dan peran IAM

   Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

   Pusat Identitas IAM

   Pilih satu atau beberapa pengguna atau grup dari daftar Pengguna dan grup.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAMB dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Cepat. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah

> **Note**
>
> Integrasi Lake Formation dengan Quick didukung hanya untuk Quick Enterprise Edition.

Akun eksternal

Untuk Akun AWS , AWS organisasi, atau IAM Principal masukkan satu atau beberapa Akun AWS ID yang valid, ID organisasi, ID unit organisasi, atau ARN untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10-32 huruf kecil atau digit.

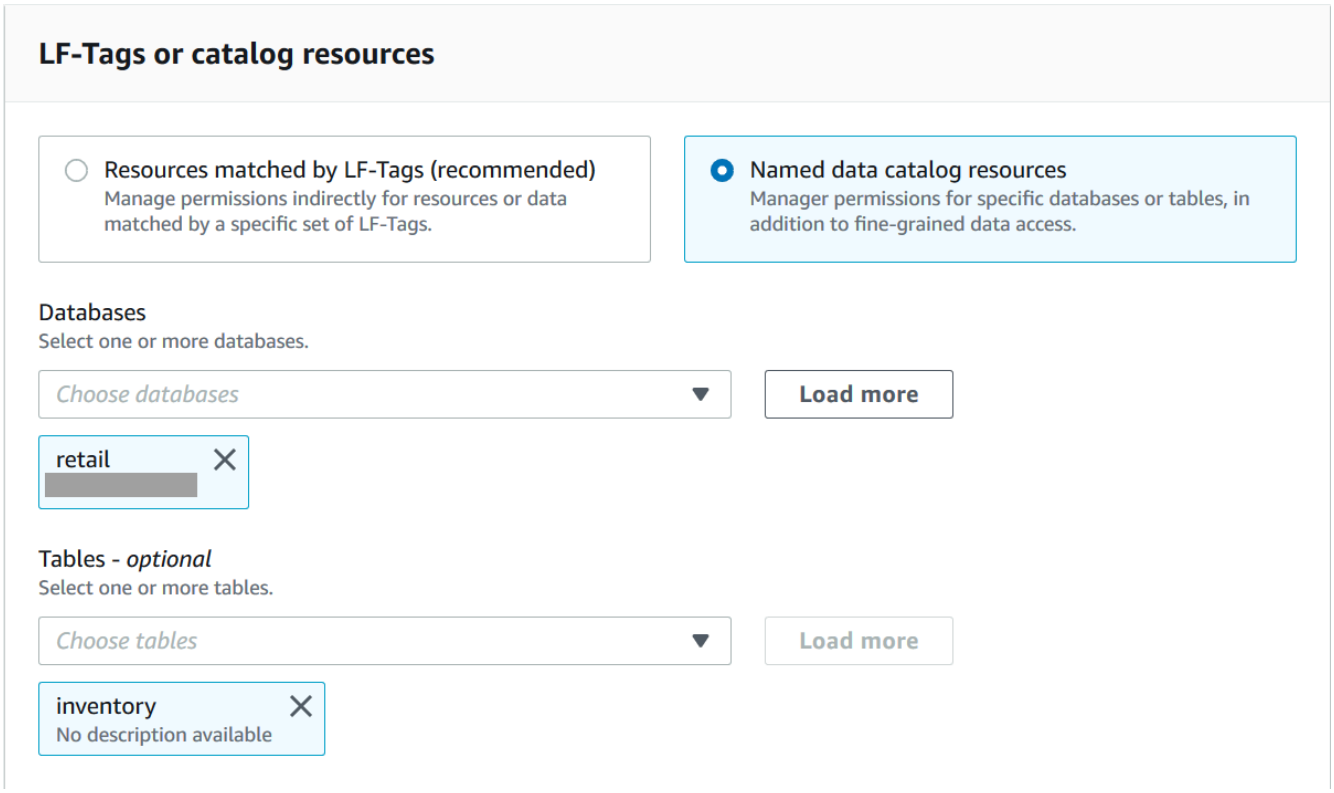
ID unit organisasi dimulai dengan “ou-” diikuti oleh 4-32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh karakter “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

Prinsipal berdasarkan atribut

Tentukan kunci atribut dan nilai (s). Jika Anda memilih lebih dari satu nilai, Anda membuat ekspresi atribut dengan operator OR. Ini berarti bahwa jika salah satu nilai tag atribut yang ditetapkan ke peran IAM atau kecocokan pengguna, role/user memperoleh izin akses pada sumber daya

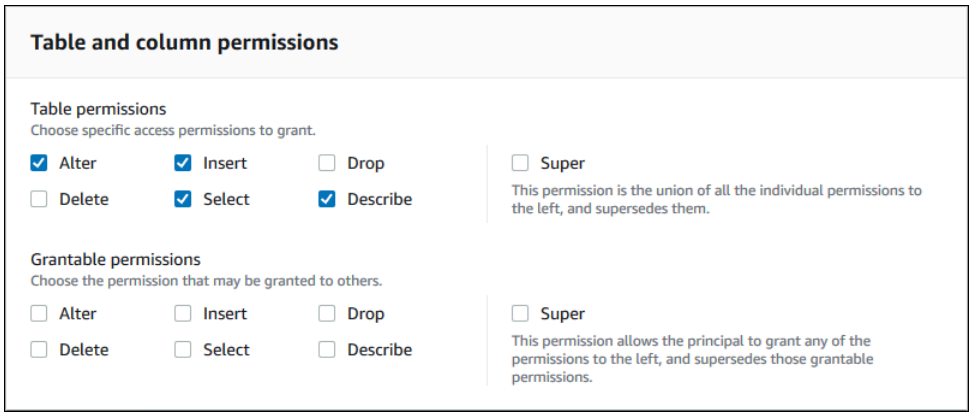
Pilih cakupan izin dengan menentukan apakah Anda memberikan izin kepada kepala sekolah dengan atribut yang cocok di akun yang sama atau di akun lain.

3. Di bagian sumber daya LF-Tags atau katalog, pilih database. Kemudian pilih satu atau lebih tabel, atau Semua tabel.

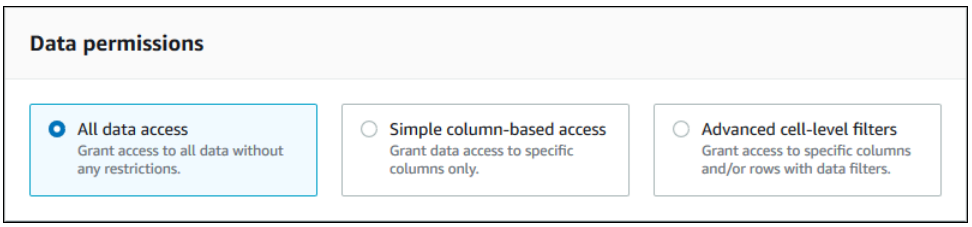


4. Tentukan izin tanpa pemfilteran data.

   Di bagian Izin, pilih izin tabel yang akan diberikan, dan secara opsional pilih izin yang dapat diberikan.



   Jika Anda memberikan Pilih, bagian Izin data muncul di bawah bagian izin Tabel dan kolom, dengan opsi Semua akses data dipilih secara default. Terima defaultnya.



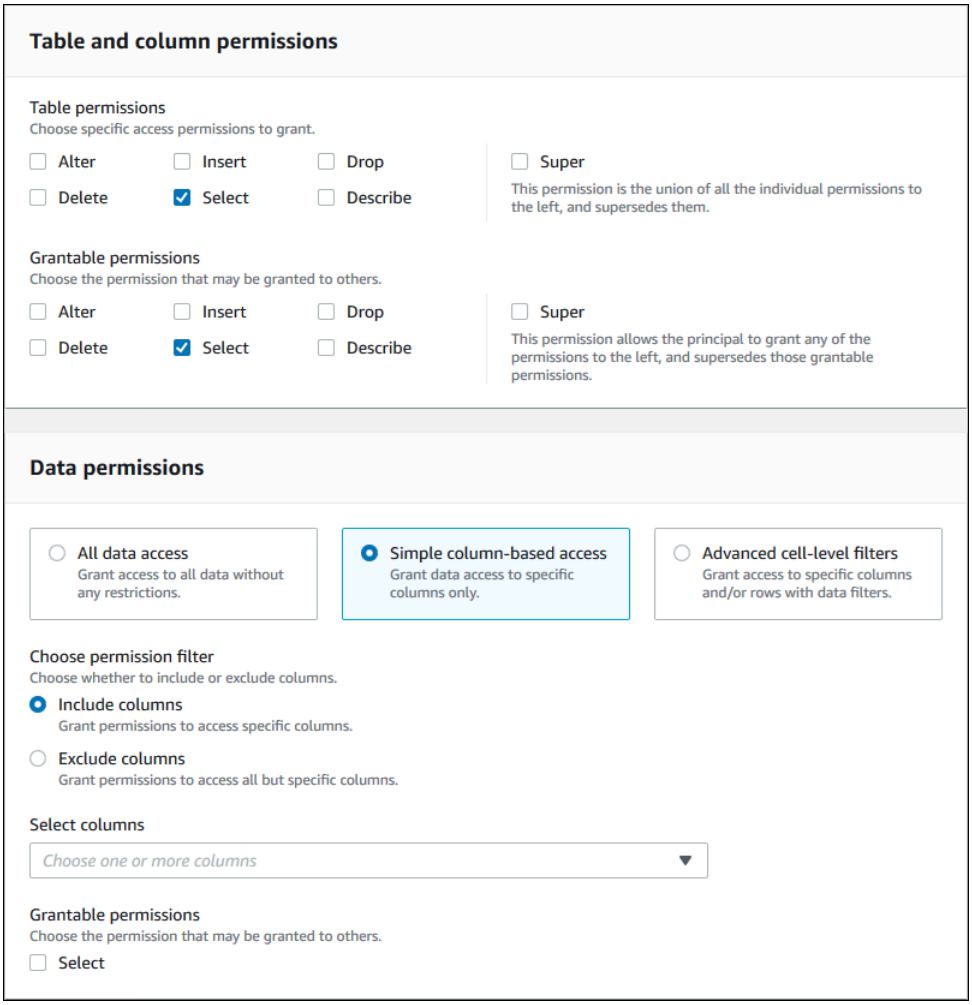
5. PilihIzin.
6. Tentukan Pilih izin dengan penyaringan data

   Pilih izin Pilih. Jangan pilih izin lainnya.

   Bagian Izin data muncul di bawah bagian Tabel dan kolom izin.
7. Lakukan salah satu tindakan berikut:

   - Terapkan penyaringan kolom sederhana saja.

     1. Pilih Akses berbasis kolom sederhana.



2. Pilih apakah akan menyertakan atau mengecualikan kolom, lalu pilih kolom yang akan disertakan atau dikecualikan.

   Hanya menyertakan daftar yang didukung saat memberikan izin ke AWS akun atau organisasi eksternal.

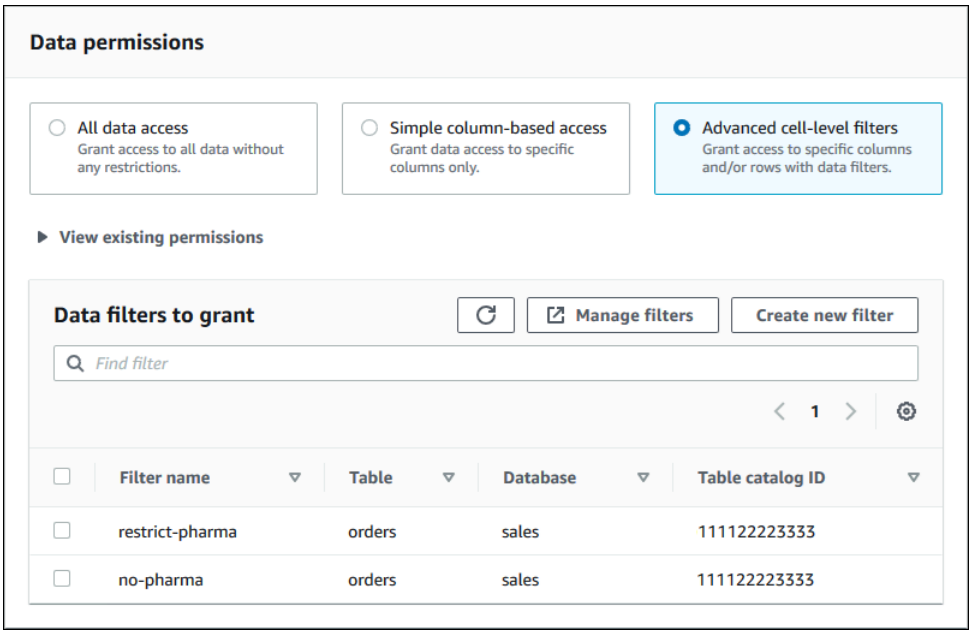
3. (Opsional) Di bawah Izin yang Dapat Diberikan, aktifkan opsi hibah untuk izin Pilih.

   Jika Anda menyertakan opsi hibah, penerima hibah hanya dapat memberikan izin pada kolom yang Anda berikan kepada mereka.

> **Note**
>
> Anda juga dapat menerapkan pemfilteran kolom hanya dengan membuat filter data yang menentukan filter kolom dan menentukan semua baris sebagai filter baris. Namun, ini membutuhkan lebih banyak langkah.

- Terapkan kolom, baris, atau penyaringan sel.

  1. Pilih Filter tingkat sel tingkat lanjut.



  2. (Opsional) Perluas Lihat izin yang ada.
  3. (Opsional) Pilih Buat filter baru.
  4. (Opsional) Untuk melihat detail filter yang tercantum, atau untuk membuat filter baru atau menghapus filter yang ada, pilih Kelola filter.

Halaman Filter data terbuka di jendela browser baru.

Setelah selesai di halaman Filter data, kembali ke halaman izin Hibah, dan jika perlu, segarkan halaman untuk melihat filter data baru yang Anda buat.

5. Pilih satu atau beberapa filter data untuk diterapkan pada hibah.

> **Note**
>
> Jika tidak ada filter data dalam daftar, itu berarti tidak ada filter data yang dibuat untuk tabel yang dipilih.

8. PilihIzin.

AWS CLI

Anda dapat memberikan izin tabel dengan menggunakan metode sumber daya bernama dan AWS Command Line Interface (AWS CLI).

Untuk memberikan izin tabel menggunakan AWS CLI

- Jalankan `grant-permissions` perintah, dan tentukan tabel sebagai sumber daya.

Example— Hibah pada satu meja - tidak ada penyaringan

Contoh berikut memberikan `SELECT` dan `ALTER` kepada pengguna `datalake_user1` di AWS akun 1111-2222-3333 pada tabel dalam 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**
>
> Jika Anda memberikan `ALTER` izin pada tabel yang memiliki data dasarnya di lokasi terdaftar, pastikan juga memberikan izin lokasi data pada lokasi tersebut kepada prinsipal. Untuk informasi selengkapnya, lihat Memberikan izin lokasi data.

Example— Hibah pada Semua Tabel dengan opsi Hibah - tanpa penyaringan

Contoh berikutnya memberikan opsi SELECT hibah pada semua tabel dalam databaseretail.

```
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- Hibah dengan penyaringan kolom sederhana

Contoh berikutnya ini memberikan SELECT subset kolom dalam tabel. persons Ini menggunakan penyaringan kolom sederhana.

```
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— Hibah dengan filter data

Contoh ini memberikan SELECT pada orders tabel dan menerapkan filter restrict-pharma data.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

Berikut ini adalah isi filegrant-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"]
}
```

Lihat juga

- Ikhtisar izin Lake Formation
- Pemfilteran data dan keamanan tingkat sel di Lake Formation
- Referensi personas Lake Formation dan izin IAM
- Memberikan izin tautan sumber daya
- Mengakses dan melihat tabel dan database Katalog Data bersama

Memberikan izin pada tampilan menggunakan metode sumber daya bernama

Langkah-langkah berikut menjelaskan cara memberikan izin pada tampilan dengan menggunakan metode sumber daya bernama dan halaman izin Hibah. Halaman ini dibagi menjadi beberapa bagian berikut:

- Jenis utama — Pengguna IAM, peran, pengguna dan grup Pusat Identitas IAM, organisasi Akun AWS, atau unit organisasi untuk memberikan izin. Anda juga dapat memberikan izin kepada prinsipal dengan atribut yang cocok.
- LF-Tags atau sumber daya katalog — Database, tabel, tampilan, atau tautan sumber daya untuk memberikan izin pada.
- Izin — Izin data lake untuk diberikan.

Buka Berikan izin halaman

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake, pembuat database, atau pengguna IAM yang memiliki izin Grantable pada database.
2. Lakukan salah satu tindakan berikut:
   - Di panel navigasi, di bawah Izin, pilih Izin data. Kemudian pilih Grant.
   - Di panel navigasi, pilih Tampilan di bawah Katalog Data. Kemudian, pada halaman Tampilan, pilih tampilan, dan dari menu Tindakan, di bawah Izin, pilih Hibah.

> **Note**
>
> Anda dapat memberikan izin pada tampilan melalui tautan sumber dayanya. Untuk melakukannya, pada halaman Tampilan, pilih tautan sumber daya, dan pada menu Tindakan, pilih Hibah sesuai target. Untuk informasi selengkapnya, lihat Cara kerja tautan sumber daya di Lake Formation.

Tentukan jenis utama

Di bagian Principal types, pilih Prinsipal atau Prinsipal berdasarkan atribut. Jika Anda memilih Prinsipal, opsi berikut tersedia:

Pengguna dan peran IAM

Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

Pusat Identitas IAM

Pilih satu atau beberapa pengguna atau grup dari daftar Pengguna dan grup.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAMB dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Amazon Quick. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah

> **Note**
>
> Integrasi Lake Formation dengan Quick hanya didukung untuk Quick Enterprise Edition.

Akun eksternal

Untuk Akun AWS, AWS organisasi, atau IAM Principal masukkan satu atau beberapa ID AWS akun, ID organisasi, ID unit organisasi, atau ARN yang valid untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10-32 huruf kecil atau digit.

ID unit organisasi dimulai dengan “ou-” diikuti oleh 4-32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

> **Lihat Juga**
>
> - Mengakses dan melihat tabel dan database Katalog Data bersama

Prinsipal berdasarkan atribut

Tentukan kunci atribut dan nilai (s). Jika Anda memilih lebih dari satu nilai, Anda membuat ekspresi atribut dengan operator OR. Ini berarti bahwa jika salah satu nilai tag atribut yang ditetapkan ke peran IAM atau kecocokan pengguna, role/user memperoleh izin akses pada sumber daya

Pilih cakupan izin dengan menentukan apakah Anda memberikan izin kepada kepala sekolah dengan atribut yang cocok di akun yang sama atau di akun lain.

Tentukan tampilan

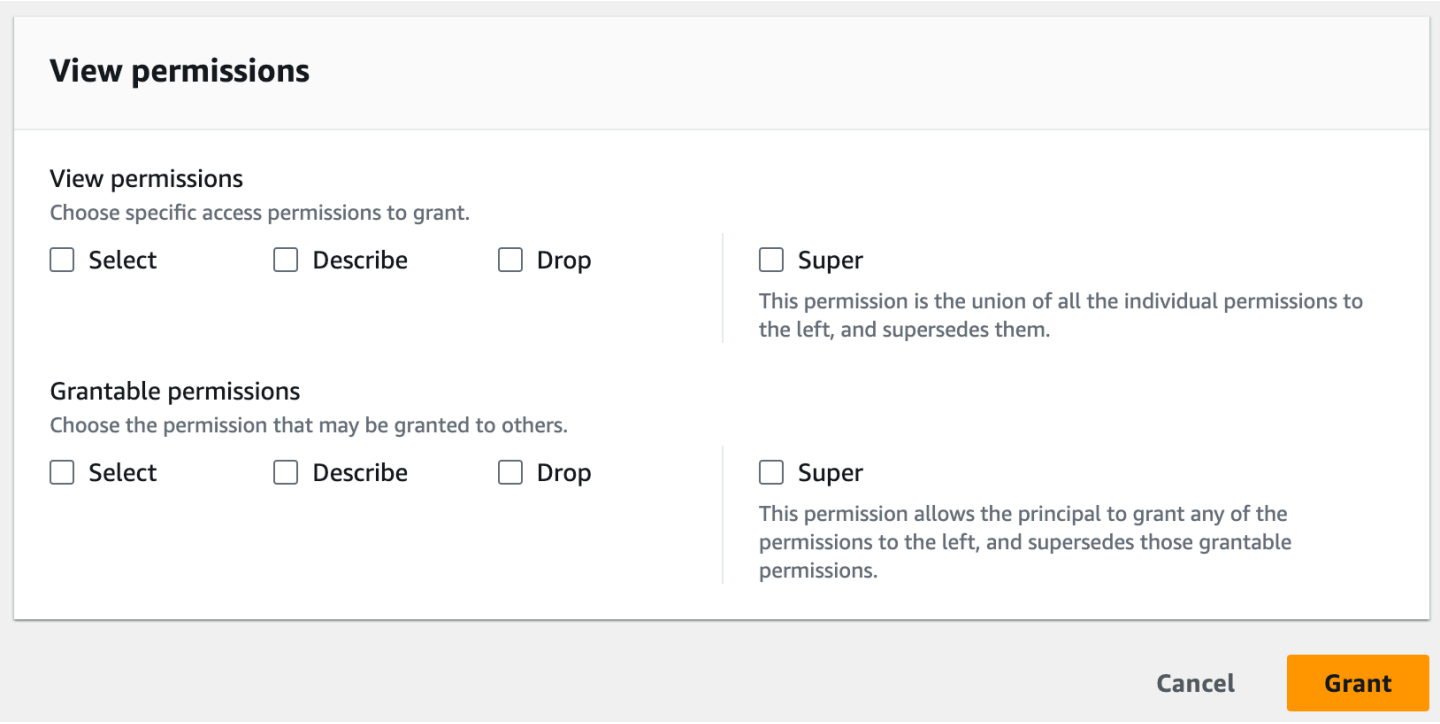
Di bagian sumber daya LF-Tags atau katalog, pilih satu atau beberapa tampilan untuk memberikan izin.

1. Pilih Sumber daya katalog data bernama.
2. Pilih satu atau beberapa tampilan dari daftar Tampilan. Anda juga dapat memilih satu atau lebih katalog, database, tabel, and/or filter data.

   Pemberian izin data lake ke `All tables` dalam database akan mengakibatkan penerima hibah memiliki izin pada semua tabel dan tampilan dalam database.

Tentukan izin

Di bagian Izin, pilih izin dan izin yang dapat diberikan.



1. Di bawah Izin tampilan, pilih satu atau beberapa izin yang akan diberikan.
2. (Opsional) Di bawah Izin yang Dapat Diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di dalamnya. Akun AWS Opsi ini tidak didukung jika Anda memberikan izin ke principal IAM dari akun eksternal.
3. PilihIzin.

Lihat Juga

- Referensi izin Lake Formation
- Memberikan izin pada database atau tabel yang dibagikan dengan akun Anda

Kontrol akses berbasis tag Lake Formation

Lake Formation berbasis tag access control (LF-TBAC) adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Dalam Lake Formation, atribut ini disebut LF-Tags. Anda dapat melampirkan LF-Tags ke sumber daya Katalog Data, dan memberikan izin ke kepala sekolah Lake Formation pada sumber daya tersebut menggunakan ini. LF-Tags Lake Formation memungkinkan operasi pada

sumber daya tersebut ketika prinsipal telah memberikan akses ke nilai tag yang cocok dengan nilai tag sumber daya.

LF-TBAC membantu dalam lingkungan yang berkembang pesat dan membantu dengan situasi di mana manajemen kebijakan menjadi rumit.

LF-TBAC adalah metode yang direkomendasikan untuk digunakan untuk memberikan izin Lake Formation ketika ada sejumlah besar objek Katalog Data termasuk katalog gabungan, database, tabel, dan tampilan. Lake Formation mendukung kontrol akses berbasis tag untuk katalog federasi tabel Amazon S3, gudang Amazon Redshift data, dan sumber data federasi seperti, SQL Server, dan Snowflake. Amazon DynamoDB

> **Note**
>
> Tag IAM tidak sama LF-Tags dengan. Tag ini tidak dapat dipertukarkan. LF-Tagsdigunakan untuk memberikan izin Lake Formation dan tag IAM digunakan untuk menentukan kebijakan IAM.

Cara kerja kontrol akses berbasis tag Lake Formation

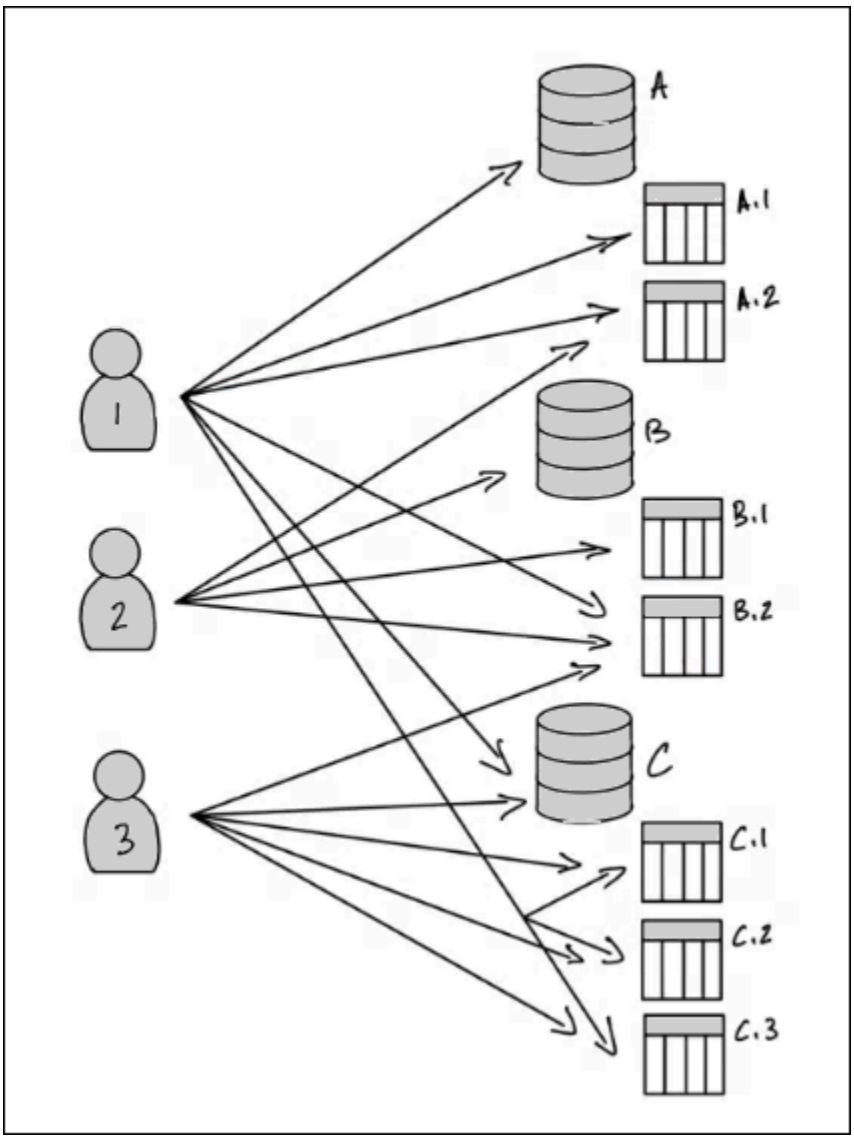
Masing-masing LF-Tag adalah pasangan kunci-nilai, seperti `department=sales` atau. `classification=restricted` Kunci dapat memiliki beberapa nilai yang ditentukan, sepertidepartment=sales,marketing,engineering,finance.

Untuk menggunakan LF-TBAC metode ini, administrator data lake dan insinyur data melakukan tugas-tugas berikut.

Tugas	Detail tugas
1. Tentukan sifat dan hubungan LF-Tags.	-
2. Buat LF-Tag pencipta di Lake Formation .	Menambahkan LF-Tag kreator
3. Buat LF-Tag Formasi Danau di.	Menciptakan LF-Tags
4. Tetapkan LF-Tags ke sumber daya Katalog Data.	Menetapkan sumber LF-Tags daya Katalog Data

Tugas	Detail tugas
5. Berikan izin kepada prinsipal lain untuk menetapkan sumber daya, secara opsional LF-Tags dengan opsi hibah.	Mengelola izin LF-Tag nilai
6. Berikan LF-Tag ekspresi kepada kepala sekolah, secara opsional dengan opsi hibah.	Memberikan izin data lake menggunakan metode LF-TBAC
7. (Disarankan) Setelah memverifikasi bahwa prinsipal memiliki akses ke sumber daya yang benar melalui LF-TBAC metode, cabut izin yang diberikan dengan menggunakan metode sumber daya bernama.	-

Pertimbangkan kasus di mana Anda harus memberikan izin kepada tiga kepala sekolah pada tiga database dan tujuh tabel.



Untuk mencapai izin yang ditunjukkan dalam diagram sebelumnya dengan menggunakan metode sumber daya bernama, Anda harus membuat 17 hibah, sebagai berikut (dalam kode semu).

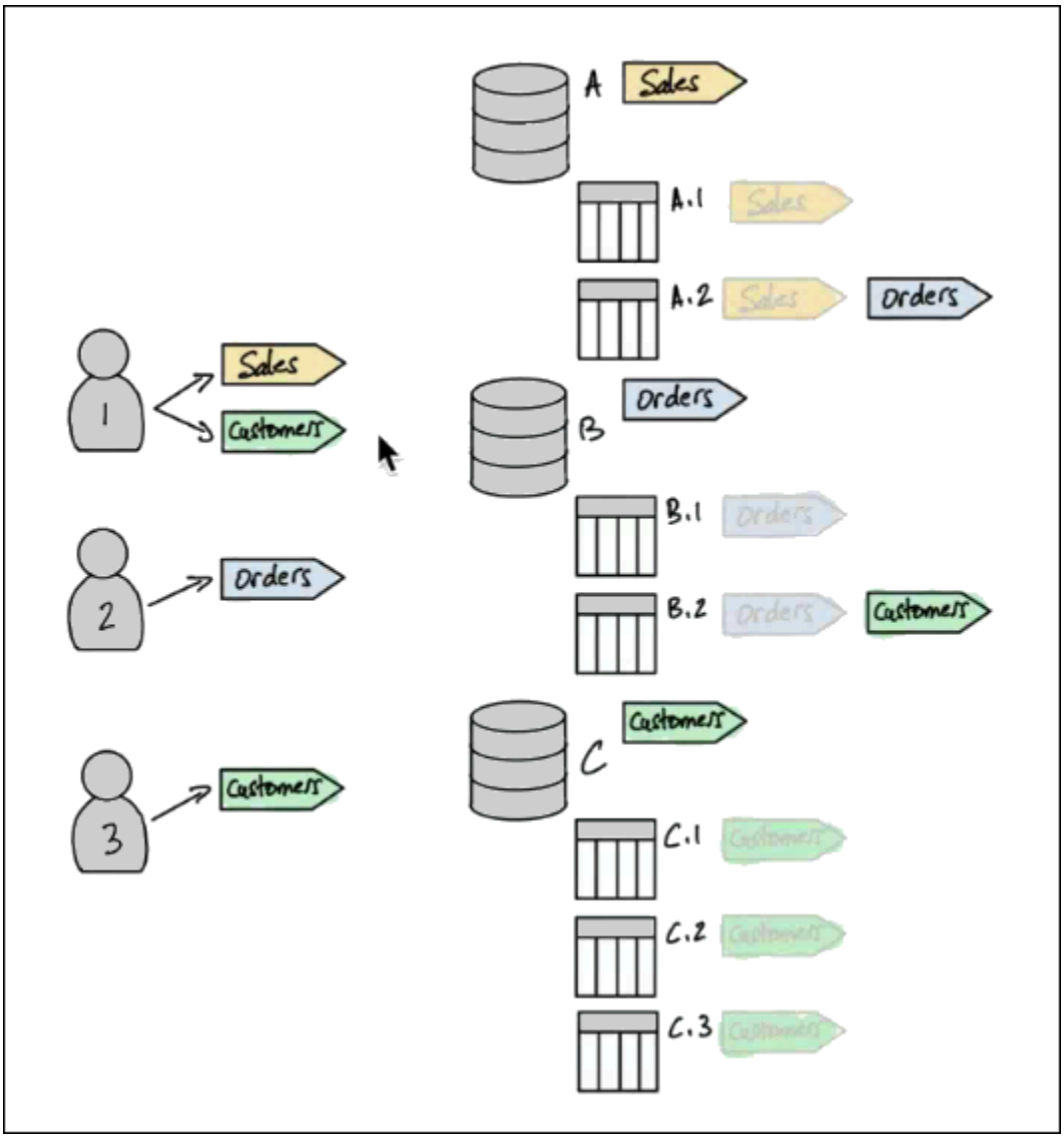
```
GRANT CREATE_TABLE ON Database A TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.1 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table B.2 TO PRINCIPAL 1
...
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 2
GRANT CREATE_TABLE ON Database B TO PRINCIPAL 2
...
GRANT SELECT, INSERT ON Table C.3 TO PRINCIPAL 3
```

Sekarang pertimbangkan bagaimana Anda akan memberikan izin dengan menggunakan LF-TBAC. Diagram berikut menunjukkan bahwa Anda telah ditetapkan LF-Tags ke database dan tabel, dan telah memberikan izin pada LF-Tags prinsipal.

Dalam contoh ini, LF-Tags mewakili area danau data yang berisi analitik untuk modul yang berbeda dari rangkaian aplikasi perencanaan sumber daya perusahaan (ERP). Anda dapat menggunakannya untuk mengontrol akses ke data analitik untuk berbagai modul. Semua LF-Tags memiliki kunci `module` dan nilai yang mungkin`Sales`,`Orders`, dan`Customers`. Contohnya LF-Tag terlihat seperti ini:

```
module=Sales
```

Diagram hanya menunjukkan LF-Tag nilai-nilai.



Menandai tugas untuk sumber daya Katalog Data dan warisan

Tabel mewarisi LF-Tags dari database dan kolom mewarisi LF-Tags dari tabel. Nilai yang diwariskan dapat diganti. Dalam diagram sebelumnya, redup LF-Tags diwariskan.

Karena pewarisan, administrator data lake hanya perlu membuat lima LF-Tag tugas berikut ke sumber daya (dalam kode semu).

```
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
```

Tag hibah untuk kepala sekolah

Setelah menetapkan LF-Tags ke database dan tabel, administrator data lake harus membuat hanya empat hibah LF-Tags untuk prinsipal, sebagai berikut (dalam pseudo-code).

```
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
```

Sekarang, prinsipal dengan `module=Sales` LF-Tag dapat mengakses sumber daya Katalog Data dengan `module=Sales` LF-Tag (misalnya, database A), prinsipal dengan `module=Customers` LF-Tag dapat mengakses sumber daya dengan `module=Customers` LF-Tag, dan sebagainya.

Perintah hibah sebelumnya tidak lengkap. Ini karena meskipun mereka menunjukkan melalui LF-Tags sumber daya Katalog Data bahwa prinsipal memiliki izin, mereka tidak menunjukkan dengan tepat izin Lake Formation mana (sepertiSELECT,ALTER) yang dimiliki prinsipal pada sumber daya tersebut. Oleh karena itu, perintah pseudo-code berikut adalah representasi yang lebih akurat tentang bagaimana izin Lake Formation diberikan pada sumber daya Katalog Data melalui. LF-Tags

```
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
```

Menyatukannya - Menghasilkan izin pada sumber daya

Mengingat LF-Tags ditugaskan ke database dan tabel dalam diagram sebelumnya, dan yang LF-Tags diberikan kepada kepala sekolah dalam diagram, tabel berikut mencantumkan izin Lake Formation yang dimiliki kepala sekolah pada database dan tabel.

Principal	Izin Diberikan Melalui LF-Tags
Prinsipal 1	• `CREATE_TABLE` pada database A • `SELECT`, `INSERT` di atas meja A.1 • `SELECT`, `INSERT` di atas meja B.2 • `CREATE_TABLE` pada database C • `SELECT`, `INSERT` di atas meja C.1 • `SELECT`, `INSERT` di atas meja C.2 • `SELECT`, `INSERT` di atas meja C.3
Kepala Sekolah 2	• `SELECT`, `INSERT` di atas meja A.2 • `CREATE_TABLE` pada database B • `SELECT`, `INSERT` di atas meja B.1
Kepala Sekolah 3	• `SELECT`, `INSERT` di atas meja B.2 • `CREATE_TABLE` pada database C • `SELECT`, `INSERT` di atas meja C.1 • `SELECT`, `INSERT` di atas meja C.2 • `SELECT`, `INSERT` di atas meja C.3

Intinya

Dalam contoh sederhana ini, menggunakan lima operasi penugasan dan delapan operasi hibah, administrator data lake dapat menentukan 17 izin. Ketika ada puluhan database dan ratusan tabel, keuntungan LF-TBAC metode dibandingkan metode sumber daya bernama menjadi jelas. Dalam kasus hipotetis kebutuhan untuk memberikan setiap akses utama ke setiap sumber daya, dan di mana `n(P)` jumlah prinsipal dan `n(R)` jumlah sumber daya:

- Dengan metode sumber daya bernama, jumlah hibah yang diperlukan adalah `n(P) ×n(R)`.
- Dengan LF-TBAC metode ini, menggunakan satu LF-Tag, total jumlah hibah kepada kepala sekolah dan penugasan ke sumber daya adalah +. `n(P) n(R)`

> Lihat juga
>
> - Mengelola LF-Tags kontrol akses metadata
> - Memberikan izin data lake menggunakan metode LF-TBAC

Topik

- Mengelola LF-Tags kontrol akses metadata
- Mengelola LF-Tag ekspresi untuk kontrol akses metadata
- Mengelola izin LF-Tag nilai

Mengelola LF-Tags kontrol akses metadata

Untuk menggunakan metode kontrol akses (LF-TBAC) berbasis tag Lake Formation untuk mengamankan objek Katalog Data seperti katalog, database, tabel, tampilan, dan kolom, Anda membuat, menetapkannya ke sumber daya LF-Tags, dan memberikan izin kepada prinsipal. LF-Tag

Sebelum Anda dapat menetapkan LF-Tags objek Data Catalog atau memberikan izin kepada prinsipal, Anda perlu menentukan. LF-Tags Hanya administrator data lake atau prinsipal dengan izin LF-Tag pembuat yang dapat membuat LF-Tags.

LF-Tag pencipta

LF-Tag creator adalah prinsipal non-admin yang memiliki izin untuk membuat dan mengelola. LF-Tags Administrator data lake dapat menambahkan LF-Tag kreator menggunakan konsol Lake Formation atau CLI. LF-Tag pembuat konten memiliki izin Lake Formation implisit untuk memperbarui, dan menghapus LF-Tags, untuk menetapkan sumber daya, dan LF-Tags untuk memberikan LF-Tag izin dan izin LF-Tag nilai kepada prinsipal lain.

Dengan peran LF-Tag pembuat, administrator data lake dapat mendelegasikan tugas manajemen tag seperti membuat dan memperbarui kunci dan nilai tag ke prinsipal non-admin. Administrator data lake juga dapat memberikan izin yang dapat diberikan kepada LF-Tag pembuat konten. `Create LF-Tag` Kemudian, LF-Tag pencipta dapat memberikan izin untuk membuat LF-Tags ke kepala sekolah lain.

Anda dapat memberikan dua jenis izin pada LF-Tags:

- LF-Tag izin -`Create LF-Tag`,`Alter`, dan`Drop`. Izin ini diperlukan untuk membuat, memperbarui, dan menghapus LF-Tags.

  Administrator dan LF-Tag pembuat data lake secara implisit memiliki izin ini pada yang LF-Tags mereka buat dan dapat memberikan izin ini secara eksplisit kepada kepala sekolah untuk mengelola tag di danau data.
- LF-Tag izin pasangan nilai kunci -`Assign`,, `Describe` dan. `Grant with LF-Tag expressions` Izin ini diperlukan untuk menetapkan LF-Tags objek Katalog Data, dan untuk memberikan izin pada sumber daya kepada kepala sekolah menggunakan kontrol akses berbasis tag Lake Formation. LF-Tag kreator secara implisit menerima izin ini saat membuat. LF-Tags

Setelah menerima `Create LF-Tag` izin dan berhasil membuat LF-Tags, LF-Tag pembuat dapat menetapkan LF-Tags sumber daya dan memberikan LF-Tag izin (`Create LF-Tag`,,, `Alter`, `Drop`, dan) kepada prinsipal non-administratif lainnya untuk mengelola tag di data lake. Anda dapat mengelola LF-Tags dengan menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

> **Note**
>
> Administrator data lake memiliki izin Lake Formation implisit untuk membuat, memperbarui, dan menghapus LF-Tags, menetapkan LF-Tags sumber daya, dan memberikan izin LF-Tag kepada kepala sekolah.

Untuk praktik dan pertimbangan terbaik, lihat Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation

Topik

- Menambahkan LF-Tag kreator
- Menciptakan LF-Tags
- Memperbarui LF-Tags
- Menghapus LF-Tags
- Daftar LF-Tags
- Menetapkan sumber LF-Tags daya Katalog Data

- Melihat LF-Tags ditugaskan ke sumber daya
- Melihat sumber daya LF-Tag yang ditugaskan
- Siklus hidup a LF-Tag
- Perbandingan kontrol akses berbasis tag Lake Formation dengan kontrol akses berbasis atribut IAM

Lihat juga

- Mengelola izin LF-Tag nilai
- Memberikan izin data lake menggunakan metode LF-TBAC
- Kontrol akses berbasis tag Lake Formation

Menambahkan LF-Tag kreator

Secara default, administrator data lake dapat membuat, memperbarui, dan menghapus LF-Tags, menetapkan tag ke objek Katalog Data, dan memberikan izin tag ke prinsipal. Jika Anda ingin mendelegasikan operasi pembuatan dan pengelolaan tag ke kepala sekolah non-admin, administrator data lake dapat membuat peran LF-Tag pembuat dan memberikan izin Lake Formation ke peran tersebut. `Create LF-Tag` Dengan `Create LF-Tag` izin yang dapat diberikan, LF-Tag pembuat konten dapat mendelegasikan tugas pembuatan dan pemeliharaan tag ke kepala sekolah non-administratif lainnya.

Agar administrator data lake dapat menetapkan LF-Tags sumber daya Katalog Data, mereka diharuskan memberikan izin asosiasi pada sumber daya LF-Tags yang tidak dibuat oleh mereka.

Note

Cross-account hibah izin hanya dapat mencakup `Describe` dan `Associate` izin. Anda tidak dapat memberikan`Create LF-Tag`,`Drop`,`Alter`, dan `Grant with LFTag expressions` izin kepada kepala sekolah di akun lain.

Topik

- Izin IAM diperlukan untuk membuat LF-Tags
- Tambahkan LF-Tag kreator

Lihat juga

- Mengelola izin LF-Tag nilai
- Memberikan izin data lake menggunakan metode LF-TBAC
- Kontrol akses berbasis tag Lake Formation

Izin IAM diperlukan untuk membuat LF-Tags

Anda harus mengonfigurasi izin untuk mengizinkan prinsipal Lake Formation dibuat LF-Tags. Tambahkan pernyataan berikut ke kebijakan izin untuk prinsipal yang perlu menjadi LF-Tag pembuat.

Note

Meskipun administrator data lake memiliki izin Lake Formation implisit untuk membuat, memperbarui, dan menghapus, menetapkan sumber daya LF-Tags, dan untuk memberikan LF-Tags LF-Tags kepada kepala sekolah, administrator danau data juga memerlukan izin IAM berikut.

Untuk informasi selengkapnya, lihat Referensi personas Lake Formation dan izin 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"
     ]
 }
```

Kepala sekolah yang menetapkan sumber daya dan memberikan LF-Tags LF-Tags kepada kepala sekolah harus memiliki izin yang sama, kecuali untuk,, dan izin. `CreateLFTag` `UpdateLFTag` `DeleteLFTag`

Tambahkan LF-Tag kreator

LF-Tag Pembuat dapat membuat LF-Tag, memperbarui kunci dan nilai tag, menghapus tag, mengaitkan tag ke sumber daya Katalog Data, dan memberikan izin pada sumber daya Katalog Data kepada prinsipal menggunakan metode. LF-TBAC LF-Tag Pembuat juga dapat memberikan izin ini kepada kepala sekolah.

Anda dapat membuat peran LF-Tag pembuat dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

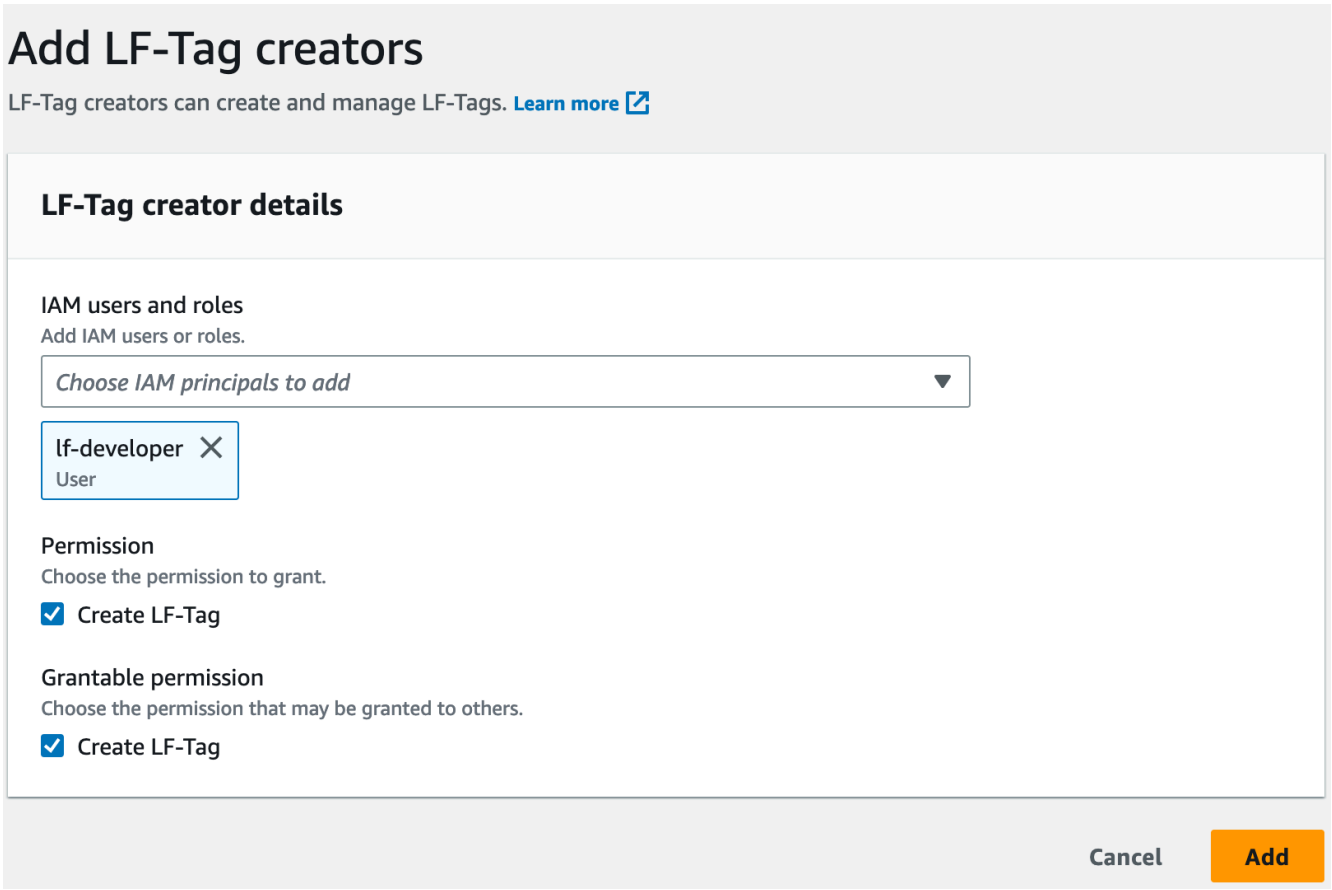
console

Untuk menambahkan LF-Tag kreator

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator datalake.
2. Di panel navigasi, di bawah Izin, pilih LF-Tags dan izin.

   Di halaman LF-Tags dan izin, pilih bagian LF-Tag pembuat dan pilih Tambahkan LF-Tag pembuat konten.



3. Pada halaman Tambah LF-Tag pembuat, pilih peran IAM atau pengguna yang memiliki izin yang diperlukan untuk membuat. LF-Tags
4. Aktifkan kotak centang `Create LF-Tag` izin.
5. (Opsional) Untuk mengaktifkan prinsipal yang dipilih untuk memberikan `Create LF-Tag` izin kepada kepala sekolah, pilih Izin yang dapat diberikan. `Create LF-Tag`
6. Pilih Tambahkan.

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"
    ]
}
```

Berikut ini adalah izin yang tersedia untuk peran LF-Tag pembuat:

Izin	Deskripsi
`Drop`	Seorang kepala sekolah dengan izin ini pada LF-Tag dapat menghapus a LF-Tag dari danau data. Prinsipal mendapat `Describe` izin implisit pada semua nilai tag LF-Tag sumber daya.
`Alter`	Prinsipal dengan izin ini pada LF-Tag dapat menambah atau menghapus nilai tag dari a LF-Tag. Prinsipal mendapat `Alter` izin implisit pada semua nilai tag dari a LF-Tag.
`Describe`	Prinsipal dengan izin ini pada a LF-Tag dapat melihat LF-Tag dan nilainya saat mereka menetapkan LF-Tags ke sumber daya atau memberikan izin pada. LF-Tags Anda dapat memberikan `Describe` pada semua nilai kunci atau pada nilai tertentu.
`Associate`	Seorang kepala sekolah dengan izin ini pada LF-Tag dapat menetapkan LF-Tag ke sumber daya Katalog Data. Memberikan hibah `Associate` implisit. `Describe`
`Grant with LF-Tag expression`	Prinsipal dengan izin ini pada a LF-Tag dapat memberikan izin pada sumber daya Katalog Data menggunakan LF-Tag kunci dan nilai. Memberikan hibah `Grant with LF-Tag expression` implisit. `Describe`

Izin ini dapat diberikan. Seorang kepala sekolah yang telah diberikan izin ini dengan opsi hibah dapat memberikannya kepada prinsipal lain.

Menciptakan LF-Tags

Semua LF-Tags harus didefinisikan dalam Lake Formation sebelum dapat digunakan. A LF-Tag terdiri dari kunci dan satu atau lebih nilai yang mungkin untuk kunci.

Setelah administrator data lake menyiapkan izin IAM yang diperlukan dan izin Lake Formation untuk peran LF-Tag pembuat, prinsipal dapat membuat file. LF-Tag LF-Tag Pembuat mendapatkan izin implisit untuk memperbarui atau menghapus nilai tag apa pun dari LF-Tag dan menghapus. LF-Tag

Anda dapat membuat LF-Tags dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

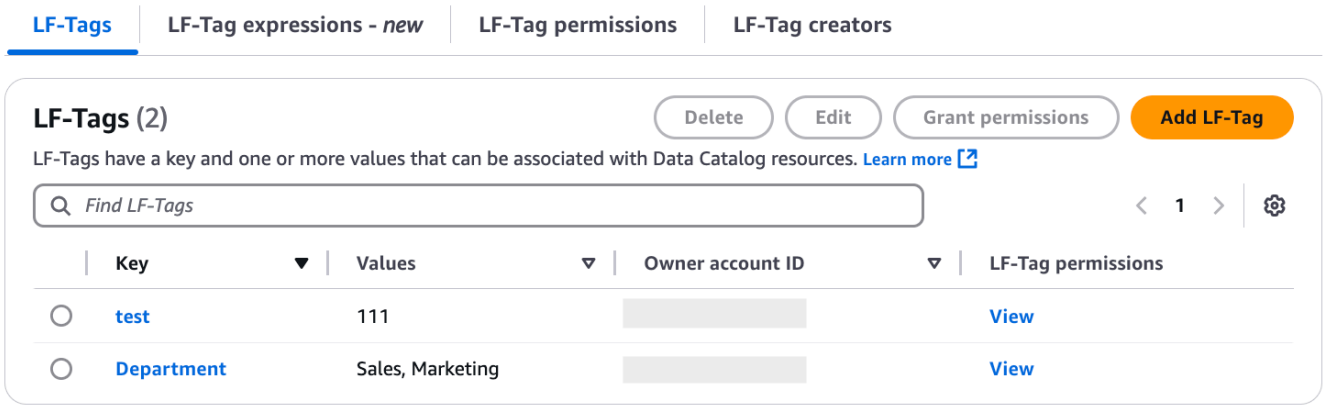
Console

Untuk membuat LF-Tag

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai prinsipal dengan izin LF-Tag pembuat atau sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, LF-Tags dan izin, pilih. LF-Tags

   Halaman LF-Tags akan muncul.



3. Pilih Tambahkan LF-Tag.
4. Dalam kotak LF-Tag dialog Tambah, masukkan kunci dan satu atau lebih nilai.

   Setiap kunci harus memiliki setidaknya satu nilai. Untuk memasukkan beberapa nilai, masukkan daftar yang dibatasi koma lalu tekan Enter, atau masukkan satu nilai pada satu waktu dan pilih Tambah setelah masing-masing. Jumlah maksimum nilai yang diizinkan adalah 1000.
5. Pilih Tambahkan tanda.

AWS CLI

Untuk membuat LF-Tag

- Masukkan `create-lf-tag` perintah.

  Contoh berikut membuat LF-Tag dengan kunci `module` dan nilai-nilai `Customers` dan`Orders`.

```
aws lakeformation create-lf-tag --tag-key module --tag-values Customers Orders
```

Sebagai pembuat tag, kepala sekolah mendapat `Alter` izin untuk ini LF-Tag dan dapat memperbarui atau menghapus nilai tag apa pun dari ini LF-Tag. Prinsipal LF-Tag pembuat juga dapat memberikan `Alter` izin kepada prinsipal lain untuk memperbarui dan menghapus nilai tag pada ini LF-Tag.

Memperbarui LF-Tags

Anda memperbarui LF-Tag yang Anda memiliki `Alter` izin dengan menambahkan atau menghapus nilai kunci yang diizinkan. Anda tidak dapat mengubah LF-Tag kuncinya. Untuk mengubah kunci, hapus LF-Tag dan tambahkan satu dengan kunci yang diperlukan. Selain `Alter` izin, Anda juga memerlukan izin `lakeformation:UpdateLFTag` IAM untuk memperbarui nilai.

Saat Anda menghapus LF-Tag nilai, tidak ada pemeriksaan yang dilakukan untuk keberadaan LF-Tag nilai tersebut pada sumber daya Katalog Data apa pun. Jika LF-Tag nilai yang dihapus dikaitkan dengan sumber daya, nilai tersebut tidak lagi terlihat untuk sumber daya, dan prinsip apa pun yang diberikan izin pada pasangan nilai kunci tersebut tidak lagi memiliki izin.

Sebelum menghapus LF-Tag nilai, Anda dapat secara opsional menggunakan `remove-lf-tags-from-resource`perintah perintah untuk menghapus sumber daya Katalog Data LF-Tag dari yang memiliki nilai yang ingin Anda hapus, lalu tag ulang sumber daya dengan nilai yang ingin Anda simpan.

Hanya administrator data lake, LF-Tag pembuat, dan kepala sekolah yang memiliki `Alter` izin di dalamnya yang dapat memperbarui file. LF-Tag LF-Tag

Anda dapat memperbarui LF-Tag dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk memperbarui LF-Tag (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, LF-Tag pembuat atau kepala sekolah dengan `Alter` izin pada file LF-Tag.
2. Di panel navigasi, di bawah Izin, LF-Tags dan izin, pilih. LF-Tags
3. Pada LF-Tagshalaman, pilih LF-Tag, lalu pilih Edit.
4. Dalam kotak LF-Tag dialog Edit, tambahkan atau hapus LF-Tag nilai.

   Untuk menambahkan beberapa nilai, di bidang Nilai, masukkan daftar yang dibatasi koma dan tekan Enter, atau masukkan satu nilai pada satu waktu atau pilih Tambah setelah masing-masing.
5. Pilih Simpan.

AWS CLI

Untuk memperbarui LF-Tag (AWS CLI)

- Masukkan `update-lf-tag` perintah. Berikan salah satu atau kedua argumen berikut:
  - `--tag-values-to-add`
  - `--tag-values-to-delete`

Example

Contoh berikut menggantikan nilai `vp` dengan nilai `vice-president` untuk LF-Tag kunci`level`.

```
aws lakeformation update-lf-tag --tag-key level --tag-values-to-add vice-president
--tag-values-to-delete vp
```

Menghapus LF-Tags

Anda dapat menghapus LF-Tags yang tidak lagi digunakan. Tidak ada pemeriksaan yang dilakukan untuk keberadaan LF-Tag sumber daya Katalog Data. Jika dihapus LF-Tag dikaitkan dengan sumber

daya, itu tidak lagi terlihat untuk sumber daya, dan prinsip apa pun yang diberikan izin yang LF-Tag tidak lagi memiliki izin.

Sebelum menghapus LF-Tag, Anda dapat secara opsional menggunakan `remove-lf-tags-from-resource`perintah untuk menghapus LF-Tag dari semua sumber daya.

Hanya administrator data lake, LF-Tag pencipta, atau kepala sekolah yang memiliki `Drop` izin LF-Tag dapat menghapus file. LF-Tag Selain `Drop` izin, kepala sekolah juga memerlukan izin `lakeformation:DeleteLFTag` IAM untuk menghapus file. LF-Tag

Anda dapat menghapus LF-Tag dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk menghapus LF-Tag (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, LF-Tags dan izin, pilih. LF-Tags
3. Pada LF-Tagshalaman, pilih LF-Tag, lalu pilih Hapus.
4. Di lingkungan Delete tag? kotak dialog, untuk mengkonfirmasi penghapusan, masukkan nilai LF-Tag kunci di bidang yang ditunjuk dan kemudian pilih Hapus.

AWS CLI

Untuk menghapus LF-Tag (AWS CLI)

- Masukkan `delete-lf-tag` perintah. Berikan kunci LF-Tag untuk menghapus.

  Example

  Contoh berikut menghapus LF-Tag dengan kunci`region`.

  ```
  aws lakeformation delete-lf-tag --tag-key region
  ```

Daftar LF-Tags

Anda dapat membuat daftar tempat Anda memiliki `Describe` atau `Associate` izin. LF-Tags Nilai yang tercantum dengan setiap LF-Tag kunci adalah nilai yang Anda memiliki izin.

LF-Tag pencipta memiliki izin implisit untuk melihat yang telah LF-Tags mereka buat.

Administrator data lake dapat melihat semua LF-Tags yang didefinisikan dalam AWS akun lokal dan semua LF-Tags yang `Associate` izinnya `Describe` telah diberikan ke akun lokal dari akun eksternal. Administrator data lake dapat melihat semua nilai untuk semua LF-Tags.

Anda dapat membuat daftar LF-Tags dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

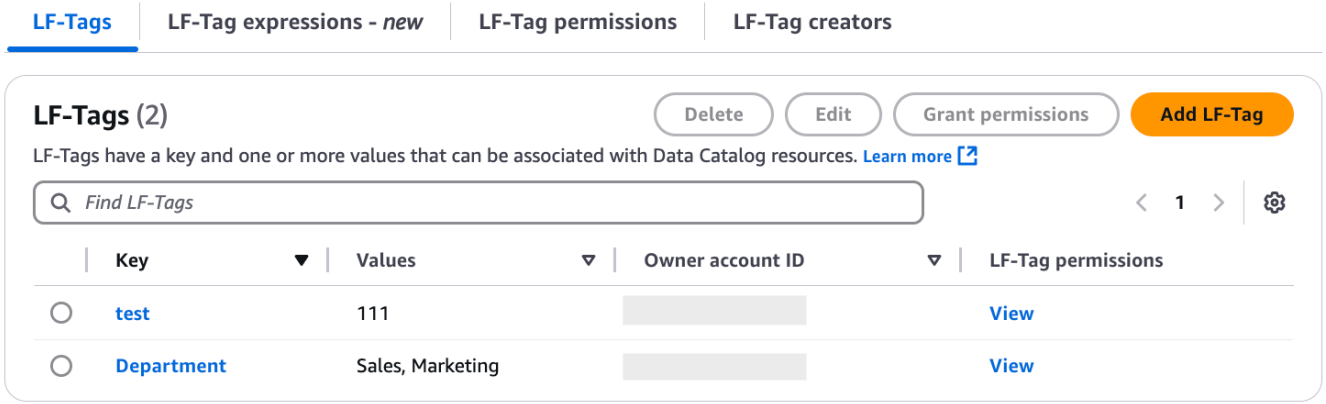
Ke daftar LF-Tags (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai LF-Tag pembuat, sebagai administrator data lake, atau sebagai prinsipal yang telah diberikan izin LF-Tags dan yang memiliki izin `lakeformation:ListLFTags` IAM.

2. Di panel navigasi, di bawah Izin, LF-Tags dan izin, pilih. LF-Tags

   Halaman LF-Tags akan muncul.



   Periksa kolom ID akun Pemilik untuk menentukan kolom LF-Tags yang dibagikan dengan akun Anda dari akun eksternal.

AWS CLI

Untuk daftar LF-Tags (AWS CLI)

- Jalankan perintah berikut sebagai administrator data lake atau sebagai prinsipal yang telah diberikan izin LF-Tags dan yang memiliki izin `lakeformation:ListLFTags` IAM.

```
aws lakeformation list-lf-tags
```

Output Anda serupa dengan yang berikut ini.

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "Orders",
                "Sales",
                "Customers"
            ]
        }
    ]
}
```

Untuk juga melihat LF-Tags bahwa diberikan dari akun eksternal, sertakan opsi perintah`--resource-share-type ALL`.

```
aws lakeformation list-lf-tags --resource-share-type ALL
```

Output Anda serupa dengan yang berikut ini. Perhatikan `NextToken` kuncinya, yang menunjukkan bahwa ada lebih banyak daftar.

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "Orders",
                "Sales",
                "Customers"
            ]
        }
    ],
    "NextToken": "eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ=="
}
```

Ulangi perintah, dan tambahkan `--next-token` argumen untuk melihat lokal yang tersisa LF-Tags dan LF-Tags yang diberikan dari akun eksternal. LF-Tags dari akun eksternal selalu di halaman terpisah.

```
aws lakeformation list-lf-tags --resource-share-type ALL
--next-token eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ==
```

```
{
    "LFTags": [
        {
            "CatalogId": "123456789012",
            "TagKey": "region",
            "TagValues": [
                "central",
                "south"
            ]
        }
```

```
    ]
}
```

API

Anda dapat menggunakan SDK yang tersedia untuk Lake Formation untuk mencantumkan tag yang diizinkan oleh pemohon untuk dilihat.

```
import boto3

client = boto3.client('lakeformation')
...

response = client.list_lf_tags(
    CatalogId='string',
    ResourceShareType='ALL',
    MaxResults=50'
)
```

Perintah ini mengembalikan `dict` objek dengan struktur berikut:

```
{
    'LFTags': [
        {
            'CatalogId': 'string',
            'TagKey': 'string',
            'TagValues': [
                'string',
            ]
        },
    ],
    'NextToken': 'string'
}
```

Untuk informasi lebih lanjut tentang izin yang diperlukan, lihat Referensi personas Lake Formation dan izin IAM.

Menetapkan sumber LF-Tags daya Katalog Data

Anda dapat menetapkan LF-Tags sumber daya Katalog Data (database, tabel, dan kolom) untuk mengontrol akses ke sumber daya tersebut. Hanya prinsipal yang diberikan pencocokan LF-Tags (dan prinsipal yang diberikan akses dengan metode sumber daya bernama) yang dapat mengakses sumber daya.

Jika tabel mewarisi LF-Tag dari database atau kolom mewarisi LF-Tag dari tabel, Anda dapat mengganti nilai yang diwarisi dengan menetapkan nilai baru ke kunci. LF-Tag

Jumlah maksimum LF-Tags yang dapat Anda tetapkan ke sumber daya adalah 50.

Topik

- Persyaratan untuk mengelola tag yang ditetapkan ke sumber daya
- Tetapkan LF-Tags ke kolom tabel
- Menetapkan LF-Tags ke sumber daya Katalog Data
- Memperbarui LF-Tags sumber daya
- Menghapus LF-Tag dari sumber daya

Persyaratan untuk mengelola tag yang ditetapkan ke sumber daya

Untuk menetapkan sumber daya Katalog Data, Anda harus: LF-Tag

- Miliki `ASSOCIATE` izin Lake Formation di LF-Tag.
- Memiliki `lakeformation:AddLFTagsToResource` izin IAM.
- Memiliki lem: GetDatabase izin pada database Glue.
- Jadilah pemilik sumber daya (pembuat), dapatkan izin `Super` Lake Formation pada sumber daya dengan `GRANT` opsi, atau miliki izin berikut dengan `GRANT` opsi:
  - Untuk database di AWS akun yang sama:`DESCRIBE`,, `CREATE_TABLE``ALTER`, dan `DROP`
  - Untuk database di akun eksternal:`DESCRIBE`, dan `CREATE_TABLE` `ALTER`
  - Untuk tabel (dan kolom):`DESCRIBE`,`ALTER`,`DROP`,`INSERT`,`SELECT`, dan `DELETE`

Selain itu, LF-Tag dan sumber daya yang ditugaskan harus berada di AWS akun yang sama.

Untuk menghapus LF-Tag dari sumber daya Katalog Data, Anda harus memenuhi persyaratan ini, dan juga memiliki izin `lakeformation:RemoveLFTagsFromResource` IAM.

Tetapkan LF-Tags ke kolom tabel

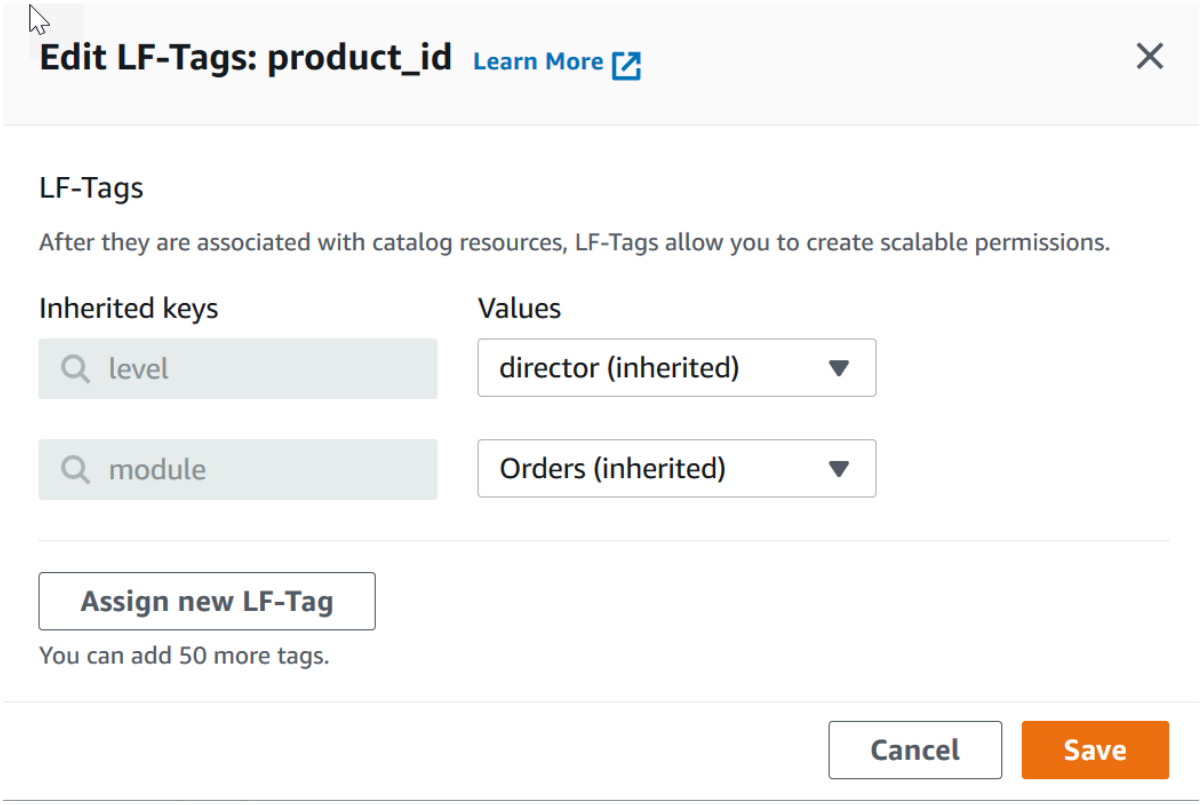
Untuk menetapkan LF-Tags ke kolom tabel (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

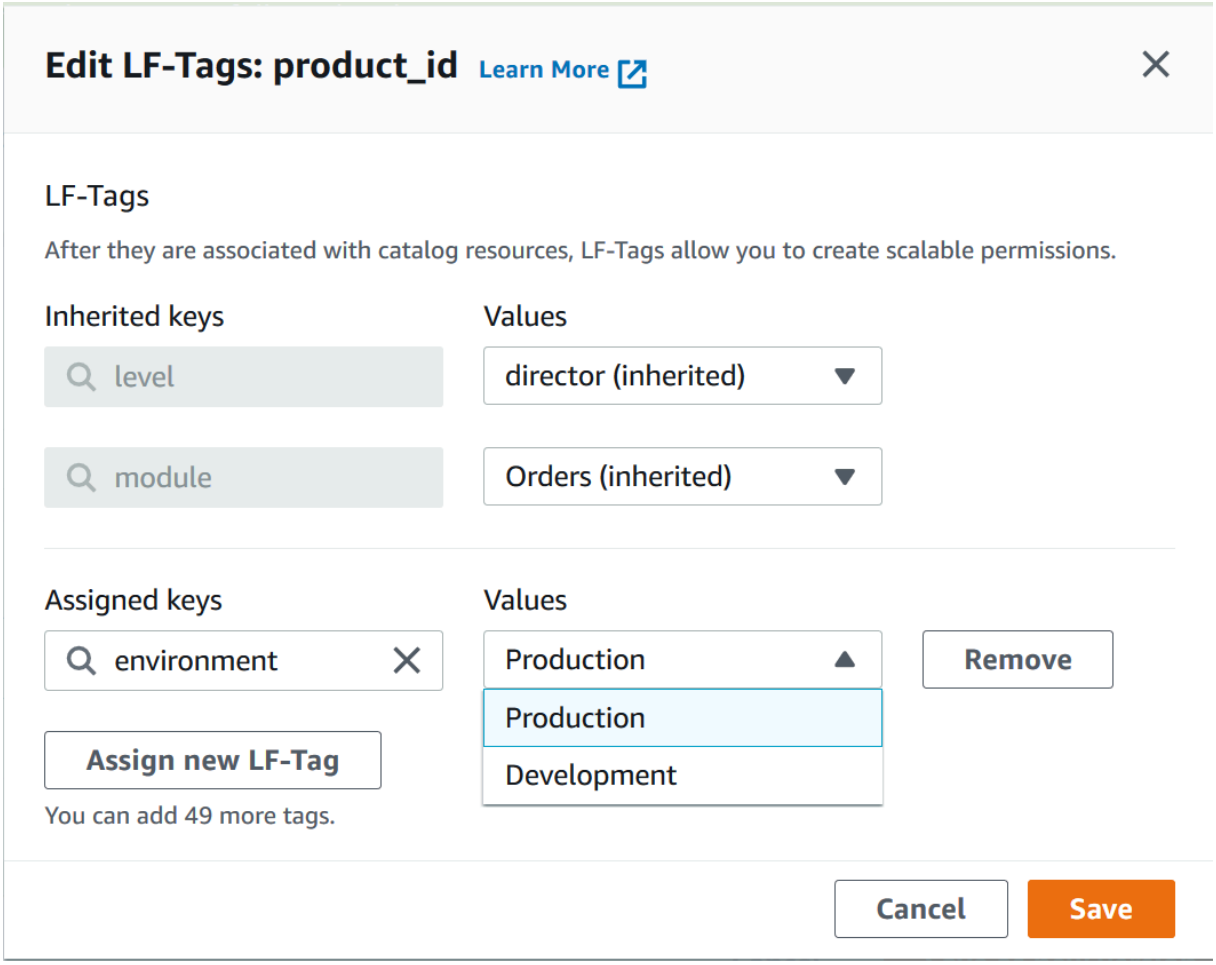
   Masuk sebagai pengguna yang memenuhi persyaratan yang tercantum di atas.
2. Di panel navigasi, pilih Tabel.
3. Pilih nama tabel (bukan tombol opsi di sebelah nama tabel).
4. Pada halaman detail tabel, di bagian Skema, pilih Edit skema.
5. Pada halaman Edit skema, pilih satu atau beberapa kolom, lalu pilih Edit LF-Tags.

> **Note**
>
> Jika Anda bermaksud menambah atau menghapus kolom dan menyimpan versi baru, lakukan itu terlebih dahulu. Kemudian edit LF-Tags.

Kotak LF-Tags dialog Edit muncul, dan menampilkan apa pun LF-Tags yang diwarisi dari tabel.



6. (Opsional) Untuk daftar Nilai di samping bidang kunci yang diwariskan, pilih nilai untuk mengganti nilai yang diwariskan.
7. (Opsional) Pilih Tetapkan baru LF-Tag. Kemudian untuk kunci yang Ditugaskan, pilih kunci, dan untuk Nilai, pilih nilai untuk kunci tersebut.



8. (Opsional) Pilih Tetapkan baru LF-Tag lagi untuk menambahkan yang lain LF-Tag.
9. Pilih Simpan.

Menetapkan LF-Tags ke sumber daya Katalog Data

Console

Untuk menetapkan LF-Tags ke database atau tabel Katalog Data

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai pengguna yang memenuhi persyaratan yang tercantum sebelumnya.

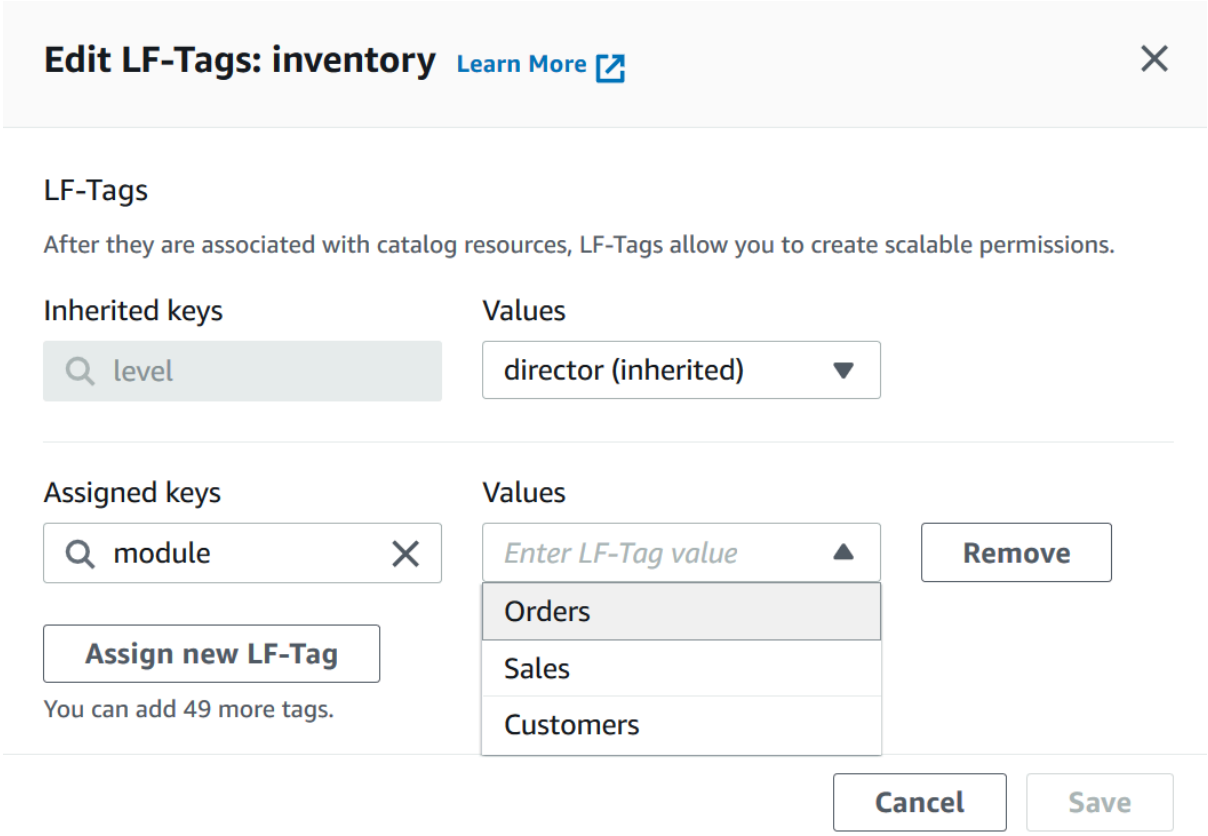
2. Di panel navigasi, di bawah Katalog data, lakukan salah satu hal berikut:

   - Untuk menetapkan LF-Tags ke database, pilih Database.
   - Untuk menetapkan LF-Tags tabel, pilih Tabel.

3. Pilih database atau tabel, dan pada menu Tindakan, pilih Edit LF-Tags.

   Kotak *resource-name* dialog Edit LF-Tags: muncul.

   Jika tabel mewarisi LF-Tags dari database yang berisi, jendela menampilkan yang diwariskan LF-Tags. Jika tidak, ini akan menampilkan teks “Tidak ada warisan yang LF-Tags terkait dengan sumber daya.”



4. (Opsional) Jika tabel telah diwarisi LF-Tags, untuk daftar Nilai di samping bidang kunci yang diwarisi, Anda dapat memilih nilai untuk mengganti nilai yang diwariskan.
5. Untuk menetapkan baru LF-Tags, lakukan langkah-langkah ini:

   a. Pilih Tetapkan baru LF-Tag.

   b. Di bidang Kunci yang ditugaskan, pilih LF-Tag kunci, dan di bidang Nilai, pilih nilai.

   c. (Opsional) Pilih Tetapkan baru LF-Tag lagi untuk menetapkan tambahan. LF-Tag

6. Pilih Simpan.

AWS CLI

Untuk menetapkan sumber LF-Tags daya Katalog Data

- Jalankan perintah `add-lf-tags-to-resource`.

  Contoh berikut menetapkan LF-Tag `module=orders` ke tabel `orders` dalam database`erp`. Ini menggunakan sintaks pintasan untuk argumen. `--lf-tags` `CatalogID`Properti untuk `--lf-tags` adalah opsional. Jika tidak disediakan, ID katalog sumber daya (dalam hal ini, tabel) diasumsikan.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":
 {"DatabaseName":"erp", "Name":"orders"}}' --lf-tags
 CatalogId=111122223333,TagKey=module,TagValues=orders
```

  Berikut ini adalah output jika perintah berhasil.

```
{
    "Failures": []
}
```

  Contoh berikutnya ini menetapkan dua LF-Tags ke `sales` tabel, dan menggunakan sintaks JSON untuk argumen. `--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"]}]'
```

  Contoh berikutnya ini menetapkan LF-Tag `level=director` ke `total` kolom tabel`sales`.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "TableWithColumns":
 {"DatabaseName":"erp", "Name":"sales", "ColumnNames":["total"]}}' --lf-tags
 TagKey=level,TagValues=director
```

Memperbarui LF-Tags sumber daya

Untuk memperbarui sumber LF-Tag daya Katalog Data (AWS CLI)

- Gunakan `add-lf-tags-to-resource` perintah, seperti yang dijelaskan dalam prosedur sebelumnya.

  Menambahkan LF-Tag dengan kunci yang sama dengan yang ada LF-Tag, tetapi dengan nilai yang berbeda memperbarui nilai yang ada.

Menghapus LF-Tag dari sumber daya

Untuk menghapus sumber LF-Tag daya Katalog Data (AWS CLI)

- Jalankan perintah `remove-lf-tags-from-resource`.

  Jika tabel memiliki LF-Tag nilai yang mengesampingkan nilai yang diwarisi dari database induk, menghapusnya LF-Tag dari tabel mengembalikan nilai yang diwariskan. Perilaku ini juga berlaku untuk kolom yang mengesampingkan nilai kunci yang diwarisi dari tabel.

  Contoh berikut menghapus LF-Tag `level=director` dari `total` kolom `sales` tabel. `CatalogID`Properti untuk `--lf-tags` adalah opsional. Jika tidak disediakan, ID katalog sumber daya (dalam hal ini, tabel) diasumsikan.

```
aws lakeformation remove-lf-tags-from-resource
--resource ' { "TableWithColumns":
{ "DatabaseName": "erp",  "Name": "sales",  "ColumnNames":[ "total"]}}'
--lf-tags  CatalogId=111122223333,TagKey=level,TagValues=director
```

Melihat LF-Tags ditugaskan ke sumber daya

Anda dapat melihat LF-Tags yang ditetapkan ke sumber daya Katalog Data. Anda harus memiliki `DESCRIBE` atau `ASSOCIATE` izin LF-Tag untuk melihatnya.

Console

Untuk melihat LF-Tags yang ditetapkan ke sumber daya (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

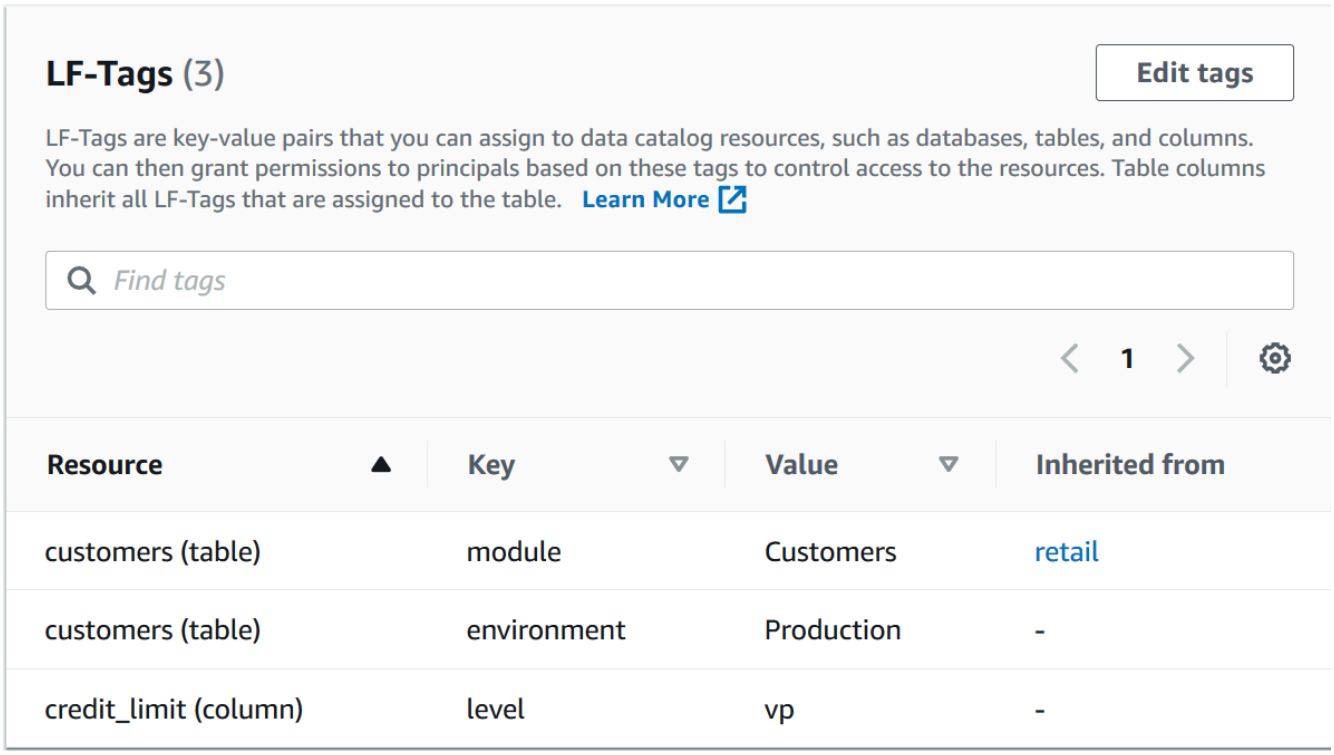
Masuk sebagai administrator data lake, pemilik sumber daya, atau pengguna yang telah diberikan izin Lake Formation pada sumber daya.

2. Di panel navigasi, di bawah judul Katalog data, lakukan salah satu hal berikut:

   - Untuk melihat LF-Tags ditetapkan ke database, pilih Database.
   - Untuk melihat LF-Tags ditetapkan ke tabel, pilih Tabel.

3. Pada halaman Tabel atau Database, pilih nama database atau tabel. Kemudian pada halaman detail, gulir ke bawah ke LF-Tagsbagian.

   Screenshot berikut menunjukkan LF-Tags ditugaskan ke `customers` tabel, yang terkandung dalam `retail` database. `module` LF-Tag Itu diwarisi dari database. `credit_limit`Kolom memiliki yang `level=vp` LF-Tag ditetapkan.



AWS CLI

Untuk melihat LF-Tags yang ditugaskan ke sumber daya (AWS CLI)

- Masukkan perintah yang serupa dengan yang berikut ini.

```
aws lakeformation get-resource-lf-tags --show-assigned-lf-tags --
resource '{ "Table": {"CatalogId":"111122223333", "DatabaseName":"erp",
 "Name":"sales"}}'
```

Perintah mengembalikan output berikut.

```
{
    "TableTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "sales"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "environment",
            "TagValues": [
                "development"
            ]
        }
    ],
    "ColumnTags": [
        {
            "Name": "total",
            "Tags": [
                {
                    "CatalogId": "111122223333",
                    "TagKey": "level",
                    "TagValues": [
                        "director"
                    ]
                }
            ]
        }
    ]
}
```

Output ini hanya menunjukkan LF-Tags yang secara eksplisit ditetapkan, tidak diwariskan. Jika Anda ingin melihat semua LF-Tags di semua kolom, termasuk yang diwariskan LF-Tags, hilangkan opsi. `--show-assigned-lf-tags`

Melihat sumber daya LF-Tag yang ditugaskan

Anda dapat melihat semua sumber daya Katalog Data tempat LF-Tag kunci tertentu ditetapkan. Untuk melakukannya, Anda memerlukan izin Lake Formation berikut:

- `Describe`atau `Associate` di LF-Tag.
- `Describe`atau izin Lake Formation lainnya pada sumber daya.

Selain itu, Anda memerlukan izin AWS Identity and Access Management (IAM) berikut:

- `lakeformation:SearchDatabasesByLFTags`
- `lakeformation:SearchTablesByLFTags`

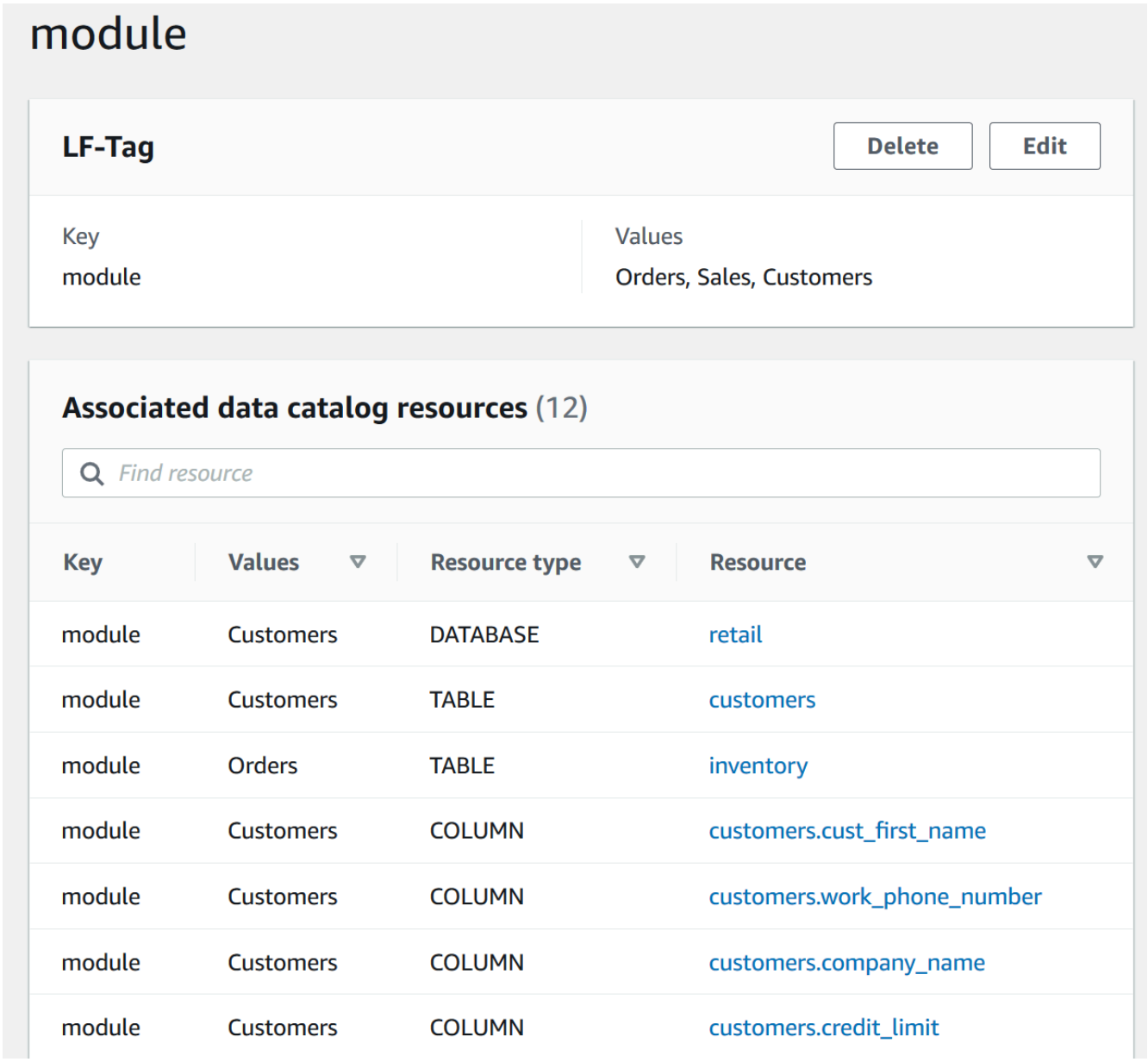
Console

Untuk melihat sumber daya yang LF-Tag ditetapkan (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake atau sebagai pengguna yang memenuhi persyaratan yang tercantum sebelumnya.
2. Di panel navigasi, di bawah Izin dan LF-Tags dan izin, pilih. LF-Tags
3. Pilih LF-Tag kunci (bukan tombol opsi di sebelah nama kunci).

   Halaman LF-Tag detail menampilkan daftar sumber daya yang LF-Tag telah ditetapkan.



AWS CLI

Untuk melihat sumber daya LF-Tag yang ditugaskan

- Jalankan `search-databases-by-lf-tags` perintah `search-tables-by-lf-tags` atau.

Example

Contoh berikut mencantumkan tabel dan kolom yang telah `level=vp` LF-Tag ditetapkan. Untuk setiap tabel dan kolom yang terdaftar, semua yang ditetapkan LF-Tags untuk tabel atau kolom adalah output, bukan hanya ekspresi pencarian.

```
aws lakeformation search-tables-by-lf-tags --expression
 TagKey=level,TagValues=vp
```

Untuk informasi lebih lanjut tentang izin yang diperlukan, lihat Referensi personas Lake Formation dan izin IAM.

Siklus hidup a LF-Tag

1. LF-Tag Pencipta Michael menciptakan a LF-Tag`module=Customers`.
2. Michael memberikannya `Associate` LF-Tag kepada insinyur data Eduardo. Memberikan hibah `Associate` implisit. `Describe`
3. Michael memberikan `Super` di atas meja `Custs` kepada Eduardo dengan opsi hibah, sehingga Eduardo dapat menetapkan LF-Tags ke meja. Untuk informasi selengkapnya, lihat Menetapkan sumber LF-Tags daya Katalog Data.
4. Eduardo menugaskan LF-Tag `module=customers` ke meja. `Custs`
5. Michael memberikan hibah berikut kepada insinyur data Sandra (dalam kode semu).

```
GRANT (SELECT, INSERT ON TABLES) ON TAGS module=customers TO Sandra WITH GRANT OPTION
```

6. Sandra memberikan hibah berikut kepada analis data Maria.

```
GRANT (SELECT ON TABLES) ON TAGS module=customers TO Maria
```

Maria sekarang dapat menjalankan kueri di atas `Custs` meja.

Lihat juga

- Kontrol akses metadata

Perbandingan kontrol akses berbasis tag Lake Formation dengan kontrol akses berbasis atribut IAM

Attribute-based Access Control (ABAC) adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Dalam AWS, atribut ini disebut tag. Anda dapat melampirkan tag ke sumber daya IAM, termasuk entitas IAM (pengguna atau peran) dan sumber daya. AWS Anda dapat membuat kebijakan ABAC tunggal atau set kecil kebijakan untuk penanggung jawab IAM Anda. Kebijakan ABAC ini dapat dirancang untuk memungkinkan operasi ketika tag penanggung jawab cocok dengan tag sumber daya. ABAC sangat membantu di lingkungan yang berkembang dengan cepat dan membantu dalam situasi ketika manajemen kebijakan menjadi rumit.

Tim keamanan dan tata kelola cloud menggunakan IAM untuk menentukan kebijakan akses dan izin keamanan untuk semua sumber daya termasuk bucket Amazon S3, instans Amazon EC2, dan sumber daya apa pun yang dapat Anda referensikan dengan ARN. Kebijakan IAM menentukan izin luas (berbutir kasar) ke sumber daya data lake Anda, misalnya, untuk mengizinkan atau menolak akses di bucket Amazon S3 atau tingkat awalan atau tingkat basis data. Untuk informasi lebih lanjut tentang IAM ABAC, lihat Untuk apa ABAC? AWS di Panduan Pengguna IAM.

Misalnya, Anda dapat membuat tiga peran dengan kunci tanda `project-access`. Mengatur nilai tanda dari peran pertama ke `Dev`, yang kedua ke `Marketing`, dan yang ketiga ke `Support`. Tetapkan tag dengan nilai yang sesuai untuk sumber daya. Anda kemudian dapat menggunakan kebijakan tunggal yang memungkinkan akses ketika peran dan sumber daya ditandai dengan nilai yang sama untuk `project-access`.

Tim tata kelola data menggunakan Lake Formation untuk menentukan izin berbutir halus ke sumber daya data lake tertentu. LF-Tags ditugaskan ke sumber daya Katalog Data (database, tabel, dan kolom) dan diberikan kepada prinsipal. Seorang kepala sekolah dengan LF-Tags LF-Tags kecocokan sumber daya dapat mengakses sumber daya itu. Izin Lake Formation adalah sekunder dari izin IAM. Misalnya, jika izin IAM tidak mengizinkan pengguna mengakses data lake, Lake Formation tidak memberikan akses ke sumber daya apa pun di dalam data lake tersebut kepada pengguna tersebut, meskipun prinsipal dan sumber daya memiliki kecocokan. LF-Tags

Kontrol akses berbasis tag Lake Formation (LF-TBAC) bekerja dengan IAM ABAC untuk memberikan tingkat izin tambahan untuk data dan sumber daya Lake Formation Anda.

- Skala izin Lake Formation TBAC dengan inovasi. Administrator idak perlu lagi memperbarui kebijakan yang ada untuk memungkinkan akses ke sumber daya baru. Misalnya, asumsikan bahwa Anda menggunakan strategi IAM ABAC dengan `project-access` tag untuk menyediakan akses ke database tertentu dalam Lake Formation. Menggunakan LF-TBAC, LF-Tag `Project=SuperApp` ditetapkan ke tabel atau kolom tertentu, dan hal yang sama LF-Tag

diberikan kepada pengembang untuk proyek itu. Melalui IAM, pengembang dapat mengakses database, dan LF-TBAC izin memberikan pengembang akses lebih lanjut ke tabel atau kolom tertentu dalam tabel. Jika tabel baru ditambahkan ke proyek, administrator Lake Formation hanya perlu menetapkan tag ke tabel baru agar pengembang diberi akses ke tabel.

- Lake Formation TBAC membutuhkan lebih sedikit kebijakan IAM. Karena Anda menggunakan kebijakan IAM untuk memberikan akses tingkat tinggi ke sumber daya Lake Formation dan Lake Formation TBAC untuk mengelola akses data yang lebih tepat, Anda membuat lebih sedikit kebijakan IAM.
- Menggunakan Lake Formation TBAC, tim dapat berubah dan tumbuh dengan cepat. Ini karena izin untuk sumber daya baru secara otomatis diberikan berdasarkan atribut. Misalnya, jika pengembang baru bergabung dengan proyek, mudah untuk memberikan akses pengembang ini dengan mengaitkan peran IAM ke pengguna dan kemudian menetapkan yang diperlukan LF-Tags untuk pengguna. Anda tidak perlu mengubah kebijakan IAM untuk mendukung proyek baru atau membuat yang baru LF-Tags.
- Finer-grained izin dimungkinkan menggunakan Lake Formation TBAC. Kebijakan IAM memberikan akses ke sumber daya tingkat atas, seperti database atau tabel Katalog Data. Menggunakan Lake Formation TBAC, Anda dapat memberikan akses ke tabel atau kolom tertentu yang berisi nilai data tertentu.

> **Note**
>
> Tag IAM tidak sama LF-Tags dengan. Tag ini tidak dapat dipertukarkan. LF-Tags digunakan untuk memberikan izin Lake Formation dan tag IAM digunakan untuk menentukan kebijakan IAM.

Mengelola LF-Tag ekspresi untuk kontrol akses metadata

LF-Tag ekspresi adalah ekspresi logis yang terdiri dari satu atau lebih LF-Tags (pasangan kunci-nilai) yang digunakan untuk memberikan izin pada sumber daya. AWS Glue Data Catalog LF-Tag ekspresi memungkinkan Anda untuk menentukan aturan yang mengatur akses ke sumber daya data Anda berdasarkan tag metadata mereka. Anda dapat menyimpan ekspresi ini dan menggunakannya kembali di beberapa hibah izin, memastikan konsistensi dan membuatnya mudah untuk mengelola perubahan pada ontologi tag dari waktu ke waktu.

Dalam LF-Tag ekspresi yang diberikan, kunci tag digabungkan menggunakan operasi AND, sedangkan nilai digabungkan menggunakan operasi OR. Misalnya, ekspresi tag `content_type:Sales AND location:US` mewakili sumber daya yang terkait dengan data penjualan di AS.

Anda dapat membuat hingga 1000 LF-Tag ekspresi dalam file Akun AWS. Ekspresi ini menyediakan cara yang fleksibel dan terukur untuk mengelola izin berdasarkan tag metadata, memastikan bahwa hanya pengguna atau aplikasi yang berwenang yang dapat mengakses sumber daya data tertentu berdasarkan aturan tag yang ditentukan.

LF-Tag ekspresi menawarkan manfaat berikut:

- Reusability — Dengan mendefinisikan dan menyimpan LF-Tag ekspresi, Anda tidak perlu lagi mereplikasi ekspresi yang sama secara manual saat menetapkan izin ke sumber daya atau prinsipal lain.
- Konsistensi — Menggunakan kembali LF-Tag ekspresi di beberapa hibah izin memastikan konsistensi dalam cara izin diberikan dan dikelola.
- Manajemen ontologi tag — LF-Tag ekspresi membantu mengelola perubahan pada ontologi tag dari waktu ke waktu, karena Anda dapat memperbarui ekspresi yang disimpan alih-alih memodifikasi hibah izin individual.

Untuk informasi lebih lanjut tentang kontrol akses berbasis tag, silakan merujuk ke. Kontrol akses berbasis tag Lake Formation

LF-Tag pencipta ekspresi

LF-Tag Pembuat ekspresi adalah kepala sekolah yang memiliki izin untuk membuat dan mengelola LF-Tag ekspresi. Administrator data lake dapat menambahkan pembuat LF-Tag ekspresi menggunakan konsol Lake Formation, CLI, API, atau SDK. LF-Tag pembuat ekspresi memiliki izin Lake Formation implisit untuk membuat, memperbarui, dan menghapus LF-Tag ekspresi, dan untuk memberikan izin LF-Tag ekspresi ke prinsipal lain.

LF-Tag pembuat ekspresi yang bukan administrator danau data menerima `Grant with LF-Tag expression` izin implisit`Alter`,, `Drop``Describe`, dan hanya untuk ekspresi yang mereka buat.

Administrator data lake juga dapat memberikan izin yang dapat diberikan kepada pembuat LF-Tag ekspresi. `Create LF-Tag expression` Kemudian, pembuat LF-Tag ekspresi dapat memberikan izin untuk membuat LF-Tag ekspresi ke prinsipal lain.

Topik

- Izin IAM diperlukan untuk membuat ekspresi LF-Tag
- Tambahkan pembuat LF-Tag ekspresi
- Membuat LF-Tag ekspresi
- Memperbarui LF-Tag ekspresi
- Menghapus ekspresi LF-Tag
- LF-Tag Ekspresi daftar

Lihat juga

- Mengelola izin LF-Tag nilai
- Memberikan izin data lake menggunakan metode LF-TBAC
- Kontrol akses berbasis tag Lake Formation

Izin IAM diperlukan untuk membuat ekspresi LF-Tag

Anda harus mengonfigurasi izin untuk mengizinkan prinsipal Lake Formation membuat LF-Tag ekspresi. Tambahkan pernyataan berikut ke kebijakan izin untuk prinsipal yang perlu menjadi pembuat LF-Tag ekspresi.

Note

Meskipun administrator data lake memiliki izin Lake Formation implisit untuk membuat, memperbarui, dan menghapus LF-Tags dan LF-Tag ekspresi, untuk menetapkan sumber daya, dan LF-Tags untuk memberikan LF-Tags dan izin LF-Tag ekspresi kepada kepala sekolah, administrator data lake juga memerlukan izin IAM berikut.

Untuk informasi selengkapnya, lihat Referensi personas Lake Formation dan izin 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"
    ]
}
```

Tambahkan pembuat LF-Tag ekspresi

LF-Tag pembuat ekspresi dapat membuat dan menyimpan LF-Tag ekspresi yang dapat digunakan kembali, memperbarui kunci dan nilai tag, menghapus ekspresi, dan memberikan izin pada sumber daya Katalog Data kepada prinsipal menggunakan metode. LF-TBAC Pembuat LF-Tag ekspresi juga dapat memberikan izin ini kepada kepala sekolah.

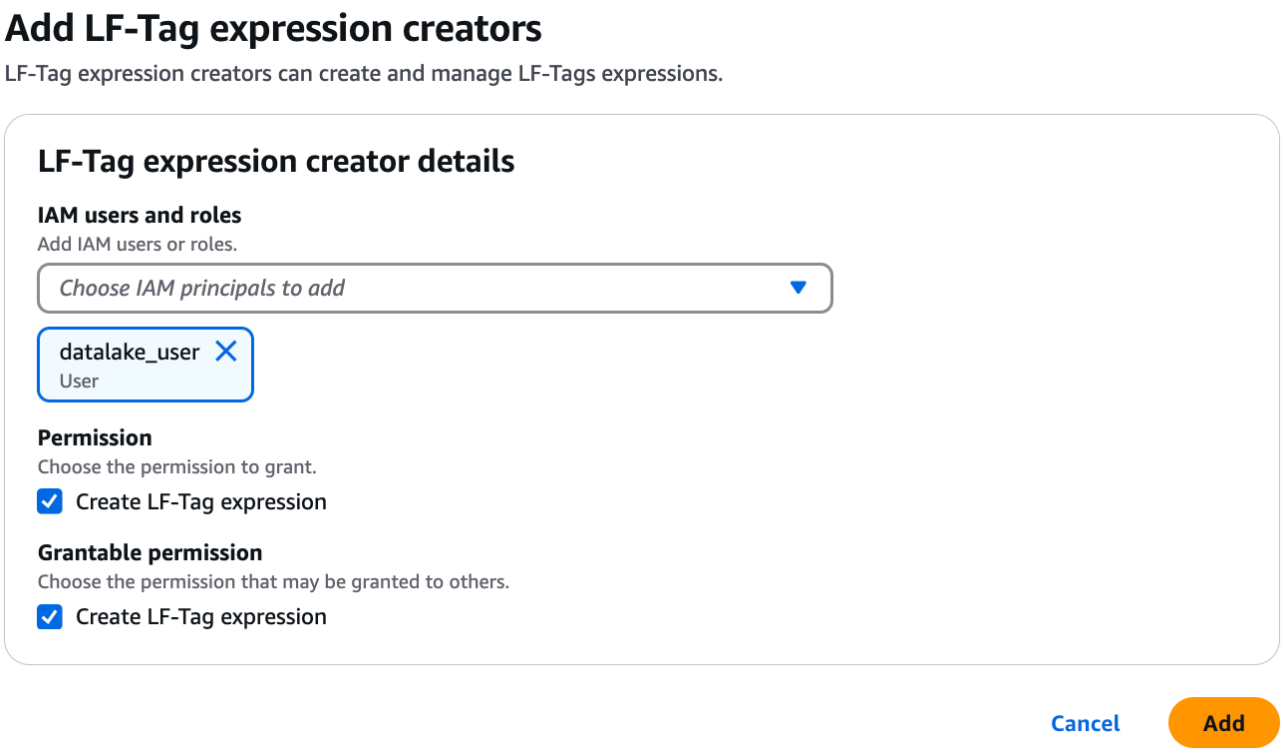
Anda dapat membuat peran pembuat LF-Tag ekspresi menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

console

Untuk menambahkan pembuat LF-Tag ekspresi

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, pilih LF-Tags dan izin.
3. Pilih tab LF-Tag ekspresi.
4. Di bagian pembuat LF-Tag ekspresi, pilih Tambahkan pembuat LF-Tag ekspresi.



5. Pada halaman Tambahkan pembuat LF-Tag ekspresi, pilih peran IAM atau pengguna yang memiliki izin yang diperlukan untuk membuat LF-Tag ekspresi.
6. Pilih kotak centang `Create LF-Tag expression` izin.
7. (Opsional) Untuk mengaktifkan prinsipal yang dipilih untuk memberikan `Create LF-Tag expression` izin kepada kepala sekolah, pilih Izin yang dapat diberikan. `Create LF-Tag expression`
8. Pilih Tambahkan.

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"
    ]
}
```

Peran pembuat LF-Tag ekspresi mendapatkan kemampuan untuk membuat, memperbarui, atau menghapus LF-Tag ekspresi.

Izin	Deskripsi
`Create`	Seorang kepala sekolah dengan izin ini dapat menambahkan LF-Tag ekspresi di danau data.
`Drop`	Kepala sekolah dengan izin ini pada LF-Tag ekspresi dapat menghapus LF-Tag ekspresi dari data lake.
`Alter`	Kepala sekolah dengan izin ini pada LF-Tag ekspresi dapat memperbarui badan ekspresi LF-Tag ekspresi.
`Describe`	Seorang kepala sekolah dengan izin ini pada LF-Tag ekspresi dapat melihat isi LF-Tag ekspresi.
`Grant with LF-Tag expression`	Izin ini memungkinkan penerima untuk menggunakan ekspresi tag sebagai sumber daya saat memberikan izin akses data atau metadata. Memberikan hibah `Grant with LF-Tag expression` implisit. `Describe`
`Super`	Untuk LF-Tag ekspresi, `Super` izin memberikan kemampuan untuk`Describe`,, `AlterDrop`, dan memberikan izin pada ekspresi tag ke prinsipal lain.

Izin ini dapat diberikan. Seorang kepala sekolah yang telah diberikan izin ini dengan opsi hibah dapat memberikannya kepada prinsipal lain.

Membuat LF-Tag ekspresi

Anda perlu mendefinisikan semua LF-Tags di Lake Formation, dan menetapkannya ke sumber daya Katalog Data sebelum dapat digunakan untuk membuat ekspresi. LF-Tag Ekspresi terdiri dari satu kunci lagi dan satu atau lebih nilai yang mungkin untuk setiap kunci.

Setelah administrator data lake menyiapkan izin IAM yang diperlukan dan izin Lake Formation untuk peran pembuat LF-Tag ekspresi, prinsipal dapat membuat ekspresi yang dapat digunakan kembali. LF-Tag Pembuat LF-Tag ekspresi mendapatkan izin implisit untuk memperbarui isi ekspresi, dan menghapus ekspresi. LF-Tag

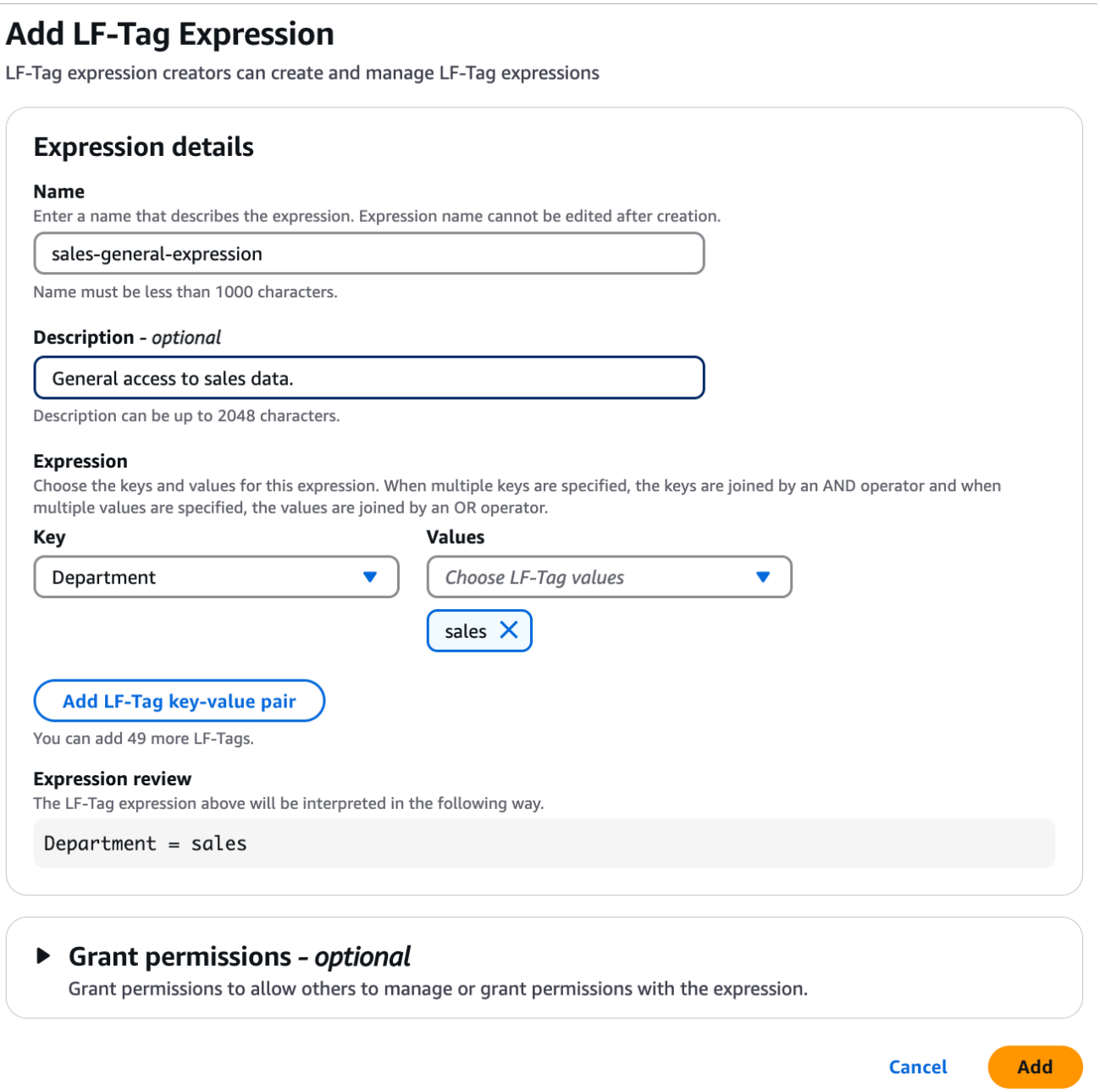
Anda dapat membuat LF-Tag ekspresi dengan menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk membuat LF-Tag ekspresi

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai prinsipal dengan izin pembuat LF-Tag ekspresi atau sebagai administrator danau data.
2. Di panel navigasi, di bawah Izin, pilih LF-Tags dan izin.
3. Pilih LF-Tag ekspresi. Halaman Tambahkan LF-Tag ekspresi muncul.



4. Masukkan informasi berikut:

   - Nama — Masukkan nama unik untuk ekspresi. Anda tidak dapat memperbarui nama ekspresi.
   - Deskripsi - Berikan deskripsi opsional untuk ekspresi dengan detail ekspresi.
   - Ekspresi - Buat ekspresi dengan menentukan kunci tag dan nilai yang terkait. Anda dapat menambahkan hingga 50 kunci per ekspresi. Anda harus memiliki izin `Grant with LF-Tags` Lake Formation pada semua tag di badan ekspresi.

Setiap kunci harus memiliki setidaknya satu nilai. Untuk memasukkan beberapa nilai, masukkan daftar yang dibatasi koma lalu tekan Enter, atau masukkan satu nilai pada satu waktu dan pilih Tambah setelah masing-masing. Jumlah maksimum nilai yang diizinkan per kunci adalah 1000.

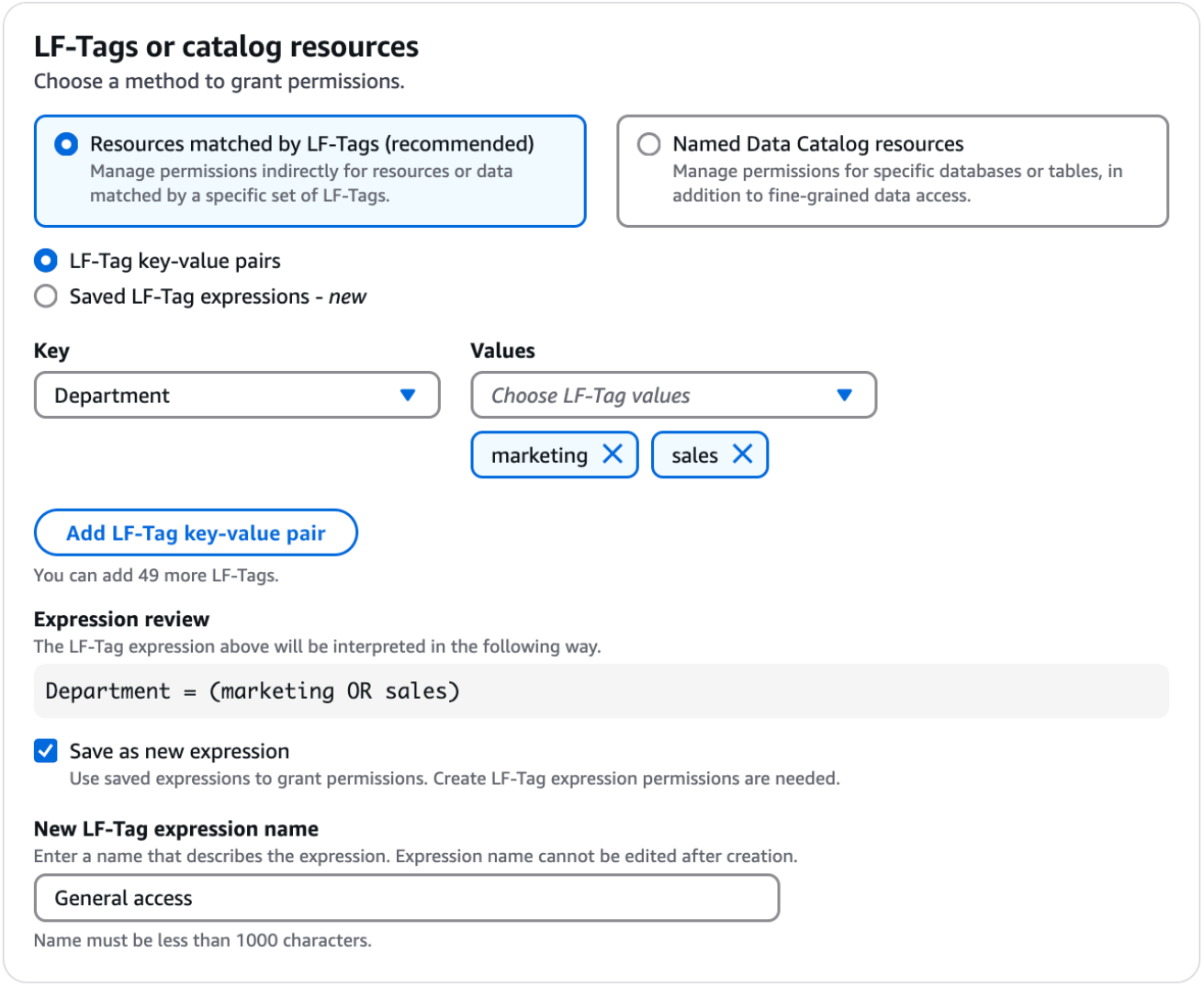
Lake Formation menggunakan AND/OR logika untuk menggabungkan beberapa kunci dan nilai dalam sebuah ekspresi. Dalam satu pasangan (kunci: daftar nilai), nilai-nilai digabungkan menggunakan operator OR logis. Misalnya, jika pasangannya (Departemen: [Penjualan, Pemasaran]), itu berarti tag cocok jika sumber daya memiliki tag Departemen dengan nilai Penjualan ATAU Pemasaran.

Ketika Anda menentukan beberapa kunci, kunci bergabung dengan operator dan logis. Jadi jika ekspresi lengkapnya adalah (Departemen: [Penjualan, Pemasaran]) DAN (Lokasi: [AS, Kanada]), itu cocok dengan sumber daya yang memiliki tag Departemen dengan nilai Penjualan ATAU Pemasaran, DAN juga memiliki tag Lokasi dengan nilai AS ATAU Kanada. Berikut ini adalah contoh lain dengan beberapa kunci dan nilai:

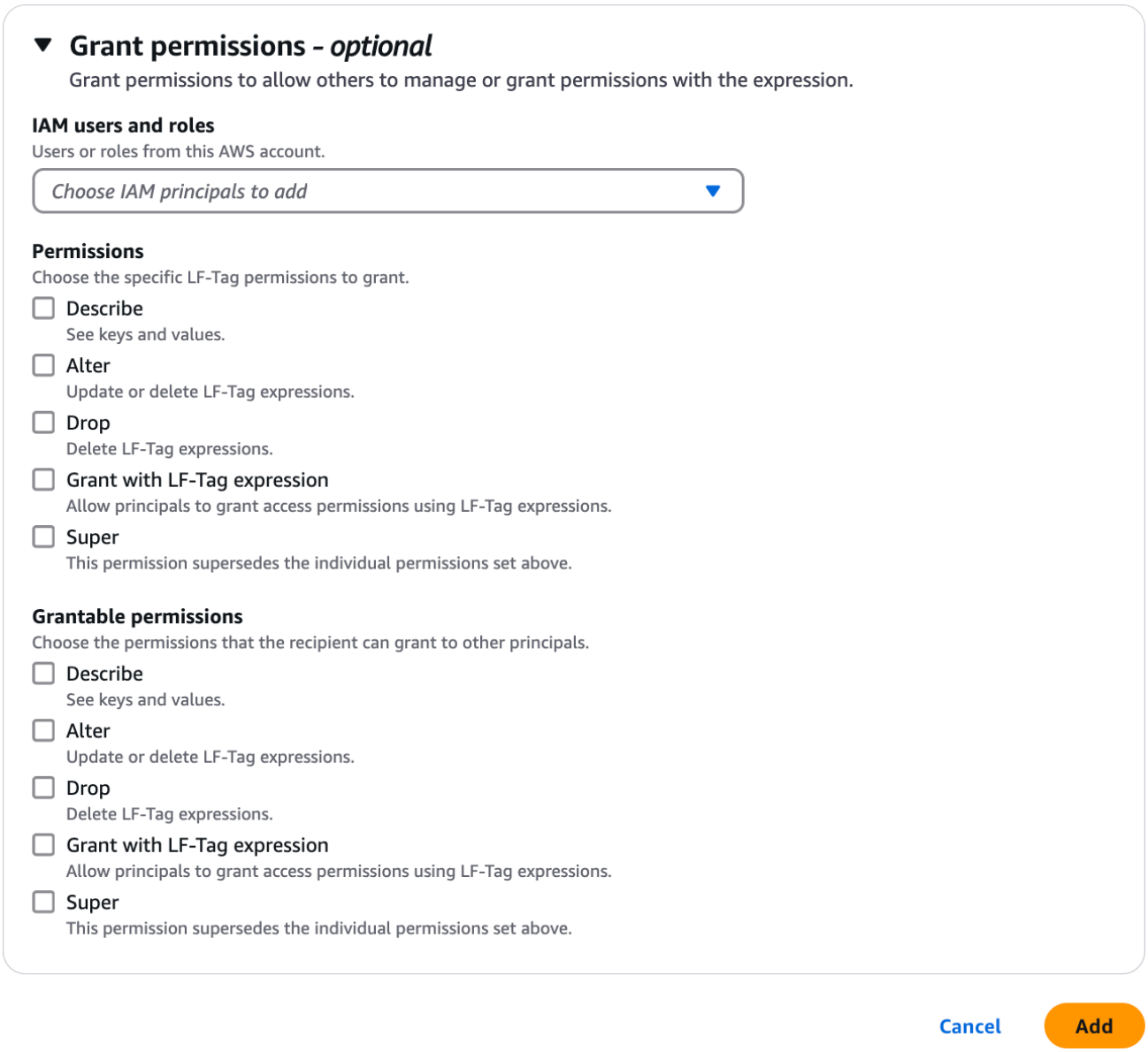
LF-Tag ekspresi: (ContentType : [Video, Audio]) DAN (Wilayah: [Eropa, Asia]) DAN (Departemen: [Teknik, ProductManagement]).

Ungkapan ini akan cocok dengan sumber daya yang memiliki: - ContentType Tag dengan nilai Video ATAU Audio DAN - Tag Wilayah dengan nilai Eropa ATAU Asia DAN - Tag Departemen dengan nilai Teknik OR ProductManagement.

Anda juga dapat menyimpan ekspresi tag saat memberikan izin data lake menggunakan. LF-Tags Pilih pasangan kunci dan nilai dan pilih opsi Simpan sebagai ekspresi baru. Masukkan nama yang menggambarkan ekspresi.



5. (Opsional) Selanjutnya users/roles, pilih, dan izin pada ekspresi yang ingin Anda berikan kepada mereka di akun. Anda juga dapat memilih izin yang dapat diberikan yang memungkinkan pengguna memberikan izin ini kepada pengguna lain di akun. Anda tidak dapat memberikan izin lintas akun pada ekspresi tag.



6. Pilih Tambahkan.

AWS CLI

Untuk membuat LF-Tag ekspresi

- Masukkan `create-lf-tag-expression` perintah.

  Contoh berikut menciptakan LF-Tag ekspresi dengan tag `Department` dengan nilai-nilai `Sales` dan`Marketing`, DAN tag `Location` dengan nilai`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"]}]}'
```

Perintah CLI ini menciptakan LF-Tag ekspresi baru di. AWS Glue Data Catalog Ekspresi dapat digunakan memberikan izin untuk sumber daya Katalog Data seperti database, tabel, tampilan atau kolom berdasarkan tag yang terkait. Dalam contoh ini, ekspresi akan cocok dengan sumber daya yang memiliki `Department` kunci dengan nilai `Sales` atau`Marketing`, dan `Location` kunci dengan nilai`US`.

Sebagai pembuat ekspresi tag, kepala sekolah mendapatkan `Alter` izin pada LF-Tag ekspresi ini dan dapat memperbarui atau menghapus ekspresi. Prinsipal pembuat LF-Tag ekspresi juga dapat memberikan `Alter` izin kepada prinsipal lain untuk memperbarui dan menghapus ekspresi ini.

Memperbarui LF-Tag ekspresi

Hanya administrator data lake, pembuat LF-Tag ekspresi, dan kepala sekolah yang memiliki `Alter` atau `Super` izin pada ekspresi yang dapat memperbarui LF-Tag ekspresi. LF-Tag Selain `Alter` izin, Anda juga memerlukan izin dan `Grant with LF-Tag` izin `lakeformation:UpdateLFTagExpression` IAM pada semua nilai kunci yang mendasari pada badan ekspresi baru untuk memperbarui ekspresi.

Anda memperbarui LF-Tag ekspresi dengan memperbarui deskripsi, isi ekspresi, dan izin yang diberikan pada ekspresi. Anda tidak dapat mengubah nama LF-Tag ekspresi. Untuk mengubah nama, hapus LF-Tag ekspresi dan tambahkan satu dengan parameter yang diperlukan.

Anda dapat memperbarui LF-Tag ekspresi menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk memperbarui LF-Tag ekspresi

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, LF-Tag pembuat atau kepala sekolah dengan `Alter` izin pada file LF-Tag.
2. Di panel navigasi, di bawah izin, pilih LF-Tags dan izin.

3. Pilih tab LF-Tag ekspresi.
4. Pada bagian LF-Tag ekspresi, pilih LF-Tag ekspresi, lalu pilih Edit.
5. Dalam kotak dialog Edit LF-Tag ekspresi, perbarui deskripsi dan perbarui badan ekspresi dengan menambahkan atau menghapus kunci dan nilai.

   Untuk menambahkan beberapa nilai, di bidang Nilai, pilih nilai dari drop-down.
6. Pilih Simpan.

AWS CLI

Perintah update-lf-tag-expression di Lake Formation memungkinkan Anda memperbarui ekspresi yang ada. LF-Tag

```
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", ...]}]}'
```

Inilah arti parameter dalam perintah yang disediakan:

- name — Nama ekspresi tag bernama yang ada yang ingin Anda perbarui.
- deskripsi — Deskripsi baru untuk ekspresi.

  catalog-id — ID Katalog Data tempat ekspresi tag bernama berada.
- expression - String ekspresi tag baru yang ingin Anda perbarui ekspresi dengan.

Menghapus ekspresi LF-Tag

Anda dapat menghapus LF-Tag ekspresi yang tidak lagi digunakan. Jika Anda telah memberikan izin kepada prinsipal pada sumber daya Katalog Data menggunakan LF-Tag ekspresi, mereka tidak akan lagi memiliki izin.

Hanya administrator data lake, pembuat LF-Tag ekspresi, atau kepala sekolah dengan `Drop` izin pada LF-Tag ekspresi yang dapat menghapus LF-Tag ekspresi. Selain `Drop` izin, kepala sekolah

juga memerlukan izin `lakeformation:DeleteLFTagExpression` IAM untuk menghapus LF-Tag ekspresi.

Anda dapat menghapus LF-Tag ekspresi menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk menghapus LF-Tag ekspresi (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, pembuat LF-Tag ekspresi, atau prinsipal yang memiliki izin untuk menghapus ekspresi.
2. Di panel navigasi, di bawah Izin, pilih LF-Tags dan izin.
3. Pilih tab LF-Tag ekspresi.
4. Pada bagian LF-Tag ekspresi, pilih LF-Tag ekspresi, lalu pilih Hapus.
5. Dalam LF-Tag ekspresi Hapus? kotak dialog, untuk mengkonfirmasi penghapusan, masukkan nama LF-Tag ekspresi di bidang yang ditunjuk dan kemudian pilih Hapus.

AWS CLI

Untuk menghapus LF-Tag (AWS CLI)

- Masukkan `delete-lf-tag-expression` perintah. Berikan nama ekspresi dan ID katalog untuk dihapus.

  Example

  Contoh berikut menghapus LF-Tag ekspresi dengan nama `my-tag-expression` dari Katalog Data dengan ID`123456789012`. `catalog-id`Parameter ini opsional jika Anda menggunakan akun yang sama dengan AWS CLI konfigurasi Anda. Setelah menghapus LF-Tag ekspresi, Lake Formation membersihkan catatan izin terkait untuk ekspresi tersebut. Ini termasuk catatan izin individu dan catatan izin agregat yang berisi ekspresi yang dihapus.

  ```
  aws lakeformation delete-lf-tag-expression \
  --name "my-tag-expression" \
  --catalog-id "123456789012"
  ```

LF-Tag Ekspresi daftar

Anda dapat mencantumkan LF-Tag ekspresi tempat Anda memiliki izin Jelaskan. Administrator data lake, pembuat LF-Tag ekspresi, dan Read-only administrator secara implisit dapat melihat semua ekspresi tag di akun mereka.

Anda dapat membuat daftar LF-Tag ekspresi menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Console

Untuk membuat daftar LF-Tag ekspresi (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai pembuat LF-Tag ekspresi, sebagai administrator data lake, atau sebagai prinsipal yang telah diberikan izin pada LF-Tag ekspresi dan yang memiliki izin `lakeformation:ListLFTagExpressions` IAM.
2. Di panel navigasi, di bawah Izin, LF-Tags dan izin.
3. Pilih tab LF-Tag ekspresi untuk melihat ekspresi. Bagian ini menunjukkan informasi tentang LF-Tag ekspresi yang ada, termasuk nama ekspresi, ekspresi itu sendiri dengan tautan ke tag yang disertakan, dan opsi untuk membuat, mengedit, atau menghapus ekspresi.

AWS CLI

Untuk daftar LF-Tags (AWS CLI)

- Untuk membuat daftar LF-Tag ekspresi menggunakan AWS CLI, Anda dapat menggunakan perintah list-lf-tag-expressions. Sintaks permintaan adalah:

  ```
  aws lakeformation list-lf-tag-expressions \
  -- catalog-id "123456789012" \
  -- max-items "100" \
  -- next-token "next-token"
  ```

  Di mana:

  - `catalog-id`adalah ID AWS akun Katalog Data yang ingin Anda cantumkan ekspresi tag.

- `max-items`menentukan jumlah maksimum ekspresi tag untuk kembali. Jika parameter ini tidak digunakan, nilai default adalah 100.
- `next-token`adalah token kelanjutan jika hasilnya terpotong dalam permintaan sebelumnya.

Respons akan mencakup daftar LF-Tag ekspresi dan token berikutnya jika berlaku.

Mengelola izin LF-Tag nilai

Anda dapat memberikan`Drop`, `Alter` izin pada prinsipal LF-Tags untuk mengelola ekspresi nilai. LF-Tag Anda juga dapat memberikan`Describe`,`Associate`, dan `Grant with LF-Tag expressions` izin LF-Tags ke kepala sekolah untuk melihat LF-Tags dan menetapkannya ke sumber daya Katalog Data (database, tabel, dan kolom). Ketika LF-Tags ditetapkan ke sumber daya Katalog Data, Anda dapat menggunakan metode kontrol akses (LF-TBAC) berbasis tag Lake Formation untuk mengamankan sumber daya tersebut. Untuk informasi selengkapnya, lihat Kontrol akses berbasis tag Lake Formation.

Anda dapat memberikan izin ini dengan opsi hibah sehingga kepala sekolah lain dapat memberikannya. `Associate`Izin `Grant with LF-Tag expressions``Describe`,, dan dijelaskan dalamTambahkan LF-Tag kreator.

Anda dapat memberikan `Associate` izin `Describe` dan pada LF-Tag AWS akun eksternal. Administrator data lake di akun itu kemudian dapat memberikan izin tersebut kepada prinsipal lain di akun tersebut. Prinsipal kepada siapa administrator data lake di akun eksternal memberikan `Associate` izin kemudian dapat menetapkan LF-Tags ke sumber daya Katalog Data yang Anda bagikan dengan akun mereka.

Saat memberikan ke akun eksternal, Anda harus menyertakan opsi hibah.

Anda dapat memberikan izin LF-Tags menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

Topik

- LF-Tag Izin daftar menggunakan konsol
- Memberikan LF-Tag izin menggunakan konsol
- Mengelola LF-Tag izin menggunakan AWS CLI

Untuk informasi lebih lanjut, lihat Mengelola LF-Tags kontrol akses metadata dan Kontrol akses berbasis tag Lake Formation.

LF-Tag Izin daftar menggunakan konsol

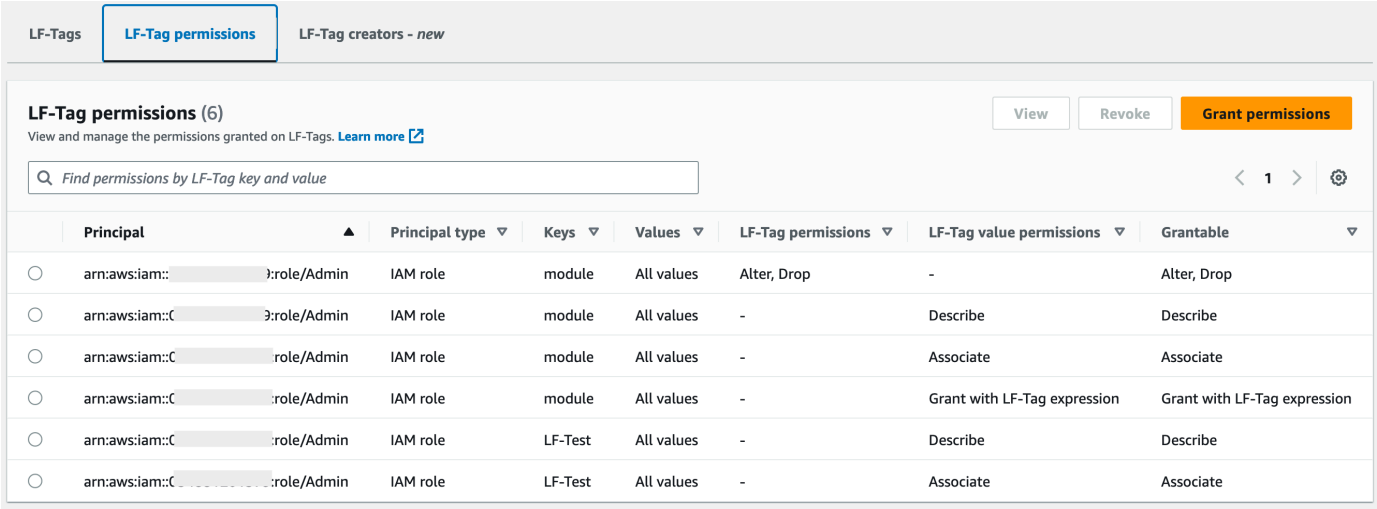
Anda dapat menggunakan konsol Lake Formation untuk melihat izin yang diberikan. LF-Tags Anda harus menjadi LF-Tag pencipta, administrator data lake, `Describe` atau memiliki atau `Associate` izin LF-Tag untuk melihatnya.

Untuk daftar LF-Tag izin (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai LF-Tag pembuat, administrator data lake, atau sebagai pengguna yang kepadanya`Drop`,`Alter`,`Associate`, atau `Describe` izin LF-Tags telah diberikan.
2. Di panel navigasi, di bawah Izin, pilih LF-Tags dan izin, dan pilih LF-Tag bagian izin.

   Bagian LF-Tag izin menunjukkan tabel yang berisi prinsipal, kunci tag, nilai, dan izin.



Memberikan LF-Tag izin menggunakan konsol

Langkah-langkah berikut menjelaskan cara memberikan izin menggunakan halaman LF-Tag izin Hibah di konsol Lake Formation. LF-Tags Halaman ini dibagi menjadi beberapa bagian ini:

- Jenis izin — Jenis izin untuk diberikan.
- Prinsipal — Pengguna atau peran IAM, atau pengguna atau peran SALL untuk memberikan izin kepada.
- LF-Tag izin pasangan nilai kunci — Pasangan LF-Tag nilai kunci untuk memberikan izin pada.

- LF-Tag izin — LF-Tags Untuk memberikan izin pada.
- LF-Tag izin ekspresi izin — LF-Tags Untuk memberikan izin pada.
- Izin — Izin untuk diberikan.

Buka LF-Tag Izin pemberian halaman

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai LF-Tag pembuat, administrator danau data, atau sebagai izin pengguna atau LF-Tag izin pasangan LF-Tag nilai kunci LF-Tags telah diberikan dengan opsi. `Grant`
2. Di panel navigasi, pilih LF-Tags dan izin, pilih LF-Tag bagian izin.
3. Pilih Berikan izin.

Tentukan jenis izin

Di bagian Jenis izin, pilih jenis izin.

LF-Tag izin

Pilih LF-Tag izin untuk mengizinkan prinsipal memperbarui LF-Tag nilai atau menghapus. LF-Tags

LF-Tag izin pasangan nilai kunci

Pilih izin pasangan LF-Tag nilai kunci untuk mengizinkan prinsipal menetapkan LF-Tags sumber daya, tampilan, dan nilai Katalog Data, LF-Tags dan berikan izin LF-Tags berdasarkan sumber daya Katalog Data kepada prinsipal.

Opsi yang tersedia di bagian berikut bergantung pada jenis Izin.

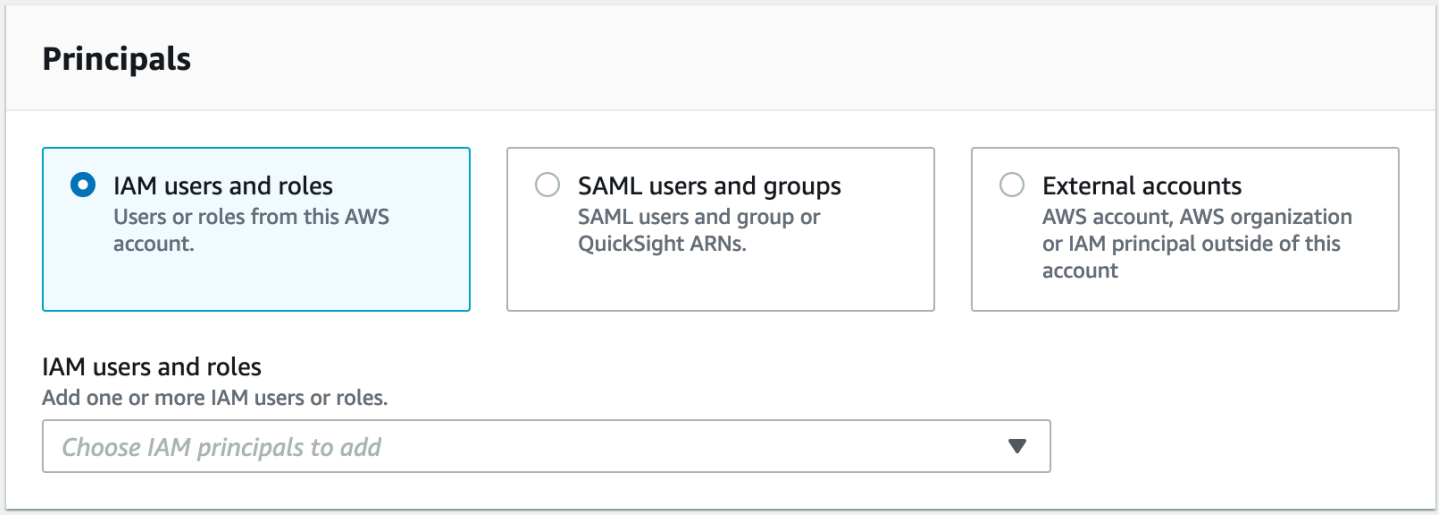
LF-Tag izin ekspresi

Pilih izin LF-Tag ekspresi untuk mengizinkan prinsipal memperbarui ekspresi atau menghapus ekspresi.

Tentukan prinsipal

> **Note**
>
> Anda tidak dapat memberikan LF-Tag izin (`AlterdanDrop`) ke akun eksternal atau prinsipal di akun lain.

Di bagian Prinsipal, pilih jenis utama dan tentukan prinsipal untuk memberikan izin.



Pengguna dan peran IAM

Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAFL dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAFL, atau ARN untuk pengguna atau grup Cepat. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. Lake Formation memberikan dan mencabut AWS CLI perintah

> **Note**
>
> Integrasi Lake Formation dengan Quick didukung hanya untuk Quick Enterprise Edition.

Akun eksternal

Untuk AWS akun, masukkan satu atau beberapa ID AWS akun yang valid. Tekan Enter setelah setiap ID.

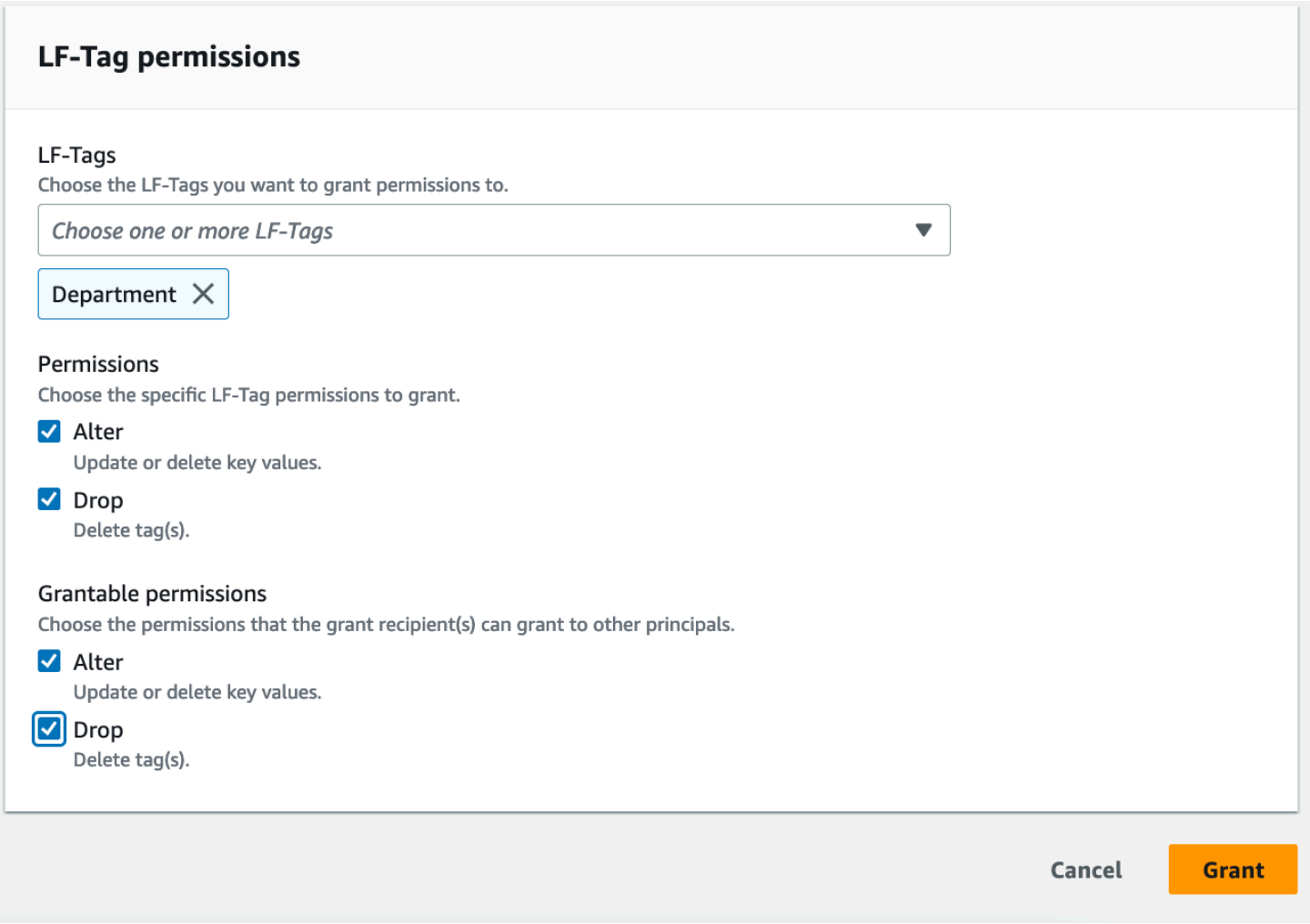
ID organisasi terdiri dari “o-” diikuti oleh 10 hingga 32 huruf kecil atau digit.

ID unit organisasi dimulai dengan “ou-” diikuti oleh 4 hingga 32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

Untuk kepala sekolah IAM, masukkan ARN untuk pengguna atau peran IAM.

Tentukan LF-Tags

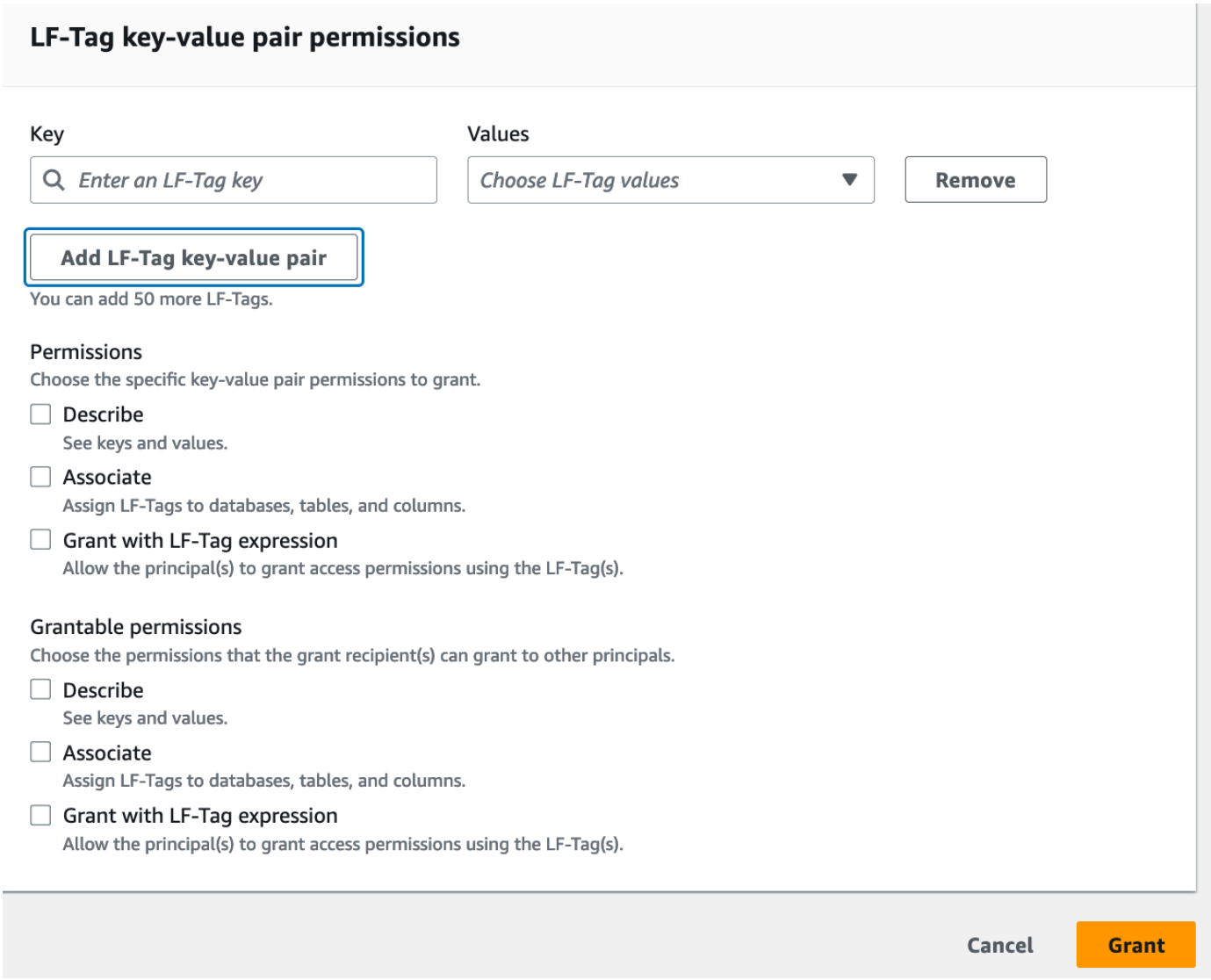
Untuk memberikan izin pada LF-Tags, di bagian LF-Tag izin, tentukan hak LF-Tags untuk memberikan izin pada.



- Pilih satu atau lebih LF-Tag menggunakan drop-down.

Tentukan pasangan LF-Tag kunci-nilai

1. Untuk memberikan izin pada pasangan LF-Tag kunci-nilai, (Anda harus terlebih dahulu memilih memilih izin pasangan LF-Tag kunci-nilai sebagai tipe Izin) pilih Tambahkan pasangan LF-Tag kunci-nilai untuk mengungkapkan baris pertama bidang untuk menentukan kunci dan nilai. LF-Tag



2. Posisikan kursor di bidang Kunci, secara opsional mulai mengetik untuk mempersempit daftar pilihan, dan pilih tombol. LF-Tag
3. Dalam daftar Nilai, pilih satu atau beberapa nilai, lalu tekan Tab atau klik atau ketuk di luar bidang untuk menyimpan nilai yang dipilih.

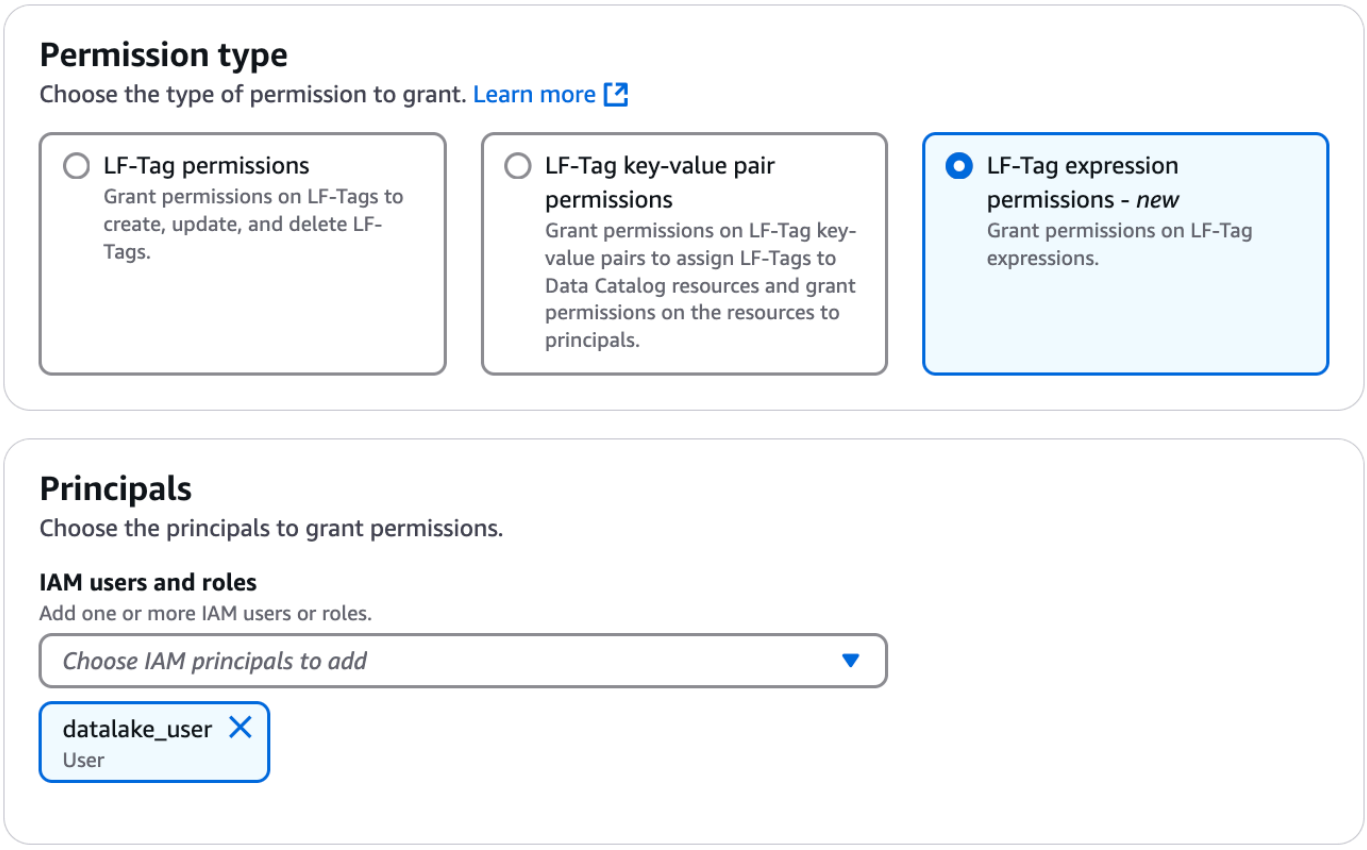
> **Note**
>
> Jika salah satu baris dalam daftar Nilai memiliki fokus, menekan Enter memilih atau menghapus kotak centang.

Nilai yang dipilih muncul sebagai ubin di bawah daftar Nilai. Pilih ✖ untuk menghapus nilai. Pilih Hapus untuk menghapus keseluruhan LF-Tag.

4. Untuk menambahkan yang lain LF-Tag, pilih Tambah LF-Tag lagi, dan ulangi dua langkah sebelumnya.

Tentukan LF-Tag ekspresi

1. Untuk memberikan izin pada LF-Tag ekspresi, (Anda harus terlebih dahulu memilih izin LF-Tag ekspresi sebagai jenis Izin).



2. Pilih LF-Tag ekspresi.

3. Ekspresi yang dipilih muncul sebagai ubin di bawah daftar LF-Tag ekspresi. Pilih ✖ untuk menghapus ekspresi.
4. Untuk menambahkan LF-Tag ekspresi lain, pilih ekspresi lain.

Tentukan izin

Bagian ini menunjukkan LF-Tag izin atau izin LF-Tag nilai berdasarkan jenis Izin yang Anda pilih pada langkah sebelumnya.

Bergantung pada jenis Izin yang Anda pilih untuk diberikan, pilih LF-Tag izin atau izin pasangan LF-Tag nilai kunci, dan izin yang dapat diberikan.

1. Di bawah LF-Tag izin, pilih izin yang akan diberikan.

   Memberikan Drop and Alter secara implisit memberikan Deskripsikan.

   Anda harus memberikan izin Alter dan Drop pada semua nilai tag.
2. Di bawah izin LT-Tag nilai kunci, pilih izin yang akan diberikan.

   Pemberian Associate secara implisit memberikan Jelaskan. Pilih Hibah dengan LF-Tag ekspresi untuk mengizinkan penerima hibah memberikan atau mencabut izin akses pada sumber daya Katalog Data menggunakan metode. LF-TBAC
3. Di bawah izin LF-Tag ekspresi, pilih izin yang akan diberikan.

   Memberikan Drop and Alter secara implisit memberikan Deskripsikan.

   Memberikan izin Super, memberikan semua izin yang tersedia.
4. (Opsional) Di bawah Izin yang Dapat Diberikan, pilih izin yang dapat diberikan oleh penerima hibah kepada prinsipal lain di akun mereka. AWS
5. PilihIzin.

Mengelola LF-Tag izin menggunakan AWS CLI

Anda dapat memberikan, mencabut, dan mencantumkan izin LF-Tags dengan menggunakan (). AWS Command Line Interface AWS CLI

Untuk membuat daftar LF-Tag izin (AWS CLI)

- Masukkan `list-permissions` perintah. Anda harus menjadi LF-Tag pencipta, administrator danau data, atau memiliki `DropAlter`,`Describe`,`Associate`,, `Grant with LF-Tag permissions` izin LF-Tag untuk melihatnya.

  Perintah berikut meminta semua izin LF-Tags yang Anda miliki.

```
aws lakeformation list-permissions --resource-type LF_TAG
```

  Berikut ini adalah contoh output untuk administrator data lake, yang melihat semua LF-Tags diberikan kepada semua kepala sekolah. Non-administrative pengguna hanya melihat LF-Tags diberikan kepada mereka. LF-Tag izin yang diberikan dari akun eksternal muncul di halaman hasil terpisah. Untuk melihatnya, ulangi perintah dan berikan `--next-token` argumen dengan token yang dikembalikan dari perintah run sebelumnya.

```
{
    "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...Wlzc2lvbnMiOnRydWV9"
}
```

Anda dapat membuat daftar semua hibah untuk LF-Tag kunci tertentu. Perintah berikut mengembalikan semua izin yang diberikan pada. LF-Tag `module`

```
aws lakeformation list-permissions --resource-type LF_TAG --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Anda juga dapat mencantumkan LF-Tag nilai yang diberikan kepada prinsipal tertentu untuk yang spesifik LF-Tag. Saat memberikan `--principal` argumen, Anda harus memberikan `--resource` argumen. Oleh karena itu, perintah hanya dapat secara efektif meminta nilai yang diberikan kepada prinsipal tertentu untuk LF-Tag kunci tertentu. Perintah berikut menunjukkan bagaimana melakukan ini untuk kepala sekolah `datalake_user1` dan LF-Tag kunci`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":["*"]}}'
```

Berikut ini adalah contoh output-nya.

```
{
    "PrincipalResourcePermissions": [
        {
            "Principal": {
                "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
            },
            "Resource": {
                "LFTag": {
                    "CatalogId": "111122223333",
                    "TagKey": "module",
                    "TagValues": [
                        "Orders",
                        "Sales"
                    ]
                }
            },
            "Permissions": [
                "ASSOCIATE"
            ],
            "PermissionsWithGrantOption": []
        }
    ]
}
```

Untuk memberikan izin pada LF-Tags (AWS CLI)

1. Masukkan perintah yang serupa dengan yang berikut ini. Contoh ini memberikan `Associate` izin `datalake_user1` kepada pengguna LF-Tag dengan kuncimodule. Ini memberikan izin untuk melihat dan menetapkan semua nilai untuk kunci itu, seperti yang ditunjukkan oleh tanda bintang (*).

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

   Pemberian `Associate` izin secara implisit memberikan izin. `Describe`

Contoh berikutnya memberikan `Associate` ke AWS akun eksternal 1234-5678-9012 pada dengan kunci, LF-Tag dengan opsi hibah. `module` Ini memberikan izin untuk melihat dan menetapkan hanya nilai dan. `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. Pemberian `GrantWithLFTagExpression` izin secara implisit memberikan izin. `Describe`

   Contoh berikutnya memberikan `GrantWithLFTagExpression` kepada pengguna dengan kuncimodule, LF-Tag dengan opsi hibah. Ini memberikan izin untuk melihat dan memberikan izin pada sumber daya Katalog Data hanya menggunakan nilai dan. `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. Contoh berikutnya memberikan `Drop` izin kepada pengguna dengan kuncimodule, LF-Tag dengan opsi hibah. Ini memberikan izin untuk menghapus file. LF-Tag Untuk menghapus LF-Tag, Anda memerlukan izin pada semua nilai untuk kunci itu.

   ```
   aws lakeformation grant-permissions --principal
    DataLakePrincipalIdentifier=111122223333 --permissions "DROP"
    --permissions-with-grant-option "DROP" --resource '{ "LFTag":
    {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
   ```

4. Contoh berikutnya memberikan `Alter` izin kepada pengguna dengan kuncimodule, LF-Tag dengan opsi hibah. Ini memberikan izin untuk menghapus file. LF-Tag Untuk memperbarui LF-Tag, Anda memerlukan izin pada semua nilai untuk kunci itu.

   ```
   aws lakeformation grant-permissions --principal
    DataLakePrincipalIdentifier=111122223333 --permissions "ALTER"
    --permissions-with-grant-option "ALTER" --resource '{ "LFTag":
    {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
   ```

Untuk mencabut izin pada (LF-Tags AWS CLI)

- Masukkan perintah yang serupa dengan yang berikut ini. Contoh ini mencabut `Associate` izin pada LF-Tag dengan kunci `module` dari penggunadatalake_user1.

```
aws lakeformation revoke-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Memberikan izin data lake menggunakan metode LF-TBAC

Anda dapat memberikan izin `DESCRIBE` dan `ASSOCIATE` Lake Formation LF-Tags ke prinsipal sehingga mereka dapat melihat LF-Tags dan menetapkannya ke sumber daya Katalog Data (database, tabel, tampilan, dan kolom). Ketika LF-Tags ditetapkan ke sumber daya Katalog Data, Anda dapat menggunakan metode kontrol akses (LF-TBAC) berbasis tag Lake Formation untuk mengamankan sumber daya tersebut. Untuk informasi selengkapnya, lihat Kontrol akses berbasis tag Lake Formation.

Pada awalnya, hanya administrator data lake yang dapat memberikan izin ini. Jika administrator data lake memberikan izin ini dengan opsi hibah, prinsipal lain dapat memberikannya. `ASSOCIATE`Izin `DESCRIBE` dan dijelaskan dalamPraktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation.

Anda dapat memberikan `ASSOCIATE` izin `DESCRIBE` dan pada LF-Tag AWS akun eksternal. Administrator data lake di akun tersebut kemudian dapat memberikan izin tersebut kepada prinsipal lain di akun tersebut. Prinsipal kepada siapa administrator data lake di akun eksternal memberikan `ASSOCIATE` izin kemudian dapat menetapkan LF-Tags ke sumber daya Katalog Data yang Anda bagikan dengan akun mereka.

Saat memberikan ke akun eksternal, Anda harus menyertakan opsi hibah.

Anda dapat memberikan izin LF-Tags menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

> **Note**
>
> Langkah-langkah berikut tidak diperlukan untuk katalog Tabel S3. Anda dapat menggunakan LF-Tags untuk memberikan izin pada katalog Tabel S3 yang ada tanpa menghapus dan membuatnya kembali.

Mengaktifkan LF-Tags dukungan untuk katalog federasi yang ada yang menggunakan izin Lake Formation

Ikuti langkah-langkah berikut, jika Anda memiliki katalog federasi yang menggunakan izin Lake Formation, seperti Amazon Redshift atau katalog yang dibuat sebelum dukungan tersedia untuk Amazon DynamoDB katalog federasi. LF-Tags

1. Hapus katalog yang ada — Panggil operasi `deleteCatalog` API untuk menghapus katalog federasi yang ada yang menggunakan izin Lake Formation.
2. Buat katalog federasi baru — Buat katalog baru dan arahkan katalog baru ke katalog yang sudah ada namespace/datashare.

   Gunakan nama baru untuk katalog — Proses ini memperbarui katalog federasi Anda yang sudah ada sebelumnya untuk mendukung fungsionalitas. LF-Tag Jika Anda ingin menggunakan nama katalog yang sama, hubungi tim AWS dukungan untuk bantuan.

Topik

- Memberikan izin Katalog Data

> **Lihat juga**
>
> - Mengelola izin LF-Tag nilai
> - Mengelola LF-Tags kontrol akses metadata
> - Kontrol akses berbasis tag Lake Formation

Memberikan izin Katalog Data

Gunakan konsol Lake Formation atau AWS CLI untuk memberikan izin Lake Formation pada database, tabel, tampilan, dan kolom Katalog Data menggunakan metode kontrol akses () LF-TBAC berbasis tag Lake Formation.

Console

Langkah-langkah berikut menjelaskan cara memberikan izin dengan menggunakan metode kontrol akses (LF-TBAC) berbasis tag Lake Formation dan halaman izin danau data Grant di konsol Lake Formation. Halaman ini dibagi menjadi beberapa bagian berikut:

- Prinsipal — Pengguna, peran, dan Akun AWS untuk memberikan izin untuk.
- LF-Tags atau sumber daya katalog — Database, tabel, atau tautan sumber daya untuk memberikan izin pada.
- Izin — Izin Lake Formation untuk diberikan.

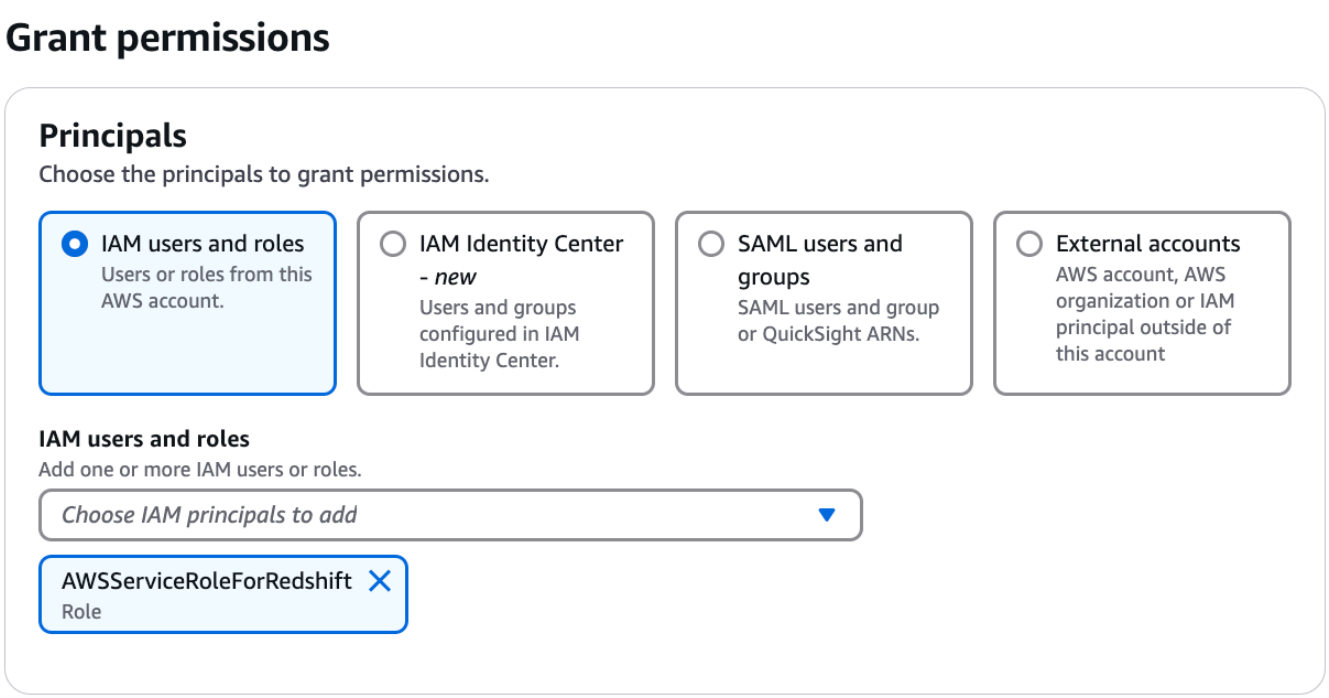
1. Buka halaman izin danau data Grant.

   Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake atau sebagai pengguna yang telah diberikan izin Lake Formation pada sumber daya Katalog Data melalui LF-TBAC opsi hibah.

   Di panel navigasi, di bawah Izin, pilih Izin danau data. Kemudian pilih Grant.

2. Tentukan prinsipal.

   Di bagian Prinsipal, pilih jenis utama dan kemudian tentukan prinsipal untuk memberikan izin.



Pengguna dan peran IAM

Pilih satu atau beberapa pengguna atau peran dari daftar pengguna dan peran IAM.

Pusat Identitas IAM

Pilih satu atau beberapa pengguna atau dari daftar Pengguna dan grup.

Pengguna dan grup SAFL

Untuk pengguna dan grup SAMB dan Cepat, masukkan satu atau beberapa Nama Sumber Daya Amazon (ARN) untuk pengguna atau grup yang digabungkan melalui SAMB, atau ARN untuk pengguna atau grup Cepat. Tekan Enter setelah setiap ARN.

Untuk informasi tentang cara membangun ARN, lihat. [Lake Formation memberikan dan mencabut AWS CLI perintah]

> **Note**
>
> Integrasi Lake Formation dengan Quick didukung hanya untuk Quick Enterprise Edition.

Akun eksternal

Untuk Akun AWS, AWS organisasi, atau prinsipal IAM masukkan satu atau beberapa Akun AWS ID, ID organisasi, ID unit organisasi, atau ARN yang valid untuk pengguna atau peran IAM. Tekan Enter setelah setiap ID.

ID organisasi terdiri dari “o-” diikuti oleh 10 hingga 32 huruf kecil atau digit.

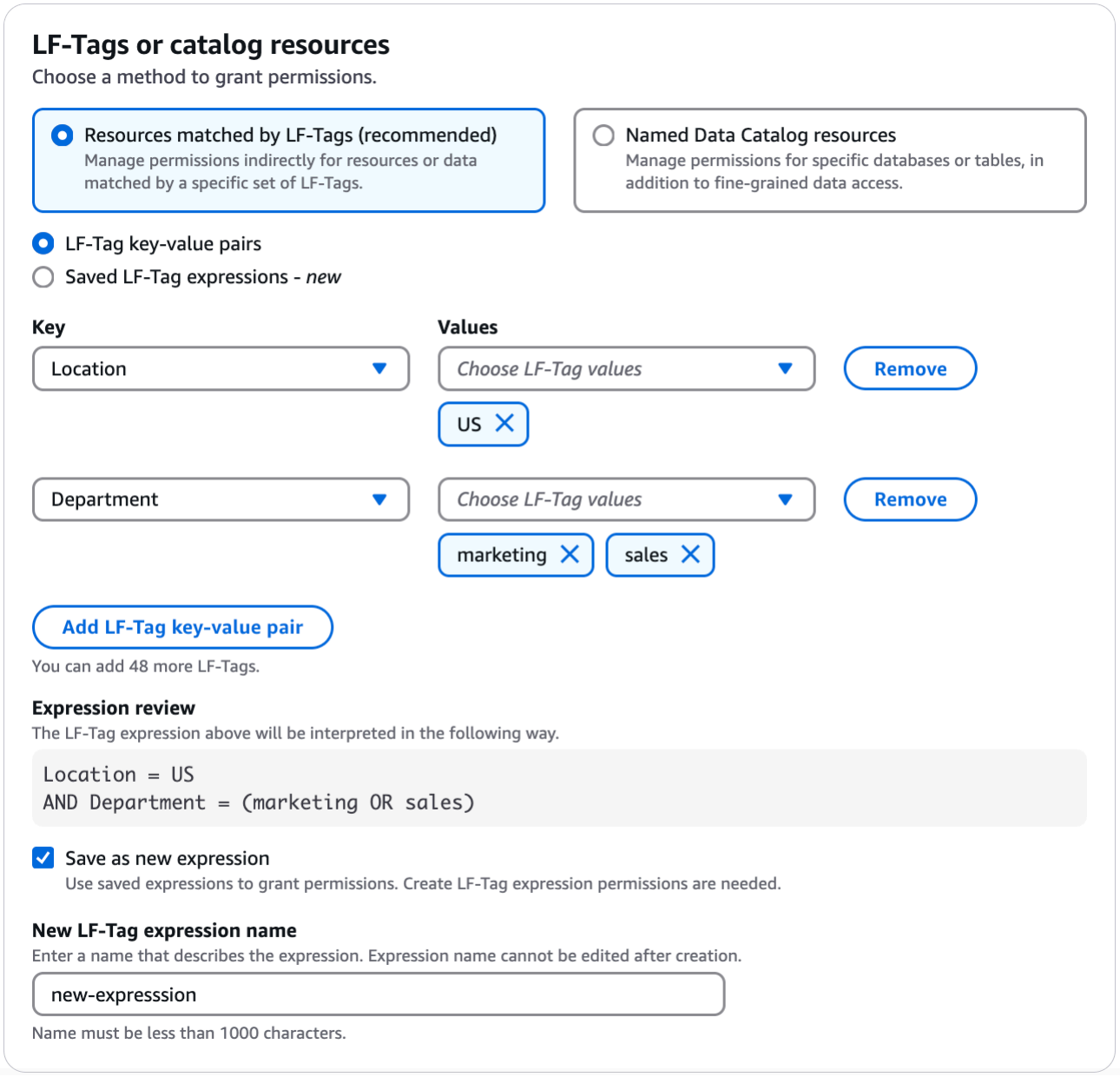
ID unit organisasi dimulai dengan “ou-” diikuti oleh 4 hingga 32 huruf kecil atau digit (ID dari root yang berisi OU). String ini diikuti oleh tanda hubung “-” kedua dan 8 hingga 32 huruf kecil atau digit tambahan.

3. Tentukan LF-Tags.

Pastikan bahwa opsi Sumber Daya yang cocok dengan LF-Tags pilihan dipilih. Pilih pasangan LF-Tag kunci-nilai atau Ekspresi tersimpan LF-Tag .

1. Jika Anda memilih opsi pasangan LF-Tag kunci-nilai, pilih kunci dan nilai.

   Jika Anda memilih lebih dari satu nilai, Anda membuat LF-Tag ekspresi dengan `OR` operator. Ini berarti bahwa jika salah satu LF-Tag nilai cocok dengan yang LF-Tag ditetapkan ke sumber daya Katalog Data, Anda diberikan izin pada sumber daya.



2. (Opsional) Pilih Tambahkan pasangan LF-Tag kunci-nilai lagi untuk menentukan yang lain. LF-Tag

   Jika Anda menentukan lebih dari satu LF-Tag, Anda membuat LF-Tag ekspresi dengan AND operator. Prinsipal diberikan izin pada sumber daya Katalog Data hanya jika sumber daya diberikan pencocokan LF-Tag untuk masing-masing LF-Tag dalam LF-Tag ekspresi.

3. Pilih Simpan sebagai opsi ekspresi baru untuk menggunakan kembali ekspresi.

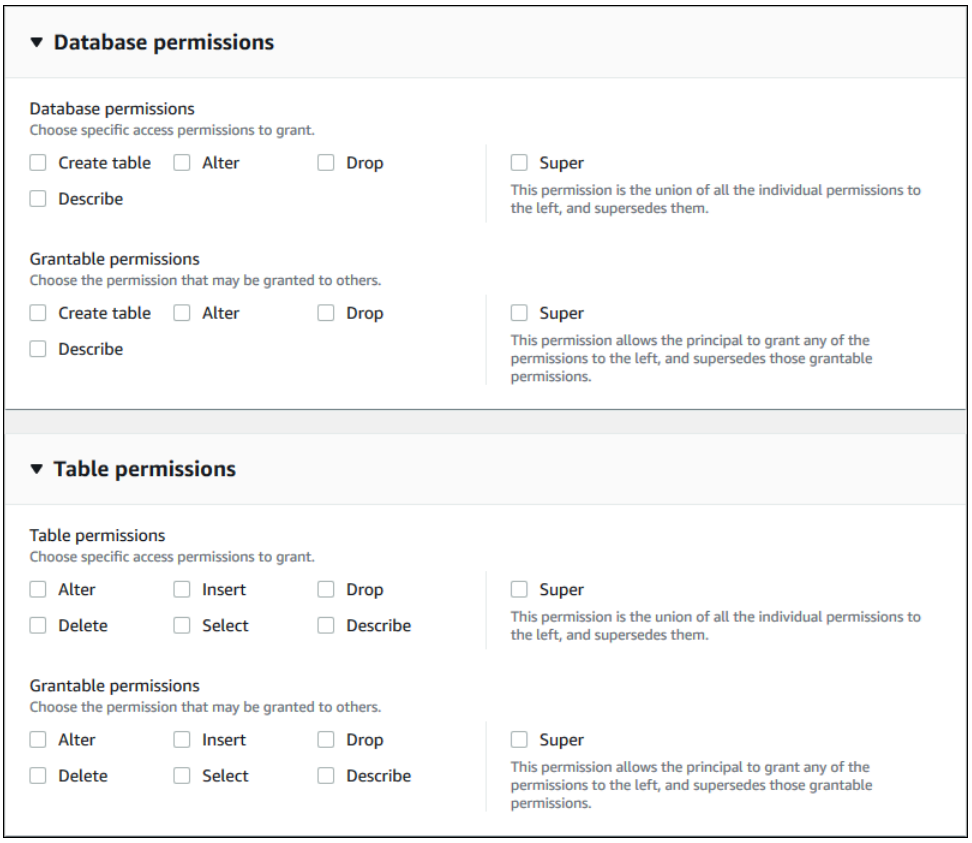
   Anda `Create LF-Tag expression` perlu menyimpan ekspresi.

Untuk informasi selengkapnya tentang LF-Tag ekspresi, lihatMengelola LF-Tag ekspresi untuk kontrol akses metadata.

4. Tentukan izin.

Tentukan izin yang ingin Anda berikan kepada prinsipal tentang pencocokan sumber daya Katalog Data. Sumber daya yang cocok adalah sumber daya yang ditugaskan LF-Tags yang cocok dengan salah satu LF-Tag ekspresi yang diberikan kepada prinsipal.

Anda dapat menentukan izin yang akan diberikan pada database yang cocok, tabel yang cocok, dan tampilan yang cocok.



Di bawah Izin database, pilih izin database yang akan diberikan kepada prinsipal pada database yang cocok.

Di bawah Izin tabel, pilih tabel atau izin tampilan yang akan diberikan kepada prinsipal pada tabel dan tampilan yang cocok.

Anda juga dapat memilih`Select`,`Describe`, dan `Drop` izin dari izin Tabel untuk diterapkan pada tampilan.

5. PilihIzin.

AWS CLI

Anda dapat menggunakan AWS Command Line Interface (AWS CLI) dan metode access control (LF-TBAC) berbasis tag Lake Formation untuk memberikan izin Lake Formation pada database, tabel, dan kolom Data Catalog.

Memberikan izin data lake menggunakan AWS CLI dan LF-TBAC metodenya

- Gunakan perintah `grant-permissions`.

  Example

  Contoh berikut memberikan LF-Tag ekspresi "`module=*`" (semua nilai LF-Tag kunci`module`) kepada penggunadatalake_user1. Pengguna tersebut akan memiliki `CREATE_TABLE` izin pada semua database yang cocok — database yang telah ditetapkan dengan kunci`module`, LF-Tag dengan nilai apa pun.

```
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

  Contoh berikutnya memberikan LF-Tag ekspresi "`(level=director) AND (region=west OR region=south)`" kepada penggunadatalake_user1. Pengguna tersebut akan memilikiSELECT,ALTER, dan DROP izin dengan opsi hibah pada tabel yang cocok—tabel yang telah ditetapkan keduanya `level=director` dan (`region=west`atau). `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

Contoh berikutnya ini memberikan LF-Tag ekspresi "`module=orders`" ke AWS akun 1234-5678-9012. Administrator data lake di akun itu kemudian dapat memberikan ekspresi "`module=orders`" kepada kepala sekolah di akun mereka. Prinsipal tersebut kemudian akan memiliki `CREATE_TABLE` izin untuk mencocokkan basis data yang dimiliki oleh akun 1111-2222-3333 dan dibagikan dengan akun 1234-5678-9012 dengan menggunakan metode sumber daya bernama atau metode. 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"]}]}}'
```

Attribute-based kontrol akses

Di AWS Lake Formation, Anda dapat memberikan akses pada AWS Glue Data Catalog objek seperti katalog, database, tabel, dan filter data menggunakan atribut yang merupakan tag IAM dan tag sesi yang terkait dengan entitas IAM seperti peran dan pengguna.

Untuk informasi selengkapnya tentang penggunaan tag sesi, lihat peran asumsi di panduan pengguna. AWS CLI

Attribute-based Access Control (ABAC) adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. AWS memanggil tag atribut ini. Anda dapat menggunakan ABAC untuk memberikan akses ke prinsipal dalam akun yang sama atau di akun lain pada sumber daya Katalog Data. Setiap prinsipal IAM dengan tag IAM yang cocok atau kunci tag sesi dan nilai mendapatkan akses ke sumber daya. Anda harus memiliki izin yang dapat diberikan pada sumber daya untuk membuat hibah ini.

ABAC memungkinkan Anda untuk memberikan akses ke beberapa pengguna secara bersamaan. Ketika pengguna baru bergabung dengan organisasi, akses mereka ke data dapat ditentukan secara otomatis berdasarkan atribut mereka, seperti fungsi pekerjaan atau departemen mereka, tanpa mengharuskan administrator untuk menetapkan peran atau izin tertentu secara manual. Dengan menggunakan atribut alih-alih peran, ABAC menyediakan cara yang lebih efisien dan dapat dipelihara untuk mengelola akses data di berbagai sistem dan lingkungan, yang pada akhirnya meningkatkan tata kelola dan kepatuhan data.

Untuk informasi selengkapnya tentang mendefinisikan atribut, lihat Mendefinisikan izin berdasarkan atribut dengan otorisasi ABAC.

Untuk informasi tentang batasan, pertimbangan, dan AWS Wilayah yang didukung, lihatPertimbangan kontrol akses berbasis atribut, batasan, dan wilayah yang didukung.

Topik

- Prasyarat untuk memberikan izin menggunakan atribut
- Memberikan izin menggunakan kontrol akses berbasis atribut

Prasyarat untuk memberikan izin menggunakan atribut

Untuk memberikan izin menggunakan kontrol akses berbasis atribut (ABAC), Anda harus menyelesaikan prasyarat berikut:

- Perbarui pengaturan Katalog Data untuk mengaktifkan izin Lake Formation untuk objek Katalog Data. Untuk informasi selengkapnya, lihat bagian Ubah model izin default atau gunakan mode akses hibrid.
- Setel pengaturan versi lintas-akun ke dua atau lebih tinggi.
- Lampirkan atribut ke entitas IAM yang memerlukan akses.
- Hanya administrator data lake atau pengguna IAM dengan izin yang diperlukan yang dapat memberikan akses pada objek Katalog Data. Untuk informasi selengkapnya tentang izin yang diperlukan, lihat Izin IAM.

Memberikan izin menggunakan kontrol akses berbasis atribut

Topik ini menjelaskan langkah-langkah yang perlu Anda ikuti untuk memberikan izin akses berbasis atribut pada sumber daya Katalog Data. Anda dapat menggunakan konsol Lake Formation atau AWS Command Line Interface (AWS CLI).

Pemberian izin menggunakan ABAC (Konsol Manajemen AWS)

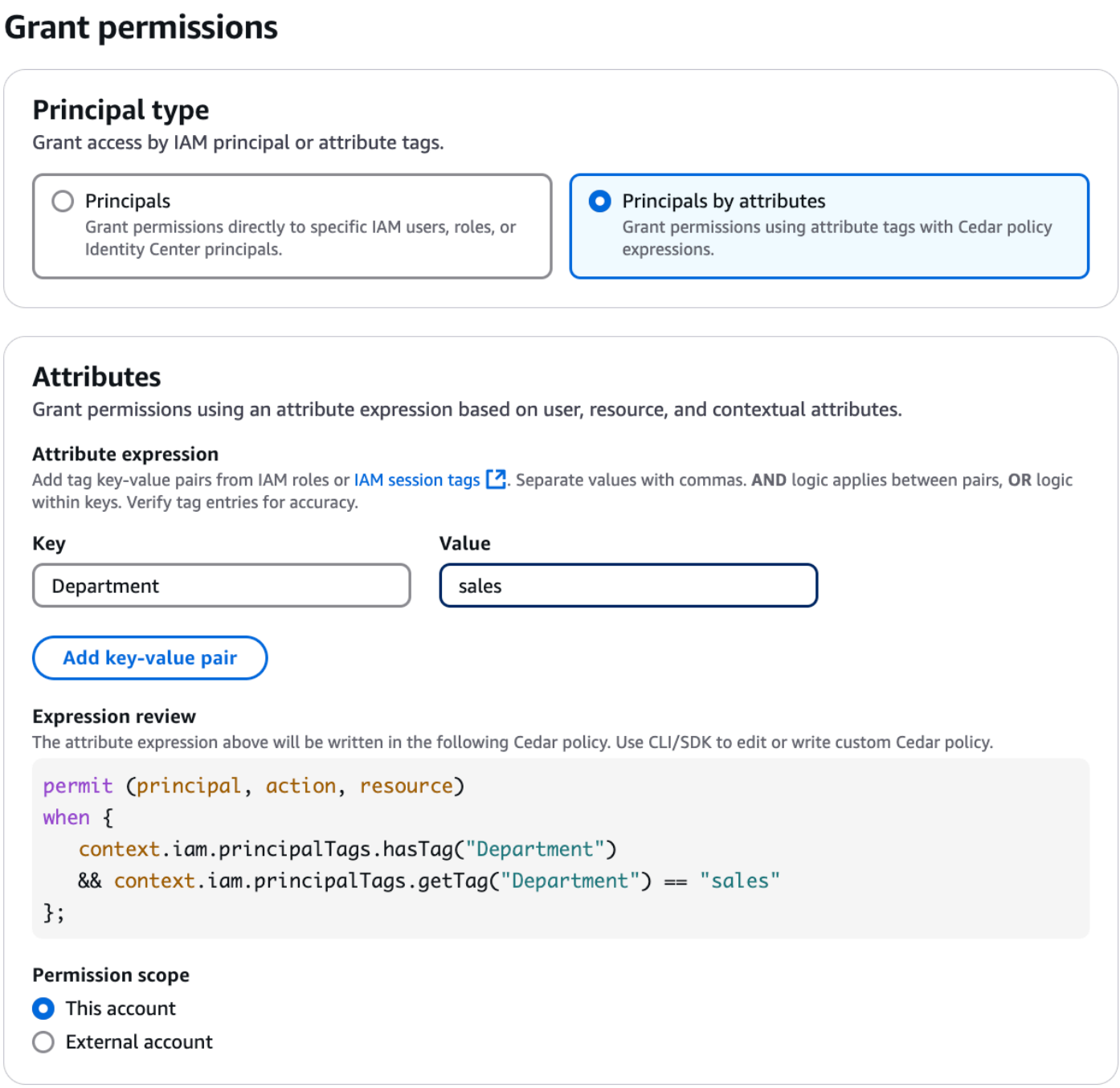
1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/, dan masuk sebagai administrator data lake, pembuat sumber daya, atau pengguna IAM yang memiliki izin Grantable pada sumber daya.
2. Lakukan salah satu tindakan berikut:

- Di panel navigasi, di bawah Izin, pilih Izin danau data. Kemudian pilih Grant.
- Di panel navigasi, pilih Katalog di bawah Katalog Data. Kemudian, pilih objek katalog (katalog, database, tabel, dan filter data), dan dari menu Tindakan di bawah Izin, dan pilih Hibah.

3. Pada halaman Hibah izin, pilih Prinsip berdasarkan atribut.
4. Tentukan kunci atribut dan nilai (s). Jika Anda memilih lebih dari satu nilai, Anda membuat ekspresi atribut dengan `OR` operator. Ini berarti bahwa jika salah satu nilai tag atribut yang ditetapkan ke peran IAM atau kecocokan pengguna, role/user memperoleh izin akses pada sumber daya.

   Jika Anda menentukan lebih dari satu tag atribut, Anda membuat ekspresi atribut dengan `AND` operator. Prinsipal diberikan izin pada sumber daya Katalog Data hanya jika IAM role/user diberi tag yang cocok untuk setiap tag atribut dalam ekspresi atribut.

   Tinjau ekspresi kebijakan Cedar yang dihasilkan yang ditampilkan di konsol.



5. Pilih ruang lingkup izin. Jika penerima hibah milik akun eksternal, pilih Akun eksternal dan masukkan ID AWS akun.
6. Selanjutnya, pilih akun Katalog Data atau di akun eksternal. Anda harus memiliki izin yang dapat diberikan terkait pada sumber daya agar berhasil menyelesaikan hibah izin.
7. Tentukan tindakan mana yang ingin Anda izinkan untuk prinsipal (pengguna atau peran) yang memiliki atribut yang cocok. Akses diberikan kepada entitas IAM yang telah diberi tag dan nilai

yang cocok dengan setidaknya salah satu ekspresi atribut yang Anda tentukan. Tinjau ekspresi kebijakan Cedar di konsol. Untuk informasi lebih lanjut tentang Cedar, lihat Apa itu Cedar? | Referensi GuideLink Bahasa Kebijakan Cedar.

8. Selanjutnya pilih sumber daya Katalog Data untuk memberikan akses. Anda dapat menentukan izin ini untuk berbagai sumber Katalog Data, termasuk katalog, database, tabel, dan filter data.
9. PilihIzin.

   Pendekatan ini memungkinkan Anda untuk mengontrol akses berdasarkan atribut, memastikan bahwa hanya pengguna atau peran dengan tag yang sesuai yang dapat melakukan tindakan spesifik pada sumber daya yang ditunjuk.

Pemberian izin menggunakan ABAC (AWS CLI)

Contoh berikut menunjukkan ekspresi atribut yang harus dipenuhi untuk menerima semua izin yang tersedia pada sumber daya. Anda juga dapat menentukan izin individual sepertiSelect,Describe, atauDrop. Ekspresi menggunakan ekspresi kebijakan Cedar. Untuk informasi lebih lanjut tentang Cedar, lihat Apa itu Cedar? | Referensi GuideLink Bahasa Kebijakan Cedar.

Kondisi ini memeriksa apakah prinsipal IAM memiliki department tag, dan nilai department tag samasales.

```
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\""}'
```

Skenario contoh izin

Skenario berikut membantu menunjukkan bagaimana Anda dapat mengatur izin untuk mengamankan akses ke data. AWS Lake Formation

Shirley adalah administrator data. Dia ingin mendirikan danau data untuk perusahaannya AnyCompany. Saat ini, semua data disimpan di Amazon S3. John adalah manajer pemasaran dan membutuhkan akses tulis ke informasi pembelian pelanggan (terkandung dalams3://customerPurchases). Seorang analis pemasaran, Diego, bergabung dengan John musim panas

ini. John membutuhkan kemampuan untuk memberikan Diego akses untuk melakukan kueri pada data tanpa melibatkan Shirley.

Mateo, dari keuangan, membutuhkan akses ke data akuntansi kueri (misalnya,`s3://transactions`). Dia ingin menanyakan data transaksi dalam tabel dalam database (`Finance_DB`) yang digunakan tim keuangan. Manajernya, Arnav, dapat memberinya akses ke. `Finance_DB` Meskipun ia seharusnya tidak dapat memodifikasi data akuntansi, ia membutuhkan kemampuan untuk mengubah data menjadi format (skema) yang cocok untuk peramalan. Data ini akan disimpan dalam bucket terpisah (`s3://financeForecasts`) yang dapat dimodifikasi.

Untuk meringkas:

- Shirley adalah administrator data lake.
- John membutuhkan `CREATE_DATABASE` dan `CREATE_TABLE` izin untuk membuat database dan tabel baru di Katalog Data.
- John juga membutuhkan`SELECT`,`INSERT`, dan `DELETE` izin pada tabel yang dia buat.
- Diego memerlukan `SELECT` izin di atas meja untuk menjalankan kueri.

Karyawan AnyCompany melakukan tindakan berikut untuk mengatur izin. Operasi API yang ditunjukkan dalam skenario ini menunjukkan sintaks yang disederhanakan untuk kejelasan.

1. Shirley mendaftarkan jalur Amazon S3 yang berisi informasi pembelian pelanggan dengan Lake Formation.

   ```
   RegisterResource(ResourcePath("s3://customerPurchases"), false, Role_ARN )
   ```

2. Shirley memberi John akses ke jalur Amazon S3 yang berisi informasi pembelian pelanggan.

   ```
   GrantPermissions(John, S3Location("s3://customerPurchases"),
    [DATA_LOCATION_ACCESS]) )
   ```

3. Shirley memberi John izin untuk membuat database.

   ```
   GrantPermissions(John, catalog, [CREATE_DATABASE])
   ```

4. John membuat database`John_DB`. John secara otomatis memiliki `CREATE_TABLE` izin pada database itu karena dia membuatnya.

```
CreateDatabase(John_DB)
```

5. John menciptakan tabel yang `John_Table` menunjuk kes`3://customerPurchases`. Karena dia membuat tabel, dia memiliki semua izin di atasnya, dan dapat memberikan izin di atasnya.

```
CreateTable(John_DB, John_Table)
```

6. John mengizinkan analisnya, Diego, akses ke meja`John_Table`.

```
 GrantPermissions(Diego, John_Table, [SELECT])
```

7. John mengizinkan analisnya, Diego, akses ke. `s3://customerPurchases/London/` Karena Shirley sudah terdaftars`3://customerPurchases`, subfoldernya terdaftar di Lake Formation.

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, [DATA_LOCATION_ACCESS], [],
S3Location("s3://customerPurchases/London/") )
```

8. John memungkinkan analisnya, Diego, untuk membuat tabel dalam database`John_DB`.

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, John_DB, [CREATE_TABLE],
[] )
```

9. Diego membuat tabel `John_DB` di `s3://customerPurchases/London/` at dan secara otomatis mendapatkan`ALTER`,,`DROP`,, `SELECT``INSERT`, dan `DELETE` izin.

```
CreateTable( 123456789012/datalake, John_DB, Diego_Table )
```

Pemfilteran data dan keamanan tingkat sel di Lake Formation

Saat Anda memberikan izin Lake Formation pada tabel Katalog Data, Anda dapat menyertakan spesifikasi pemfilteran data untuk membatasi akses ke data tertentu dalam hasil kueri dan mesin yang terintegrasi dengan Lake Formation. Lake Formation menggunakan pemfilteran data untuk mencapai keamanan tingkat kolom, keamanan tingkat baris, dan keamanan tingkat sel. Anda dapat menentukan dan menerapkan filter data pada kolom bersarang jika data sumber Anda berisi struktur bersarang.

Dengan kemampuan penyaringan data Lake Formation, Anda dapat menerapkan tingkat keamanan data berikut.

Column-level keamanan

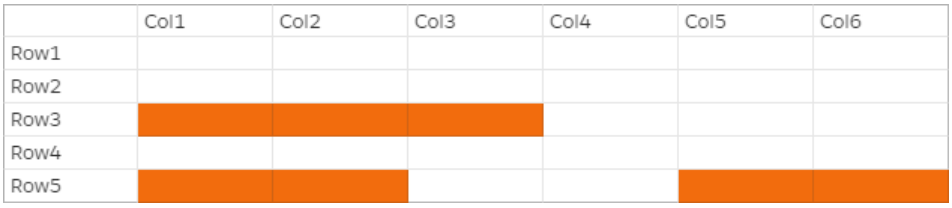
Pemberian izin pada tabel Katalog Data dengan keamanan tingkat kolom (pemfilteran kolom) memungkinkan pengguna untuk hanya melihat kolom tertentu dan kolom bersarang yang dapat mereka akses dalam tabel. Pertimbangkan `persons` tabel yang digunakan dalam beberapa aplikasi untuk perusahaan komunikasi multi-wilayah besar. Pemberian izin pada tabel Katalog Data dengan pemfilteran kolom dapat membatasi pengguna yang tidak bekerja di departemen SDM untuk melihat informasi identitas pribadi (PII) seperti nomor jaminan sosial atau tanggal lahir. Anda juga dapat menentukan kebijakan keamanan dan memberikan akses ke hanya sebagian sub-struktur kolom bersarang.

Row-level keamanan

Pemberian izin pada tabel Katalog Data dengan keamanan tingkat baris (pemfilteran baris) memungkinkan pengguna untuk melihat hanya baris data tertentu yang dapat mereka akses dalam tabel. Pemfilteran didasarkan pada nilai satu atau lebih kolom. Anda dapat menyertakan struktur kolom bersarang saat mendefinisikan ekspresi baris-filter. Misalnya, jika kantor regional yang berbeda dari perusahaan komunikasi memiliki departemen SDM mereka sendiri, Anda dapat membatasi catatan orang yang dapat dilihat karyawan SDM hanya untuk catatan untuk karyawan di wilayah mereka.

Cell-level keamanan

Cell-level keamanan menggabungkan pemfilteran baris dan pemfilteran kolom untuk model izin yang sangat fleksibel. Jika Anda melihat baris dan kolom tabel sebagai kisi, dengan menggunakan keamanan tingkat sel, Anda dapat membatasi akses ke elemen individual (sel) kisi di mana saja dalam dua dimensi. Artinya, Anda dapat membatasi akses ke kolom yang berbeda tergantung pada baris. Ini diilustrasikan oleh diagram berikut, di mana kolom terbatas diarsir.



Melanjutkan contoh tabel orang, Anda dapat membuat filter data di tingkat sel yang membatasi akses ke kolom alamat jalan jika baris memiliki kolom negara yang disetel ke “UK”, tetapi mengizinkan akses ke kolom alamat jalan jika baris memiliki kolom negara yang disetel ke “AS”.

Filter hanya berlaku untuk operasi baca. Oleh karena itu, Anda hanya dapat memberikan izin `SELECT` Lake Formation dengan filter.

Cell-level keamanan pada kolom bersarang

Lake Formation memungkinkan Anda untuk menentukan dan menerapkan filter data dengan keamanan tingkat sel pada kolom bersarang. Namun, mesin analitik terintegrasi seperti Amazon Athena, Amazon EMR, dan Amazon Redshift Spectrum mendukung eksekusi kueri terhadap tabel bersarang yang dikelola Lake Formation dengan keamanan tingkat baris dan kolom.

Untuk batasan, lihat Batasan penyaringan data.

Topik

- Filter data di Lake Formation
- Dukungan PartiQL dalam ekspresi filter baris
- Izin diperlukan untuk menanyakan tabel dengan pemfilteran tingkat sel
- Mengelola filter data

Filter data di Lake Formation

Anda dapat menerapkan keamanan tingkat kolom, tingkat baris, dan tingkat sel dengan membuat filter data. Anda memilih filter data saat Anda memberikan izin `SELECT` Lake Formation pada tabel. Jika tabel berisi struktur kolom bersarang, Anda dapat menentukan filter data dengan menyertakan atau mengecualikan kolom turunan dan menentukan ekspresi filter tingkat baris pada atribut bersarang.

Setiap filter data milik tabel tertentu dalam Katalog Data Anda. Filter data mencakup informasi berikut:

- Nama filter
- ID Katalog dari tabel yang terkait dengan filter
- Nama tabel
- Nama database yang berisi tabel
- Spesifikasi kolom - daftar kolom dan kolom bersarang (dengan `struct` tipe data) untuk menyertakan atau mengecualikan dalam hasil kueri.
- Ekspresi filter baris - ekspresi yang menentukan baris untuk disertakan dalam hasil kueri. Dengan beberapa batasan, ekspresi memiliki sintaks `WHERE` klausa dalam bahasa PartiQL. Untuk

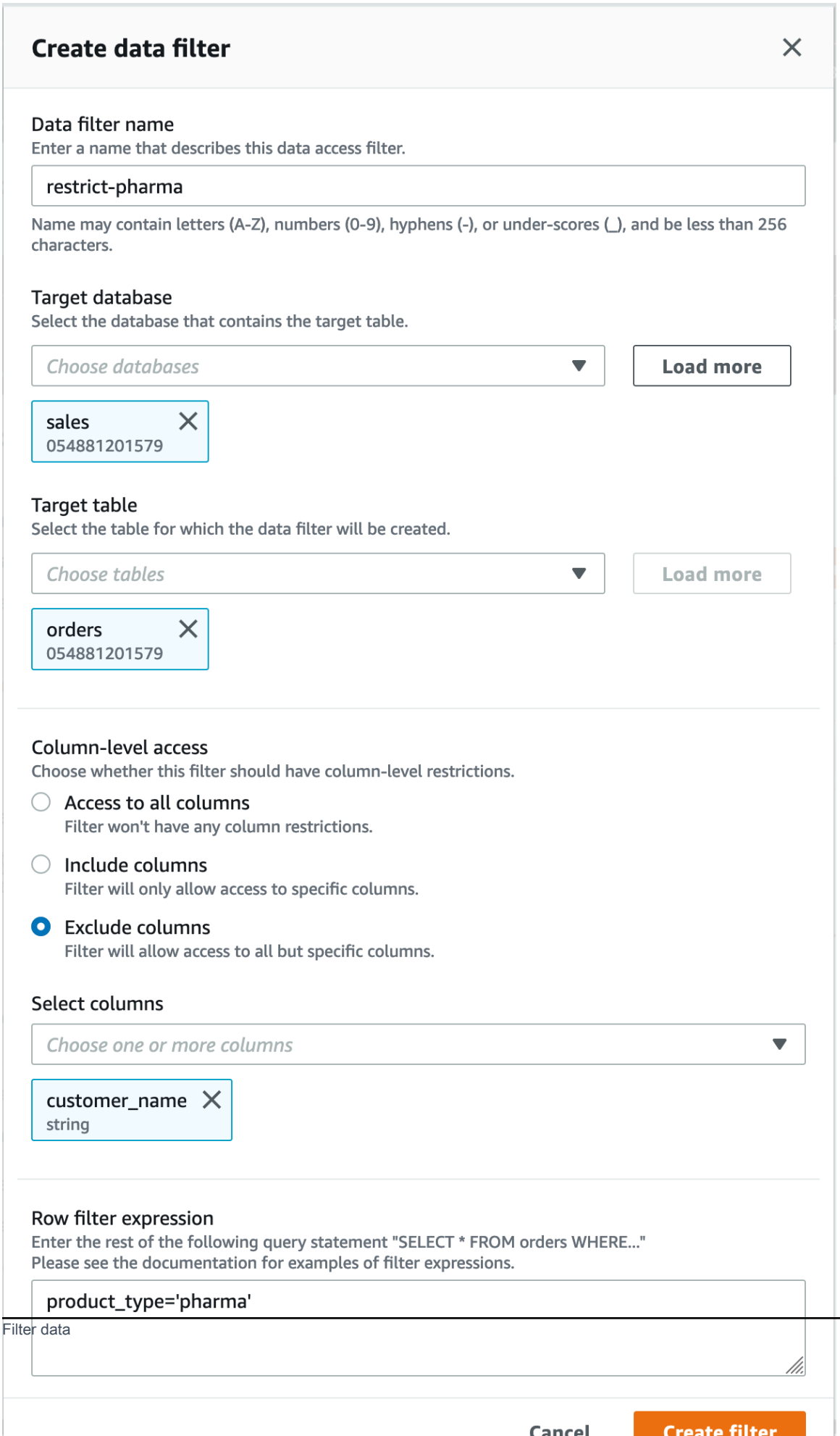
menentukan semua baris, pilih Akses ke semua baris di bawah Row-level akses di konsol atau gunakan `AllRowsWildcard` dalam panggilan API.

Untuk informasi selengkapnya tentang apa yang didukung dalam ekspresi filter baris, lihatDukungan PartiQL dalam ekspresi filter baris.

Tingkat penyaringan yang Anda dapatkan tergantung pada bagaimana Anda mengisi filter data.

- Saat Anda menentukan wildcard “semua kolom” dan memberikan ekspresi filter baris, Anda hanya membuat keamanan tingkat baris (pemfilteran baris).
- Saat Anda menyertakan atau mengecualikan kolom tertentu dan kolom bersarang, dan menentukan “semua baris” menggunakan wildcard semua baris, Anda hanya membuat keamanan tingkat kolom (pemfilteran kolom).
- Saat Anda menyertakan atau mengecualikan kolom tertentu dan juga memberikan ekspresi filter baris, Anda membangun keamanan tingkat sel (penyaringan sel).

Tangkapan layar berikut dari konsol Lake Formation menunjukkan filter data yang melakukan pemfilteran tingkat sel. Untuk kueri terhadap `orders` tabel, ini membatasi akses ke `customer_name` kolom dan hasil kueri hanya mengembalikan baris di mana `product_type` kolom berisi 'pharma'.


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
Row filter expression
Enter the rest of the following query statement "SELECT * FROM orders WHERE..."
Please see the documentation for examples of filter expressions.
product_type='pharma'
Cancel
Create filter

Perhatikan penggunaan tanda kutip tunggal untuk melampirkan string literal, `'pharma'`.

Anda dapat menggunakan konsol Lake Formation untuk membuat filter data ini, atau Anda dapat menyediakan objek permintaan berikut ke operasi `CreateDataCellsFilter` API.

```
{
     "Name": "restrict-pharma",
     "DatabaseName": "sales",
     "TableName": "orders",
     "TableCatalogId": "111122223333",
     "RowFilter": {"FilterExpression": "product_type='pharma'"},
     "ColumnWildcard": {
         "ExcludedColumnNames": ["customer_name"]
     }
}
```

Anda dapat membuat filter data sebanyak yang Anda butuhkan untuk sebuah tabel. Untuk melakukannya, Anda memerlukan `SELECT` izin dengan opsi hibah di atas meja. Administrator Data Lake secara default memiliki izin untuk membuat filter data pada semua tabel di akun itu. Anda biasanya hanya menggunakan subset dari filter data yang mungkin saat memberikan izin pada tabel kepada prinsipal. Misalnya, Anda dapat membuat filter data kedua untuk `orders` tabel yang merupakan filter data hanya keamanan baris. Mengacu pada tangkapan layar sebelumnya, Anda dapat memilih opsi Akses ke semua kolom dan menyertakan ekspresi filter baris dari. `product_type<>pharma` Nama filter data ini bisa jadi`no-pharma`. Ini membatasi akses ke semua baris yang memiliki `product_type` kolom diatur ke 'pharma'.

Objek permintaan untuk operasi `CreateDataCellsFilter` API untuk filter data ini adalah sebagai berikut.

```
{
     "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"]
}
```

Anda kemudian dapat memberikan SELECT pada `orders` tabel dengan filter `restrict-pharma` data ke pengguna administratif, dan SELECT pada `orders` tabel dengan filter `no-pharma` data untuk pengguna non-administratif. Untuk pengguna di sektor perawatan kesehatan, Anda akan memberikan SELECT di `orders` atas meja dengan akses penuh ke semua baris dan kolom (tanpa filter data), atau mungkin dengan filter data lain yang membatasi akses ke informasi harga.

Anda dapat menyertakan atau mengecualikan kolom bersarang saat menentukan keamanan tingkat kolom dan tingkat baris dalam filter data. Dalam contoh berikut, akses ke `product.offer` bidang ditentukan menggunakan nama kolom yang memenuhi syarat (dibungkus dengan tanda kutip ganda). Hal ini penting untuk bidang bersarang untuk menghindari kesalahan yang terjadi ketika nama kolom berisi karakter khusus, dan untuk mempertahankan kompatibilitas mundur dengan definisi keamanan tingkat kolom tingkat atas.

```
{
     "Name": "example_dcf",
     "DatabaseName": "example_db",
     "TableName": "example_table",
     "TableCatalogId": "111122223333",
     "RowFilter": { "FilterExpression": "customer.customerName <> 'John'" },
     "ColumnNames": ["customer", "\"product\".\"offer\""]
}
```

Lihat juga

- Mengelola filter data

Dukungan PartiQL dalam ekspresi filter baris

Anda dapat membuat ekspresi filter baris menggunakan subset tipe data, operator, dan agregasi PartiQL. Lake Formation tidak mengizinkan fungsi PartiQL yang ditentukan pengguna atau standar dalam ekspresi filter. Anda dapat menggunakan operator perbandingan untuk membandingkan kolom dengan konstanta (misalnya,`views >= 10000`), tetapi Anda tidak dapat membandingkan kolom dengan kolom lain.

Ekspresi filter baris dapat berupa ekspresi sederhana atau ekspresi komposit. Total panjang ekspresi harus kurang dari 2048 karakter.

Ekspresi sederhana

Ekspresi sederhana akan berupa format: `<column name > <comparison operator ><value >`

- Nama kolom

  Ini bisa berupa kolom data tingkat atas, kolom partisi, atau kolom bersarang yang ada dalam skema tabel dan harus termasuk dalam yang Jenis data yang didukung tercantum di bawah ini.
- Operator perbandingan

  Berikut ini adalah operator yang didukung: `=, >, <, >=, <=, <>,!=, BETWEEN, IN, LIKE, NOT, IS [NOT] NULL`
- Semua perbandingan string dan kecocokan `LIKE` pola peka huruf besar/kecil. Anda tidak dapat menggunakan operator IS [NOT] NULL pada kolom partisi.
- Nilai kolom

  Nilai Kolom harus sesuai dengan tipe data dari nama kolom.

Ekspresi komposit

Ekspresi komposit akan berupa format:`( <simple expression >) <AND/OR >(<simple expression >)`. Ekspresi komposit dapat digabungkan lebih lanjut menggunakan operator logis`AND/OR`.

Jenis data yang didukung

Filter baris yang merujuk ke AWS Glue Data Catalog tabel yang berisi tipe data yang tidak didukung akan menghasilkan kesalahan. Berikut ini adalah tipe data yang didukung untuk kolom tabel dan konstanta, yang dipetakan ke tipe Amazon Redshift data:

- `STRING, CHAR, VARCHAR`
- `INT, LONG, BIGINT, FLOAT, DECIMAL, DOUBLE`
- `BOOLEAN`
- `STRUCT`

Untuk informasi selengkapnya tentang tipe data di Amazon Redshift, lihat Jenis data di Panduan Pengembang Database Amazon Redshift.

Ekspresi filter baris

Example

Berikut ini adalah contoh ekspresi filter baris yang valid untuk tabel dengan kolom: `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

Berikut ini adalah contoh yang valid dari ekspresi baris filter untuk tabel dengan kolom bersarang: `year > 2010 and customer.customerId <> 1`

Bidang bersarang di bawah kolom partisi tidak boleh direferensikan saat mendefinisikan ekspresi tingkat baris bersarang.

Konstanta string harus diapit dalam tanda kutip tunggal.

Kata kunci terpesan

Jika ekspresi filter baris Anda berisi kata kunci PartiQL, Anda akan menerima kesalahan penguraian karena nama kolom mungkin bertentangan dengan kata kunci. Ketika ini terjadi, lepaskan nama kolom dengan menggunakan tanda kutip ganda. Beberapa contoh kata kunci yang dicadangkan adalah “pertama”, “terakhir”, “asc”, “hilang”. Lihat spesifikasi PartiQL untuk daftar kata kunci yang dicadangkan.

Referensi PartiQL

Untuk informasi lebih lanjut tentang PartiQL, lihat. https://partiql.org/

Izin diperlukan untuk menanyakan tabel dengan pemfilteran tingkat sel

Izin berikut AWS Identity and Access Management (IAM) diperlukan untuk menjalankan kueri terhadap tabel dengan pemfilteran tingkat sel.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:StartQueryPlanning",
                "lakeformation:GetQueryState",
                "lakeformation:GetWorkUnits",
                "lakeformation:GetWorkUnitResults"
            ],
            "Resource": "*"
        }
    ]
}
```

Untuk informasi selengkapnya tentang izin Lake Formation, lihatReferensi personas Lake Formation dan izin IAM.

Mengelola filter data

Untuk menerapkan keamanan tingkat kolom, tingkat baris, dan tingkat sel, Anda dapat membuat dan memelihara filter data. Setiap filter data milik tabel Katalog Data. Anda dapat membuat beberapa filter data untuk tabel, dan kemudian menggunakan satu atau lebih dari mereka saat memberikan izin pada tabel. Anda juga dapat menentukan dan menerapkan filter data pada kolom bersarang yang memiliki `struct` tipe data yang memungkinkan pengguna mengakses hanya sub-struktur kolom bersarang.

Anda memerlukan `SELECT` izin dengan opsi hibah untuk membuat atau melihat filter data. Untuk mengizinkan prinsipal di akun Anda melihat dan menggunakan filter data, Anda dapat memberikan `DESCRIBE` izin di dalamnya.

> **Note**
>
> Lake Formation tidak mendukung pemberian `Describe` izin pada filter data, yang dibagikan dari akun lain.

Anda dapat mengelola filter data menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Untuk informasi tentang filter data, lihat Filter data di Lake Formation

Membuat filter data

Anda dapat membuat satu atau beberapa filter data untuk setiap tabel Katalog Data.

Untuk membuat filter data untuk tabel Katalog Data (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator danau data, pemilik tabel target, atau kepala sekolah yang memiliki izin Lake Formation pada tabel target.
2. Di panel navigasi, di bawah Katalog data, pilih Filter data.
3. Pada halaman Filter data, pilih Buat filter baru.
4. Dalam Buat filter data kotak dialog, masukkan informasi berikut:

   - Nama filter data
   - Database target - Tentukan database yang berisi tabel.
   - Tabel target
   - Column-level akses - Biarkan set ini ke Akses ke semua kolom untuk menentukan pemfilteran baris saja. Pilih Sertakan kolom atau Kecualikan kolom untuk menentukan pemfilteran kolom atau sel, lalu tentukan kolom yang akan disertakan atau dikecualikan.

     Kolom bersarang - Jika Anda menerapkan filter pada tabel yang berisi kolom bersarang, Anda dapat secara eksplisit menentukan sub-struktur kolom struct bersarang dalam filter data.

     Ketika Anda memberikan izin SELECT kepada kepala sekolah pada filer ini, prinsipal yang menjalankan kueri berikut, hanya akan melihat data untuk `customer.customerName` dan tidak. `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'
```

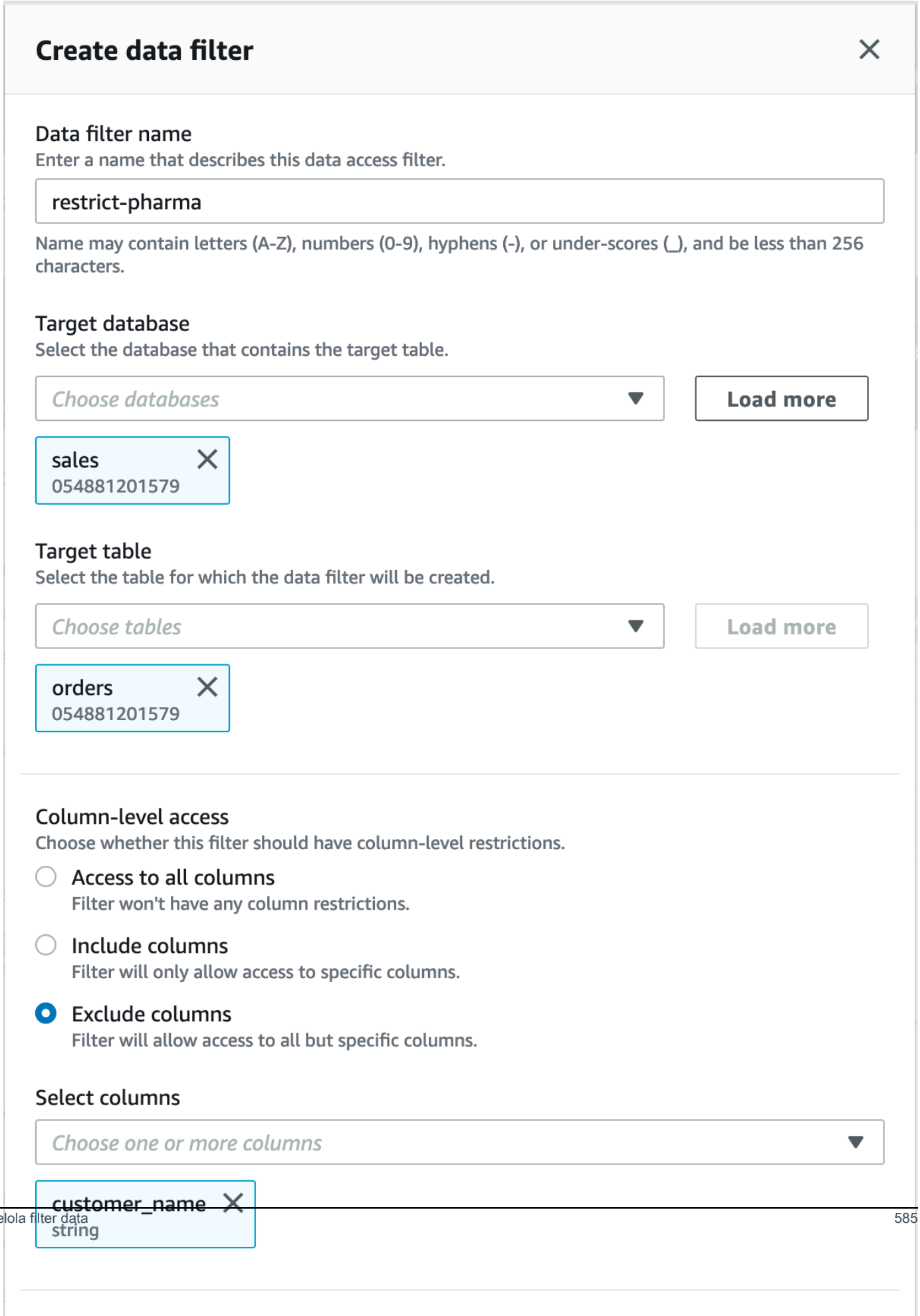
Saat Anda memberikan izin ke `customer` kolom, prinsipal menerima akses ke kolom dan bidang bersarang di bawah kolom (`customerName`dan`customerID`).

- Ekspresi filter baris - Masukkan ekspresi filter untuk menentukan pemfilteran baris atau sel. Untuk tipe dan operator data yang didukung, lihatDukungan PartiQL dalam ekspresi filter baris. Pilih Akses ke semua baris untuk memberikan akses ke semua.

  Anda dapat menyertakan struct kolom sebagian dari kolom bersarang dalam ekspresi filter baris untuk memfilter baris yang berisi nilai tertentu.

  Ketika prinsipal diberikan izin ke tabel dengan ekspresi filter baris`Select * from example_nestedtable where customer.customerName <>'John'`, dan Column-levelakses diatur ke Akses ke semua kolom, hasil kueri hanya menampilkan baris yang `customerName <>'John'` mengevaluasi ke true.

Tangkapan layar berikut menunjukkan filter data yang mengimplementasikan penyaringan sel. Dalam kueri terhadap `orders` tabel, ia menolak akses ke `customer_name` kolom dan hanya menampilkan baris yang memiliki 'pharma' di kolom. `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. Pilih Buat filter.

Untuk membuat filter data dengan kebijakan filter sel pada bidang bersarang

Bagian ini menggunakan skema contoh berikut untuk menunjukkan cara membuat filter sel data:

```
[
    { 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. Pada Buat filter data, halaman masukkan nama untuk filter data.
2. Selanjutnya, gunakan drop-down untuk memilih nama database dan nama tabel.
3. Di bagian Column-level akses, pilih kolom Termasuk, dan pilih kolom bersarang (`customer.customerName`).
4. Di bagian Row-level akses, pilih opsi Akses ke semua baris.
5. Pilih Buat filter.

   Saat Anda memberikan `SELECT` izin pada filter ini, kepala sekolah mendapatkan akses ke semua baris di `customerName` kolom.
6. Selanjutnya, tentukan filter data lain untuk hal yang sama database/table.
7. Di bagian Column-level akses, pilih kolom Termasuk, dan pilih kolom bersarang lainnya (`customer.customerid`).
8. Di bagian Row-level akses, pilih Filter baris, dan masukkan ekspresi filter baris (`customer.customerid <> 5`).
9. Pilih Buat filter.

   Saat Anda memberikan `SELECT` izin pada filter ini, prinsipal menerima akses ke semua baris di`customerName`, dan `customerId` bidang kecuali sel yang nilainya 5 di `customerId` kolom.

Memberikan izin filter data

Anda dapat memberikan izin`SELECT`, `DESCRIBE` dan `DROP` Lake Formation pada filter data kepada kepala sekolah.

Pada awalnya, hanya Anda yang dapat melihat filter data yang Anda buat untuk sebuah tabel. Untuk mengaktifkan prinsipal lain untuk melihat filter data dan memberikan izin Katalog Data dengan filter data, Anda harus:

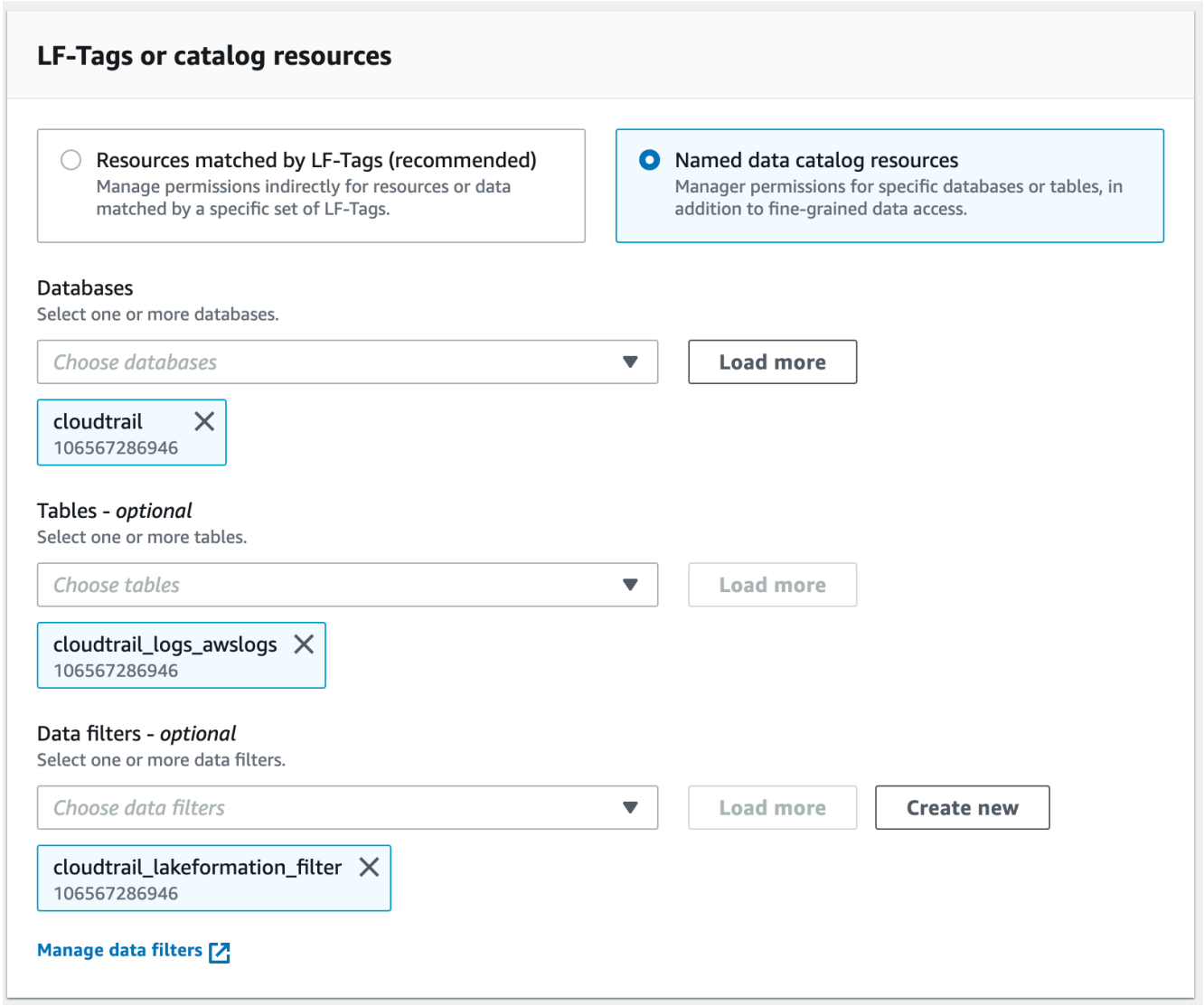
- Berikan `SELECT` pada tabel kepada kepala sekolah dengan opsi hibah, dan terapkan filter data ke hibah.
- Berikan `DESCRIBE` atau `DROP` izin pada filter data kepada kepala sekolah.

Anda dapat memberikan `SELECT` izin ke AWS akun eksternal. Administrator data lake di akun itu kemudian dapat memberikan izin itu kepada kepala sekolah lain di akun tersebut. Saat memberikan ke akun eksternal, Anda harus menyertakan opsi hibah sehingga administrator akun eksternal dapat lebih lanjut memasukkan izin ke pengguna lain di akun. his/her Saat memberikan kepada kepala sekolah di akun Anda, pemberian dengan opsi hibah adalah opsional.

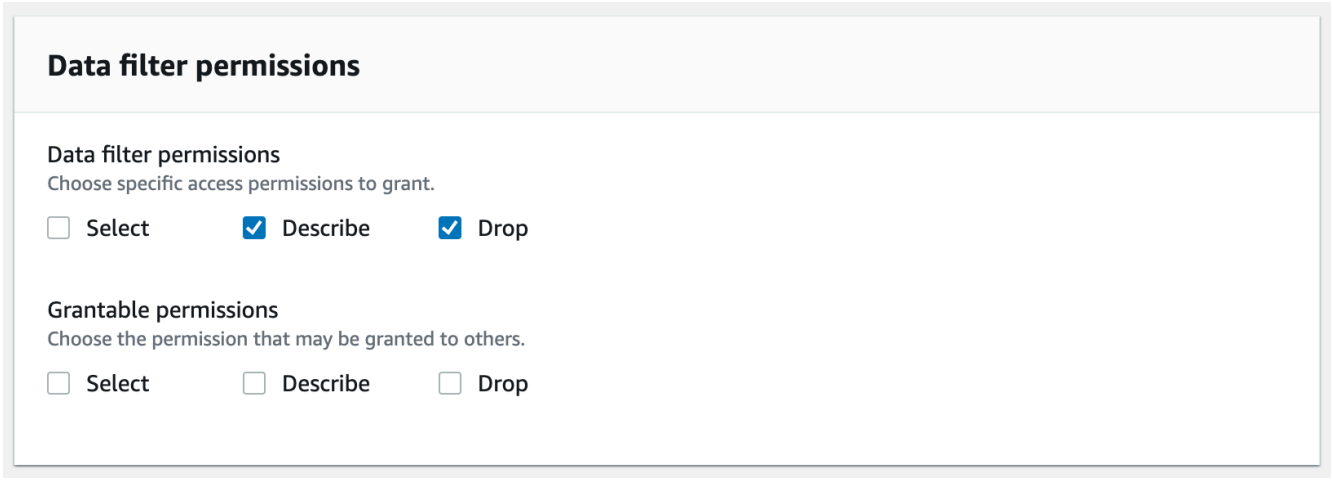
Anda dapat memberikan dan mencabut izin pada filter data menggunakan AWS Lake Formation konsol, API, atau (). AWS Command Line Interface AWS CLI

Console

1. Masuk ke Konsol Manajemen AWS dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, di bawah Izin, pilih Izin danau data.
3. Pada halaman Izin, di bagian Izin data, pilih Hibah.
4. Pada halaman Berikan izin data, pilih prinsipal untuk memberikan izin.
5. Di bagian sumber daya LF-Tags atau katalog, pilih Sumber daya katalog data bernama. Kemudian pilih database, tabel, dan filter data yang ingin Anda berikan izin.



6. Di bagian Izin filter data, pilih izin yang ingin Anda berikan kepada prinsipal yang dipilih.



AWS CLI

- Masukkan `grant-permissions` perintah. Tentukan `DataCellsFilter` untuk `resource` argumen, dan tentukan `DESCRIBE` atau `DROP` untuk `Permissions` argumen dan, secara opsional, untuk `PermissionsWithGrantOption` argumen.

  Contoh berikut memberikan opsi `DESCRIBE` hibah kepada pengguna `datalake_user1` pada filter data`restrict-pharma`, yang termasuk dalam `orders` tabel di `sales` database di AWS akun 1111-2222-3333.

  ```
  aws lakeformation grant-permissions --cli-input-json file://grant-params.json
  ```

  Berikut ini adalah isi 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"]
  }
  ```

Memberikan izin data yang disediakan oleh filter data

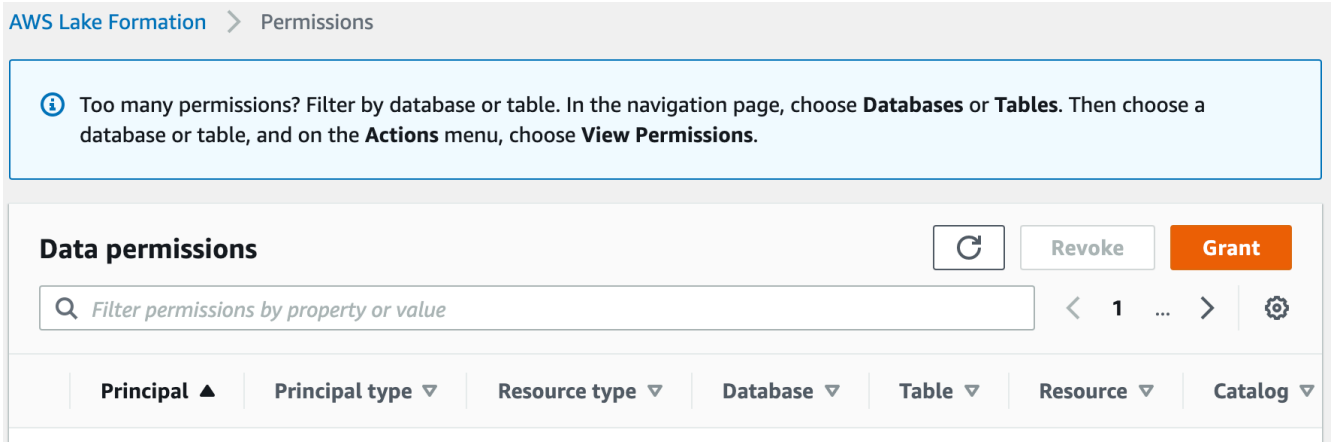
Filter data mewakili subset data dalam tabel. Untuk menyediakan akses data ke kepala sekolah, `SELECT` izin harus diberikan kepada prinsipal tersebut. Dengan izin ini para kepala sekolah dapat:

- Lihat nama tabel yang sebenarnya dalam daftar tabel yang dibagikan dengan akun mereka.
- Buat filter data pada tabel bersama dan berikan izin kepada penggunanya pada filter data tersebut.

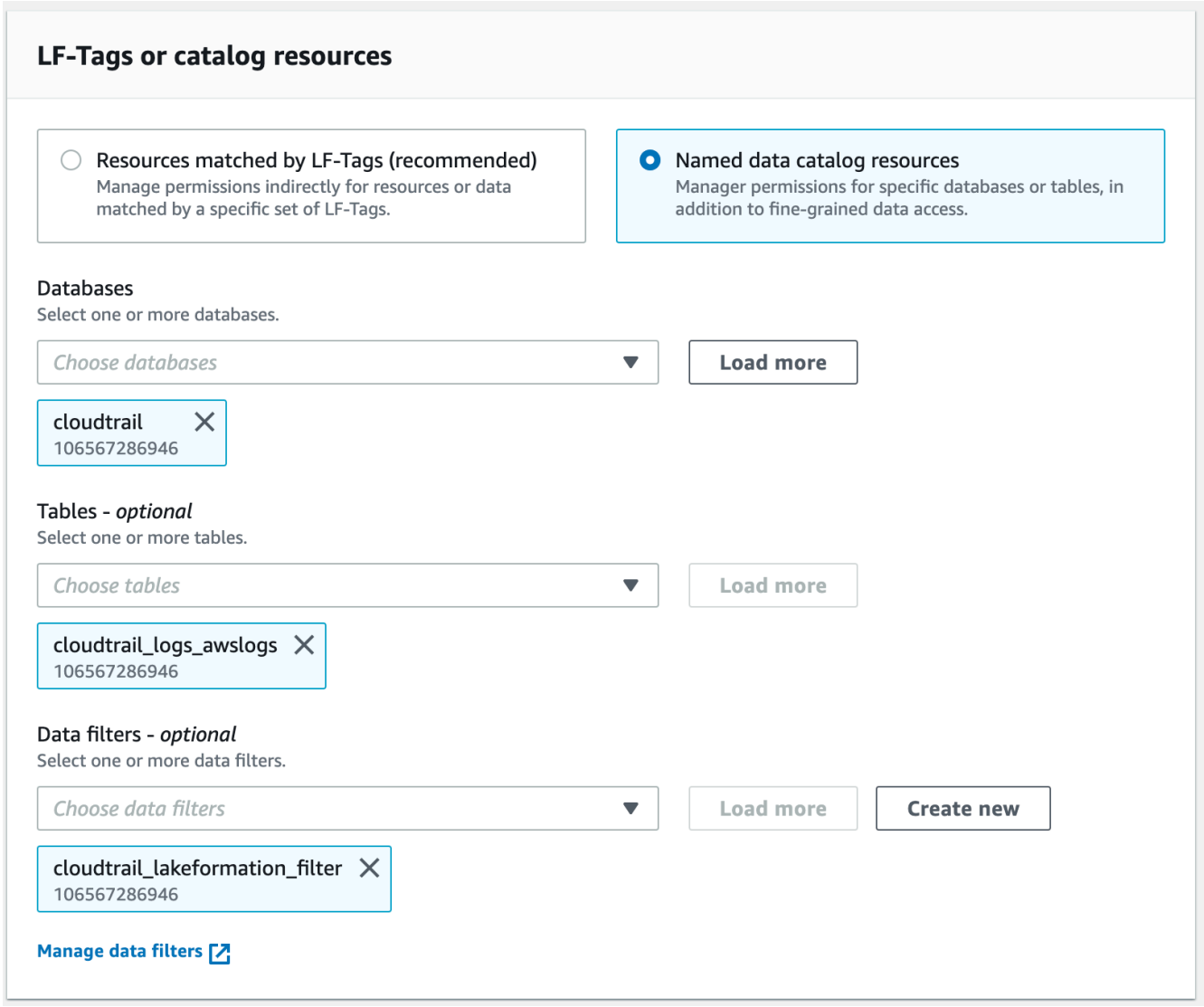
Console

Untuk memberikan izin SELECT

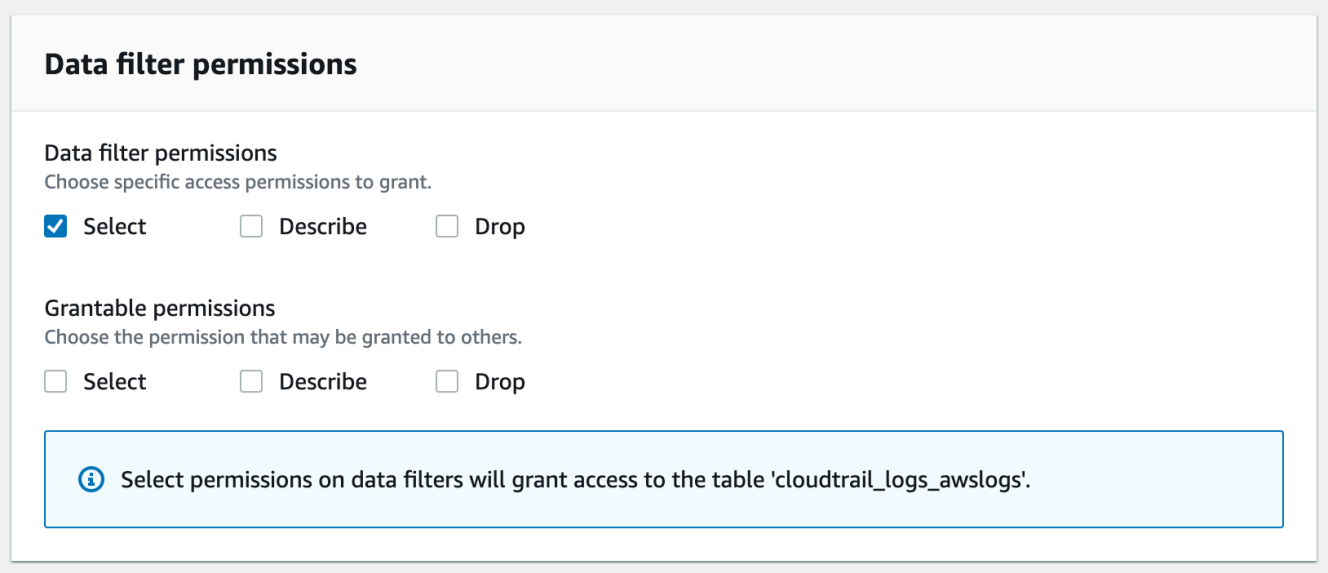
1. Buka halaman Izin di konsol Lake Formation, lalu pilih Grant.



2. Pilih prinsipal yang ingin Anda berikan aksesnya, dan pilih Sumber daya katalog data bernama.



3. Untuk memberikan akses ke data yang diwakili oleh filter, pilih Pilih di bawah Izin filter data.



CLI

Masukkan `grant-permissions` perintah. Tentukan `DataCellsFilter` argumen sumber daya, dan tentukan `SELECT` untuk argumen Izin.

Contoh berikut memberikan opsi hibah kepada pengguna `datalake_user1` pada filter data`restrict-pharma`, yang termasuk dalam `orders` tabel dalam `sales` database di Akun AWS `1111-2222-3333`. `SELECT`

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

Berikut ini adalah isi 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"]
}
```

Melihat filter data

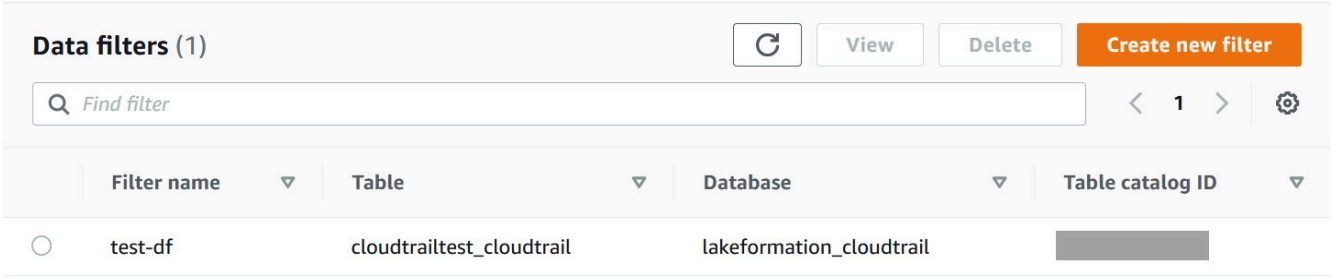
Anda dapat menggunakan konsol Lake Formation AWS CLI, atau Lake Formation API untuk melihat filter data.

Untuk melihat filter data, Anda harus menjadi administrator Data Lake atau memiliki izin yang diperlukan pada filter data.

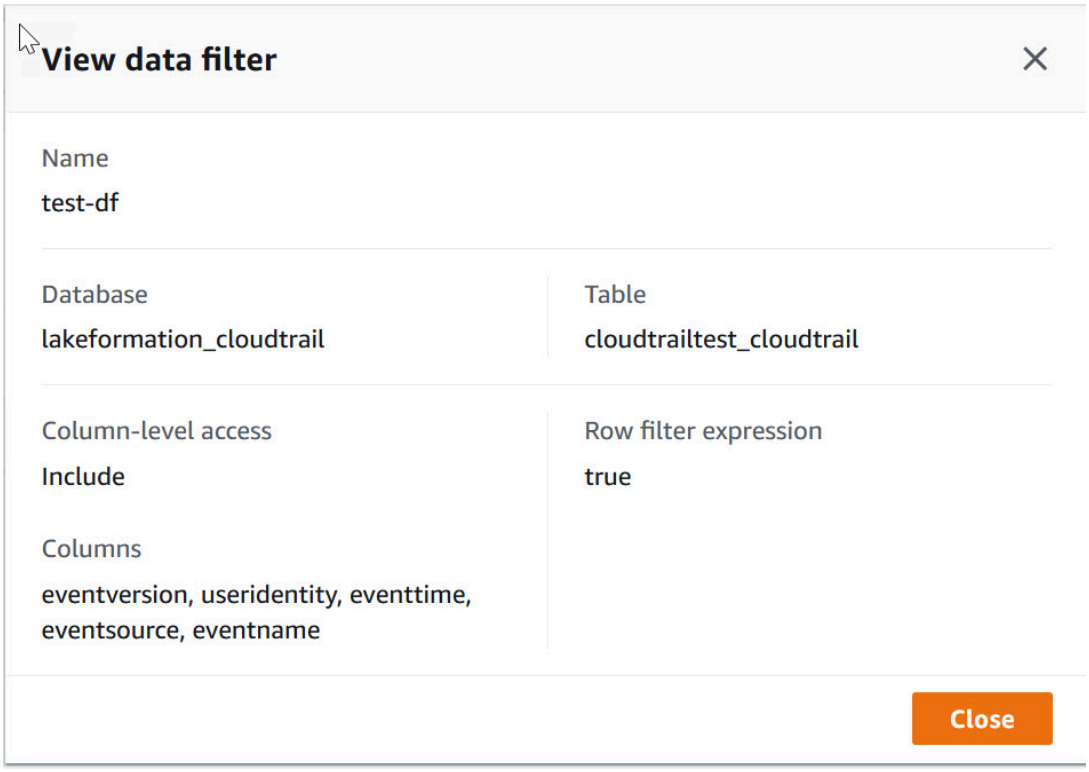
Console

1. Masuk ke Konsol Manajemen AWS dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, di bawah Katalog data, pilih Filter data.

   Halaman ini menampilkan filter data yang dapat Anda akses.



3. Untuk melihat detail filter data, pilih filter data, lalu pilih Lihat. Jendela baru muncul dengan informasi rinci filter data.



AWS CLI

Masukkan `list-data-cells-filter` perintah dan tentukan sumber daya tabel.

Contoh berikut mencantumkan filter data untuk `cloudtrailtest_cloudtrail` tabel.

```
aws lakeformation list-data-cells-filter --table '{ "CatalogId":"123456789012",
"DatabaseName":"lakeformation_cloudtrail", "Name":"cloudtrailtest_cloudtrail"}'
```

API/SDK

Gunakan `ListDataCellsFilter` API dan tentukan sumber daya tabel.

Contoh berikut menggunakan Python untuk daftar 20 filter data pertama untuk tabel. `myTable`

```
response = client.list_data_cells_filter(
    Table = {
        'CatalogId': '111122223333',
        'DatabaseName': 'mydb',
        'Name': 'myTable'
    },
```

```
    MaxResults=20
)
```

Izin filter data daftar

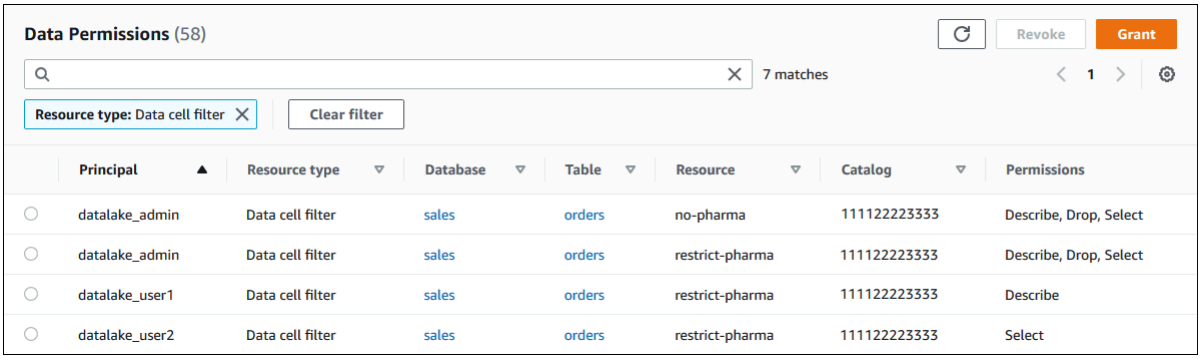
Anda dapat menggunakan konsol Lake Formation untuk melihat izin yang diberikan pada filter data.

Untuk melihat izin pada filter data, Anda harus menjadi administrator Data Lake atau memiliki izin yang diperlukan pada filter data.

Console

1. Masuk ke Konsol Manajemen AWS dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di panel navigasi, di bawah Izin, pilih Izin data.
3. Pada halaman Izin Data, klik atau ketuk di bidang pencarian, dan pada menu Properti, pilih Jenis sumber daya.
4. Pada menu Jenis sumber daya, pilih Jenis sumber daya: Filter sel data.

   Filter data yang Anda miliki izin terdaftar. Anda mungkin harus menggulir secara horizontal untuk melihat kolom Izin dan Grantable.



AWS CLI

- Masukkan `list-permissions` perintah. Tentukan `DataCellsFilter` untuk `resource` argumen, dan tentukan `DESCRIBE` atau `DROP` untuk `Permissions` argumen dan, secara opsional, untuk `PermissionsWithGrantOption` argumen.

Contoh berikut mencantumkan DESCRIBE izin dengan opsi hibah pada filter `restrict-pharma` data. Hasilnya terbatas pada izin yang diberikan untuk kepala sekolah `datalake_user1` dan `orders` tabel dalam `sales` database di AWS akun 1111-2222-3333.

```
aws lakeformation list-permissions --cli-input-json file://list-params.json
```

Berikut ini adalah isi 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"]
}
```

Melihat izin database dan tabel di Lake Formation

Anda dapat melihat izin Lake Formation yang diberikan pada database atau tabel Katalog Data. Anda dapat melakukannya dengan menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

Menggunakan konsol, Anda dapat melihat izin mulai dari halaman Database atau Tabel, atau dari halaman izin Data.

> **Note**
>
> Jika Anda bukan administrator database atau pemilik sumber daya, Anda dapat melihat izin yang dimiliki prinsipal lain pada sumber daya hanya jika Anda memiliki izin Lake Formation pada sumber daya dengan opsi hibah.

Selain izin Lake Formation yang diperlukan, Anda memerlukan izin AWS Identity and Access Management (IAM)`glue:GetDatabases`,, `glue:GetDatabaseglue:GetTables`, `glue:GetTable` dan. `lakeformation:ListPermissions`

Untuk melihat izin pada database (konsol, mulai dari halaman Database)

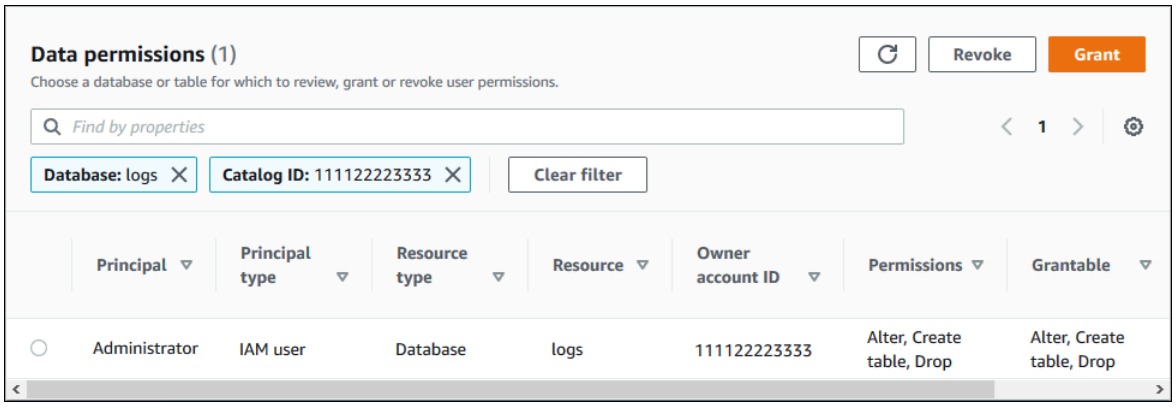
1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, pembuat database, atau sebagai pengguna yang memiliki izin Lake Formation di database dengan opsi hibah.

2. Di panel navigasi, pilih Basis data.
3. Pilih database, dan pada menu Tindakan, pilih Lihat izin.

> **Note**
>
> Jika Anda memilih tautan sumber daya basis data, Lake Formation menampilkan izin pada tautan sumber daya, bukan pada basis data target tautan sumber daya.

Halaman izin Data mencantumkan semua izin Lake Formation untuk database. Nama database dan ID katalog (ID AWS akun) pemilik database muncul sebagai label di bawah kotak pencarian. Ubin menunjukkan bahwa filter telah diterapkan untuk daftar izin hanya untuk database itu. Anda dapat menyesuaikan filter dengan menutup ubin atau memilih Clear filter.



Untuk melihat izin pada database (konsol, mulai dari halaman izin Data)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

Masuk sebagai administrator data lake, pembuat database, atau sebagai pengguna yang memiliki izin Lake Formation di database dengan opsi hibah.

2. Di panel navigasi, pilih Izin data.
3. Posisikan kursor di kotak pencarian, dan pada menu Properties yang muncul, pilih Database.
4. Pada menu Database yang muncul, pilih database.

> **Note**
>
> Jika Anda memilih tautan sumber daya basis data, Lake Formation menampilkan izin pada tautan sumber daya, bukan pada basis data target tautan sumber daya.

Halaman izin Data mencantumkan semua izin Lake Formation untuk database. Nama database muncul sebagai ubin di bawah kotak pencarian. Ubin menunjukkan bahwa filter telah diterapkan untuk daftar izin hanya untuk database itu. Anda dapat menghapus filter dengan menutup ubin atau memilih Hapus filter.

Untuk melihat izin pada tabel (konsol, mulai dari halaman Tabel)

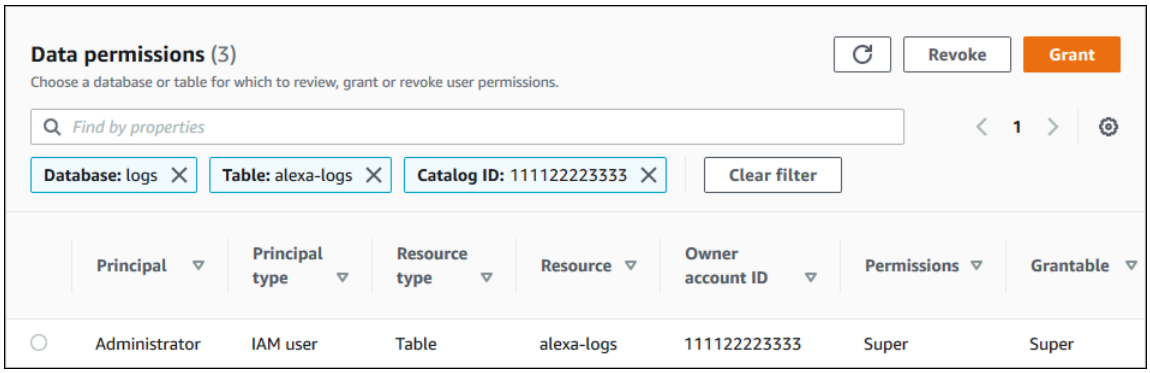
1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, pembuat tabel, atau sebagai pengguna yang memiliki izin Lake Formation di atas tabel dengan opsi hibah.
2. Di panel navigasi, pilih Tabel.
3. Pilih tabel, dan pada menu Tindakan, pilih Lihat izin.

> **Note**
>
> Jika Anda memilih tautan sumber daya tabel, Lake Formation menampilkan izin pada tautan sumber daya, bukan pada tabel target tautan sumber daya.

Halaman izin Data mencantumkan semua izin Lake Formation untuk tabel. Nama tabel, nama database database yang berisi tabel, dan ID katalog (ID AWS akun) pemilik tabel muncul sebagai label di bawah kotak pencarian. Label menunjukkan bahwa filter telah diterapkan untuk

daftar izin hanya untuk tabel itu. Anda dapat menyesuaikan filter dengan menutup label atau memilih Hapus filter.



Untuk melihat izin pada tabel (konsol, mulai dari halaman izin Data)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake, pembuat tabel, atau sebagai pengguna yang memiliki izin Lake Formation di atas tabel dengan opsi hibah.
2. Di panel navigasi, pilih Izin data.
3. Posisikan kursor di kotak pencarian, dan pada menu Properties yang muncul, pilih Database.
4. Pada menu Database yang muncul, pilih database.

   > **Important**
   >
   > Jika Anda ingin melihat izin pada tabel yang dibagikan dengan AWS akun Anda dari akun eksternal, Anda harus memilih database di akun eksternal yang berisi tabel, bukan tautan sumber daya ke database.

   Halaman izin Data mencantumkan semua izin Lake Formation untuk database.
5. Posisikan kursor di kotak pencarian lagi, dan pada menu Properties yang muncul, pilih Tabel.
6. Pada menu Tabel yang muncul, pilih tabel.

   Halaman izin Data mencantumkan semua izin Lake Formation untuk tabel. Nama tabel dan nama database database yang berisi tabel muncul sebagai ubin di bawah kotak pencarian. Ubin menunjukkan bahwa filter telah diterapkan untuk daftar izin hanya untuk tabel itu. Anda dapat menyesuaikan filter dengan menutup ubin atau memilih Clear filter.

Untuk melihat izin pada tabel (AWS CLI)

- Masukkan `list-permissions` perintah.

  Contoh berikut mencantumkan izin pada tabel yang dibagikan dari akun eksternal. `CatalogId`Properti adalah ID AWS akun dari akun eksternal, dan nama database mengacu pada database di akun eksternal yang berisi tabel.

```
aws lakeformation list-permissions  --resource-type TABLE --resource '{ "Table":
 {"DatabaseName":"logs", "Name":"alexa-logs", "CatalogId":"123456789012"}}'
```

Mencabut izin menggunakan konsol Lake Formation

Anda dapat menggunakan konsol untuk mencabut semua jenis Izin Lake Formation — izin Katalog Data, izin tag kebijakan, izin filter data, dan izin lokasi.

Untuk mencabut izin Lake Formation pada sumber daya (konsol)

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake atau sebagai pengguna yang telah diberikan izin dengan opsi hibah pada sumber daya.
2. Di panel navigasi, di bawah Izin, pilih Izin data lake, LF-Tags dan izin, atau Lokasi data.
3. Pilih izin atau lokasi, lalu pilih Batalkan.
4. Di kotak dialog yang terbuka, pilih Cabut.

Cross-account berbagi data di Lake Formation

Kemampuan lintas akun Lake Formation memungkinkan pengguna untuk berbagi data lake terdistribusi dengan aman di beberapa AWS organisasi Akun AWS, atau langsung dengan prinsipal IAM di akun lain yang menyediakan akses halus ke metadata Katalog Data dan data yang mendasarinya. Perusahaan besar biasanya menggunakan beberapa Akun AWS, dan banyak dari akun tersebut mungkin memerlukan akses ke danau data yang dikelola oleh satu Akun AWS. Pengguna dan pekerjaan AWS Glue ekstrak, transformasi, dan muat (ETL) dapat melakukan kueri dan menggabungkan tabel di beberapa akun dan tetap memanfaatkan perlindungan data tingkat tabel dan tingkat kolom Lake Formation.

Saat Anda memberikan izin Lake Formation pada sumber daya Katalog Data ke akun eksternal atau langsung ke kepala IAM di akun lain, Lake Formation menggunakan layanan AWS Resource Access Manager (AWS RAM) untuk membagikan sumber daya. Jika akun penerima hibah berada di organisasi yang sama dengan akun pemberi hibah, sumber daya bersama segera tersedia untuk penerima hibah. Jika akun penerima hibah tidak berada di organisasi yang sama, AWS RAM kirimkan undangan ke akun penerima hibah untuk menerima atau menolak hibah sumber daya. Kemudian, untuk membuat sumber daya bersama tersedia, administrator data lake di akun penerima hibah harus menggunakan AWS RAM konsol atau AWS CLI untuk menerima undangan.

Lake Formation mendukung berbagi sumber daya Katalog Data dengan akun eksternal dalam mode akses hybrid. Mode akses hibrida memberikan fleksibilitas untuk mengaktifkan izin Lake Formation secara selektif untuk database dan tabel di situs Anda. AWS Glue Data Catalog
Dengan mode akses Hybrid, Anda sekarang memiliki jalur tambahan yang memungkinkan Anda mengatur izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu kebijakan izin pengguna atau beban kerja lain yang ada.

Untuk informasi selengkapnya, lihat Mode akses hibrid.

Pembagian lintas-akun langsung

Prinsipal resmi dapat berbagi sumber daya secara eksplisit dengan prinsipal IAM di akun eksternal. Fitur ini berguna ketika pemilik akun ingin memiliki kendali atas siapa di akun eksternal yang dapat mengakses sumber daya. Izin yang diterima oleh kepala sekolah IAM adalah gabungan hibah langsung dan hibah tingkat akun yang dialirkan ke kepala sekolah. Hanya penerima hibah izin yang dapat melihat hibah lintas akun langsung. Kepala sekolah yang menerima pembagian sumber daya tidak dapat berbagi sumber daya dengan kepala sekolah lain.

Metode untuk berbagi sumber daya Katalog Data

Dengan satu operasi hibah Lake Formation, Anda dapat memberikan izin lintas akun pada sumber daya Katalog Data berikut.

- Database
- Tabel individual (dengan penyaringan kolom opsional)
- Beberapa tabel yang dipilih
- Semua tabel dalam database (dengan menggunakan wildcard All Tables)

Ada dua opsi untuk berbagi database dan tabel Anda dengan prinsipal lain Akun AWS atau IAM di akun lain.

- Kontrol akses berbasis tag Lake Formation (LF-TBAC) (disarankan)

  Kontrol akses berbasis tag Lake Formation adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Anda dapat menggunakan kontrol akses berbasis tag untuk berbagi sumber daya Katalog Data (database, tabel, dan kolom) dengan prinsipal IAM eksternal Akun AWS, Organizations and organization units (OU). Dalam Lake Formation, atribut ini disebut LF-tags. Untuk informasi selengkapnya, lihat Mengelola data lake menggunakan kontrol akses berbasis tag Lake Formation.

  > **Note**
  >
  > LF-TBAC Metode pemberian izin Katalog Data digunakan AWS Resource Access Manager untuk hibah lintas akun.
  > Lake Formation sekarang mendukung pemberian izin lintas akun ke Organizations dan unit organisasi menggunakan metode. LF-TBAC
  > Untuk mengaktifkan kemampuan ini, Anda perlu memperbarui pengaturan versi Cross account ke Versi 3 atau lebih tinggi.
  > Untuk informasi selengkapnya, lihat Memperbarui pengaturan versi berbagi data lintas akun.

- Lake Formation bernama sumber daya

  Berbagi data lintas akun Lake Formation menggunakan metode sumber daya bernama memungkinkan Anda memberikan izin Lake Formation dengan opsi hibah pada tabel dan database Katalog Data ke eksternal Akun AWS, kepala sekolah IAM, organisasi, atau unit organisasi. Operasi hibah secara otomatis membagikan sumber daya tersebut.

> **Note**
>
> Anda juga dapat mengizinkan AWS Glue crawler mengakses penyimpanan data di akun yang berbeda menggunakan kredenal Lake Formation. Untuk informasi selengkapnya, lihat Cross-account crawling di Panduan AWS Glue Pengembang.

Layanan terintegrasi seperti Athena dan Amazon Redshift Spectrum memerlukan tautan sumber daya untuk dapat menyertakan sumber daya bersama dalam kueri. Untuk informasi selengkapnya tentang tautan sumber daya, lihatCara kerja tautan sumber daya di Lake Formation.

Untuk pertimbangan dan batasan, lihatPraktik dan pertimbangan terbaik berbagi data lintas akun.

Topik

- Prasyarat
- Memperbarui pengaturan versi berbagi data lintas akun
- Berbagi tabel dan database Katalog Data Akun AWS atau prinsip IAM dari akun eksternal
- Memberikan izin pada database atau tabel yang dibagikan dengan akun Anda
- Memberikan izin tautan sumber daya
- Mengakses data dasar tabel bersama
- Cross-account CloudTrail penebangan
- Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau
- Melihat semua hibah lintas akun menggunakan operasi API GetResourceShares

Topik terkait

- Ikhtisar izin Lake Formation
- Mengakses dan melihat tabel dan database Katalog Data bersama
- Membuat tautan sumber daya
- Memecahkan masalah akses lintas akun

Prasyarat

Sebelum AWS akun Anda dapat membagikan sumber daya Katalog Data (katalog, database, dan tabel) dengan akun atau kepala sekolah lain di akun lain, dan sebelum Anda dapat mengakses sumber daya yang dibagikan dengan akun Anda, prasyarat berikut harus dipenuhi.

Persyaratan berbagi data lintas akun umum

- Untuk berbagi database dan tabel Katalog Data dalam mode akses hibrida dan berbagi objek dalam katalog federasi, Anda perlu memperbarui pengaturan versi Cross account ke Versi 4.
- Sebelum memberikan izin lintas akun pada sumber daya Katalog Data, Anda harus mencabut semua izin Lake Formation dari `IAMAllowedPrincipals` grup untuk sumber daya tersebut. Jika prinsipal panggilan memiliki izin lintas akun untuk mengakses sumber

daya dan `IAMAllowedPrincipals` izin ada di sumber daya, maka Lake Formation melemparAccessDeniedException.

Persyaratan ini hanya berlaku ketika Anda mendaftarkan lokasi data yang mendasarinya dalam mode Lake Formation. Jika Anda mendaftarkan lokasi data dalam mode hibrida, izin `IAMAllowedPrincipals` grup dapat ada di database atau tabel bersama.

- Untuk database yang berisi tabel yang ingin Anda bagikan, Anda harus mencegah tabel baru memiliki hibah default `Super` toIAMAllowedPrincipals. Pada konsol Lake Formation, edit database dan matikan Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini atau masukkan AWS CLI perintah berikut, ganti `database` dengan nama database. Jika lokasi data yang mendasari terdaftar dalam mode akses hybrid, Anda tidak perlu mengubah pengaturan default ini. Dalam mode akses hybrid, Lake Formation memungkinkan Anda untuk secara selektif menerapkan izin Lake Formation dan kebijakan izin IAM untuk Amazon S3 dan pada sumber daya yang sama. AWS Glue

```
aws glue update-database --name database --database-input
 '{"Name":"database","CreateTableDefaultPermissions":[]}'
```

- Untuk memberikan izin lintas akun, pemberi harus memiliki izin dan layanan yang diperlukan AWS Identity and Access Management (IAM). AWS Glue AWS RAM Kebijakan AWS terkelola `AWSLakeFormationCrossAccountManager` memberikan izin yang diperlukan.

  Administrator data lake di akun yang menerima pembagian sumber daya menggunakan AWS RAM harus memiliki kebijakan tambahan berikut. Hal ini memungkinkan administrator untuk menerima undangan berbagi AWS RAM sumber daya. Ini juga memungkinkan administrator untuk mengaktifkan berbagi sumber daya dengan organisasi.

  JSON

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

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

- Jika Anda ingin berbagi sumber daya Katalog Data dengan AWS Organizations atau unit organisasi, berbagi dengan organisasi harus diaktifkan AWS RAM.

  Untuk informasi tentang cara mengaktifkan berbagi dengan organisasi, lihat Mengaktifkan berbagi dengan AWS organisasi di Panduan AWS RAM Pengguna.

  Anda harus memiliki `ram:EnableSharingWithAwsOrganization` izin untuk mengaktifkan berbagi dengan organisasi.
- Untuk berbagi sumber daya secara langsung dengan prinsipal IAM di akun lain, Anda perlu memperbarui pengaturan versi Cross account ke Versi 3. Pengaturan ini tersedia di halaman Pengaturan katalog data. Jika Anda menggunakan Versi 1, lihat petunjuk untuk memperbarui pengaturanMemperbarui pengaturan versi berbagi data lintas akun.
- Anda tidak dapat membagikan sumber daya Katalog Data yang dienkripsi dengan kunci terkelola AWS Glue layanan dengan akun lain. Anda hanya dapat membagikan sumber daya Katalog Data yang dienkripsi dengan kunci enkripsi pelanggan, dan akun yang menerima pembagian sumber daya harus memiliki izin pada kunci enkripsi Katalog Data untuk mendekripsi objek.

Cross-account berbagi data menggunakan LF-TBAC persyaratan

- Untuk berbagi sumber daya Katalog Data dengan AWS Organizations dan unit organisasi (OU), Anda perlu memperbarui setelan versi Cross account ke Versi 3 atau lebih tinggi.
- Untuk membagikan sumber daya Katalog Data dengan versi 3 dari setelan versi Cross account, pemberi harus memiliki izin IAM yang ditentukan dalam kebijakan AWS `AWSLakeFormationCrossAccountManager` terkelola di akun Anda.
- Jika Anda menggunakan versi 1 atau versi 2 dari pengaturan versi Cross account, Anda harus memiliki kebijakan sumber daya Katalog Data (`glue:PutResourcePolicy`) yang memungkinkan LF-TBAC. Untuk informasi selengkapnya, lihat Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau.
- Jika saat ini Anda menggunakan kebijakan sumber daya Katalog AWS Glue Data untuk berbagi sumber daya, dan Anda ingin memberikan izin lintas akun menggunakan versi 3 atau lebih

tinggi dari setelan versi Cross account, Anda harus menambahkan `glue:ShareResource` izin di Pengaturan Katalog Data menggunakan operasi `glue:PutResourcePolicy` API seperti yang ditunjukkan di bagianMengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau. Kebijakan ini tidak diperlukan jika akun Anda tidak membuat hibah lintas akun menggunakan kebijakan sumber daya Katalog AWS Glue Data (`glue:PutResourcePolicy`izin penggunaan versi 1 dan versi 2) untuk memberikan akses lintas akun.

```
{
      "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"
      ]
    }
```

- Jika akun Anda telah membuat pembagian lintas akun menggunakan kebijakan sumber daya Katalog AWS Glue Data, dan saat ini Anda menggunakan metode sumber daya bernama atau LF-TBAC dengan setelan lintas akun versi 3 untuk berbagi sumber daya, yang digunakan AWS RAM untuk berbagi sumber daya, Anda harus menyetel `EnableHybrid` argumen `'true'` saat menjalankan operasi `glue:PutResourcePolicy` API. Untuk informasi selengkapnya, lihat Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau.

Penyiapan diperlukan di setiap akun yang mengakses sumber daya bersama

- Jika Anda berbagi sumber daya dengan Akun AWS, setidaknya satu pengguna di akun konsumen harus menjadi administrator data lake untuk melihat sumber daya bersama. Untuk informasi tentang cara membuat administrator data lake, lihatBuat administrator danau data.

  Administrator data lake dapat memberikan izin Lake Formation pada sumber daya bersama ke kepala sekolah lain di akun. Prinsipal lain tidak dapat mengakses sumber daya bersama sampai administrator data lake memberi mereka izin pada sumber daya.

- Layanan terintegrasi seperti Athena dan Redshift Spectrum memerlukan tautan sumber daya untuk dapat menyertakan sumber daya bersama dalam kueri. Prinsipal perlu membuat tautan sumber daya di Katalog Data mereka ke sumber daya bersama dari yang lain. Akun AWS Untuk informasi selengkapnya tentang tautan sumber daya, lihatCara kerja tautan sumber daya di Lake Formation.
- Ketika sumber daya dibagikan langsung dengan prinsipal IAM, untuk menanyakan tabel menggunakan Athena, prinsipal perlu membuat tautan sumber daya. Untuk membuat tautan sumber daya, kepala sekolah memerlukan Formasi Danau `CREATE_TABLE` atau `CREATE_DATABASE` izin, dan izin `glue:CreateTable` atau `glue:CreateDatabase` IAM.

  Jika akun produsen berbagi tabel yang berbeda di bawah database yang sama dengan prinsipal yang sama atau lainnya, prinsipal tersebut dapat segera menanyakan tabel tersebut.

> **Note**
>
> Untuk administrator data lake dan untuk prinsipal yang telah diberikan izin oleh administrator danau data, sumber daya bersama muncul di Katalog Data seolah-olah sumber daya lokal (dimiliki). Pekerjaan ekstrak, transformasi, dan muat (ETL) dapat mengakses data dasar sumber daya bersama.
> Untuk sumber daya bersama, halaman Tabel dan Database di konsol Lake Formation menampilkan ID akun pemilik.
> Saat data dasar sumber daya bersama diakses, peristiwa CloudTrail log dibuat di akun penerima sumber daya bersama dan akun pemilik sumber daya. CloudTrail Peristiwa dapat berisi ARN dari prinsipal yang mengakses data, tetapi hanya jika akun penerima memilih untuk memasukkan ARN utama dalam log. Untuk informasi selengkapnya, lihat Cross-account CloudTrail penebangan.

Memperbarui pengaturan versi berbagi data lintas akun

Dari waktu ke waktu, AWS Lake Formation memperbarui pengaturan berbagi data lintas akun untuk membedakan perubahan yang dilakukan pada AWS RAM penggunaan dan untuk mendukung pembaruan yang dilakukan pada fitur berbagi data lintas akun. Ketika Lake Formation melakukan ini, itu membuat versi baru dari pengaturan versi akun Cross.

Perbedaan utama antara pengaturan versi lintas akun

Untuk informasi selengkapnya tentang cara kerja berbagi data lintas akun di bawah pengaturan versi Cross account yang berbeda, lihat bagian berikut.

> **Note**
>
> Untuk berbagi data dengan akun lain, pemberi harus memiliki izin kebijakan IAM yang `AWSLakeFormationCrossAccountManager` dikelola. Ini adalah prasyarat untuk semua versi.
> Memperbarui pengaturan versi Cross account tidak memengaruhi izin yang dimiliki penerima pada sumber daya bersama. Ini berlaku saat memperbarui dari versi 1 ke versi 2, versi 2 ke versi 3, dan versi 1 ke versi 3. Lihat pertimbangan yang tercantum di bawah ini saat memperbarui versi.

Versi 1

Metode sumber daya bernama: Memetakan setiap hibah izin Lake Formation lintas akun ke satu pembagian AWS RAM sumber daya. Pengguna (peran pemberi atau prinsipal) tidak memerlukan izin tambahan.

LF-TBAC Metode: Hibah izin Cross-account Lake Formation tidak digunakan AWS RAM untuk berbagi data. Pengguna harus memiliki `glue:PutResourcePolicy` izin.

Manfaat dari memperbarui versi: Versi awal - tidak berlaku.

Pertimbangan saat memperbarui versi: Versi awal - tidak berlaku

Versi 2

Metode sumber daya bernama: Mengoptimalkan jumlah pembagian AWS RAM sumber daya dengan memetakan beberapa hibah izin lintas akun dengan satu pembagian sumber daya. AWS RAM Pengguna tidak memerlukan izin tambahan.

LF-TBAC Metode: Hibah izin Cross-account Lake Formation tidak digunakan AWS RAM untuk berbagi data. Pengguna harus memiliki `glue:PutResourcePolicy` izin.

Manfaat dari memperbarui versi: Penyiapan lintas akun yang dapat diskalakan dengan pemanfaatan kapasitas yang optimal. AWS RAM

Pertimbangan saat memperbarui versi: Pengguna yang ingin memberikan izin Lake Formation lintas akun harus memiliki izin dalam kebijakan terkelola. `AWSLakeFormationCrossAccountManager` AWS Jika tidak, Anda harus memiliki `ram:AssociateResourceShare` dan `ram:DisassociateResourceShare` izin untuk berhasil berbagi sumber daya dengan akun lain.

Versi 3

Metode sumber daya bernama: Mengoptimalkan jumlah pembagian AWS RAM sumber daya dengan memetakan beberapa hibah izin lintas akun dengan satu pembagian sumber daya. AWS RAM Pengguna tidak memerlukan izin tambahan.

LF-TBAC Metode: Lake Formation digunakan AWS RAM untuk hibah lintas akun. Pengguna harus menambahkan lem: ShareResource pernyataan untuk `glue:PutResourcePolicy` izin. Penerima harus menerima undangan berbagi sumber daya dari. AWS RAM

Manfaat dari memperbarui versi: Mendukung kemampuan berikut:

- Memungkinkan berbagi sumber daya secara eksplisit dengan prinsipal IAM di akun eksternal.

  Untuk informasi selengkapnya, lihat Memberikan izin pada sumber daya Katalog Data.
- Memungkinkan saham lintas akun menggunakan LF-TBAC metode untuk Organizations atau unit organisasi (OU).
- Menghapus overhead pemeliharaan AWS Glue kebijakan tambahan untuk hibah lintas akun.

Pertimbangan saat memperbarui versi: Saat Anda menggunakan LF-TBAC metode untuk berbagi sumber daya, jika pemberi menggunakan versi yang lebih rendah dari versi 3, dan penerima menggunakan versi 3 atau lebih tinggi, pemberi akan menerima pesan galat berikut: “Permintaan hibah lintas akun tidak valid. Akun konsumen memiliki opt-in untuk versi lintas akun: v3. Harap `CrossAccountVersion` perbarui `DataLakeSetting` ke versi minimal v3 (Layanan: AmazonDataCatalog; Kode Status: 400; Kode Kesalahan: InvalidInputException)”. Namun, jika pemberi menggunakan versi 3 dan penerima menggunakan versi 1 atau versi 2, hibah lintas akun berhasil digunakan LF-Tags .

Cross-account hibah yang dibuat menggunakan metode sumber daya bernama kompatibel di berbagai versi. Bahkan jika akun pemberi menggunakan versi yang lebih lama (versi 1 atau 2) dan akun penerima menggunakan versi yang lebih baru (versi 3 atau lebih tinggi), fungsi akses lintas akun beroperasi dengan mulus tanpa masalah kompatibilitas atau kesalahan.

Untuk berbagi sumber daya secara langsung dengan kepala sekolah IAM di akun lain, hanya pemberi yang perlu menggunakan versi 3.

Cross-account hibah yang dibuat menggunakan LF-TBAC metode mengharuskan pengguna untuk memiliki kebijakan AWS Glue Data Catalog sumber daya di akun. Saat Anda memperbarui ke versi 3, LF-TBAC hibah menggunakan AWS RAM. Agar hibah lintas akun AWS RAM berbasis berhasil, Anda harus menambahkan `glue:ShareResource` pernyataan ke kebijakan sumber daya Katalog Data yang ada seperti yang ditunjukkan di bagian iniMengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau.

Versi 4

Pemberi membutuhkan versi 4 atau lebih tinggi untuk berbagi sumber daya Katalog Data dalam mode akses hibrida atau berbagi objek dalam katalog federasi.

Versi 5

Cross Account Versi 5 meningkatkan berbagi sumber daya lintas akun yang memungkinkan Anda berbagi jumlah tabel yang tidak terbatas ke akun lain, menghilangkan batas asosiasi sumber daya sebelumnya per jenis sumber daya. Untuk memulai, tingkatkan ke lintas akun versi 5 melalui konsol Lake Formation atau API. Setiap hibah izin lintas akun baru akan secara otomatis menggunakan pola wildcard dalam pembagian sumber daya, bukan asosiasi sumber daya individu. Semua saham lintas akun yang ada terus berfungsi, dan semua API Lake Formation yang ada tetap kompatibel.

Manfaat dari memperbarui versi: Cross-account v5 meningkatkan berbagi lintas akun, memungkinkan Anda berbagi ratusan ribu tabel di seluruh akun.

Pertimbangan saat memperbarui versi: Hibah baru setelah peningkatan versi 5 akan menambahkan pola sumber daya wildcard ke pembagian sumber daya Manajer AWS Sumber Daya yang ada atau membuat pembagian baru dengan pola wildcard. Setelah ditingkatkan ke versi 5, downgrade tidak didukung.

Pengoptimalan AWS RAM saham sumber daya

Versi baru (versi 2 ke atas) hibah lintas akun secara optimal memanfaatkan AWS RAM kapasitas untuk memaksimalkan penggunaan lintas akun. Saat Anda berbagi sumber daya dengan kepala eksternal Akun AWS atau IAM, Lake Formation dapat membuat pembagian sumber daya baru atau mengaitkan sumber daya dengan bagian yang ada. Dengan berasosiasi dengan saham yang ada, Lake Formation mengurangi jumlah undangan pembagian sumber daya yang harus diterima konsumen. Versi 5 lebih mengoptimalkan penggunaan RAM dengan menggunakan pola sumber daya berbasis wildcard alih-alih asosiasi sumber daya individu, sehingga secara signifikan mengurangi asosiasi sumber daya per pembagian sumber daya.

Aktifkan AWS RAM berbagi melalui TBAC atau berbagi sumber daya langsung ke kepala sekolah

Untuk berbagi sumber daya secara langsung dengan prinsipal IAM di akun lain atau untuk mengaktifkan pembagian lintas akun TBAC ke Organizations atau unit organisasi, Anda perlu memperbarui pengaturan versi Cross account ke versi 3. Untuk informasi selengkapnya tentang batas AWS RAM sumber daya, lihatPraktik dan pertimbangan terbaik berbagi data lintas akun.

Izin yang diperlukan untuk memperbarui pengaturan versi lintas akun

Jika pemberi izin lintas akun telah `AWSLakeFormationCrossAccountManager` mengelola izin kebijakan IAM, maka tidak ada pengaturan izin tambahan yang diperlukan untuk peran pemberi izin lintas akun atau prinsipal. Namun, jika pemberi lintas akun tidak menggunakan kebijakan terkelola, maka peran pemberi atau prinsipal harus mengikuti izin IAM yang diberikan agar versi baru hibah lintas akun berhasil.

JSON

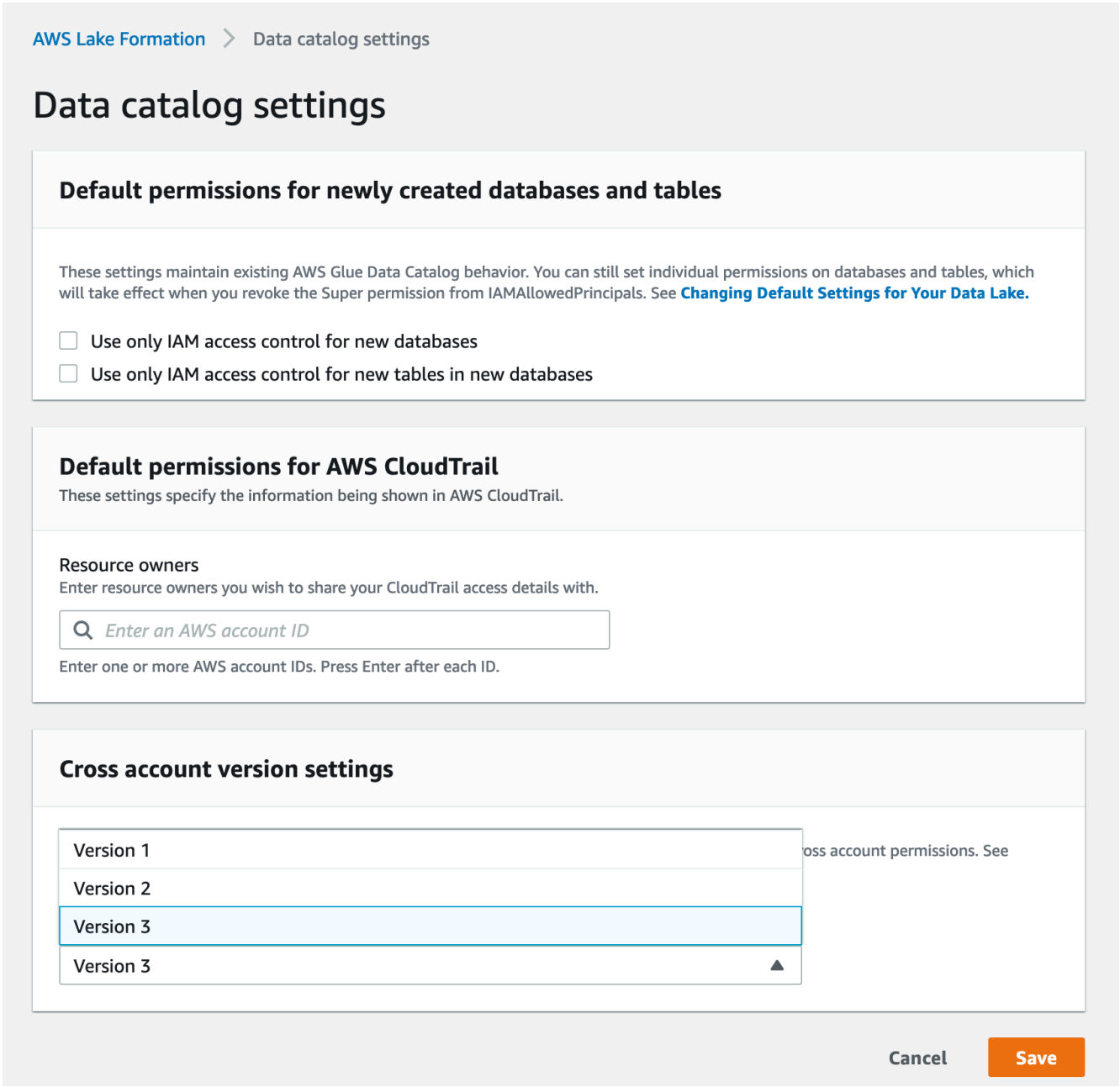
```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
         "ram:AssociateResourceShare",
         "ram:DisassociateResourceShare",
         "ram:GetResourceShares"
       ],
     "Resource": "*",
     "Condition": {
       "StringLike": {
         "ram:ResourceShareName": "LakeFormation*"
        }
      }
    }
  ]
}
```

Untuk mengaktifkan versi baru

Ikuti langkah-langkah ini untuk memperbarui pengaturan versi Cross account melalui AWS Lake Formation konsol atau AWS CLI.

Console

1. Pilih Versi 2, Versi 3, Versi 4, atau Versi 5 di bawah Pengaturan versi Cross account pada halaman Pengaturan katalog data. Jika Anda memilih Versi 1, Lake Formation akan menggunakan mode berbagi sumber daya default.



2. Pilih Simpan.

AWS Command Line Interface (AWS CLI)

Gunakan `put-data-lake-settings` AWS CLI perintah untuk mengatur `CROSS_ACCOUNT_VERSION` parameter. Nilai yang diterima adalah 1, 2, 3, 4, dan 5.

```
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

Setelah Anda memilih Versi 2 atau Versi 3, semua hibah sumber daya bernama baru akan melalui mode hibah lintas akun baru. Untuk menggunakan AWS RAM kapasitas secara optimal untuk saham lintas akun Anda yang ada, kami sarankan Anda untuk mencabut hibah yang dibuat dengan versi lama, dan memberikan kembali dalam mode baru.

Berbagi tabel dan database Katalog Data Akun AWS atau prinsip IAM dari akun eksternal

Bagian ini mencakup petunjuk tentang cara memberikan izin lintas akun pada sumber daya Katalog Data ke AWS akun eksternal, kepala IAM, AWS organisasi, atau unit organisasi. Operasi hibah secara otomatis membagikan sumber daya tersebut.

Topik

- Berbagi data menggunakan kontrol akses berbasis tag
- Cross-account berbagi data menggunakan metode sumber daya bernama

Berbagi data menggunakan kontrol akses berbasis tag

AWS Lake Formation kontrol akses berbasis tag (LF-TBAC) adalah strategi otorisasi yang mendefinisikan izin berdasarkan atribut. Langkah-langkah berikut menjelaskan cara memberikan izin lintas akun dengan menggunakan. LF-Tags

Mengatur diperlukan pada producer/grantor akun

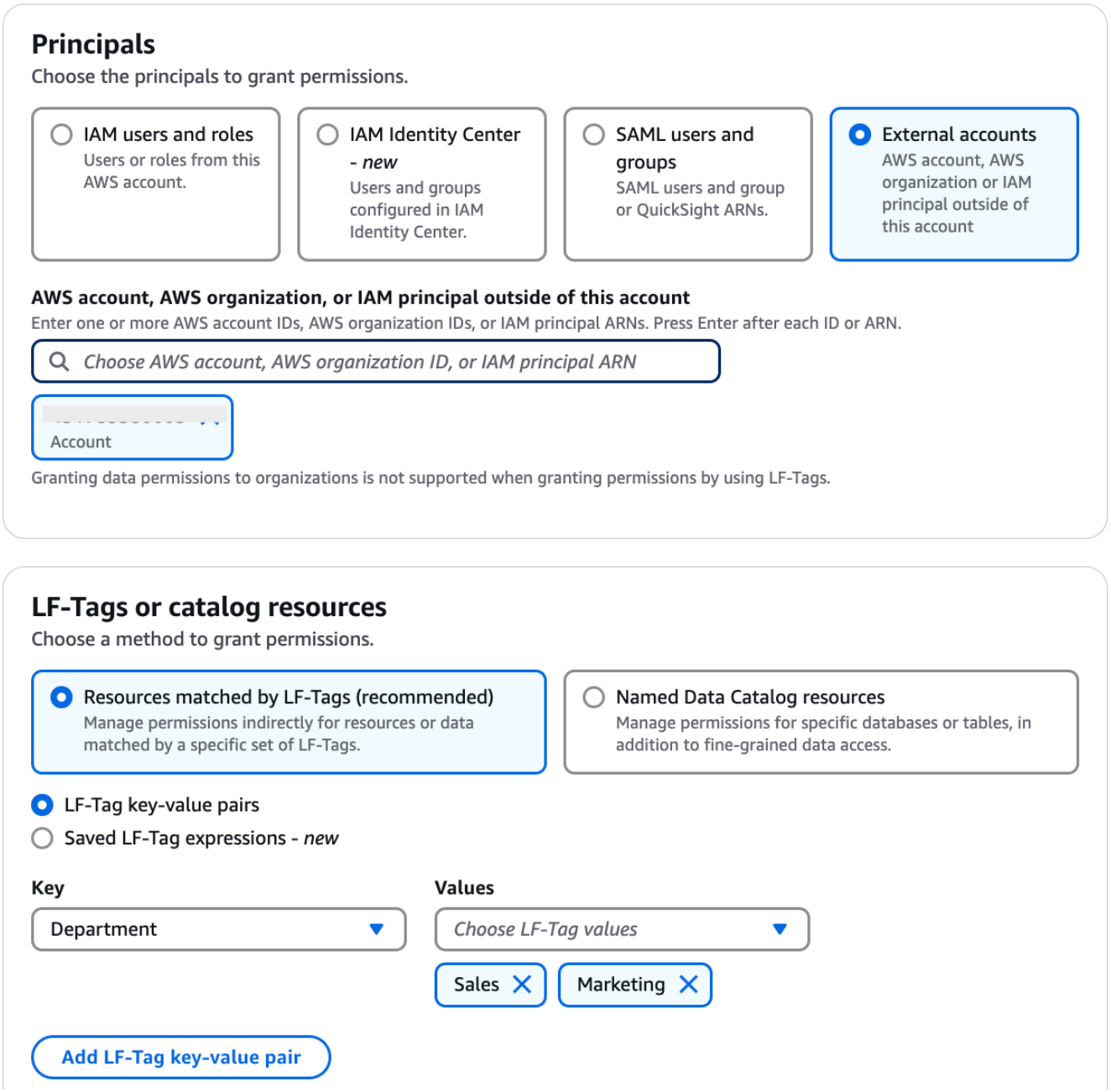
1. Tambahkan LF-Tags.

    1. Masuk ke konsol Lake Formation sebagai administrator danau data atau LF-Tag pembuat.
    2. Di bilah navigasi kiri, pilih Izin, dan LF-Tags dan izin.
    3. Pilih Tambahkan LF-Tag.

    Untuk petunjuk rinci untuk membuat LF-Tags, lihatMenciptakan LF-Tags.
2. Berikan Jelaskan pasangan LF-Tag nilai kunci izin and/or Rekanan ke prinsipal IAM di akun atau akun eksternal Anda.

    Pemberian izin pada pasangan LF-Tag nilai kunci memungkinkan prinsipal untuk melihat, dan menetapkannya ke sumber daya Katalog Data (database LF-Tags, tabel, dan kolom).
3. Selanjutnya, administrator data lake atau kepala IAM dengan izin Associate dapat menetapkan database, tabel, atau kolom LF-Tag ke. Untuk informasi selengkapnya, lihat Menetapkan sumber LF-Tags daya Katalog Data.
4. Selanjutnya, berikan izin data ke akun eksternal menggunakan LF-Tag ekspresi. Hal ini memungkinkan penerima atau penerima izin untuk mengakses sumber daya Katalog Data yang ditandai dengan kunci dan nilai yang sama.

    1. Di panel navigasi, pilih Izin dan Izin data.
    2. PilihIzin.
    3. Pada halaman izin Hibah, untuk Prinsipal, pilih Akun eksternal, dan masukkan Akun AWS ID penerima hibah atau peran IAM kepala sekolah atau Nama Sumber Daya Amazon (ARN) untuk prinsipal (ARN utama) jika membuat hibah lintas akun langsung ke prinsipal eksternal. Anda perlu menekan Enter setelah memasukkan ID akun.



4. Untuk sumber daya LF-Tags atau katalog, pilih Sumber daya yang cocok dengan LF-Tags (disarankan).

   a. Pilih opsi pasangan LF-Tag kunci-nilai atau Ekspresi tersimpan LF-Tag .

   b. Jika Anda memilih, pasangan LF-Tag kunci-nilai, masukkan kunci dan nilai LF-Tagyang terkait dengan sumber daya Katalog Data yang dibagikan dengan akun penerima hibah.

      Penerima hibah diberikan izin pada sumber daya Katalog Data yang diberikan pencocokan LF-Tag dalam ekspresi. LF-Tag Jika LF-Tag ekspresi menentukan beberapa nilai per kunci tag, salah satu nilai tag bisa cocok.

5. Pilih izin tingkat database atau tingkat tabel untuk diberikan pada sumber daya yang cocok dengan ekspresi. LF-Tag

> **Important**
>
> Karena administrator data lake harus memberikan izin pada sumber daya bersama kepada prinsipal di akun penerima hibah, Anda harus selalu memberikan izin lintas akun dengan opsi hibah.

Untuk informasi selengkapnya, lihat Memberikan LF-Tag izin menggunakan konsol.

> **Note**
>
> Kepala sekolah yang menerima hibah lintas akun langsung tidak akan memiliki opsi izin yang Dapat Diberikan.

Mengatur diperlukan pada receiving/grantee akun

1. Masuk ke konsol Lake Formation sebagai administrator data lake dari akun konsumen.
2. Selanjutnya, terima bagian sumber daya di akun konsumen.
    1. Buka AWS RAM konsol.
    2. Di panel navigasi, di bawah Dibagikan dengan saya, pilih Pembagian sumber daya.
    3. Pilih pembagian sumber daya, pilih Terima pembagian sumber daya.
3. Ketika Anda berbagi sumber daya dengan akun lain, sumber daya masih milik akun produsen dan tidak terlihat dalam konsol Athena. Untuk membuat sumber daya terlihat di konsol Athena, Anda perlu membuat tautan sumber daya yang menunjuk ke sumber daya bersama. Untuk petunjuk tentang cara membuat tautan sumber daya, lihat Membuat tautan sumber daya ke tabel Katalog Data bersama dan Membuat tautan sumber daya ke database Katalog Data bersama
    1. Pilih Database atau Tabel di bawah Katalog Data.
    2. Pada Databases/Tables halaman, pilih Buat, tautan Sumber Daya.
    3. Masukkan informasi berikut untuk tautan sumber daya database:
        - Nama tautan sumber daya — Nama unik untuk tautan sumber daya.

- Katalog tujuan - Katalog tempat Anda membuat tautan sumber daya.
- Wilayah basis data bersama — Wilayah database yang dibagikan dengan Anda jika Anda membuat tautan sumber daya di Wilayah yang berbeda.
- Database bersama - Pilih database bersama.
- ID katalog database bersama — Masukkan ID katalog untuk database bersama.

4. Pilih Buat. Anda dapat melihat tautan sumber daya yang baru dibuat di daftar database.

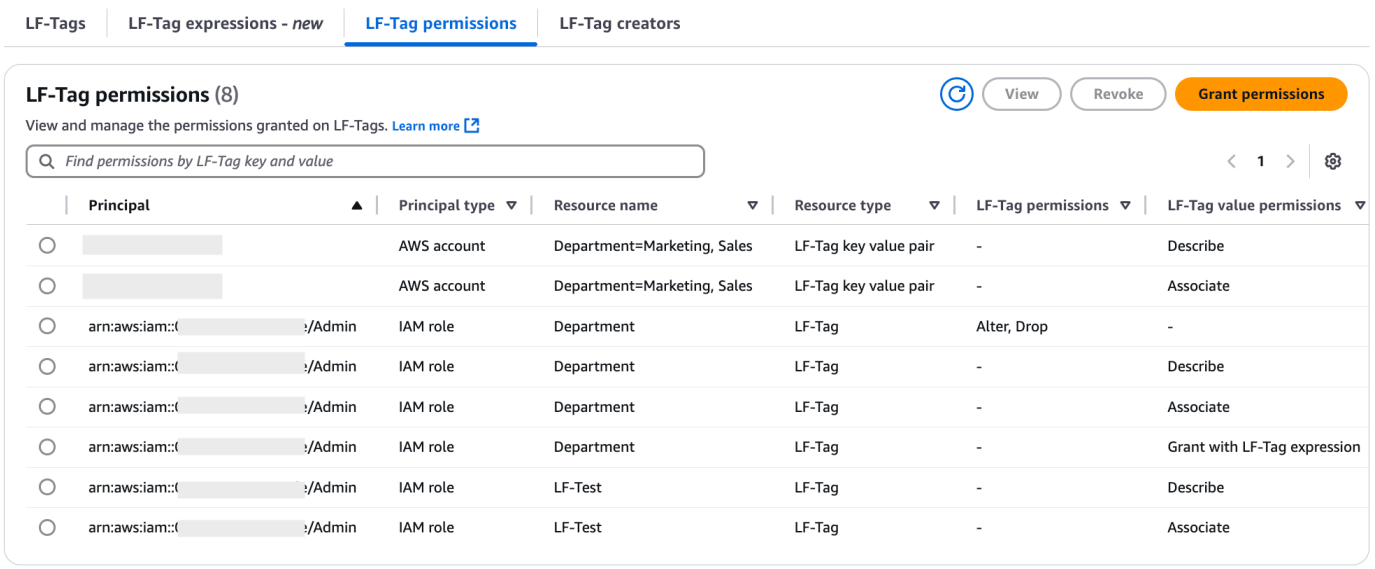
Demikian pula, Anda dapat membuat tautan sumber daya ke tabel bersama.

4. Sekarang berikan izin Jelaskan pada tautan sumber daya ke prinsipal IAM yang Anda bagikan sumber daya.

   1. Pada Databases/Tableshalaman, pilih tautan sumber daya, dan pada menu Tindakan, pilih Hibah.
   2. Di bagian Hibah izin, pilih pengguna dan peran IAM.
   3. Pilih peran IAM yang ingin Anda berikan akses ke tautan sumber daya.
   4. Di bagian Izin tautan sumber daya, pilih Jelaskan.
   5. PilihIzin.

5. Selanjutnya, berikan izin LF-Tag nilai kunci kepada prinsipal di akun konsumen.

   Anda harus dapat menemukan LF-Tags yang dibagikan dengan Anda di akun konsumen di konsol Lake Formation, di bawah Izin, LF-Tagsdan izin. Anda dapat mengaitkan tag yang dibagikan dari pemberi pada sumber daya yang dibagikan dari akun pemberi yang mencakup: database, tabel, dan kolom. Anda selanjutnya dapat memberikan izin pada sumber daya ke prinsipal lain.



1. Di panel navigasi, di bawah Izin, Izin data, pilih Hibah.
2. Pada halaman Hibah izin, pilih pengguna dan peran IAM.
3. Selanjutnya, pilih pengguna dan peran IAM di akun Anda untuk memberikan akses ke yang dibagikan databases/tables.
4. Selanjutnya, untuk LF-Tags atau katalog sumber daya, pilih Sumber daya yang cocok dengan LF-Tags.
5. Selanjutnya, pilih kunci dan nilai LF-Tag yang dibagikan dengan Anda.
6. Selanjutnya, pilih izin database dan tabel yang ingin Anda berikan kepada pengguna dan peran IAM. Anda juga dapat memilih izin yang dapat diberikan yang memungkinkan pengguna dan peran IAM untuk memberikan izin kepada orang lain. users/roles
7. PilihIzin.
8. Anda dapat melihat hibah izin di bawah Izin data di konsol Lake Formation.

Cross-account berbagi data menggunakan metode sumber daya bernama

Anda dapat memberikan izin untuk langsung ke kepala sekolah di AWS akun lain, atau ke eksternal atau. Akun AWS AWS Organizations Pemberian izin Lake Formation ke Organizations atau unit organisasi setara dengan memberikan izin kepada setiap orang Akun AWS di organisasi atau unit organisasi tersebut.

Saat Anda memberikan izin ke akun atau organisasi eksternal, Anda harus menyertakan opsi Izin yang Dapat Diberikan. Hanya administrator data lake di akun eksternal yang dapat mengakses

sumber daya bersama hingga administrator memberikan izin pada sumber daya bersama ke prinsipal lain di akun eksternal.

> **Note**
>
> Opsi izin yang dapat diberikan tidak didukung saat memberikan izin langsung ke prinsipal IAM dari akun eksternal.

Ikuti petunjuk Memberikan izin database menggunakan metode sumber daya bernama untuk memberikan izin lintas akun menggunakan metode sumber daya bernama.

Memberikan izin pada database atau tabel yang dibagikan dengan akun Anda

Setelah sumber daya Katalog Data milik AWS akun lain dibagikan dengan AWS akun Anda, sebagai administrator data lake, Anda dapat memberikan izin pada sumber daya bersama ke prinsipal lain di akun Anda. Namun, Anda tidak dapat memberikan izin pada sumber daya ke AWS akun atau organisasi lain.

Anda dapat menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI) untuk memberikan izin.

Untuk memberikan izin pada database bersama (bernama metode sumber daya, konsol)

- Ikuti petunjuk dalam Memberikan izin database menggunakan metode sumber daya bernama. Dalam daftar Database di bawah LF-Tags atau sumber daya katalog, pastikan bahwa Anda memilih database di akun eksternal, bukan link sumber daya untuk database.

  Jika Anda tidak melihat database dalam daftar database, pastikan bahwa Anda telah menerima undangan berbagi sumber daya AWS Resource Access Manager (AWS RAM) untuk database. Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

  Juga, untuk izin `CREATE_TABLE` dan `ALTER` izin, ikuti instruksi diMemberikan izin lokasi data (akun yang sama), dan pastikan untuk memasukkan ID akun pemilik di bidang Lokasi akun Terdaftar.

Untuk memberikan izin pada tabel bersama (bernama metode sumber daya, konsol)

- Ikuti petunjuk dalam Memberikan izin tabel menggunakan metode sumber daya bernama. Dalam daftar Database di bawah LF-Tags atau sumber daya katalog, pastikan bahwa Anda memilih database di akun eksternal, bukan link sumber daya untuk database.

  Jika Anda tidak melihat tabel dalam daftar tabel, pastikan bahwa Anda telah menerima undangan berbagi AWS RAM sumber daya untuk tabel. Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

  Juga, untuk `ALTER` izin, ikuti instruksi diMemberikan izin lokasi data (akun yang sama), dan pastikan untuk memasukkan ID akun pemilik di bidang Lokasi akun Terdaftar.

Untuk memberikan izin pada sumber daya bersama (LF-TBAC metode, konsol)

- Ikuti petunjuk dalam Memberikan izin Katalog Data. Di bagian sumber daya LF-Tags atau katalog, berikan LF-Tag ekspresi persis bahwa akun eksternal diberikan ke akun Anda, atau subset dari ekspresi tersebut.

  Misalnya, jika akun eksternal memberikan LF-Tag ekspresi `module=customers AND environment=production` ke akun Anda dengan opsi hibah, sebagai administrator data lake, Anda dapat memberikan ekspresi yang sama, atau `module=customers` atau `environment=production` kepada prinsipal di akun Anda. Anda hanya dapat memberikan izin yang sama atau subset dari Lake Formation (misalnya,, `SELECT``ALTER`, dan seterusnya) yang diberikan pada sumber daya melalui ekspresi. LF-Tag

Untuk memberikan izin pada tabel bersama (bernama metode sumber daya, AWS CLI)

- Masukkan perintah yang serupa dengan yang berikut ini. Dalam contoh ini:
  - ID AWS akun Anda adalah 1111-2222-3333.
  - Akun yang memiliki tabel dan yang memberikannya ke akun Anda adalah 1234-5678-9012.
  - `SELECT`Izin diberikan pada tabel bersama `pageviews` kepada penggunadatalake_user1. Pengguna itu adalah prinsipal di akun Anda.
  - `pageviews`Tabel ada di `analytics` database, yang dimiliki oleh akun 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"}}'
```

Perhatikan bahwa akun pemilik harus ditentukan dalam `CatalogId` properti dalam `resource` argumen.

Memberikan izin tautan sumber daya

Ikuti langkah-langkah ini untuk memberikan AWS Lake Formation izin pada satu atau beberapa tautan sumber daya ke prinsipal di AWS akun Anda.

Setelah Anda membuat tautan sumber daya, hanya Anda yang dapat melihat dan mengaksesnya. (Ini mengasumsikan bahwa Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini tidak diaktifkan untuk database.) Untuk mengizinkan kepala sekolah lain di akun Anda mengakses tautan sumber daya, berikan setidaknya izin. `DESCRIBE`

> **Important**
>
> Pemberian izin pada tautan sumber daya tidak memberikan izin pada database atau tabel target (tertaut). Anda harus memberikan izin pada target secara terpisah.

Anda dapat memberikan izin menggunakan konsol Lake Formation, API, atau AWS Command Line Interface (AWS CLI).

console

Untuk memberikan izin tautan sumber daya menggunakan konsol Lake Formation

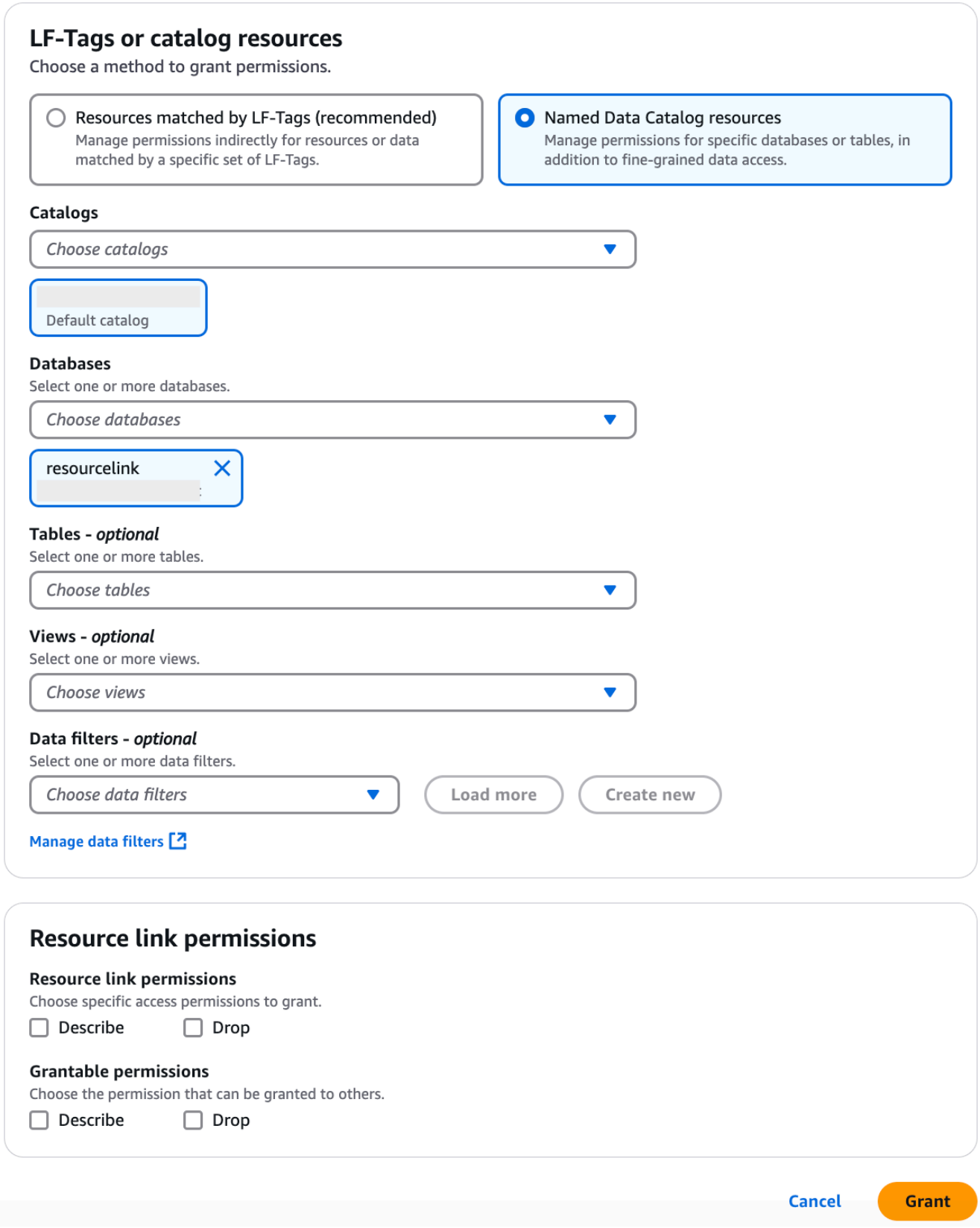
1. Lakukan salah satu tindakan berikut:

   - Untuk tautan sumber daya basis data, ikuti langkah-langkah di[Memberikan izin database menggunakan metode sumber daya bernama](). untuk melakukan hal berikut:
     1. Pilih tautan sumber daya dari daftar database di bawah Katalog Data, Database.
     2. Pilih Hibah untuk membuka halaman izin Hibah.

3. Tentukan prinsipal untuk memberikan izin.
4. Bidang Katalog dan Database diisi.

- Untuk tautan sumber daya tabel, ikuti langkah-langkah Memberikan izin tabel menggunakan metode sumber daya bernama untuk melakukan hal berikut:
  1. Pilih tautan sumber daya dari daftar tabel di bawah Katalog Data, Tabel.
  2. Buka halaman izin Hibah.
  3. Tentukan kepala sekolah.
  4. Bidang Katalog, Database, Tabel diisi.
  5. Tentukan kepala sekolah.

2. Di bawah Izin, pilih izin yang akan diberikan. Secara opsional, pilih izin yang dapat diberikan.


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. PilihIzin.

AWS CLI

Untuk memberikan izin tautan sumber daya menggunakan AWS CLI

- Jalankan `grant-permissions` perintah, tentukan tautan sumber daya sebagai sumber daya.

  Example

  Contoh ini memberikan `DESCRIBE` kepada pengguna `datalake_user1` pada tautan sumber daya tabel di database `incidents-link` `issues` di AWS akun 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"}}'
```

Lihat Juga:

- Membuat tautan sumber daya
- Referensi izin Lake Formation

Mengakses data dasar tabel bersama

Asumsikan bahwa AWS akun A membagikan tabel Katalog Data dengan akun B—misalnya, `SELECT` dengan memberikan opsi hibah pada tabel ke akun B. Agar prinsipal di akun B dapat membaca data dasar tabel bersama, kondisi berikut harus dipenuhi:

- Administrator data lake di akun B harus menerima pembagian. (Ini tidak diperlukan jika akun A dan B berada di organisasi yang sama atau jika hibah dibuat dengan metode kontrol akses berbasis tag Lake Formation.)
- Administrator data lake harus memberikan kembali kepada kepala sekolah `SELECT` izin Lake Formation yang diberikan akun A pada tabel bersama.
- Kepala sekolah harus memiliki izin IAM berikut di atas tabel, database yang berisi itu, dan akun A Data Catalog.

> **Note**
>
> Dalam kebijakan IAM berikut:
>
> - Ganti *<account-id-A>* dengan ID AWS akun akun A.
> - Ganti *<region>* dengan Region yang valid.
> - Ganti *<database>* dengan nama database di akun A yang berisi tabel bersama.
> - Ganti *<table>* dengan nama tabel bersama.

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": [
            "*"
           ]
    }
```

```
    ]
}
```

Lihat Juga:

- Menerima undangan berbagi sumber daya dari AWS RAM

Cross-account CloudTrail penebangan

Lake Formation menyediakan jejak audit terpusat dari semua akses lintas akun ke data di danau data Anda. Saat AWS akun penerima mengakses data dalam tabel bersama, Lake Formation menyalin CloudTrail peristiwa tersebut ke log akun pemilik. CloudTrail Peristiwa yang disalin mencakup kueri terhadap data oleh layanan terintegrasi seperti Amazon Athena dan Amazon Redshift Spectrum, dan akses data berdasarkan pekerjaan. AWS Glue

CloudTrail peristiwa untuk operasi lintas akun pada sumber daya Katalog Data disalin dengan cara yang sama.

Sebagai pemilik sumber daya, jika Anda mengaktifkan pencatatan tingkat objek di Amazon S3, Anda dapat menjalankan kueri yang menggabungkan CloudTrail peristiwa S3 dengan peristiwa CloudTrail Lake Formation untuk menentukan akun yang telah mengakses bucket S3 Anda.

Topik

- Termasuk identitas utama dalam log lintas akun CloudTrail
- Menanyakan CloudTrail log untuk akses lintas akun Amazon S3

Termasuk identitas utama dalam log lintas akun CloudTrail

Secara default, CloudTrail peristiwa lintas akun yang ditambahkan ke log penerima sumber daya bersama dan disalin ke log pemilik sumber daya hanya berisi ID AWS utama prinsip akun eksternal—bukan Nama Sumber Daya Amazon (ARN) yang dapat dibaca manusia dari prinsipal (ARN utama). Saat berbagi sumber daya dalam batas-batas tepercaya, seperti dalam organisasi atau tim yang sama, Anda dapat memilih untuk memasukkan ARN utama dalam acara tersebut CloudTrail . Akun pemilik sumber daya kemudian dapat melacak prinsipal di akun penerima yang mengakses sumber daya milik mereka.

> **Important**
>
> Sebagai penerima sumber daya bersama, untuk melihat ARN utama dalam peristiwa di CloudTrail log Anda sendiri, Anda harus memilih untuk membagikan ARN utama dengan akun pemilik.
> Jika akses data terjadi melalui tautan sumber daya, dua peristiwa dicatat di akun penerima sumber daya bersama: satu untuk akses tautan sumber daya dan satu untuk akses sumber daya target. Acara untuk akses tautan sumber daya memang mencakup ARN utama. Acara untuk akses sumber daya target tidak termasuk ARN utama tanpa keikutsertaan. Acara akses tautan sumber daya tidak disalin ke akun pemilik.

Berikut ini adalah kutipan dari CloudTrail acara lintas akun default (tanpa keikutsertaan). Akun yang melakukan akses data adalah 1111-2222-3333. Ini adalah log yang ditampilkan di akun panggilan dan akun pemilik sumber daya. Lake Formation mengisi log di kedua akun dalam kasus lintas akun.

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "AWSAccount",
        "principalId": "AROAQGFTBBBGOBWV2EMZA:GlueJobRunnerSession",
        "accountId": "111122223333"
    },
    "eventSource": "lakeformation.amazonaws.com",
    "eventName": "GetDataAccess",
...
...
    "additionalEventData": {
        "requesterService": "GLUE_JOB",
        "lakeFormationRoleSessionName": "AWSLF-00-GL-111122223333-G13T0Rmng2"
    },
...
}
```

Sebagai konsumen sumber daya bersama, ketika Anda memilih untuk memasukkan ARN utama, kutipannya menjadi sebagai berikut. `lakeFormationPrincipal`Bidang mewakili peran akhir atau pengguna yang melakukan kueri melalui Amazon Athena, Amazon Redshift Spectrum, atau pekerjaan. AWS Glue

```
{
```

```
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "AWSAccount",
        "principalId": "AROAQGFTBBBGOBWV2EMZA: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"
    },
...
}
```

Untuk ikut serta untuk menyertakan ARN utama dalam log lintas akun CloudTrail

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai `Administrator` pengguna, atau pengguna dengan kebijakan `Administrator Access` IAM.
2. Pada panel navigasi, silakan pilih Pengaturan.
3. Pada halaman Setelan katalog data, di bagian Izin default untuk AWS CloudTrail bagian, untuk pemilik Sumber Daya, masukkan satu atau beberapa ID akun pemilik AWS sumber daya.

   Tekan Enter setelah setiap ID akun.
4. Pilih Simpan.

   Sekarang CloudTrail peristiwa lintas akun yang disimpan di log untuk penerima sumber daya bersama dan pemilik sumber daya berisi ARN utama.

Menanyakan CloudTrail log untuk akses lintas akun Amazon S3

Sebagai pemilik sumber daya bersama, Anda dapat melakukan kueri CloudTrail log S3 untuk menentukan akun yang telah mengakses bucket Amazon S3 Anda (asalkan Anda mengaktifkan pencatatan tingkat objek di Amazon S3). Ini hanya berlaku untuk lokasi S3 yang Anda daftarkan di Lake Formation. Jika konsumen sumber daya bersama memilih untuk menyertakan ARN

utama dalam CloudTrail log Lake Formation, Anda dapat menentukan peran atau pengguna yang mengakses bucket.

Saat menjalankan kueri Amazon Athena, Anda dapat bergabung dengan acara Lake Formation dan CloudTrail acara S3 CloudTrail di properti nama sesi. Kueri juga dapat memfilter acara Lake Formation`eventName="GetDataAccess"`, dan acara S3 pada `eventName="Get Object"` atau. `eventName="Put Object"`

Berikut ini adalah kutipan dari CloudTrail acara lintas akun Lake Formation di mana data di lokasi S3 terdaftar diakses.

```
{
  "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"
   }
}
```

Nilai `lakeFormationRoleSessionName` kunci,`AWSLF-00-GL-111122223333-B8JSAjo5QA`, dapat digabungkan dengan nama sesi di `principalId` kunci CloudTrail acara S3. Berikut ini adalah kutipan dari acara CloudTrail S3. Ini menunjukkan lokasi nama sesi.

```
{
   "eventSource": "s3.amazonaws.com",
   "eventName": "Get Object"
   ..............
   ..............
   "principalId": "AROAQSOX5XXUR7D6RMYLR: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": "AROAQSOX5XXUR7D6RMYLR",
       "arn": "arn:aws:iam::111122223333:role/aws-service-role/
lakeformation.amazonaws.com/Deformationally",
```

```
       "accountId": "111122223333",
       "user Name": "Deformationally"
     },
   ..............
   ..............
}
```

Nama sesi diformat sebagai berikut:

```
AWSLF-<version-number>-<query-engine-code>-<account-id->-<suffix>
```

`version-number`

Versi format ini, saat ini`00`. Jika format nama sesi berubah, versi berikutnya adalah`01`.

`query-engine-code`

Menunjukkan entitas yang mengakses data. Nilai saat ini adalah:

`GL`	AWS GluePekerjaan ETL
`AT`	Athena
`RE`	Amazon Redshift Spectrum

`account-id`

ID AWS akun yang meminta kredensi dari Lake Formation.

`suffix`

String yang dihasilkan secara acak.

Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau

Dimungkinkan untuk memberikan akses lintas akun ke sumber daya Katalog Data dan data yang mendasarinya dengan menggunakan salah satu AWS Glue atau AWS Lake Formation.

DiAWS Glue, Anda memberikan izin lintas akun dengan membuat atau memperbarui kebijakan sumber daya Katalog Data. Di Lake Formation, Anda memberikan izin lintas akun dengan menggunakan model `GRANT/REVOKE` izin Lake Formation dan operasi API. `Grant Permissions`

> **Tip**
>
> Kami menyarankan agar hanya mengandalkan izin Lake Formation untuk mengamankan data lake Anda.

Anda dapat melihat hibah lintas akun Lake Formation dengan menggunakan konsol Lake Formation atau konsol AWS Resource Access Manager (AWS RAM). Namun, halaman konsol tersebut tidak menampilkan izin lintas akun yang diberikan oleh kebijakan sumber daya Katalog AWS Glue Data. Demikian pula, Anda dapat melihat hibah lintas akun dalam kebijakan sumber daya Katalog Data menggunakan halaman Pengaturan AWS Glue konsol, tetapi halaman tersebut tidak menampilkan izin lintas akun yang diberikan menggunakan Lake Formation.

Untuk memastikan bahwa Anda tidak melewatkan hibah apa pun saat melihat dan mengelola izin lintas akun, Lake Formation dan AWS Glue meminta Anda untuk melakukan tindakan berikut untuk menunjukkan bahwa Anda mengetahui dan mengizinkan hibah lintas akun oleh Lake Formation dan. AWS Glue

Saat memberikan izin lintas akun menggunakan AWS Glue Kebijakan sumber daya Katalog Data

Jika akun Anda (akun pemberi atau akun produsen) tidak membuat hibah lintas akun yang digunakan AWS RAM untuk berbagi sumber daya, Anda dapat menyimpan kebijakan sumber daya Katalog Data seperti biasa di. AWS Glue Namun, jika hibah yang melibatkan pembagian AWS RAM sumber daya telah dibuat, Anda harus melakukan salah satu hal berikut untuk memastikan bahwa menyimpan kebijakan sumber daya berhasil:

- Saat Anda menyimpan kebijakan sumber daya di halaman Pengaturan AWS Glue konsol, konsol mengeluarkan peringatan yang menyatakan bahwa izin dalam kebijakan akan ditambahkan ke izin apa pun yang diberikan menggunakan konsol Lake Formation. Anda harus memilih Lanjutkan untuk menyimpan kebijakan.
- Saat menyimpan kebijakan sumber daya menggunakan operasi `glue:PutResourcePolicy` API, Anda harus menyetel `EnableHybrid` bidang ke 'TRUE' (type = string).

Untuk memperbarui kebijakan sumber daya yang ada, gunakan operasi `glue:GetResourcePolicy` API untuk mengambil kebijakan Anda saat ini terlebih dahulu, lalu ubah sesuai kebutuhan sebelum memanggil`glue:PutResourcePolicy`.

> **Note**
>
> Saat membuat kebijakan AWS Glue sumber daya untuk akses lintas akun, berikan hanya izin minimum yang diperlukan untuk kasus penggunaan spesifik Anda.

Untuk informasi selengkapnya, lihat PutResourcePolicy Tindakan (Python: put_resource_policy) di Panduan Pengembang.AWS Glue

Saat memberikan izin lintas akun menggunakan metode sumber daya bernama Lake Formation

Jika tidak ada kebijakan sumber daya Katalog Data di akun Anda (akun produser), hibah lintas akun Lake Formation yang Anda lakukan melanjutkan seperti biasa. Namun, jika kebijakan sumber daya Katalog Data ada, Anda harus menambahkan pernyataan berikut untuk mengizinkan hibah lintas akun Anda berhasil jika dibuat dengan metode sumber daya bernama. Ganti *<region>* dengan nama Region yang valid dan *<account-id>* dengan ID AWS akun Anda (ID akun produsen).

```
{
  "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"
  ]
}
```

Tanpa pernyataan tambahan ini, hibah Lake Formation berhasil, tetapi diblokir AWS RAM, dan akun penerima tidak dapat mengakses sumber daya yang diberikan.

> **Important**
>
> Saat menggunakan metode kontrol akses (LF-TBAC) berbasis tag Lake Formation untuk membuat hibah lintas akun, Anda harus memiliki kebijakan sumber daya Katalog Data dengan setidaknya izin yang ditentukan. Prasyarat

> **Lihat Juga:**
>
> - Kontrol akses metadata(untuk diskusi tentang metode sumber daya bernama versus metode kontrol akses (LF-TBAC) berbasis tag Lake Formation).
> - Melihat tabel dan database Katalog Data bersama
> - Bekerja dengan Pengaturan Katalog Data di AWS Glue Konsol di Panduan AWS Glue Pengembang
> - Memberikan Cross-Account Akses di Panduan AWS Glue Pengembang (untuk contoh kebijakan sumber daya Katalog Data)

Melihat semua hibah lintas akun menggunakan operasi API GetResourceShares

Jika perusahaan Anda memberikan izin lintas akun menggunakan kebijakan AWS Glue Data Catalog sumber daya dan hibah Lake Formation, satu-satunya cara untuk melihat semua hibah lintas akun di satu tempat adalah dengan menggunakan operasi API. `glue:GetResourceShares`

Saat Anda memberikan izin Lake Formation di seluruh akun dengan menggunakan metode sumber daya bernama, AWS Resource Access Manager (AWS RAM) membuat kebijakan sumber daya AWS Identity and Access Management (IAM) dan menyimpannya di akun Anda AWS . Kebijakan memberikan izin yang diperlukan untuk mengakses sumber daya. AWS RAM membuat kebijakan sumber daya terpisah untuk setiap hibah lintas akun. Anda dapat melihat semua kebijakan ini dengan menggunakan operasi `glue:GetResourceShares` API.

> **Note**
>
> Operasi ini juga mengembalikan kebijakan sumber daya Katalog Data. Namun, jika Anda mengaktifkan enkripsi meta data dalam pengaturan Katalog Data, dan Anda tidak memiliki

izin pada AWS KMS kunci, operasi tidak akan menampilkan kebijakan sumber daya Katalog Data.

Untuk melihat semua hibah lintas akun

- Masukkan AWS CLI perintah berikut.

```
aws glue get-resource-policies
```

Berikut ini adalah contoh kebijakan sumber daya yang AWS RAM membuat dan menyimpan saat Anda memberikan izin pada tabel `t` dalam database `db1` ke AWS akun 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"
     ]
    }
  ]
}
```

Lihat juga:

- GetResourceShares Tindakan (Python: get_resource_policies) di Panduan Pengembang AWS Glue

Mengakses dan melihat tabel dan database Katalog Data bersama

Untuk administrator data lake dan untuk kepala sekolah yang telah diberikan izin, sumber daya yang dibagikan dengan AWS akun Anda muncul di Katalog Data seolah-olah itu adalah sumber daya di akun Anda. Konsol menampilkan akun yang memiliki sumber daya.

Anda dapat melihat sumber daya yang dibagikan dengan akun Anda menggunakan konsol Lake Formation. Anda juga dapat menggunakan konsol AWS Resource Access Manager (AWS RAM) untuk melihat sumber daya yang dibagikan dengan akun dan sumber daya yang telah Anda bagikan dengan AWS akun lain menggunakan metode sumber daya bernama.

Important

Saat seseorang menggunakan metode sumber daya bernama untuk memberikan izin lintas akun pada sumber daya Katalog Data ke akun atau AWS organisasi Anda, Lake Formation menggunakan layanan AWS Resource Access Manager (AWS RAM) untuk membagikan sumber daya. Jika akun Anda berada di AWS organisasi yang sama dengan akun pemberi, sumber daya bersama segera tersedia untuk Anda.
Namun, jika akun Anda tidak berada di organisasi yang sama, AWS RAM kirimkan undangan ke akun Anda untuk menerima atau menolak pembagian sumber daya. Kemudian, untuk membuat sumber daya bersama tersedia, administrator data lake di akun Anda harus menggunakan AWS RAM konsol atau CLI untuk menerima undangan.
Konsol Lake Formation menampilkan peringatan jika ada undangan berbagi AWS RAM sumber daya yang menunggu untuk diterima. Hanya pengguna yang berwenang untuk melihat AWS RAM undangan yang menerima peringatan.

Lihat Juga:

- Berbagi tabel Katalog Data dan database di seluruh AWS Akun
- Cross-account berbagi data di Lake Formation

- Mengakses data dasar tabel bersama
- Kontrol akses metadata(untuk informasi tentang metode sumber daya bernama versus LF-TBAC metode untuk berbagi sumber daya.)

Topik

- Menerima undangan berbagi sumber daya dari AWS RAM
- Melihat tabel dan database Katalog Data bersama

Menerima undangan berbagi sumber daya dari AWS RAM

Jika sumber daya Katalog Data dibagikan dengan AWS akun Anda dan akun Anda tidak berada di AWS organisasi yang sama dengan akun berbagi, Anda tidak memiliki akses ke sumber daya bersama hingga Anda menerima undangan berbagi sumber daya dari AWS Resource Access Manager (AWS RAM). Sebagai administrator data lake, Anda harus terlebih dahulu AWS RAM meminta undangan yang tertunda dan kemudian menerima undangan.

Anda dapat menggunakan AWS RAM konsol, API, atau AWS Command Line Interface (AWS CLI) untuk melihat dan menerima undangan.

Untuk melihat dan menerima undangan berbagi sumber daya dari AWS RAM (konsol)

1. Pastikan Anda memiliki izin yang diperlukan AWS Identity and Access Management (IAM) untuk melihat dan menerima undangan berbagi sumber daya.

   Untuk informasi tentang kebijakan IAM yang disarankan untuk administrator data lake, lihat. the section called “Izin administrator danau data”
2. Ikuti petunjuk dalam Menerima dan Menolak Undangan di Panduan Pengguna.AWS RAM

Untuk melihat dan menerima undangan berbagi sumber daya dari AWS RAM (AWS CLI)

1. Pastikan Anda memiliki izin yang diperlukan AWS Identity and Access Management (IAM) untuk melihat dan menerima undangan berbagi sumber daya.

   Untuk informasi tentang kebijakan IAM yang disarankan untuk administrator data lake, lihat. the section called “Izin administrator danau data”
2. Masukkan perintah berikut untuk melihat undangan berbagi sumber daya yang tertunda.

```
aws ram get-resource-share-invitations
```

Output harus serupa dengan yang berikut ini.

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

Perhatikan status`PENDING`.

3. Salin nilai `resourceShareInvitationArn` kunci ke clipboard.
4. Tempelkan nilai ke perintah berikut, ganti*<invitation-arn>*, dan masukkan perintah.

```
aws ram accept-resource-share-invitation --resource-share-invitation-
arn <invitation-arn>
```

Output harus serupa dengan yang berikut ini.

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

Perhatikan status`ACCEPTED`.

Melihat tabel dan database Katalog Data bersama

Anda dapat melihat sumber daya yang dibagikan dengan akun Anda menggunakan konsol Lake Formation atau AWS CLI. Anda juga dapat menggunakan konsol AWS Resource Access Manager (AWS RAM) atau CLI untuk melihat kedua sumber daya yang dibagikan dengan akun Anda dan sumber daya yang telah Anda bagikan dengan akun lain AWS .

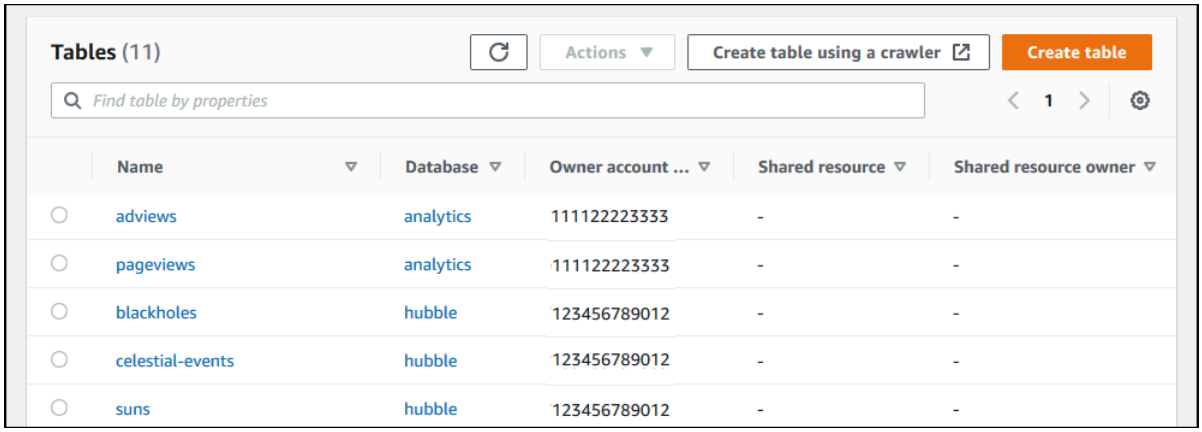
Untuk melihat sumber daya bersama menggunakan konsol Lake Formation

1. Buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.

   Masuk sebagai administrator data lake atau pengguna yang telah diberikan izin pada tabel bersama.
2. Untuk melihat sumber daya yang dibagikan dengan AWS akun Anda, lakukan salah satu hal berikut:

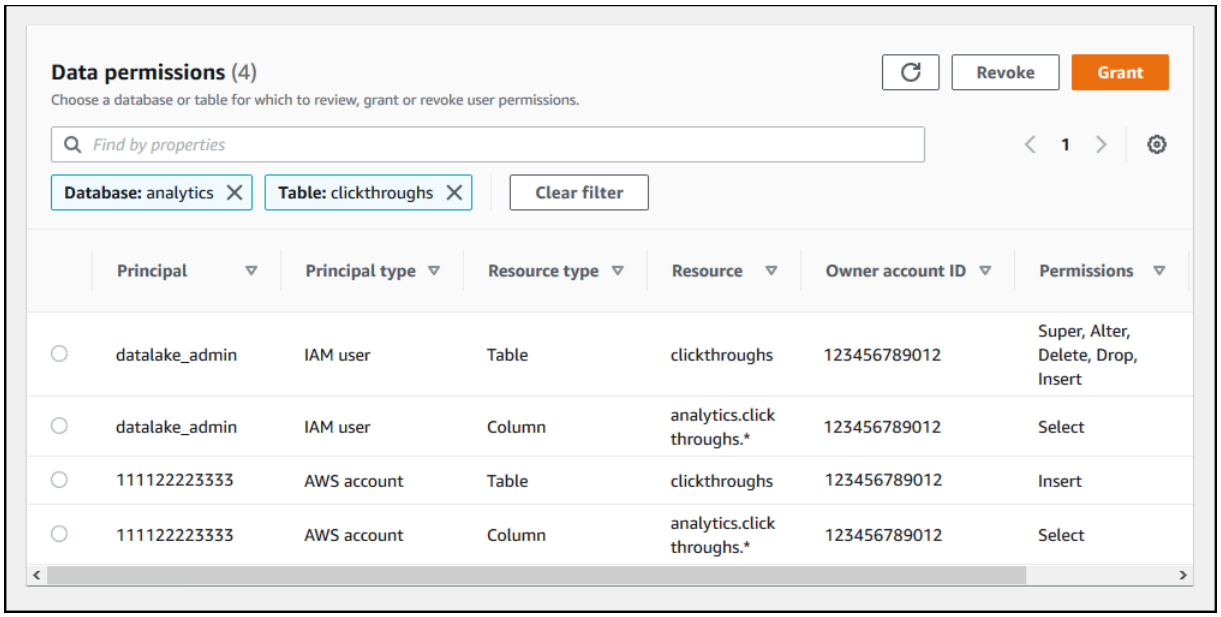
   - Untuk melihat tabel yang dibagikan dengan akun Anda, di panel navigasi, pilih Tabel.
   - Untuk melihat database yang dibagikan dengan akun Anda, di panel navigasi, pilih Database.

   Konsol menampilkan daftar database atau tabel baik di akun Anda dan dibagikan dengan akun Anda. Untuk sumber daya yang dibagikan dengan akun Anda, konsol menampilkan ID AWS akun pemilik di bawah kolom ID akun Pemilik (kolom ketiga pada tangkapan layar berikut).



3. Untuk melihat sumber daya yang Anda bagikan dengan AWS akun atau organisasi lain, di panel navigasi, pilih Izin data.

   Sumber daya yang Anda bagikan tercantum di halaman Izin data dengan nomor akun eksternal yang ditampilkan di kolom Utama, seperti yang ditunjukkan pada gambar berikut.



Untuk melihat sumber daya bersama menggunakan AWS RAM konsol

1. Pastikan Anda memiliki izin yang diperlukan AWS Identity and Access Management (IAM) untuk melihat sumber daya bersama menggunakan. AWS RAM

   Minimal, Anda harus memiliki izin. `ram:ListResources` Izin ini disertakan dalam kebijakan yang dikelola AWS. `AWSLakeFormationCrossAccountManager`

2. Masuk ke Konsol Manajemen AWS dan buka AWS RAM konsol di https://console.aws.amazon.com/ram/home.
3. Lakukan salah satu tindakan berikut:
   - Untuk melihat sumber daya yang Anda bagikan, di panel navigasi, di bawah Dibagikan oleh saya, pilih Sumber daya bersama.
   - Untuk melihat sumber daya yang dibagikan dengan Anda, di panel navigasi, di bawah Dibagikan dengan saya, pilih Sumber daya bersama.

Membuat tautan sumber daya

Tautan sumber daya adalah objek Katalog Data yang merupakan tautan ke database dan tabel metadata—biasanya ke database dan tabel bersama dari akun lain. AWS Mereka membantu mengaktifkan akses lintas akun ke data di danau data di semua AWS Wilayah.

> **Note**
>
> Lake Formation mendukung kueri tabel Katalog Data di seluruh AWS Wilayah. Anda dapat mengakses database dan tabel Katalog Data dari AWS Wilayah mana pun dengan membuat tautan sumber daya di wilayah yang mengarah ke database dan tabel bersama di Wilayah yang berbeda.

Topik

- Cara kerja tautan sumber daya di Lake Formation
- Membuat tautan sumber daya ke tabel Katalog Data bersama
- Membuat tautan sumber daya ke database Katalog Data bersama
- Penanganan tautan sumber daya di AWS Glue API

Cara kerja tautan sumber daya di Lake Formation

Tautan sumber daya adalah objek Katalog Data yang merupakan tautan ke database atau tabel lokal atau bersama. Setelah Anda membuat link sumber daya ke database atau tabel, Anda dapat menggunakan nama link sumber daya di mana pun Anda akan menggunakan database atau nama tabel. Bersama dengan tabel yang Anda miliki atau tabel yang dibagikan dengan Anda, tautan

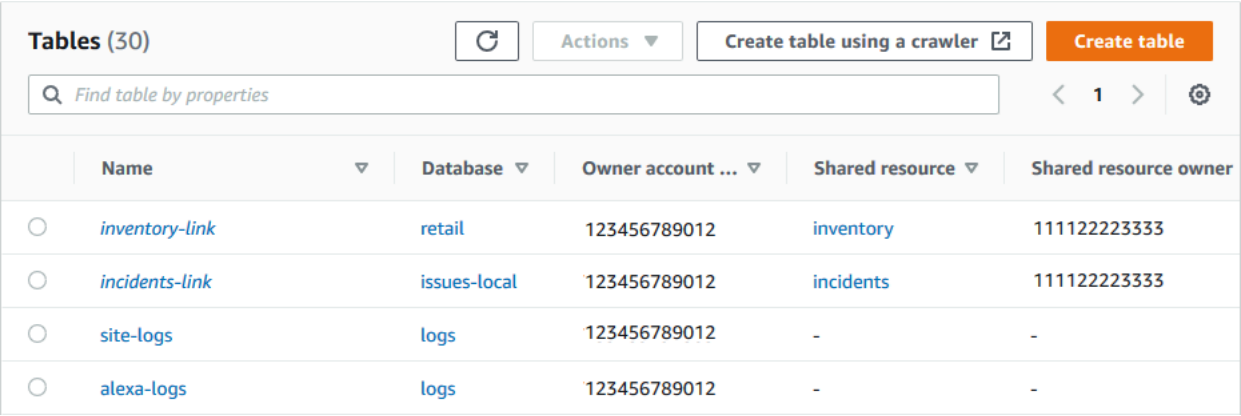
sumber daya tabel dikembalikan oleh `glue:GetTables()` dan muncul sebagai entri di halaman Tabel konsol Lake Formation. Tautan sumber daya ke database bertindak dengan cara yang sama.

Membuat link sumber daya ke database atau tabel memungkinkan Anda untuk melakukan hal berikut:

- Tetapkan nama yang berbeda ke database atau tabel di Katalog Data Anda. Ini sangat berguna jika AWS akun yang berbeda berbagi database atau tabel dengan nama yang sama, atau jika beberapa database di akun Anda memiliki tabel dengan nama yang sama.
- Akses database dan tabel Katalog Data dari AWS Wilayah mana pun dengan membuat tautan sumber daya di wilayah tersebut yang menunjuk ke database dan tabel di wilayah lain. Anda dapat menjalankan kueri di wilayah mana pun dengan tautan sumber daya ini menggunakan Athena, Amazon EMR, dan AWS Glue menjalankan pekerjaan ETL Spark, tanpa menyalin data sumber atau metadata di Katalog Data Glue.
- Gunakan AWS layanan terintegrasi seperti Amazon Athena dan Amazon Redshift Spectrum untuk menjalankan kueri yang mengakses database atau tabel bersama. Beberapa layanan terintegrasi tidak dapat langsung mengakses database atau tabel di seluruh akun. Namun, mereka dapat mengakses tautan sumber daya di akun Anda ke database dan tabel di akun lain.

> **Note**
>
> Anda tidak perlu membuat tautan sumber daya untuk mereferensikan database atau tabel bersama dalam skrip AWS Glue ekstrak, transformasi, dan muat (ETL). Namun, untuk menghindari ambiguitas ketika beberapa AWS akun berbagi database atau tabel dengan nama yang sama, Anda dapat membuat dan menggunakan tautan sumber daya atau menentukan ID katalog saat menjalankan operasi ETL.

Contoh berikut menunjukkan halaman Tabel konsol Lake Formation, yang mencantumkan dua tautan sumber daya. Nama tautan sumber daya selalu ditampilkan dalam huruf miring. Setiap tautan sumber daya ditampilkan bersama dengan nama dan pemilik sumber daya bersama yang ditautkan. Dalam contoh ini, administrator data lake di AWS akun 1111-2222-3333 berbagi tabel dan dengan akun 1234-5678-9012. `inventory` `incidents` Seorang pengguna di akun itu kemudian membuat tautan sumber daya ke tabel bersama tersebut.



Berikut ini adalah catatan dan batasan pada tautan sumber daya:

- Tautan sumber daya diperlukan untuk mengaktifkan layanan terintegrasi seperti Athena dan Redshift Spectrum untuk menanyakan data dasar tabel bersama. Kueri dalam layanan terintegrasi ini dibangun terhadap nama tautan sumber daya.
- Dengan asumsi bahwa pengaturan Gunakan hanya kontrol akses IAM untuk tabel baru dalam database ini dimatikan untuk database yang berisi, hanya prinsipal yang membuat tautan sumber daya yang dapat melihat dan mengaksesnya. Untuk mengaktifkan prinsipal lain di akun Anda untuk mengakses tautan sumber daya, berikan `DESCRIBE` izin di atasnya. Untuk memungkinkan orang lain menjatuhkan tautan sumber daya, berikan `DROP` izin di atasnya. Administrator data lake dapat mengakses semua tautan sumber daya di akun. Untuk menjatuhkan tautan sumber daya yang dibuat oleh kepala sekolah lain, administrator data lake harus terlebih dahulu memberikan `DROP` izin pada tautan sumber daya. Untuk informasi selengkapnya, lihat Referensi izin Lake Formation.

  > **Important**
  >
  > Pemberian izin pada tautan sumber daya tidak memberikan izin pada database atau tabel target (tertaut). Anda harus memberikan izin pada target secara terpisah.

- Untuk membuat tautan sumber daya, Anda memerlukan Formasi Danau `CREATE_TABLE` atau `CREATE_DATABASE` izin, serta izin `glue:CreateTable` atau `glue:CreateDatabase` AWS Identity and Access Management (IAM).
- Anda dapat membuat tautan sumber daya ke sumber daya Katalog Data lokal (milik), serta sumber daya yang dibagikan dengan AWS akun Anda.
- Saat Anda membuat tautan sumber daya, tidak ada pemeriksaan yang dilakukan untuk melihat apakah sumber daya bersama target ada atau apakah Anda memiliki izin lintas akun pada sumber

daya tersebut. Ini memungkinkan Anda untuk membuat tautan sumber daya dan sumber daya bersama dalam urutan apa pun.

- Jika Anda menghapus tautan sumber daya, sumber daya bersama yang ditautkan tidak akan dihapus. Jika Anda menjatuhkan sumber daya bersama, tautan sumber daya ke sumber daya tersebut tidak akan dihapus.
- Dimungkinkan untuk membuat rantai tautan sumber daya. Namun, tidak ada nilai dalam melakukannya, karena API hanya mengikuti tautan sumber daya pertama.

Lihat juga:

- Memberikan izin pada sumber daya Katalog Data

Membuat tautan sumber daya ke tabel Katalog Data bersama

Anda dapat membuat tautan sumber daya ke tabel bersama di AWS Wilayah mana pun menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Untuk membuat tautan sumber daya ke tabel bersama (konsol)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai kepala sekolah yang memiliki `CREATE_TABLE` izin Lake Formation di database untuk memuat tautan sumber daya.
2. Di panel navigasi, pilih Tabel di bawah Katalog Data, lalu pilih Buat, tautan Sumber Daya.
3. Pada halaman tautan Buat sumber daya, berikan informasi berikut:

   Nama tautan sumber daya

   Masukkan nama yang mematuhi aturan yang sama dengan nama tabel. Namanya bisa sama dengan tabel bersama target.

   Basis Data

   Database dalam Katalog Data lokal berisi link sumber daya.

   Pemilik tabel bersama Wilayah

   Jika Anda membuat tautan sumber daya di Wilayah yang berbeda, pilih wilayah tabel bersama target.

Tabel bersama

Pilih tabel bersama dari daftar, atau masukkan nama tabel lokal (dimiliki) atau bersama.

Daftar ini berisi semua tabel yang dibagikan dengan akun Anda. Perhatikan database dan ID akun pemilik yang tercantum dengan setiap tabel. Jika Anda tidak melihat tabel yang Anda tahu telah dibagikan dengan akun Anda, periksa hal berikut:

- Jika Anda bukan administrator data lake, periksa apakah administrator danau data memberi Anda izin Lake Formation di atas tabel.
- Jika Anda adalah administrator data lake, dan akun Anda tidak berada dalam AWS organisasi yang sama dengan akun pemberian, pastikan bahwa Anda telah menerima undangan berbagi sumber daya AWS Resource Access Manager (AWS RAM) untuk tabel. Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

Database tabel bersama

Jika Anda memilih tabel bersama dari daftar, bidang ini diisi dengan database tabel bersama di akun eksternal. Jika tidak, masukkan database lokal (untuk tautan sumber daya ke tabel lokal) atau database tabel bersama di akun eksternal.

Pemilik meja bersama

Jika Anda memilih tabel bersama dari daftar, bidang ini diisi dengan ID akun pemilik tabel bersama. Jika tidak, masukkan ID AWS akun Anda (untuk tautan sumber daya ke tabel lokal) atau ID AWS akun yang membagikan tabel.

4. Pilih Buat untuk membuat tautan sumber daya.

   Anda kemudian dapat melihat nama link sumber daya di bawah kolom Nama pada halaman Tabel.

5. (Opsional) Berikan `DESCRIBE` izin Lake Formation pada tautan sumber daya ke kepala sekolah yang harus dapat melihat tautan dan mengakses tabel target.

   Namun, pemberian izin pada tautan sumber daya tidak memberikan izin pada database atau tabel target (tertaut). Anda harus memberikan izin pada database target secara terpisah agar table/resource tautan dapat terlihat di Athena.

Untuk membuat tautan sumber daya ke tabel bersama di Wilayah yang sama (AWS CLI)

1. Masukkan perintah yang serupa dengan yang berikut ini.

```
aws glue create-table --database-name myissues --table-input
 '{"Name":"my_customers","TargetTable":
{"CatalogId":"111122223333","DatabaseName":"issues","Name":"customers"}}'
```

   Perintah ini membuat tautan sumber daya bernama `my_customers` ke tabel bersama`customers`, yang ada di database `issues` di AWS akun 1111-2222-3333. Tautan sumber daya disimpan dalam database lokal`myissues`.

2. (Opsional) Berikan `DESCRIBE` izin Lake Formation pada tautan sumber daya ke kepala sekolah yang harus dapat melihat tautan dan mengakses tabel target.

   Namun, pemberian izin pada tautan sumber daya tidak memberikan izin pada tabel target (tertaut). Anda harus memberikan izin pada database target secara terpisah agar table/resource tautan dapat terlihat di Athena.

Untuk membuat tautan sumber daya ke tabel bersama di Wilayah yang berbeda (AWS CLI)

1. Masukkan perintah yang serupa dengan yang berikut ini.

```
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"
        }
    }
}‘
```

   Perintah ini membuat tautan sumber daya bernama `rl_useast1salestb_ireland` di Wilayah Eropa (Irlandia) ke `useast1_salestb` tabel bersama, yang ada di `useast1_salesdb`

database di akun 444455556666 di AWS Wilayah AS Timur (Virginia N.). Tautan sumber daya disimpan dalam database lokal`ireland_db`.

2. Berikan `DESCRIBE` izin Lake Formation kepada kepala sekolah yang harus dapat melihat tautan dan mengakses target tautan melalui tautan.

   Namun, pemberian izin pada tautan sumber daya tidak memberikan izin pada tabel target (tertaut). Anda harus memberikan izin pada tabel target secara terpisah agar table/resource tautan dapat terlihat di Athena.

> Lihat juga:
>
> - Cara kerja tautan sumber daya di Lake Formation
> - `MENJELASKAN`

Membuat tautan sumber daya ke database Katalog Data bersama

Anda dapat membuat tautan sumber daya ke database bersama menggunakan AWS Lake Formation konsol, API, atau AWS Command Line Interface (AWS CLI).

Untuk membuat tautan sumber daya ke database bersama (konsol)

1. Buka AWS Lake Formation konsol di https://console.aws.amazon.com/lakeformation/. Masuk sebagai administrator data lake atau sebagai pembuat basis data.

   Pembuat database adalah kepala sekolah yang telah diberikan `CREATE_DATABASE` izin Lake Formation.

2. Di panel navigasi, pilih Databases, lalu pilih Create, Resource link.
3. Pada halaman tautan Buat sumber daya, berikan informasi berikut:

   Nama tautan sumber daya

   Masukkan nama yang mematuhi aturan yang sama dengan nama database. Namanya bisa sama dengan basis data bersama target.

   Katalog tujuan

   Pilih katalog tujuan untuk tautan sumber daya database.

Pemilik basis data bersama Wilayah

Jika Anda membuat tautan sumber daya di Wilayah yang berbeda, pilih Wilayah basis data bersama target.

Database bersama

Pilih database dari daftar, atau masukkan nama database lokal (dimiliki) atau bersama.

Daftar ini berisi semua database yang dibagikan dengan akun Anda. Perhatikan ID akun pemilik yang terdaftar dengan setiap database. Jika Anda tidak melihat database yang Anda tahu telah dibagikan dengan akun Anda, periksa hal berikut:

- Jika Anda bukan administrator data lake, periksa apakah administrator danau data memberi Anda izin Lake Formation pada database.
- Jika Anda adalah administrator data lake, dan akun Anda tidak berada dalam AWS organisasi yang sama dengan akun pemberian, pastikan bahwa Anda telah menerima undangan berbagi sumber daya AWS Resource Access Manager (AWS RAM) untuk database. Untuk informasi selengkapnya, lihat Menerima undangan berbagi sumber daya dari AWS RAM.

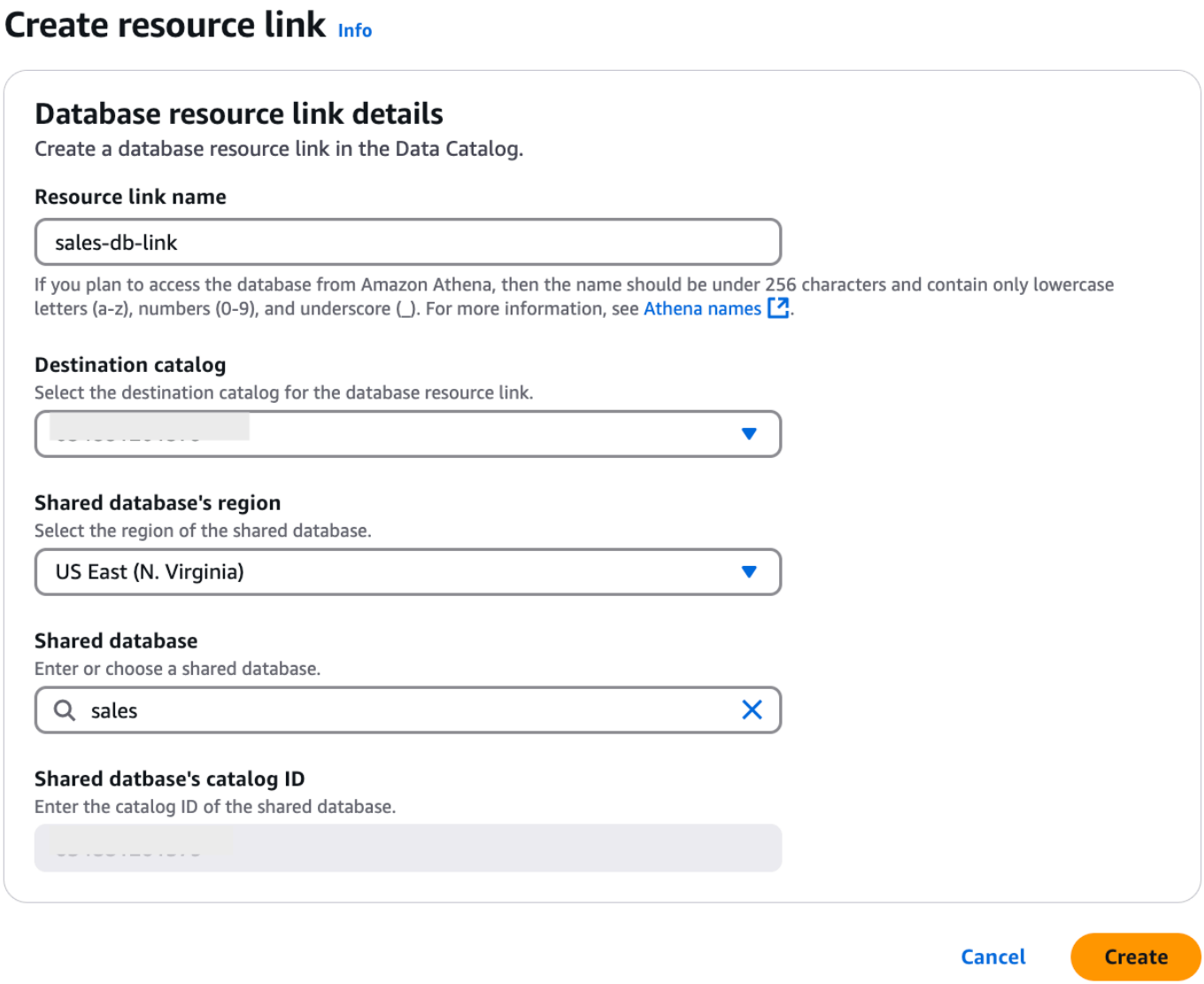
Pemilik basis data bersama

Jika Anda memilih database bersama dari daftar, bidang ini diisi dengan ID akun pemilik database bersama. Jika tidak, masukkan ID AWS akun Anda (untuk tautan sumber daya ke database lokal) atau ID AWS akun yang berbagi database.

ID katalog database bersama

Masukkan ID katalog untuk database bersama. Saat membuat tautan sumber daya ke database yang dibagikan dari AWS akun lain, Anda perlu menentukan ID katalog ini untuk mengidentifikasi Katalog Data akun mana yang berisi basis data sumber.

Ketika Anda memilih database bersama dari menu tarik-turun, sistem secara otomatis mengisi ID katalog akun yang memiliki dan telah membagikan database tersebut dengan Anda.



4. Pilih Buat untuk membuat tautan sumber daya.

   Anda kemudian dapat melihat nama tautan sumber daya di bawah kolom Nama pada halaman Database.

5. (Opsional) Berikan `DESCRIBE` izin Formasi Danau pada tautan sumber daya ke kepala sekolah dari Wilayah Eropa (Irlandia) yang harus dapat melihat tautan dan mengakses basis data target.

   Namun, pemberian izin pada tautan sumber daya tidak memberikan izin pada database atau tabel target (tertaut). Anda harus memberikan izin pada database target secara terpisah agar table/resource tautan dapat terlihat di Athena.

Untuk membuat tautan sumber daya ke database bersama di Wilayah yang sama (AWS CLI)

1. Masukkan perintah yang serupa dengan yang berikut ini.

```
aws glue create-database --database-input '{"Name":"myissues","TargetDatabase":
{"CatalogId":"111122223333","DatabaseName":"issues"}}'
```

   Perintah ini membuat tautan sumber daya bernama `myissues` ke database bersama`aissues`, yang ada di AWS akun 1111-2222-3333.

2. (Opsional) Berikan `DESCRIBE` izin Lake Formation kepada kepala sekolah pada tautan sumber daya yang harus dapat melihat tautan dan mengakses database atau tabel target.

   Namun, pemberian izin pada tautan sumber daya tidak memberikan izin pada database atau tabel target (tertaut). Anda harus memberikan izin pada database target secara terpisah agar table/resource tautan dapat terlihat di Athena.

Untuk membuat tautan sumber daya ke database bersama di Wilayah yang berbeda (AWS CLI)

1. Masukkan perintah yang serupa dengan yang berikut ini.

```
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"
       }
    }
}'
```

   Perintah ini membuat tautan sumber daya yang disebutkan `rl_useast1shared_irelanddb` di AWS akun 111122223333 di Wilayah Eropa (Irlandia) ke database bersama`useast1shared_db`, yang ada di AWS akun 444455556666 di Wilayah AS Timur (Virginia N.).

2. Berikan `DESCRIBE` izin Formasi Danau kepada kepala sekolah dari Wilayah Eropa (Irlandia) yang harus dapat melihat tautan dan mengakses target tautan melalui tautan.

Lihat juga:

- Cara kerja tautan sumber daya di Lake Formation
- MENJELASKAN

Penanganan tautan sumber daya di AWS Glue API

Tabel berikut menjelaskan cara API Katalog AWS Glue Data menangani database dan tautan sumber daya tabel. Untuk semua operasi `Get*` API, hanya database dan tabel yang memiliki izin pemanggil untuk dikembalikan. Selain itu, saat mengakses database atau tabel target melalui tautan sumber daya, Anda harus memiliki izin AWS Identity and Access Management (IAM) dan Lake Formation pada target dan tautan sumber daya. Izin Lake Formation yang diperlukan pada tautan sumber daya adalah`DESCRIBE`. Untuk informasi selengkapnya, lihat MENJELASKAN.

Operasi API basis data

Operasi API	Penanganan tautan sumber daya
`CreateDatabase`	Jika database adalah link sumber daya, membuat link sumber daya ke database target yang ditunjuk.
`UpdateDatabase`	Jika database yang ditunjuk adalah tautan sumber daya, ikuti tautan dan perbarui basis data target. Jika tautan sumber daya harus dimodifikasi untuk ditautkan ke database yang berbeda, Anda harus menghapusnya dan membuat yang baru.
`DeleteDatabase`	Menghapus tautan sumber daya. Itu tidak menghapus database tertaut (target).
`GetDatabase`	Jika pemanggil memiliki izin pada target, ikuti tautan untuk mengembalikan properti target. Jika tidak, ia mengembalikan properti tautan.
`GetDatabases`	Mengembalikan daftar database, termasuk link sumber daya. Untuk setiap tautan sumber daya dalam kumpulan hasil, operasi mengikuti tautan untuk mendapatkan properti target tautan. Anda

Operasi API	Penanganan tautan sumber daya
	harus menentukan `ResourceShareType = ALL` untuk melihat database yang dibagikan dengan akun Anda.

Operasi API tabel

Operasi API	Penanganan tautan sumber daya
`CreateTable`	Jika database adalah link sumber daya, ikuti link database dan membuat tabel dalam database target. Jika tabel adalah link sumber daya, operasi membuat link sumber daya dalam database yang ditunjuk. Membuat link sumber daya tabel melalui link sumber daya database tidak didukung.
`UpdateTable`	Jika tabel atau database yang ditunjuk adalah tautan sumber daya, perbarui tabel target. Jika kedua tabel dan database adalah link sumber daya, operasi gagal.
`DeleteTable`	Jika database yang ditunjuk adalah tautan sumber daya, ikuti tautan dan hapus tautan sumber daya tabel atau tabel di basis data target. Jika tabel adalah link sumber daya, operasi menghapus link sumber daya tabel dalam database yang ditunjuk. Menghapus tautan sumber daya tabel tidak menghapus tabel target.
`BatchDeleteTable`	Sama seperti`DeleteTable` .
`GetTable`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan mengembalikan tabel atau tabel sumber daya link dari database target. Jika tidak, jika tabel adalah link sumber daya, operasi mengikuti link dan mengembalikan properti tabel target.
`GetTables`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan mengembalikan tabel dan link sumber daya tabel dari database target. Jika database target adalah database bersama dari AWS akun lain, operasi hanya mengembalikan tabel bersama dalam database tersebut. Itu tidak mengikuti link sumber daya tabel dalam database target. Jika tidak, jika database yang ditunjuk

Operasi API	Penanganan tautan sumber daya
	adalah database lokal (dimiliki), operasi mengembalikan semua tabel dalam database lokal, dan mengikuti setiap link sumber daya tabel untuk mengembalikan properti tabel target.
`SearchTables`	Mengembalikan tabel dan link sumber daya tabel. Itu tidak mengikuti tautan untuk mengembalikan properti tabel target. Anda harus menentukan `ResourceShareType = ALL` untuk melihat tabel yang dibagikan dengan akun Anda.
`GetTableVersion`	Sama seperti`GetTable`.
`GetTableVersions`	Sama seperti`GetTable`.
`DeleteTableVersion`	Sama seperti`DeleteTable` .
`BatchDeleteTableVersion`	Sama seperti`DeleteTable` .

Operasi API partisi

Operasi API	Penanganan tautan sumber daya
`CreatePartition`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan membuat partisi dalam tabel yang ditunjuk dalam database target. Jika tabel adalah tautan sumber daya, operasi mengikuti tautan sumber daya dan membuat partisi di tabel target. Membuat partisi melalui tautan sumber daya tabel dan tautan sumber daya basis data tidak didukung.
`BatchCreatePartition`	Sama seperti`CreatePartition` .
`UpdatePartition`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan update partisi dalam tabel yang ditunjuk dalam database target. Jika tabel adalah tautan sumber daya, operasi mengikuti tautan sumber daya dan memperbarui partisi di tabel

Operasi API	Penanganan tautan sumber daya
	target. Memperbarui partisi melalui tautan sumber daya tabel dan tautan sumber daya basis data tidak didukung.
`DeletePartition`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan menghapus partisi dalam tabel yang ditunjuk dalam database target. Jika tabel adalah tautan sumber daya, operasi mengikuti tautan sumber daya dan menghapus partisi di tabel target. Menghapus partisi melalui tautan sumber daya tabel dan tautan sumber daya basis data tidak didukung.
`BatchDeletePartition`	Sama seperti`DeletePartition` .
`GetPartition`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan mengembalikan informasi partisi dari tabel yang ditunjuk. Jika tidak, jika tabel adalah tautan sumber daya, operasi mengikuti tautan dan mengembalikan informasi partisi. Jika kedua tabel dan database adalah link sumber daya, ia mengembalikan set hasil kosong.
`GetPartitions`	Jika database yang ditunjuk adalah link sumber daya, ikuti link database dan mengembalikan informasi partisi untuk semua partisi dalam tabel yang ditunjuk. Jika tidak, jika tabel adalah tautan sumber daya, operasi mengikuti tautan dan mengembalikan informasi partisi. Jika kedua tabel dan database adalah link sumber daya, ia mengembalikan set hasil kosong.
`BatchGetPartition`	Sama seperti`GetPartition` .

User-defined fungsi operasi API

Operasi API	Penanganan Tautan Sumber Daya
(Semua operasi API)	Jika database adalah tautan sumber daya, ikuti tautan sumber daya dan lakukan operasi pada basis data target.

Lihat juga:

- Cara kerja tautan sumber daya di Lake Formation

Mengakses tabel di seluruh Wilayah

Lake Formation mendukung kueri tabel Katalog Data di seluruh AWS Wilayah. Anda dapat mengakses data di Wilayah dari Wilayah lain menggunakan Amazon Athena, Amazon EMR, dan AWS Glue ETL dengan membuat tautan sumber daya di Wilayah lain yang menunjuk ke database dan tabel sumber. Dengan akses tabel lintas wilayah, Anda dapat mengakses data di seluruh Wilayah tanpa menyalin data dasar atau metadata ke dalam Katalog Data.

Misalnya, Anda dapat membagikan database atau tabel di akun produsen ke akun konsumen di Wilayah A. Setelah menerima undangan berbagi sumber daya di Wilayah A, administrator data lake akun konsumen dapat membuat tautan sumber daya ke sumber daya bersama di Wilayah B. Administrator akun konsumen dapat memberikan izin pada sumber daya bersama ke prinsipal IAM di akun tersebut di Wilayah A dan dapat memberikan izin tautan sumber daya di Wilayah B. Dengan menggunakan tautan sumber daya, prinsipnya cipal di akun konsumen dapat menanyakan data bersama dari Wilayah B.

Anda juga dapat meng-host sumber data Amazon S3 di Wilayah A di akun produsen, dan mendaftarkan lokasi data di akun pusat di Wilayah B. Anda dapat membuat sumber daya Katalog Data di akun pusat, mengatur izin Lake Formation, dan berbagi data dengan konsumen di akun Anda atau dengan akun eksternal di Wilayah B. Fitur Lintas wilayah memungkinkan pengguna mengakses tabel Katalog Data ini dari Wilayah C menggunakan tautan sumber daya.

Dengan menggunakan fitur ini, Anda dapat menanyakan database federasi di Apache Hive Metastores di seluruh Wilayah, dan juga menggabungkan tabel di Wilayah lokal dengan tabel di Wilayah lain saat menjalankan kueri.

Lake Formation mendukung fitur-fitur berikut dengan akses tabel lintas wilayah:

- LF-Tag kontrol akses berbasis
- Fine-grained izin kontrol akses
- Menulis operasi pada database atau tabel bersama dengan izin yang sesuai
- Cross-account berbagi data di tingkat akun dan langsung dengan tingkat kepala sekolah IAM

Non-administrative pengguna dengan `Create_Database` dan `Create_Table` izin dapat membuat tautan sumber daya lintas wilayah.

> **Note**
>
> Anda dapat membuat tautan sumber daya lintas wilayah di Wilayah mana pun dan mengakses data tanpa menerapkan izin Lake Formation. Untuk data sumber di Amazon S3 yang tidak terdaftar di Lake Formation, akses ditentukan oleh kebijakan izin IAM untuk Amazon S3 dan tindakan. AWS Glue

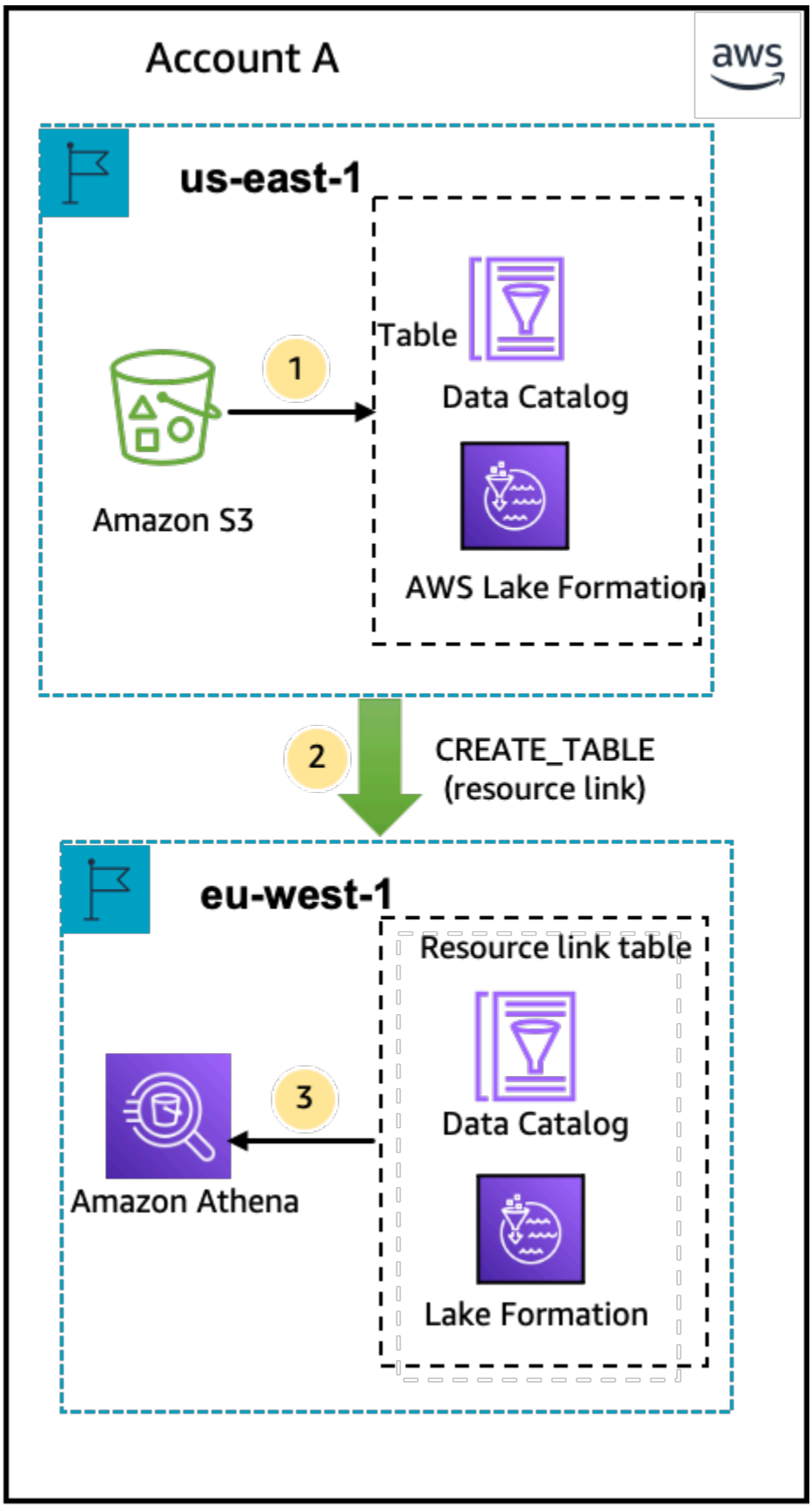
Untuk batasan, lihat Keterbatasan akses data Lintas Wilayah.

Alur kerja

Diagram berikut menunjukkan alur kerja untuk mengakses data di seluruh AWS Wilayah dari AWS akun yang sama dan dari akun eksternal.

Alur kerja untuk mengakses tabel yang dibagikan dalam tabel yang sama AWS akun

Pada diagram di bawah ini, data dibagikan dengan pengguna di AWS akun yang sama di Wilayah AS Timur (Virginia N.), dan pengguna menanyakan data bersama dari Wilayah Eropa (Irlandia).


Account A
aws
us-east-1
Table
Data Catalog
1
Amazon S3
AWS Lake Formation
2
CREATE_TABLE
(resource link)
eu-west-1
Resource link table
Data Catalog
3
Amazon Athena
Lake Formation

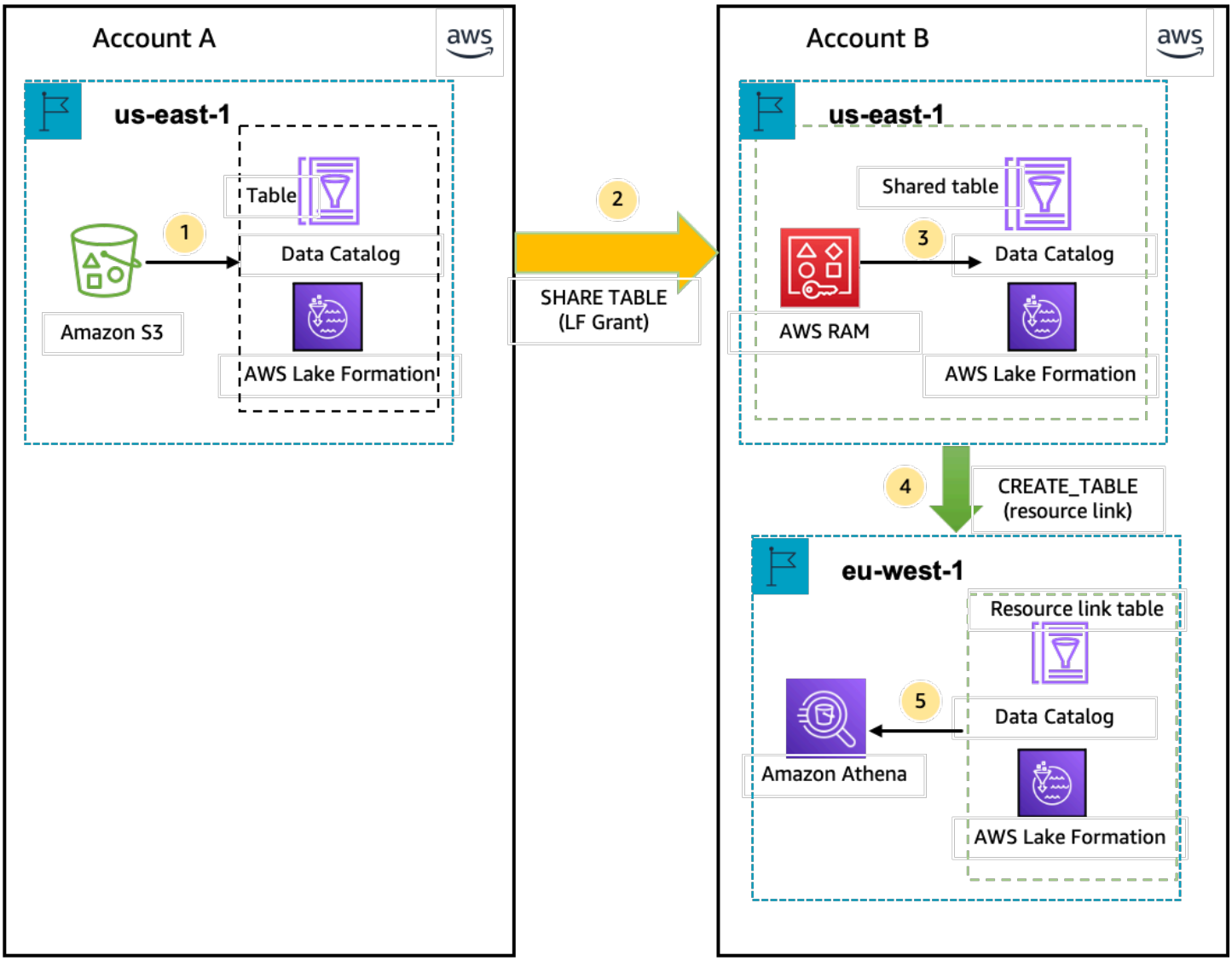
Administrator data lake melakukan kegiatan berikut (langkah 1-2):

1. Administrator data lake menyiapkan AWS akun dengan database dan tabel Katalog Data dan mendaftarkan lokasi data Amazon S3 dengan Lake Formation di Wilayah AS Timur (Virginia N.).

   Memberikan `Select` izin pada sumber daya Katalog Data (tabel produk dalam diagram) kepada prinsipal (pengguna) di akun yang sama.
2. Membuat tautan sumber daya di Wilayah Eropa (Irlandia) yang menunjuk ke tabel sumber di Wilayah AS Timur (Virginia N.). Memberikan `DESCRIBE` izin pada tautan sumber daya dari Wilayah Eropa (Irlandia) kepada kepala sekolah.
3. Pengguna menanyakan tabel dari Wilayah Eropa (Irlandia) menggunakan Athena.

Alur kerja untuk mengakses tabel yang dibagikan dengan eksternal AWS akun

Pada diagram di bawah ini, akun produsen (Akun A) menghosting bucket Amazon S3, mendaftarkan lokasi data, dan membagikan tabel Katalog Data dengan akun konsumen (Akun B) di Wilayah AS Timur (Virginia N.) dan pengguna dari akun konsumen (Akun B) menanyakan tabel dari Wilayah Eropa (Irlandia).



1. Administrator data lake menyiapkan AWS akun (akun produsen) dengan sumber daya Katalog Data dan lokasi data Amazon S3 yang terdaftar di Lake Formation di Wilayah AS Timur (Virginia N.).
2. Administrator data lake dari akun produsen membagikan tabel Katalog Data ke akun konsumen.
3. Administrator data lake dari akun konsumen menerima undangan pembagian data di Wilayah AS Timur (Virginia N.) dan Memberikan `Select` izin pada tabel bersama kepada kepala sekolah dari Wilayah yang sama.
4. Administrator data lake dari akun konsumen membuat tautan sumber daya di Wilayah Eropa (Irlandia) yang menunjuk ke tabel bersama target di Wilayah AS Timur (Virginia N.) dan memberikan izin `DESCRIBE` pengguna pada tautan sumber daya dari Wilayah Eropa (Irlandia).
5. Pengguna menanyakan data dari Wilayah Eropa (Irlandia) menggunakan Athena.

Menyiapkan akses tabel lintas wilayah

Untuk mengakses data dari Wilayah yang berbeda, Anda harus terlebih dahulu menyiapkan database dan tabel Katalog Data di Wilayah tempat Anda mendaftarkan lokasi data Amazon S3. Anda dapat membagikan database dan tabel Katalog Data dengan prinsipal di akun Anda atau di akun lain. Kemudian, Anda perlu membuat administrator data lake yang dapat membuat tautan sumber daya yang menunjuk ke lokasi data bersama target di Wilayah tempat pengguna menanyakan data.

Untuk kueri data yang dibagikan dalam akun yang sama dari Wilayah yang berbeda

Di bagian ini, tabel bersama target Wilayah disebut sebagai Wilayah A dan pengguna menjalankan kueri dari Wilayah B.

1. Pengaturan akun di Wilayah A (tempat Anda membuat dan membagikan data)

   Administrator data lake perlu menyelesaikan tindakan berikut:

   a. Daftarkan lokasi data Amazon S3.

      Untuk informasi selengkapnya, lihat Menambahkan lokasi Amazon S3 ke danau data Anda.

   b. Buat database dan tabel di akun. Ini juga dapat dilakukan oleh pengguna non-administratif yang memiliki izin untuk membuat database dan tabel.

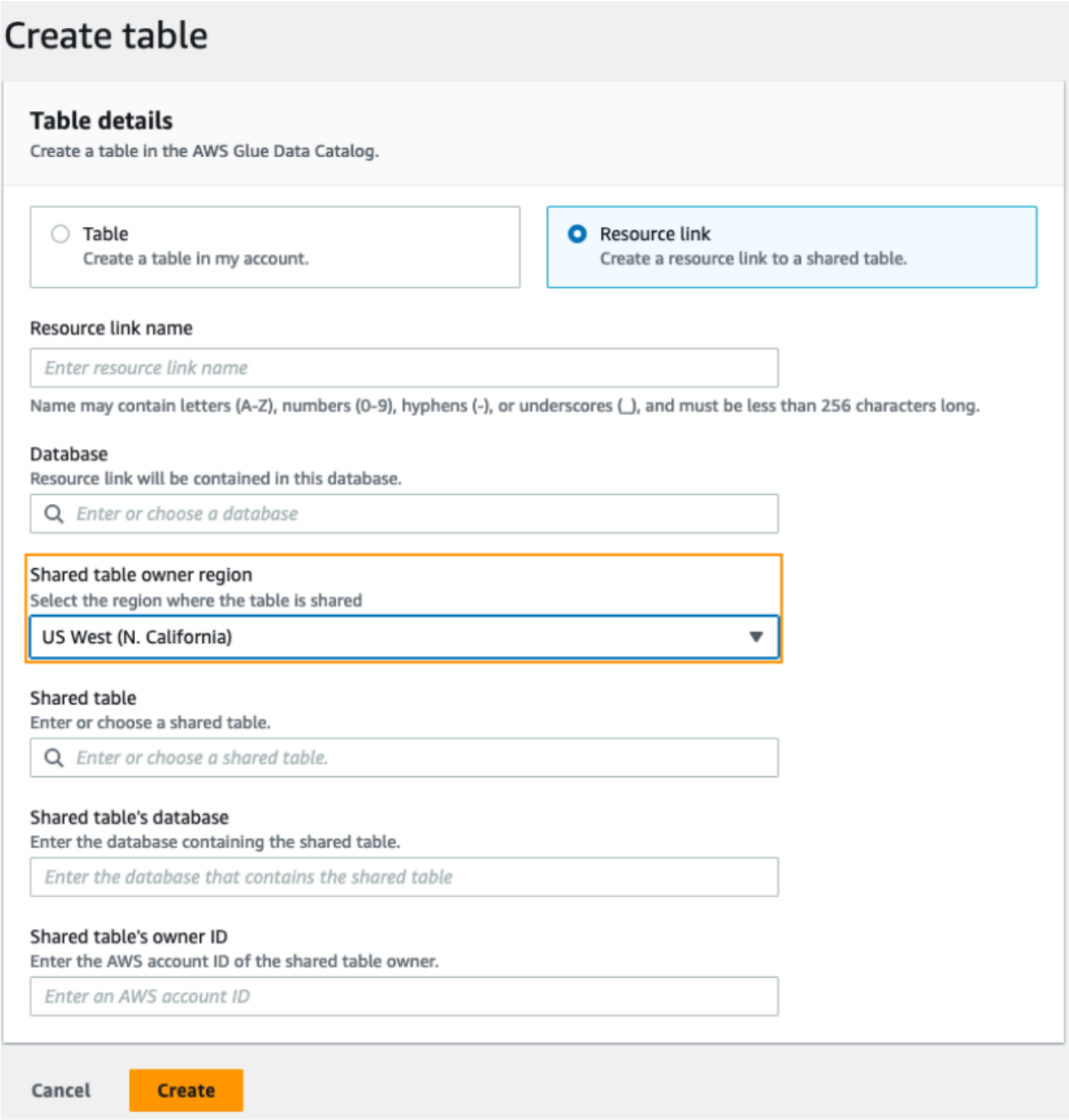
   c. Berikan izin data pada tabel ke kepala sekolah dengan. `Grantable permissions`

      Untuk informasi selengkapnya, lihat Memberikan izin pada sumber daya Katalog Data.

2. Pengaturan akun di Wilayah B (tempat Anda mengakses data)

   Administrator data lake perlu menyelesaikan tindakan berikut:

   a. Buat tautan sumber daya di Wilayah B yang menunjuk ke tabel bersama target di Wilayah A. Tentukan Wilayah pemilik tabel bersama di Buat tabel layar.



Untuk petunjuk tentang cara membuat tautan sumber daya ke database dan tabel, lihatMembuat tautan sumber daya.

b. Berikan `Describe` izin kepada kepala sekolah IAM pada tautan sumber daya di Wilayah B.

Untuk informasi selengkapnya tentang pemberian izin pada tautan sumber daya, lihat. Memberikan izin tautan sumber daya

Prinsipal IAM di Wilayah B dapat menanyakan tabel target melalui tautan menggunakan Athena.

Untuk mengakses data lintas akun dari Wilayah yang berbeda

1. Producer/grantor pengaturan akun

   Administrator data lake perlu menyelesaikan tindakan berikut:

   a. Siapkan producer/grantor akun di Wilayah A.
   b. Daftarkan lokasi data Amazon S3 di Wilayah A.
   c. Buat database dan tabel. Ini dapat dilakukan oleh pengguna non-administratif yang memiliki izin untuk membuat tabel.
   d. Berikan izin data ke consumer/grantee akun pada tabel di Wilayah A dengan`Grantable permissions`.

      Untuk informasi selengkapnya, lihat Berbagi tabel dan database Katalog Data Akun AWS atau prinsip IAM dari akun eksternal.

2. Consumer/grantee pengaturan akun

   Administrator data lake perlu menyelesaikan tindakan berikut:

   a. Terima undangan berbagi sumber daya dari AWS RAM di Wilayah A.
   b. Buat tautan sumber daya di Wilayah B yang menunjuk ke tabel bersama. Wilayah B adalah tempat pengguna ingin menanyakan tabel.
   c. Berikan izin data pada tabel bersama kepada prinsipal IAM di Wilayah A.

      > **Note**
      >
      > Anda harus memberikan izin ke tabel bersama di Wilayah yang sama tempat tabel dibagikan.

   d. Berikan izin kepada kepala sekolah pada tautan sumber daya di Wilayah B.

      Prinsipal di akun konsumen di Wilayah B kemudian menanyakan tabel bersama dari Wilayah B menggunakan Athena.

Keamanan di AWS Lake Formation

Keamanan cloud di AWS adalah prioritas tertinggi. Sebagai AWS pelanggan, Anda mendapat manfaat dari pusat data dan arsitektur jaringan yang dibangun untuk memenuhi persyaratan organisasi yang paling sensitif terhadap keamanan.

Keamanan adalah tanggung jawab bersama antara Anda AWS dan Anda. Model tanggung jawab bersama menggambarkan hal ini sebagai keamanan dari cloud dan keamanan di cloud:

- Keamanan cloud — AWS bertanggung jawab untuk melindungi infrastruktur yang menjalankan AWS layanan di AWS Cloud. AWS juga memberi Anda layanan yang dapat Anda gunakan dengan aman. Auditor pihak ketiga secara berkala menguji dan memverifikasi efektivitas keamanan kami sebagai bagian dari Program kepatuhan AWS. Untuk mempelajari tentang program kepatuhan yang berlaku untuk AWS Lake Formation, lihat AWS Layanan dalam Lingkup berdasarkan Program Kepatuhan.
- Keamanan di cloud — Tanggung jawab Anda ditentukan oleh AWS layanan yang Anda gunakan. Anda juga bertanggung jawab atas faktor lain, yang mencakup sensitivitas data Anda, persyaratan perusahaan Anda, serta undang-undang dan peraturan yang berlaku.

Dokumentasi ini membantu Anda memahami cara menerapkan model tanggung jawab bersama saat menggunakan Lake Formation. Topik berikut menunjukkan cara mengonfigurasi Lake Formation untuk memenuhi tujuan keamanan dan kepatuhan Anda. Anda juga belajar cara menggunakan AWS layanan lain yang membantu Anda memantau dan mengamankan sumber daya Lake Formation Anda.

Topik

- Perlindungan Data di Lake Formation
- Keamanan Infrastruktur di AWS Lake Formation
- Pencegahan "confused deputy" lintas layanan
- Pencatatan peristiwa keamanan di AWS Lake Formation

Perlindungan Data di Lake Formation

Model tanggung jawab AWS bersama model berlaku untuk perlindungan data di. Seperti yang dijelaskan dalam model AWS ini, bertanggung jawab untuk melindungi infrastruktur global yang

menjalankan semua AWS Cloud. Anda bertanggung jawab untuk mempertahankan kendali atas konten yang di-host pada infrastruktur ini. Anda juga bertanggung jawab atas tugas-tugas konfigurasi dan manajemen keamanan untuk Layanan AWS yang Anda gunakan. Untuk informasi selengkapnya tentang privasi data, lihat FAQ Privasi Data AWS . Untuk informasi tentang perlindungan data di Eropa, lihat Pusat Peraturan Umum Perlindungan Data (GDPR).

Untuk tujuan perlindungan data, kami menyarankan Anda melindungi Akun AWS kredensil dan mengatur pengguna individu dengan AWS IAM Identity Center atau AWS Identity and Access Management (IAM). Dengan cara itu, setiap pengguna hanya diberi izin yang diperlukan untuk memenuhi tanggung jawab tugasnya. Kami juga menyarankan supaya Anda mengamankan data dengan cara-cara berikut:

- Gunakan autentikasi multi-faktor (MFA) pada setiap akun.
- Gunakan SSL/TLS untuk berkomunikasi dengan AWS sumber daya. Kami mensyaratkan TLS 1.2 dan menganjurkan TLS 1.3.
- Siapkan API dan pencatatan aktivitas pengguna dengan AWS CloudTrail. Untuk informasi tentang penggunaan CloudTrail jejak untuk menangkap AWS aktivitas, lihat Bekerja dengan CloudTrail jejak di AWS CloudTrail Panduan Pengguna.
- Gunakan solusi AWS enkripsi, bersama dengan semua kontrol keamanan default di dalamnya Layanan AWS.
- Gunakan layanan keamanan terkelola tingkat lanjut seperti Amazon Macie, yang membantu menemukan dan mengamankan data sensitif yang disimpan di Amazon S3.
- Jika Anda memerlukan modul kriptografi tervalidasi FIPS 140-3 saat mengakses AWS melalui antarmuka baris perintah atau API, gunakan titik akhir FIPS. Lihat informasi selengkapnya tentang titik akhir FIPS yang tersedia di Standar Pemrosesan Informasi Federal (FIPS) 140-3.

Kami sangat merekomendasikan agar Anda tidak pernah memasukkan informasi identifikasi yang sensitif, seperti nomor rekening pelanggan Anda, ke dalam tanda atau bidang isian bebas seperti bidang Nama. Ini termasuk saat Anda bekerja dengan atau lainnya Layanan AWS menggunakan konsol, API AWS CLI, atau AWS SDK. Data apa pun yang Anda masukkan ke dalam tanda atau bidang isian bebas yang digunakan untuk nama dapat digunakan untuk log penagihan atau log diagnostik. Saat Anda memberikan URL ke server eksternal, kami sangat menganjurkan supaya Anda tidak menyertakan informasi kredensial di dalam URL untuk memvalidasi permintaan Anda ke server itu.

Enkripsi saat Istirahat

AWS Lake Formation mendukung enkripsi data di bidang-bidang berikut:

- Data di danau data Amazon Simple Storage Service (Amazon S3).

  Lake Formation mendukung enkripsi data dengan AWS Key Management Service(AWS KMS). Data biasanya ditulis ke data lake dengan cara AWS Glue mengekstrak, mengubah, dan memuat (ETL) pekerjaan. Untuk informasi tentang cara mengenkripsi data yang ditulis oleh AWS Glue lowongan, lihat Mengenkripsi Data yang Ditulis oleh Crawler, Pekerjaan, dan Titik Akhir Pengembangan di Panduan Pengembang.AWS Glue
- The AWS Glue Data Catalog, yang merupakan tempat Lake Formation menyimpan tabel metadata yang menggambarkan data di danau data.

  Untuk informasi selengkapnya, lihat Mengenkripsi Katalog Data Anda di Panduan AWS Glue Pengembang.

Untuk menambahkan lokasi Amazon S3 sebagai penyimpanan di danau data Anda, Anda mendaftarkan lokasi dengan AWS Lake Formation. Anda kemudian dapat menggunakan izin Lake Formation untuk kontrol akses berbutir halus ke AWS Glue Data Catalog objek yang mengarah ke lokasi ini, dan ke data yang mendasarinya di lokasi.

Lake Formation mendukung pendaftaran lokasi Amazon S3 yang berisi data terenkripsi. Lihat informasi yang lebih lengkap di Mendaftarkan lokasi Amazon S3 terenkripsi.

Keamanan Infrastruktur di AWS Lake Formation

Sebagai layanan terkelola, AWS Lake Formation dilindungi oleh prosedur keamanan jaringan AWS global yang dijelaskan dalam whitepaper Amazon Web Services: Overview of Security Processes.

Anda menggunakan panggilan API yang AWS dipublikasikan untuk mengakses Lake Formation melalui jaringan. Klien harus mendukung Keamanan Lapisan Pengangkutan (TLS) 1.0 atau versi yang lebih baru. Kami merekomendasikan TLS 1.2 atau versi yang lebih baru. Klien juga harus mendukung suite cipher dengan perfect forward secrecy (PFS) seperti Ephemeral Diffie-Hellman (DHE) atau Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). Sebagian besar sistem modern seperti Java 7 dan sistem yang lebih baru mendukung mode ini.

Selain itu, permintaan harus ditandatangani menggunakan ID kunci akses dan kunci akses rahasia yang terkait dengan principal IAM. Atau Anda bisa menggunakan AWS Security Token Service (AWS STS) untuk membuat kredensial keamanan sementara guna menandatangani permintaan.

Pencegahan "confused deputy" lintas layanan

Masalah "confused deputy" adalah masalah keamanan di mana entitas yang tidak memiliki izin untuk melakukan tindakan dapat memengaruhi entitas yang memiliki hak akses lebih tinggi untuk melakukan tindakan. Pada tahun AWS, peniruan lintas layanan dapat mengakibatkan masalah wakil yang membingungkan. Peniruan identitas lintas layanan dapat terjadi ketika satu layanan (layanan yang dipanggil) memanggil layanan lain (layanan yang dipanggil). Layanan pemanggilan dapat dimanipulasi menggunakan izinnya untuk bertindak pada sumber daya pelanggan lain dengan cara yang seharusnya tidak dilakukannya kecuali bila memiliki izin untuk mengakses. Untuk mencegah hal ini, AWS sediakan alat yang membantu Anda melindungi data Anda untuk semua layanan dengan prinsip layanan yang telah diberikan akses ke sumber daya di akun Anda.

Kami merekomendasikan penggunaan kunci konteks kondisi global `aws:SourceArn` dan `aws:SourceAccount` dalam kebijakan sumber daya untuk membatasi izin yang diberikan oleh AWS Lake Formation ke layanan lain untuk mengakses sumber daya. Jika Anda menggunakan kedua kunci konteks kondisi global, `aws:SourceAccount` nilai dan akun dalam `aws:SourceArn` nilai harus menggunakan ID akun yang sama saat digunakan dalam pernyataan kebijakan yang sama.

Saat ini, Lake Formation hanya mendukung `aws:SourceArn` dalam format berikut:

```
arn:aws:lakeformation:aws-region:account-id:*
```

Contoh berikut menunjukkan bagaimana Anda dapat menggunakan kunci konteks kondisi `aws:SourceAccount` global `aws:SourceArn` dan global di Lake Formation untuk mencegah masalah wakil yang membingungkan.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "ConfusedDeputyPreventionExamplePolicy",
      "Effect": "Allow",
```

```
      "Principal": {
        "Service": "lakeformation.amazonaws.com"
      },
      "Action": [
        "sts:AssumeRole"
      ],
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "ArnEquals": {
          "aws:SourceArn": "arn:aws:lakeformation:aws-region:123456789012:*"
        }
      }
    }
  ]
}
```

Pencatatan peristiwa keamanan di AWS Lake Formation

AWS Lake Formation terintegrasi dengan AWS CloudTrail, layanan yang menyediakan catatan tindakan yang diambil oleh pengguna, peran, atau AWS layanan di Lake Formation. CloudTrail menangkap semua panggilan API untuk Lake Formation sebagai peristiwa. Panggilan yang diambil termasuk panggilan dari konsol Lake Formation, panggilan AWS Command Line Interface, dan kode ke operasi Lake Formation API.

Untuk informasi lebih lanjut tentang pencatatan peristiwa di Lake Formation, lihatPencatatan panggilan API AWS Lake Formation menggunakan AWS CloudTrail.

> **Note**
>
> `GetTableObjects`,`UpdateTableObjects`, dan `GetWorkUnitResults` merupakan operasi pesawat data volume tinggi. Panggilan ke APIs ini saat ini tidak masuk ke CloudTrail. Untuk informasi selengkapnya tentang operasi bidang data CloudTrail, lihat Mencatat peristiwa data untuk jejak di Panduan AWS CloudTrail Pengguna.
> Perubahan dalam Lake Formation untuk mendukung CloudTrail acara tambahan akan didokumentasikan diRiwayat dokumen untuk AWS Lake Formation.

Mengintegrasikan layanan pihak ketiga dengan Lake Formation

Mengintegrasikan dengan AWS Lake Formation memungkinkan layanan pihak ketiga untuk mengakses data dengan aman di danau data berbasis Amazon S3 mereka. Anda dapat menggunakan Lake Formation sebagai mesin otorisasi untuk mengelola atau menerapkan izin ke data lake Anda dengan AWS layanan terintegrasi seperti AWS Glue ETL, Amazon Athena, Amazon EMR, dan Redshift Spectrum. Lake Formation menyediakan dua opsi untuk mengintegrasikan layanan:

1. Pengaturan integrasi aplikasi Lake Formation: Lake Formation dapat menjual kredensi sementara cakupan dalam bentuk token STS AWS ke lokasi Amazon S3 terdaftar berdasarkan izin efektif, sehingga aplikasi yang berwenang dapat mengakses data atas nama pengguna.
2. Penegakan pusat: Operasi API kueri Lake Formation mengambil data dari Amazon S3 dan memfilter hasilnya berdasarkan izin yang efektif. Mesin atau aplikasi yang terintegrasi dengan operasi API kueri dapat bergantung pada Lake Formation untuk mengevaluasi izin identitas panggilan dan memfilter data dengan aman berdasarkan izin ini. Third-party mesin kueri hanya melihat dan beroperasi pada data yang difilter.

Penjual kredenal Lake Formation tidak terintegrasi dengan kueri spark sql. Penjual kredenal hanya berfungsi dengan kueri yang berjalan melalui pustaka ETL. AWS Glue

Topik

- Menggunakan integrasi aplikasi Lake Formation

Menggunakan integrasi aplikasi Lake Formation

Lake Formation memungkinkan layanan pihak ketiga untuk berintegrasi dengan Lake Formation dan mendapatkan akses sementara ke data Amazon S3 atas nama penggunanya dengan menggunakan GetTemporaryGlueTableCredentialsdan GetTemporaryGluePartitionCredentialsmengoperasikannya. Ini memungkinkan layanan pihak ketiga untuk menggunakan fitur otorisasi dan penjual kredenal yang sama dengan yang digunakan oleh layanan AWS analitik lainnya. Bagian ini menjelaskan cara menggunakan operasi API ini untuk mengintegrasikan mesin kueri pihak ketigaLake Formation.

Operasi API ini dinonaktifkan secara default. Ada dua opsi untuk mengotorisasi Lake Formation untuk mengintegrasikan aplikasi:

- Konfigurasikan tag sesi IAM yang divalidasi setiap kali operasi API integrasi aplikasi dipanggil

  Untuk informasi selengkapnya, lihat Mengaktifkan izin untuk mesin kueri pihak ketiga untuk memanggil operasi API integrasi aplikasi.
- Aktifkan opsi yang Memungkinkan mesin eksternal mengakses data di lokasi Amazon S3 dengan akses tabel penuh

  Opsi ini memungkinkan mesin kueri dan aplikasi untuk mendapatkan kredensional tanpa tag sesi IAM jika pengguna memiliki akses tabel penuh. Ini memberikan mesin kueri dan manfaat kinerja aplikasi serta menyederhanakan akses data. Amazon EMR di Amazon EC2 dapat memanfaatkan pengaturan ini.

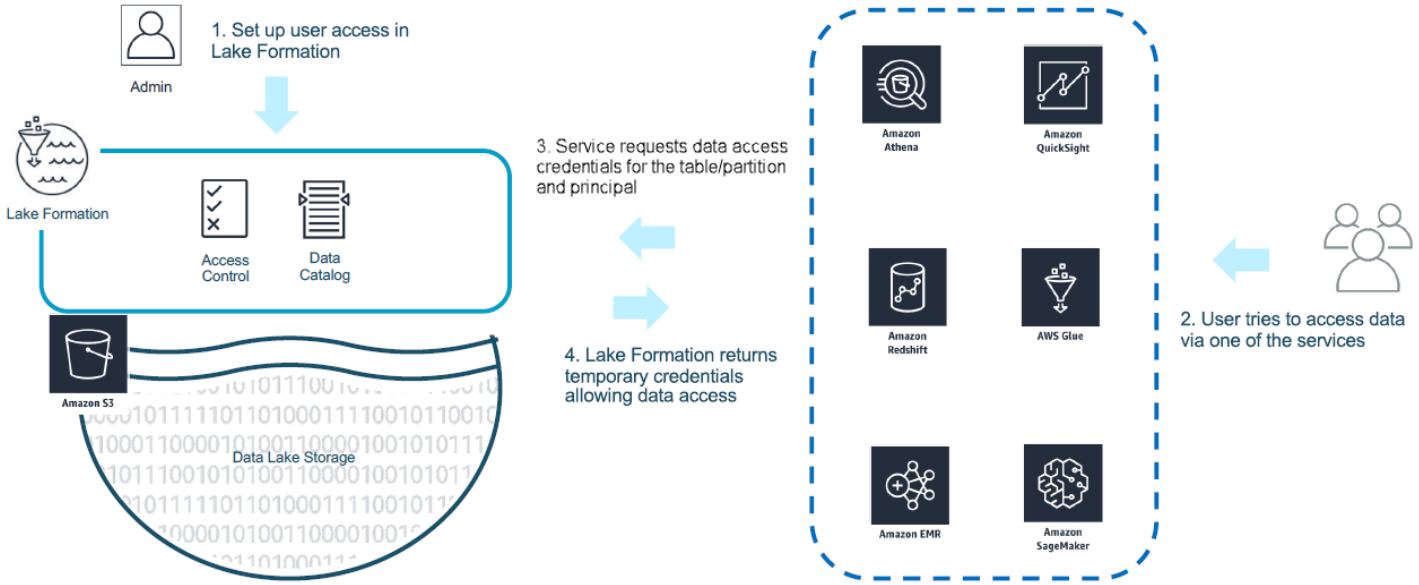
  Untuk informasi selengkapnya, lihat Integrasi aplikasi untuk akses tabel penuh.

Topik

- Cara kerja integrasi aplikasi Lake Formation
- Peran dan tanggung jawab dalam integrasi aplikasi Lake Formation
- Lake Formation alur kerja untuk operasi API integrasi aplikasi
- Mendaftarkan mesin kueri pihak ketiga
- Mengaktifkan izin untuk mesin kueri pihak ketiga untuk memanggil operasi API integrasi aplikasi
- Integrasi aplikasi untuk akses tabel penuh

Cara kerja integrasi aplikasi Lake Formation

Bagian ini menjelaskan cara menggunakan operasi API integrasi aplikasi untuk mengintegrasikan aplikasi pihak ketiga (mesin kueri) denganLake Formation.



1. Lake FormationAdministrator melakukan kegiatan berikut:
   - Mendaftarkan lokasi Amazon S3 dengan Lake Formation dengan menyediakan peran IAM (digunakan untuk kredenal penjual) yang memiliki izin yang sesuai untuk mengakses data dalam lokasi Amazon S3
   - Mendaftarkan aplikasi pihak ketiga untuk dapat memanggil operasi API penjual kredenal Lake Formation. Lihat the section called “Mendaftarkan mesin kueri pihak ketiga”
   - Memberikan izin bagi pengguna untuk mengakses database dan tabel

     Misalnya, jika Anda ingin mempublikasikan kumpulan data sesi pengguna yang menyertakan beberapa kolom yang berisi informasi identitas pribadi (PII), untuk membatasi akses, Anda menetapkan kolom ini LF-TBACtag bernama “klasifikasi” dengan nilai “sensitif”. Selanjutnya, Anda menentukan izin yang memungkinkan analis bisnis mengakses data sesi pengguna, tetapi mengecualikan kolom yang ditandai dengan klasifikasi = sensitif.
2. Seorang prinsipal (pengguna) mengirimkan kueri ke layanan terintegrasi.
3. Aplikasi terintegrasi mengirimkan permintaan ke Lake Formation yang meminta informasi tabel dan kredensional untuk mengakses tabel.
4. Jika prinsipal kueri diizinkan untuk mengakses tabel, Lake Formation mengembalikan kredensialnya ke aplikasi terintegrasi, yang memungkinkan akses data.

> **Note**
>
> Lake Formation tidak mengakses data yang mendasarinya saat menjual kredensyal.

5. Layanan terintegrasi membaca data dari Amazon S3, memfilter kolom berdasarkan kebijakan yang diterimanya, dan mengembalikan hasilnya kembali ke prinsipal.

> **Important**
>
> Lake Formationoperasi API penjual kredenal memungkinkan penegakan terdistribusi dengan model penolakan eksplisit pada kegagalan (fail-close). Ini memperkenalkan model keamanan tiga pihak antara pelanggan, layanan pihak ketiga, dan Lake Formation. Layanan terintegrasi dipercaya untuk menegakkan Lake Formation izin dengan benar (penegakan terdistribusi).

Layanan terintegrasi bertanggung jawab untuk memfilter data yang dibaca dari Amazon S3 berdasarkan kebijakan yang dikembalikan Lake Formation sebelum data yang difilter dikembalikan kembali ke pengguna. Layanan terintegrasi mengikuti model fail-close, yang berarti bahwa mereka harus gagal dalam kueri jika mereka tidak dapat menerapkan izin yang diperlukan. Lake Formation

Peran dan tanggung jawab dalam integrasi aplikasi Lake Formation

Berikut ini adalah peran dan tanggung jawab terkait untuk memungkinkan integrasi aplikasi pihak ketiga dengan AWS Lake Formation.

Peran	Tanggung Jawab
Pelanggan	• Aktifkan pengaturan integrasi aplikasi Lake Formation (lihatthe section called “Mendaftarkan mesin kueri pihak ketiga”). • Secara eksplisit mendaftarkan pihak ketiga yang disetujui dengan Lake Formation (lihat). the section called “Mendaftarkan mesin kueri pihak ketiga” • Menguji dan memvalidasi solusi pihak ketiga dengan izin Lake Formation. • Memantau dan mengaudit penggunaan pihak ketiga dari operasi API penjual kredenal Lake Formation.

Peran	Tanggung Jawab
Pihak ketiga	• Secara publik mendokumentasikan kemampuan yang didukung untuk setiap revisi perangkat lunak dan memberikan instruksi untuk mengaktifkannya dengan benar. • Secara akurat mengiklankan kemampuan yang didukung saat memanggil operasi API penjual kredenal Lake Formation (sesuai dengan dokumentasi). • Menyimpan dan menangani kredensi yang dijual dengan aman untuk menghindari kebocoran kredensi dan eskalasi hak istimewa. • Menerapkan izin berdasarkan kemampuan yang didukung dan hanya mengembalikan data yang difilter ke pengguna • Gagal kueri saat tidak dapat menerapkan izin yang diperlukan dengan benar
`AWS Lake Formation`	• Dengan benar memperoleh dan mengembalikan izin yang efektif untuk prinsipal tertentu. • Memvalidasi kemampuan pihak ketiga yang didukung berdasarkan panggilan per panggilan operasi API. • Mengembalikan kredensi IAM yang tercakup ke bawah hanya jika kemampuan mesin yang diiklankan cocok dengan yang ditentukan pada sumber daya katalog, jika tidak, akan mengembalikan kesalahan.

Lake Formation alur kerja untuk operasi API integrasi aplikasi

Berikut ini adalah alur kerja untuk operasi API integrasi aplikasi:

1. Pengguna mengirimkan kueri atau permintaan data menggunakan mesin kueri pihak ketiga yang terintegrasi. Mesin kueri mengasumsikan peran IAM yang mewakili pengguna atau sekelompok pengguna, dan mengambil kredensi tepercaya untuk digunakan saat memanggil operasi API integrasi aplikasi.
2. Mesin kueri memanggil`GetUnfilteredTableMetadata`, dan jika itu adalah tabel yang dipartisi, mesin kueri memanggil `GetUnfilteredPartitionsMetadata` untuk mengambil metadata dan informasi kebijakan dari Katalog Data.

3. Lake Formation melakukan otorisasi untuk permintaan tersebut. Jika pengguna tidak memiliki izin yang sesuai di atas meja, maka AccessDeniedExceptiondilemparkan.
4. Sebagai bagian dari permintaan, mesin kueri mengirimkan penyaringan yang didukungnya. Ada dua flag yang dapat dikirim dalam array: COLUMN_PERMISSIONS dan CELL_FILTER_PERMISSION. Jika mesin kueri tidak mendukung salah satu fitur ini, dan kebijakan ada di tabel untuk fitur tersebut, maka a PermissionTypeMismatchExceptiondilemparkan dan kueri gagal. Hal ini untuk menghindari kebocoran data.
5. Respons yang dikembalikan berisi yang berikut:
   - Seluruh skema untuk tabel sehingga mesin query dapat menggunakannya untuk mengurai data dari penyimpanan.
   - Daftar kolom resmi yang dapat diakses pengguna. Jika daftar kolom resmi kosong, ini menunjukkan bahwa pengguna memiliki `DESCRIBE` izin, tetapi tidak memiliki `SELECT` izin, dan kueri gagal.
   - Sebuah bendera,`IsRegisteredWithLakeFormation`, yang menunjukkan apakah Lake Formation dapat menjual kredensi ke data sumber daya ini. Jika ini mengembalikan false, maka kredensi pelanggan harus digunakan untuk mengakses Amazon S3.
   - Daftar `CellFilters` jika ada yang harus diterapkan ke baris data. Daftar ini berisi kolom dan ekspresi untuk mengevaluasi setiap baris. Ini seharusnya hanya diisi jika CELL_FILTER_PERMISSION dikirim sebagai bagian dari permintaan dan ada filter data terhadap tabel untuk pengguna pemanggil.
6. Setelah metadata diambil, mesin kueri memanggil `GetTemporaryGlueTableCredentials` atau `GetTemporaryGluePartitionCredentials` untuk mendapatkan AWS kredensional untuk mengambil data dari lokasi Amazon S3.
7. Mesin kueri membaca objek yang relevan dari Amazon S3, memfilter data berdasarkan kebijakan yang diterimanya di langkah 2, dan mengembalikan hasilnya kepada pengguna.

Operasi API integrasi aplikasi untuk Lake Formation berisi konten tambahan untuk mengonfigurasi integrasi dengan mesin kueri pihak ketiga. Anda dapat melihat detail operasi di bagian operasi Credential vending API.

`QuerySessionContext`Ini adalah struktur yang juga dapat dikirim oleh mesin kueri Lake Formation untuk operasi API integrasi aplikasi ini. Hal ini memungkinkan Lake Formation untuk menyimpan dan memanfaatkan konteks tambahan untuk query yang diberikan. Berikut ini memberikan contoh bagaimana QuerySessionContextharus digunakan:

1. Mesin kueri membuat `GetInternalUnfilteredMetadata` panggilan, meneruskan struktur QSC yang berisi id kueri unik dalam permintaan:

```
{
    "QuerySessionContext": {
        "QueryId": "your-unique-identifier-here"
    }
}
```

2. `GetInternalUnfilteredMetadata`Panggilan akan mengembalikan `QueryAuthorizationId` string dalam respons. Pada panggilan kueri berikutnya (dan berikutnya) yang menerima struktur QSC dalam input, mesin kueri melewati struktur QSC yang sama yang sekarang juga berisi yang dikembalikan oleh`QueryAuthorizationId`. Lake Formation Misalkan panggilan berikutnya adalah`GetTemporaryGlueTableCredentials`; permintaan akan berisi:

```
{
    "QuerySessionContext": {
        "QueryAuthorizationId": "lf-returned-query-authz-id-here",
        "QueryId": "your-unique-identifier-here"
    },
}
```

Mendaftarkan mesin kueri pihak ketiga

Sebelum mesin kueri pihak ketiga dapat menggunakan operasi API integrasi aplikasi, Anda perlu mengaktifkan izin secara eksplisit untuk mesin kueri untuk memanggil operasi API atas nama Anda. Ini dilakukan dalam beberapa langkah:

1. Anda perlu menentukan AWS akun dan tag sesi IAM yang memerlukan izin untuk memanggil operasi API integrasi aplikasi melalui AWS Lake Formation konsol, AWS CLI atau. API/SDK
2. Saat mesin kueri pihak ketiga mengasumsikan peran eksekusi di akun Anda, mesin kueri harus melampirkan tag sesi yang terdaftar di Lake Formation yang mewakili mesin pihak ketiga. Lake Formationmenggunakan tag ini untuk memvalidasi bahwa jika permintaan berasal dari mesin yang disetujui. Untuk informasi selengkapnya tentang tag sesi, lihat Tag sesi di Panduan Pengguna IAM.
3. Saat menyiapkan peran eksekusi mesin kueri pihak ketiga, Anda harus memiliki kumpulan izin minimum berikut dalam kebijakan 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. Siapkan kebijakan kepercayaan peran pada peran eksekusi mesin kueri untuk memiliki kontrol akses yang baik pada pasangan nilai kunci tag sesi mana yang dapat dilampirkan ke peran ini. Dalam contoh berikut, peran ini hanya diperbolehkan untuk memiliki kunci tag sesi `"LakeFormationAuthorizedCaller"` dan nilai `"engine1"` tag sesi yang akan dilampirkan, dan tidak ada pasangan nilai kunci tag sesi lainnya yang diizinkan.

```
{
    "Sid": "AllowPassSessionTags",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/query-execution-role"
    },
    "Action": "sts:TagSession",
    "Condition": {
    "StringLike": {
        "aws:RequestTag/LakeFormationAuthorizedCaller": "engine1"        }
    }
}
```

Saat `LakeFormationAuthorizedCaller` memanggil operasi STS:AssumeRole API untuk mengambil kredensil untuk mesin kueri yang akan digunakan, tag sesi harus disertakan dalam permintaan. AssumeRole Kredensi sementara yang dikembalikan dapat digunakan untuk membuat permintaan API integrasi Lake Formation aplikasi.

Lake Formationoperasi API integrasi aplikasi memerlukan prinsipal panggilan untuk menjadi peran IAM. Peran IAM harus menyertakan tag sesi dengan nilai yang telah ditentukan sebelumnya yang telah didaftarkan. Lake Formation Tag ini memungkinkan Lake Formation untuk memverifikasi bahwa peran yang digunakan untuk memanggil operasi API integrasi aplikasi diizinkan untuk melakukannya.

Mengaktifkan izin untuk mesin kueri pihak ketiga untuk memanggil operasi API integrasi aplikasi

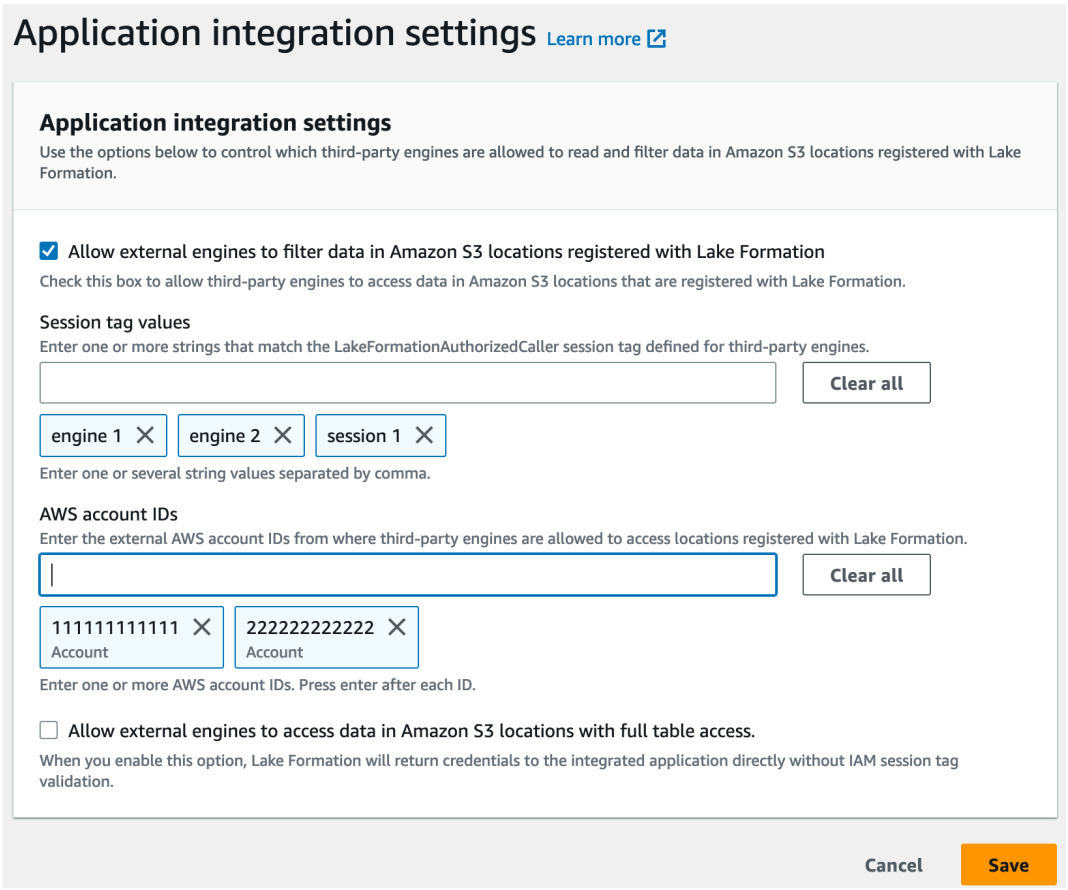
Ikuti langkah-langkah ini untuk mengizinkan mesin kueri pihak ketiga memanggil operasi API integrasi aplikasi melalui AWS Lake Formation konsol, AWS CLI atau API/SDK.

Console

Untuk mendaftarkan akun Anda untuk pemfilteran data eksternal:

1. Masuk ke Konsol Manajemen AWS, dan buka konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di navigasi sisi kiri, perluas Administrasi, lalu pilih Pengaturan integrasi aplikasi.
3. Pada halaman pengaturan Integrasi aplikasi, pilih opsi Izinkan mesin eksternal memfilter data di lokasi Amazon S3 yang terdaftar. Lake Formation
4. Masukkan tag sesi yang Anda buat untuk mesin pihak ketiga. Untuk informasi tentang tag sesi, lihat Melewati tag sesi di AWS STS di Panduan AWS Identity and Access Management Pengguna.
5. Masukkan ID akun untuk pengguna yang dapat menggunakan mesin pihak ketiga untuk mengakses informasi metadata tanpa filter dan kredensi akses data sumber daya di akun saat ini.

   Anda juga dapat menggunakan bidang ID AWS akun untuk mengonfigurasi akses lintas akun.



CLI

Gunakan perintah `put-data-lake-settings` CLI untuk mengatur parameter berikut.

Ada tiga bidang untuk dikonfigurasi saat menggunakan AWS CLI perintah ini:

- `allow-external-data-filtering` — (boolean) Menunjukkan bahwa mesin pihak ketiga dapat mengakses informasi metadata tanpa filter dan akses data dari sumber daya di akun saat ini.
- `external-data-filtering-allow-list`— (array) Daftar ID akun yang dapat mengakses informasi metadata tanpa filter dan akses data dari sumber daya di akun saat ini saat menggunakan mesin pihak ketiga. Bila AllowExternalDataFiltering disetel ke true, ExternalDataFilteringAllowList properti harus menyertakan setidaknya satu ID akun. Daftar kosong tidak diperbolehkan.
- `authorized-sessions-tag-value-list`— (array) Daftar nilai tag sesi resmi (string). Jika kredensi peran IAM telah dilampirkan dengan pasangan nilai kunci resmi, maka jika tag sesi disertakan dalam daftar, sesi diberikan akses ke informasi metadata tanpa filter dan kredenal

akses data pada sumber daya di akun yang dikonfigurasi. Kunci tag sesi resmi didefinisikan sebagai`*LakeFormationAuthorizedCaller*`.

- `AllowFullTableExternalDataAccess`- (boolean) Apakah akan mengizinkan mesin kueri pihak ketiga untuk mendapatkan kredensional akses data tanpa tag sesi ketika pemanggil memiliki izin akses data penuh.

Contoh:

```
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

Gunakan operasi `PutDataLakeSetting` API untuk mengatur parameter berikut.

Ada tiga bidang yang harus dikonfigurasi saat menggunakan operasi API ini:

- `AllowExternalDataFiltering`— (Boolean) Menunjukkan apakah mesin pihak ketiga dapat mengakses informasi metadata tanpa filter dan kredenal akses data sumber daya di akun saat ini.

- `ExternalDataFilteringAllowList`— (array) Daftar ID akun yang dapat mengakses informasi metadata tanpa filter dan kredensi akses data sumber daya di akun saat ini menggunakan mesin pihak ketiga.
- `AuthorizedSectionsTagValueList`— (array) Daftar nilai tag resmi (string). Jika kredensi peran IAM telah dilampirkan dengan tag resmi, maka sesi diberikan akses ke informasi metadata tanpa filter dan kredenal akses data pada sumber daya di akun yang dikonfigurasi. Kunci tag sesi resmi didefinisikan sebagai`*LakeFormationAuthorizedCaller*`.
- `AllowFullTableExternalDataAccess`- (boolean) Apakah akan mengizinkan mesin kueri pihak ketiga untuk mendapatkan kredensional akses data tanpa tag sesi ketika pemanggil memiliki izin akses data penuh.

Contoh:

```
//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));
}
```

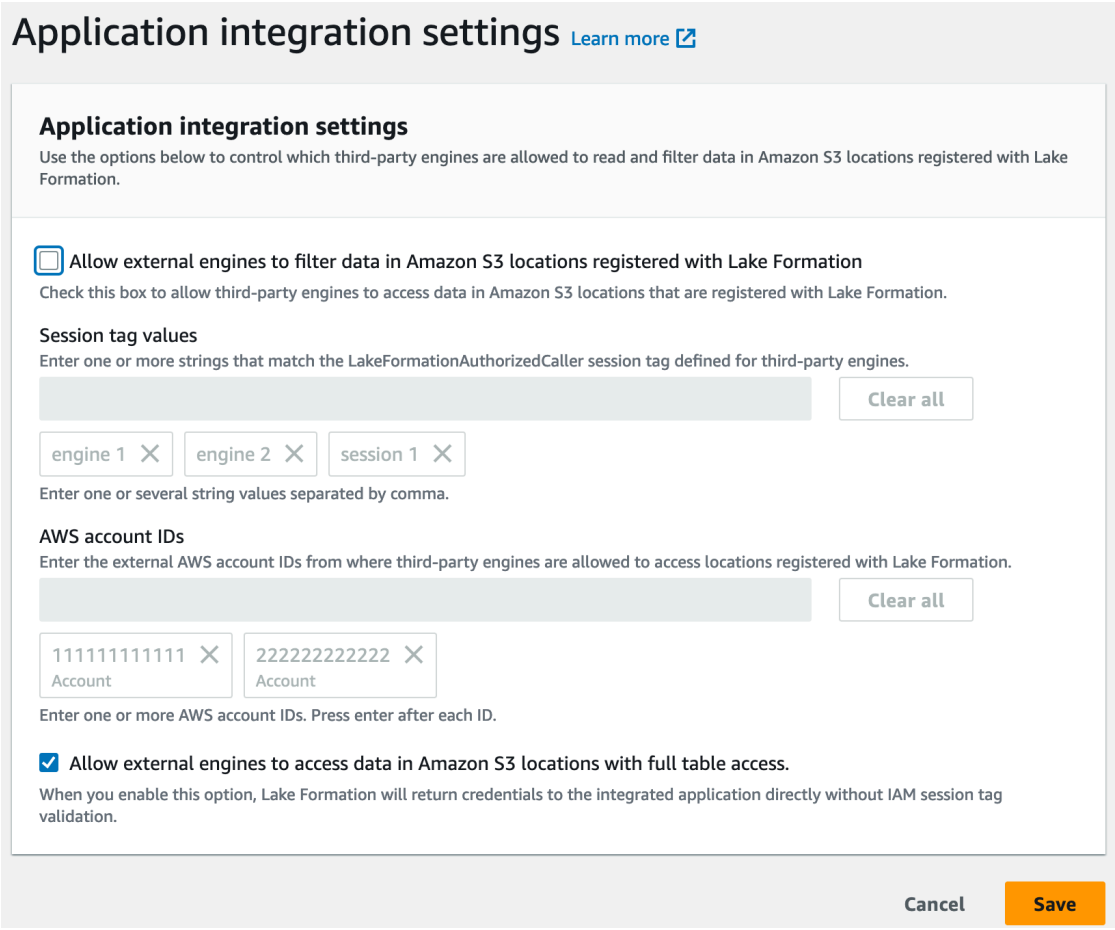
Integrasi aplikasi untuk akses tabel penuh

Ikuti langkah-langkah berikut untuk mengaktifkan mesin kueri pihak ketiga mengakses data tanpa validasi tag sesi IAM:

Console

1. Masuk ke konsol Lake Formation di https://console.aws.amazon.com/lakeformation/.
2. Di navigasi sisi kiri, perluas Administrasi, dan pilih Pengaturan integrasi aplikasi.
3. Pada halaman Pengaturan integrasi aplikasi, pilih opsi Izinkan mesin eksternal mengakses data di lokasi Amazon S3 dengan akses tabel penuh.

   Saat Anda mengaktifkan opsi ini, Lake Formation mengembalikan kredensi ke aplikasi kueri secara langsung tanpa validasi tag sesi IAM.



AWS CLI

Gunakan perintah `put-data-lake-settings` CLI untuk mengatur parameter. `AllowFullTableExternalDataAccess`

```
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
    }
}
```

Bekerja dengan AWS layanan lain

AWS layanan seperti Amazon Athena, Amazon Redshift AWS Glue Spectrum, dan Amazon EMR dapat AWS Lake Formation digunakan untuk mengakses data dengan aman di lokasi Amazon S3 yang terdaftar di Lake Formation. Dengan Lake Formation, Anda dapat menentukan dan mengelola izin kontrol akses berbutir halus (FGAC) untuk tabel Anda di. AWS Glue Data Catalog Masing-masing AWS layanan ini adalah penelepon tepercaya ke Lake Formation, dan Lake Formation menyediakan akses ke data yang disimpan di Amazon S3 melalui kredenal sementara. Untuk informasi selengkapnya, lihat Cara kerja integrasi aplikasi Lake Formation.

Untuk memanfaatkan kemampuan ini, Lake Formation mengharuskan Anda mendaftarkan lokasi Amazon S3 terlebih dahulu, dan menetapkan izin yang sesuai ke kepala IAM untuk mengakses tabel, database, dan lokasi Amazon S3. Untuk informasi selengkapnya, lihat Mengelola izin Lake Formation.

Tabel berikut mencantumkan jenis izin Lake Formation yang didukung oleh Amazon Athena AWS Glue, Amazon EMR, dan Amazon Redshift Spectrum untuk mengakses data AWS Glue dari tabel standar dan tabel transaksional (Apache Iceberg, Apache Hudi, dan Linux foundation Delta Lake) dengan data yang disimpan di Amazon S3 dan metadata tabel di Katalog Data.

AWS layanan dan jenis izin yang didukung untuk tabel dan tampilan AWS Glue standar

AWS layanan	Izin tingkat tabel	Izin tingkat kolom	Izin baris dan tingkat sel
Athena SQL	Akses baca/tulis	Akses baca	Akses baca
Athena Spark	Tidak didukung	Tidak didukung	Tidak didukung
Redshift Spectrum pada cluster yang disediakan atau Amazon Redshift tanpa server	Akses baca/tulis	Akses baca	Akses baca
Apache Spark di Amazon EMR (EC2)	Akses baca/tulis	Akses baca	Akses baca

AWS layanan	Izin tingkat tabel	Izin tingkat kolom	Izin baris dan tingkat sel
Sarang Apache di Amazon EMR (EC2)	Akses baca/tulis	Akses baca	Tidak didukung
Apache Spark di EMR Tanpa Server	Akses baca/tulis	Akses baca	Akses baca
Apache Hive di EMR Tanpa Server	Tidak didukung	Tidak didukung	Tidak didukung
Amazon EMR di EKS	Tidak didukung	Tidak didukung	Tidak didukung
AWS Glue ETL	Akses baca/tulis	AWS Glue 5.0 atau lebih tinggi mendukung akses baca.	AWS Glue 5.0 atau lebih tinggi mendukung akses baca.

Pertimbangan dan batasan

- Athena Spark tidak mendukung kueri tabel Katalog Data dengan izin Lake Formation.
- Pengguna yang berbasis di Athena SAMP dapat membaca sumber data yang diamankan menggunakan izin Lake Formation dengan mengaktifkan federasi berbasis SAMP 2.0. Pengguna SALL dapat memasukkan data ke dalam tabel Parket.
- Apache Spark di EMR Tanpa Server tidak mendukung kueri tampilan Katalog Data.
- Apache Hive di EMR Serverless tidak mendukung tabel kueri dengan izin Lake Formation.
- AWS Glue 5.0 atau yang lebih tinggi mendukung kontrol akses berbutir halus pada tabel Iceberg dan Hive di Katalog Data yang didukung oleh S3. Kemampuan ini memungkinkan Anda mengonfigurasi kontrol akses tingkat tabel, baris, kolom, dan sel untuk kueri baca dalam pekerjaan Apache Spark Anda AWS Glue .

  Untuk informasi selengkapnya, lihat AWS Glue versi.

AWS layanan dan jenis izin yang didukung untuk format tabel transaksional

AWS layanan	Gunung es	Hudi	Danau Delta (asli)	Danau Delta (tabel symlink)
Athena SQL	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis memerlukan akses tabel penuh.	Mendukung membaca dan membuat operasi pada tabel dengan tabel, kolom, baris, dan izin tingkat sel. Operasi tulis tidak didukung.	Athena (versi mesin 3) mendukung membaca tabel Delta Lake asli dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis tidak didukung.	Athena (versi mesin 3) mendukung membaca tabel symlink Delta Lake dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis tidak didukung.
Spektrum Pergeseran Merah pada cluster yang disediakan	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis tidak didukung.	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis tidak didukung.	Tidak didukung	Mendukung membaca tabel Delta Lake melalui manifes symlink dengan tabel, kolom, baris, dan izin tingkat sel. Operasi tulis tidak didukung.
Apache Spark di Amazon EMR (EC2)	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis memerlukan akses tabel penuh.	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis memerlukan akses tabel penuh.	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis tidak didukung.	Mendukung tabel bacaan dengan izin tabel, kolom, baris, dan tingkat sel. Operasi tulis memerlukan akses tabel penuh.
AWS Glue ETL	AWS Glue 5.0 atau lebih tinggi	Mendukung read/write pada	Mendukung read/write pada	Mendukung read/write pada

AWS layanan	Gunung es	Hudi	Danau Delta (asli)	Danau Delta (tabel symlink)
	mendukung tabel bacaan dengan tabel, kolom, baris, dan izin tingkat sel.	tabel dengan izin tingkat tabel.	tabel dengan izin tingkat tabel.	tabel dengan izin tingkat tabel.

Topik

- Menggunakan AWS Lake Formation dengan Amazon Athena
- Menggunakan AWS Lake Formation dengan Amazon Redshift Spectrum
- Menggunakan AWS Lake Formation dengan AWS Glue
- Menggunakan AWS Lake Formation dengan Amazon EMR
- Menggunakan AWS Lake Formation dengan Cepat
- Menggunakan AWS Lake Formation dengan AWS CloudTrail Danau

Menggunakan AWS Lake Formation dengan Amazon Athena

Amazon Athena adalah layanan kueri tanpa server yang membantu Anda menganalisis data terstruktur, semi-terstruktur, dan tidak terstruktur yang disimpan di Amazon S3. Anda dapat menggunakan Athena SQL untuk menanyakan data dari format data CSV, JSON, Parket, dan Avro. Athena SQL juga mendukung format tabel seperti Apache Hive, Apache Hudi, dan Apache Iceberg. Athena terintegrasi dengan metadata AWS Glue Data Catalog untuk menyimpan kumpulan data Anda di Amazon S3. Athena dapat menggunakan Lake Formation untuk mendefinisikan dan memelihara kebijakan kontrol akses pada kumpulan data tersebut.

Berikut adalah beberapa kasus penggunaan umum di mana Anda dapat menggunakan Lake Formation dengan Athena.

- Gunakan izin Lake Formation untuk mengakses sumber daya Katalog Data (database dan tabel) dari Athena. Anda dapat menggunakan salah satu metode sumber daya bernama atau LF-tag untuk menentukan izin pada database dan tabel. Untuk informasi lebih lanjut, lihat:
  - Memberikan izin database menggunakan metode sumber daya bernama
  - Kontrol akses berbasis tag Lake Formation

> **Note**
>
> Izin Lake Formation hanya berlaku saat menggunakan Athena SQL untuk menanyakan data sumber dari Amazon S3 dan metadata di Katalog Data.
> Athena Spark tidak mendukung kueri tabel Katalog Data dengan izin Lake Formation. Izin Lake Formation mendukung operasi baca dan tulis pada database dan tabel.

> **Note**
>
> Anda tidak dapat menerapkan filter data saat menggunakan LF-tag untuk mengelola izin pada sumber daya Katalog Data.

- Kontrol hasil kueri menggunakan Filter data di Lake Formation untuk mengamankan tabel di data lake Amazon S3 Anda dengan memberikan izin di kolom, baris, dan tingkat sel. Lihat batasan proyeksi partisi di Panduan Pengguna Amazon Athena.
- Menerapkan kontrol akses berbutir halus pada data yang tersedia untuk pengguna Athena berbasis SAMP saat menjalankan kueri federasi.

  Driver Athena JDBC dan ODBC mendukung konfigurasi akses federasi ke sumber data Anda menggunakan Penyedia Identitas berbasis SAML (iDP). Gunakan Cepat terintegrasi dengan Lake Formation dengan peran IAM Anda yang ada atau pengguna atau grup SAMP untuk memvisualisasikan hasil kueri Athena.

  > **Note**
  >
  > Izin Lake Formation untuk pengguna dan grup SAMP hanya akan berlaku ketika Anda mengirimkan kueri ke Athena menggunakan driver JDBC atau ODBC.

  Untuk informasi lebih lanjut, lihat Menggunakan Lake Formation dan Athena JDBC dan ODBC driver untuk akses federasi ke Athena.

  > **Note**
  >
  > Saat ini, otorisasi akses ke identitas SAMP di Lake Formation tidak didukung di wilayah berikut:

- Timur Tengah (Bahrain) - me-south-1
- Asia Pasifik (Hong Kong) - ap-east-1
- Afrika (Cape Town) - af-south-1
- Tiongkok (Ningxia) - cn-barat laut-1
- Asia Pasifik (Osaka) - ap-northeast-3

- Gunakan Cross-account berbagi data di Lake Formation untuk menanyakan tabel di akun lain.

> **Note**
>
> Untuk informasi selengkapnya tentang batasan saat menggunakan izin Lake FormationViews, lihat Pertimbangan dan Batasan.

Support untuk format tabel transaksional

Menerapkan izin Lake Formation memungkinkan Anda mengamankan data transaksional di danau data berbasis Amazon S3 Anda. Tabel di bawah ini mencantumkan format tabel transaksional yang didukung di Athena dan izin Lake Formation. Lake Formation memberlakukan izin ini saat pengguna Athena menjalankan kueri mereka.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Athena
Apache Hudi	Format yang digunakan untuk menyederhanakan pemrosesan data inkremental dan pengembangan pipa data. Athena mendukung operasi buat dan baca menggunak an format tabel Apache Hudi pada kumpulan data Amazon S3 untuk tipe tabel Copy on Write (CoW) dan Merge On Read (MoR) Hudi. Athena	Gunakan Pemfilteran data dan keamanan tingkat sel di Lake Formation untuk mengamank an tabel Hudi menggunakan tabel, kolom, baris, dan izin tingkat sel.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Athena
	tidak mendukung operasi menulis pada tabel Hudi. Gunakan Athena untuk menanyakan kumpulan data Hudi.	
Gunung Es Apache	Format tabel terbuka yang mengelola koleksi besar file sebagai tabel, dan mendukung operasi danau data analitik modern seperti penyisipan tingkat catatan, pembaruan , penghapusan, dan kueri perjalanan waktu. Untuk informasi lebih lanjut tentang dukungan Athena untuk tabel Iceberg, lihat Menggunakan tabel Iceberg.	Izin tabel, kolom, baris, dan tingkat sel didukung. Saat ini, Lake Formation tidak mendukung pengelolaan izin pada operasi tulis seperti `VACUUMMERGE`, `UPDATE` dan `OPTIMIZE` pada tabel dalam Format Tabel Terbuka.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Athena
Yayasan Linux Delta Lake	Delta Lake adalah proyek sumber terbuka yang membantu mengimple mentasikan arsitektur data lake modern yang biasanya dibangun di Amazon S3 atau Hadoop Distributed File System (HDFS). Athena mendukung tabel danau Delta yang dibuat menggunakan definisi tabel manifes berbasis symlink dari tabel Delta Lake. AWS Glue Data Catalog Untuk informasi selengkapnya, lihat tabel Crawl Delta Lake menggunakan AWS Glue crawler. Athena (versi mesin 3) mendukung membaca tabel Danau Delta asli. Untuk informasi selengkap nya, lihat Memperkenalkan dukungan tabel Delta Lake asli dengan AWS Glue crawler.	Izin tabel, kolom, baris, dan tingkat sel didukung untuk tabel symlink dan tabel Delta Lake asli.

Sumber daya tambahan

Posting blog, video, dan lokakarya

- Kueri kumpulan data Apache Hudi di danau data Amazon S3 dengan Amazon Athena

- Bangun data lake Apache Iceberg menggunakan Amazon Athena, Amazon EMR, dan AWS Glue
- Masukkan, perbarui, hapus di Amazon S3 dengan Athena dan Apache Iceberg
- LF-Tag berbasis lokakarya Lake Formation kontrol akses untuk menanyakan data lake.

Menggunakan AWS Lake Formation dengan Amazon Redshift Spectrum

Amazon Redshift Spectrum memungkinkan Anda untuk menanyakan dan mengambil data di data lake Amazon S3 tanpa memuat data ke node cluster Amazon Redshift.

Redshift Spectrum mendukung dua cara mendaftarkan katalog AWS Glue data eksternal yang diaktifkan dengan Lake Formation.

- Menggunakan peran IAM terlampir cluster yang memiliki izin ke Katalog Data

  Untuk membuat peran IAM, ikuti langkah-langkah yang diuraikan dalam prosedur di bawah ini.

  Mengontrol akses ke AWS Glue Data Catalog
- Menggunakan identitas IAM federasi yang dikonfigurasi untuk mengelola akses ke sumber daya eksternal AWS Glue Data Catalog

  Redshift Spectrum mendukung kueri tabel Lake Formation menggunakan identitas IAM federasi. Identitas IAM dapat berupa pengguna IAM atau peran IAM. Untuk informasi selengkapnya tentang federasi identitas IAM di Redshift Spectrum, lihat Menggunakan identitas federasi untuk mengelola akses Amazon Redshift ke sumber daya lokal dan tabel eksternal Redshift Spectrum.

Dengan integrasi Lake Formation dengan Redshift Spectrum, Anda dapat menentukan baris, kolom, dan izin kontrol akses tingkat sel pada tabel setelah data Anda terdaftar di Lake Formation.

Untuk informasi lebih lanjut, lihat Menggunakan Spektrum Pergeseran Merah dengan. AWS Lake Formation

Redshift Spectrum mendukung pembacaan atau `SELECT` kueri pada tabel skema eksternal yang dikelola Lake Formation.

Untuk informasi selengkapnya, lihat Membuat skema eksternal untuk Spektrum Pergeseran Merah.

Support untuk tipe tabel transaksional

Tabel ini mencantumkan format tabel transaksional yang didukung dalam Redshift Spectrum dan izin Lake Formation yang berlaku.

Format tabel yang didukung

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Redshift Spectrum
Apache Hudi	Format yang digunakan untuk menyederhanakan pemrosesan data inkremental dan pengembangan pipa data. Redshift Spectrum mendukung operasi insert, delete, dan upsert write menggunakan format tabel Apache Hudi Copy on Write (CoW) di Amazon S3. Untuk informasi selengkapnya, lihat Membuat tabel eksternal untuk data yang dikelola di Apache Hudi.	Gunakan Pemfilteran data dan keamanan tingkat sel di Lake Formation untuk mengamankan tabel Hudi menggunakan tabel, kolom, baris, dan izin tingkat sel.
Gunung Es Apache	Format tabel terbuka yang mengelola koleksi besar file sebagai tabel dan mendukung operasi danau data analitik modern seperti penyisipan tingkat catatan, pembaruan, penghapusan, dan kueri perjalanan waktu. Untuk informasi selengkapnya, lihat Menggunakan tabel	Redshift Spectrum mendukung tabel Apache Iceberg untuk kueri.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Redshift Spectrum
	Apache Iceberg dengan Amazon Redshift.	
Yayasan Linux Delta Lake	Delta Lake adalah proyek sumber terbuka yang membantu mengimple mentasikan arsitektur data lake modern yang biasanya dibangun di Amazon S3 atau Hadoop Distributed File System (HDFS). Redshift Spectrum mendukung kueri tabel Delta Lake. Untuk informasi selengkapnya, lihat Membuat tabel eksternal untuk data yang dikelola di Delta Lake.	Izin tabel, kolom, baris, dan tingkat sel didukung.

Sumber daya tambahan

Posting blog dan lokakarya

- Memusatkan tata kelola untuk data lake Anda menggunakan AWS Lake Formation sambil mengaktifkan arsitektur data modern dengan Amazon Redshift Spectrum
- Gunakan Redshift Spectrum untuk menanyakan tabel Apache HUDI Copy On Write (CoW) di danau data Amazon S3

Menggunakan AWS Lake Formation dengan AWS Glue

Insinyur data dan DevOps profesional menggunakan AWS Glue Extract, Transform and Load (ETL) dengan Apache Spark untuk melakukan transformasi pada kumpulan data mereka di Amazon S3 dan memuat data yang diubah ke dalam data lake dan gudang data untuk analitik, pembelajaran mesin,

dan pengembangan aplikasi. Dengan tim yang berbeda mengakses kumpulan data yang sama di Amazon S3, sangat penting untuk memberikan dan membatasi izin berdasarkan peran mereka.

AWS Lake Formation dibangun di atas AWS Glue, dan layanan berinteraksi dengan cara-cara berikut:

- Lake Formation dan AWS Glue berbagi Katalog Data yang sama.
- Fitur konsol Lake Formation berikut memanggil AWS Glue konsol:
  - Lowongan — Untuk informasi selengkapnya, lihat Menambahkan Lowongan di Panduan AWS Glue Pengembang.
  - Crawler — Untuk informasi selengkapnya, lihat Katalogisasi Tabel dengan Crawler di Panduan Pengembang.AWS Glue
- Alur kerja yang dihasilkan saat Anda menggunakan cetak biru Lake Formation adalah alur kerja. AWS Glue Anda dapat melihat dan mengelola alur kerja ini di konsol Lake Formation dan AWS Glue konsol.
- Transformasi pembelajaran mesin disediakan dengan Lake Formation dan dibangun di atas operasi AWS Glue API. Anda membuat dan mengelola transformasi pembelajaran mesin di AWS Glue konsol. Untuk informasi selengkapnya, lihat Transformasi Machine Learning di Panduan AWS Glue Pengembang.

Anda dapat menggunakan kontrol akses berbutir halus Lake Formation untuk mengelola sumber daya Katalog Data dan lokasi data Amazon S3 yang ada.

> **Note**
>
> AWS Glue 5.0 atau lebih tinggi mendukung kontrol akses berbutir halus pada tabel Iceberg dan Hive yang didukung oleh S3. Kemampuan ini memungkinkan Anda mengonfigurasi kontrol akses tingkat tabel, baris, kolom, dan sel untuk kueri baca dalam pekerjaan Apache Spark Anda AWS Glue .

Support untuk tipe tabel transaksional

Menerapkan izin Lake Formation memungkinkan Anda mengamankan data transaksional di danau data berbasis Amazon S3 Anda. Tabel di bawah ini mencantumkan format tabel transaksional yang didukung dalam AWS Glue dan izin Lake Formation. Lake Formation memberlakukan izin ini untuk AWS Glue operasi.

Format tabel yang didukung

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di AWS Glue
Apache Hudi	Format tabel terbuka yang digunakan untuk menyederh anakan pemrosesan data inkremental dan pengemban gan pipa data. Sebagai contoh, lihat Menggunakan kerangka kerja Hudi di. AWS Glue	Izin tingkat tabel tersedia untuk tabel Hudi. Untuk informasi selengkapnya, lihat Batas.
Gunung Es Apache	Format tabel terbuka yang mengelola koleksi besar file sebagai tabel. Sebagai contoh, lihat Menggunakan kerangka Iceberg di. AWS Glue	AWS Glue versi 5.0 dan yang lebih tinggi memungkinkan Anda mengonfigurasi kontrol akses tingkat tabel, baris, kolom, dan sel untuk kueri baca dalam pekerjaan Apache Spark Anda AWS Glue untuk tabel Iceberg. Untuk informasi selengkapnya, lihat Batas.
Yayasan Linux Delta Lake	Delta Lake adalah proyek sumber terbuka yang membantu mengimple mentasikan arsitektur data lake modern yang biasanya dibangun di Amazon S3 atau Hadoop Distributed File System (HDFS).	Izin tingkat tabel tersedia untuk tabel Delta Lake. Untuk informasi selengkapnya, lihat Batas.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di AWS Glue
	Sebagai contoh, lihat Menggunakan kerangka Delta Lake di AWS Glue.	

Sumber daya tambahan

Posting blog dan repositori

- Gunakan AWS Glue konektor untuk membaca dan menulis tabel Apache Iceberg dengan transaksi ACID dan melakukan perjalanan waktu
- Menulis ke tabel Apache Hudi menggunakan konektor khusus AWS Glue
- AWS repositori template Cloudformation dan contoh kode pyspark untuk menganalisis data streaming menggunakan, Apache Hudi AWS Glue, dan Amazon S3.

Menggunakan AWS Lake Formation dengan Amazon EMR

Amazon EMR adalah platform cluster AWS terkelola yang fleksibel tempat Anda dapat menjalankan kode khusus apa pun pada kerangka kerja data besar yang didukung seperti Hadoop Map-Reduce, Spark, Hive, Presto, dll. Organizations juga menggunakan Amazon EMR untuk menjalankan aplikasi pemrosesan data batch dan streaming di seluruh cluster yang sangat terdistribusi. Menggunakan Apache Spark di Amazon EMR, Anda dapat menjalankan transformasi data dan kode kustom pada database dan tabel yang izinnya dikelola oleh Lake Formation.

Ada tiga opsi untuk menyebarkan Amazon EMR:

- EMR pada EC2
- EMR Tanpa Server
- Amazon EMR di EKS

Untuk informasi selengkapnya, lihat Mengintegrasikan Amazon EMR dengan Lake Formation atau Menggunakan EMR Tanpa Server dengan kontrol akses berbutir halus AWS Lake Formation

Support untuk format tabel transaksional

Amazon EMR merilis 6.15.0 dan yang lebih tinggi termasuk dukungan untuk tabel Lake Formation, baris, kolom, dan izin kontrol akses tingkat sel pada format tabel Apache Hudi, Apache Iceberg, dan Delta Lake saat Anda membaca dan menulis data dengan Spark SQL.

Untuk batasan, lihat Pertimbangan untuk Amazon EMR dengan Lake Formation.

Format tabel yang didukung

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Amazon EMR
Apache Hudi	Format tabel terbuka yang digunakan untuk menyederhanakan pemrosesan data inkremental dan pengembangan pipa data. Untuk daftar operasi yang didukung, lihat Apache Hudi dan Lake Formation.	Amazon EMR mendukung tabel, baris, kolom, dan kontrol akses tingkat sel dengan Apache Hudi.
Gunung Es Apache	Format tabel terbuka yang mengelola koleksi besar file sebagai tabel. Untuk daftar operasi yang didukung, lihat Apache Iceberg and Lake Formation.	Amazon EMR mendukung tabel, baris, kolom, dan kontrol akses tingkat sel dengan Apache Iceberg.
Yayasan Linux Delta Lake	Delta Lake adalah proyek sumber terbuka yang membantu mengimplementasikan arsitektur data lake modern yang biasanya dibangun di Amazon S3 atau Hadoop Distributed File System (HDFS).	Amazon EMR mendukung tabel, baris, kolom, dan kontrol akses tingkat sel dengan tabel Delta Lake.

Format tabel	Deskripsi dan operasi yang diizinkan	Izin Lake Formation didukung di Amazon EMR
	Untuk daftar operasi yang didukung, lihat Delta Lake and Lake Formation.	

Sumber daya tambahan

Panduan pengguna, posting blog, dan lokakarya

- Integrasi dengan Amazon EMR menggunakan Peran Runtime
- Mulai cepat dengan Apache Hudi, Apache Iceberg, dan Delta Lake dengan Amazon EMR di EKS
- Menggunakan Delta Lake OSS dengan EMR Tanpa Server

Menggunakan AWS Lake Formation dengan Cepat

Quick mendukung penjelajahan kumpulan data yang dikelola oleh izin Lake Formation di Amazon S3 menggunakan Athena.

Baik pengguna edisi Standard dan Enterprise Quick berintegrasi dengan Lake Formation, tetapi sedikit berbeda.

- Edisi perusahaan — Berikan izin kontrol akses berbutir halus (FGAC) kepada pengguna dan grup Cepat individu untuk mengakses database dan tabel.
- Edisi standar - Berikan izin untuk peran IAM untuk mengakses database dan tabel.

> **Note**
>
> Secara default, Quick menggunakan peran bernama`aws-quicksight-service-role-v0`. Anda juga dapat menentukan peran kustom dengan izin yang diperlukan yang memungkinkan Cepat untuk mengakses Athena.

Untuk informasi selengkapnya, lihat Mengotorisasi koneksi melalui AWS Lake Formation

Sumber daya tambahan

Unggahan blog

- Aktifkan izin berbutir halus untuk penulis Cepat di AWS Lake Formation
- Analisis data Anda dengan aman AWS Lake Formation dan Cepat

Menggunakan AWS Lake Formation dengan AWS CloudTrail Danau

AWS CloudTrail Lake mendukung penjelajahan penyimpanan data acara menggunakan Amazon Athena izin berbutir halus di. AWS Lake Formation

> **Note**
>
> CloudTrail Danau hanya bisa ditanyakan melalui Amazon Athena.

Untuk mendaftarkan penyimpanan data acara CloudTrail Danau Anda dengan Lake Formation, lihat Federasi penyimpanan data acara.

Pencatatan panggilan API AWS Lake Formation menggunakan AWS CloudTrail

AWS Lake Formation terintegrasi dengan AWS CloudTrail, layanan yang menyediakan catatan tindakan yang diambil oleh pengguna, peran, atau AWS layanan di Lake Formation. CloudTrail menangkap semua panggilan Lake Formation API sebagai peristiwa. Panggilan yang diambil termasuk panggilan dari konsol Lake Formation, panggilan AWS Command Line Interface, dan kode ke tindakan Lake Formation API. Jika Anda membuat jejak, Anda dapat mengaktifkan pengiriman CloudTrail acara secara terus menerus ke bucket Amazon S3, termasuk acara untuk Lake Formation. Jika Anda tidak mengonfigurasi jejak, Anda masih dapat melihat peristiwa terbaru di CloudTrail konsol dalam Riwayat acara. Dengan menggunakan informasi yang dikumpulkan oleh CloudTrail, Anda dapat menentukan permintaan yang dibuat untuk Lake Formation, alamat IP dari mana permintaan itu dibuat, siapa yang membuat permintaan, kapan dibuat, dan detail tambahan.

Untuk mempelajari selengkapnya CloudTrail, lihat Panduan AWS CloudTrail Pengguna.

Informasi Lake Formation di CloudTrail

CloudTrail diaktifkan secara default saat Anda membuat AWS akun baru. Ketika aktivitas terjadi di Lake Formation, aktivitas tersebut dicatat sebagai CloudTrail peristiwa bersama dengan peristiwa AWS layanan lainnya dalam sejarah Peristiwa. Peristiwa mewakili permintaan tunggal dari sumber manapun dan mencakup informasi tentang tindakan yang diminta, tanggal dan waktu tindakan, dan parameter permintaan. Selain itu, setiap acara atau entri log berisi informasi tentang siapa yang membuat permintaan. Informasi identitas membantu Anda menentukan hal berikut ini:

- Apakah permintaan itu dibuat dengan kredenal pengguna root atau AWS Identity and Access Management (IAM).
- Apakah permintaan tersebut dibuat dengan kredensial keamanan sementara untuk satu peran atau pengguna gabungan.
- Apakah permintaan itu dibuat oleh AWS layanan lain.

Untuk informasi selengkapnya, lihat Elemen userIdentity CloudTrail .

Anda dapat melihat, mencari, dan mengunduh acara terbaru untuk AWS akun Anda. Untuk informasi selengkapnya, lihat Melihat peristiwa dengan Riwayat CloudTrail acara.

Untuk catatan acara yang sedang berlangsung di AWS akun Anda, termasuk acara untuk Lake Formation, buat jejak. Jejak memungkinkan CloudTrail untuk mengirimkan file log ke bucket Amazon S3. Secara default, ketika Anda membuat jejak pada konsol tersebut, jejak diterapkan ke semua Wilayah AWS . Jejak mencatat peristiwa dari semua Wilayah di AWS partisi dan mengirimkan file log ke bucket Amazon S3 yang Anda tentukan. Selain itu, Anda dapat mengonfigurasi AWS layanan lain, seperti Amazon Athena, untuk menganalisis lebih lanjut dan menindaklanjuti data peristiwa yang dikumpulkan dalam CloudTrail log. CloudTrail juga dapat mengirimkan file log ke Amazon CloudWatch Logs and CloudWatch Events.

Untuk informasi selengkapnya, lihat berikut:

- Gambaran umum untuk membuat jejak
- CloudTrail layanan dan integrasi yang didukung
- Mengonfigurasi notifikasi Amazon SNS untuk CloudTrail
- Menerima file CloudTrail log dari beberapa wilayah dan Menerima file CloudTrail log dari beberapa akun

Memahami peristiwa Lake Formation

Semua tindakan Lake Formation API dicatat oleh CloudTrail dan didokumentasikan dalam Panduan AWS Lake Formation Pengembang. Misalnya, panggilan kePutDataLakeSettings,GrantPermissions, dan RevokePermissions tindakan menghasilkan entri dalam file CloudTrail log.

Contoh berikut menunjukkan CloudTrail peristiwa untuk GrantPermissions tindakan tersebut. Entri termasuk pengguna yang memberikan izin (datalake_admin), kepala sekolah yang izin diberikan kepada (datalake_user1), dan izin yang diberikan (CREATE_TABLE). Entri juga menunjukkan bahwa hibah gagal karena database target tidak ditentukan dalam resource argumen.

```
{
    "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": "8d2ccef0-55f3-42d3-9ede-3a6faedaa5c1",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "eventCategory": "Management",
    "recipientAccountId": "111122223333"
}
```

Contoh berikutnya menunjukkan entri CloudTrail log untuk `GetDataAccess` tindakan tersebut. Prinsipal tidak langsung memanggil API ini. Sebaliknya, `GetDataAccess` dicatat setiap kali AWS layanan utama atau terintegrasi meminta kredensil sementara untuk mengakses data di lokasi danau data yang terdaftar di Lake Formation.

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "AWSAccount",
        "principalId": "AROAQGFTBBBGOBWV2EMZA: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"
    },
...
}
```

Menelusuri akses data CloudTrail dengan Lake Formation

Lake Formation mengeluarkan kredensil sementara yang tercakup ke bawah ke mesin kueri—seperti, Amazon EMR, dan—untuk mengakses lokasi data tabel di Amazon S3. AWS Glue Lake Formation menggunakan peran IAM yang Anda tentukan saat mendaftarkan lokasi data dengan Lake Formation untuk menjual kredensil ini.

Perubahan perilaku CloudTrail logging (Februari 2026)

Sebelum Februari 2026, ketika mesin kueri melakukan `s3:GetObject` panggilan untuk mengambil data kueri, CloudTrail log menampilkan peran IAM pendaftaran Lake Formation sebagai kepala sekolah.

Mulai Februari 2026, CloudTrail log untuk `s3:GetObject` panggilan juga menyertakan pengidentifikasi unik IAM (dimulai denganAR0A) dari peran IAM yang memulai kueri. Pengidentifikasi ini sesuai dengan peran IAM yang memulai kueri, pekerjaan EMR Amazon, atau pekerjaan ETL. AWS Glue

Menelusuri akses data dari inisiasi kueri ke Amazon S3

Setiap `s3:GetObject` panggilan CloudTrail sesuai dengan `lakeformation:GetDataAccess` panggilan. Anda dapat menggunakan pengenal peran IAM yang ada di kedua entri log untuk melacak peran IAM pemanggil dari inisiasi kueri hingga akses data di Amazon S3.

Untuk mengaktifkan informasi identitas sumber di CloudTrail log peristiwa data Amazon S3, ikuti langkah-langkah di bawah ini. Untuk mengaktifkan peristiwa data Amazon S3, lihat CloudTrail Mengaktifkan pencatatan peristiwa CloudTrail untuk bucket dan objek Amazon S3.

> **Important**
>
> Mengaktifkan peristiwa data Amazon S3 CloudTrail di menimbulkan biaya tambahan.

Untuk mengaktifkan pelacakan identitas sumber di CloudTrail log

1. Menggunakan, jalankan `get-data-lake-settings` perintah. Sebagai tanggapan, edit `Parameters` bidang yang akan ditambahkan`"SET_SOURCE_IDENTITY": "TRUE"`.

   Jalankan perintah `put-data-lake-settings`.

   ```
   "Parameters": {
       "CROSS_ACCOUNT_VERSION": "4",
       "SET_SOURCE_IDENTITY": "TRUE"
   },
   ```

2. Dalam kebijakan kepercayaan peran IAM yang digunakan untuk mendaftarkan lokasi data Amazon S3 Anda dengan Lake Formation, tambahkan: `sts:SetSourceIdentity`

   ```
   {
       "Version": "2012-10-17",
       "Statement": [
           {
               "Effect": "Allow",
               "Principal": {
                   "Service": [
                       "lakeformation.amazonaws.com"
                   ]
               },
               "Action": [
                   "sts:AssumeRole",
                   "sts:SetSourceIdentity"
               ]
           }
       ]
   }
   ```

Log berikut menunjukkan informasi identitas sumber dalam `s3:GetObject` CloudTrail log contoh.

```
"sessionContext": {
```

```
    "sessionIssuer": {
        "type": "Role",
        "principalId": "AROA2EXAMPLEXYZ3AB6CDKLM",
        "arn": "arn:aws:iam::111122223333:role/LFRegistrationRole",
        "accountId": "111122223333",
        "userName": "LFRegistrationRole"
    },
    "sourceIdentity": "V2.111122223333.AROA3EXAMPLEPQR7ST9UVWXY"
},
"invokedBy": "glue.amazonaws.com"
```

Lihat Juga

- Untuk informasi tentang melacak konteks pengguna Pusat Identitas IAM di CloudTrail log, lihatMengintegrasikan Pusat Identitas IAM.

Pemantauan akses katalog federasi Amazon Redshift

Katalog federasi Amazon Redshift di AWS Glue Data Catalog tidak menghasilkan peristiwa di log `GetDataAccess` Lake Formation. CloudTrail Untuk melacak akses data untuk tabel Amazon Redshift di Katalog Data, monitor, `GetTable` dan `BatchGetTable` peristiwa, yang menangkap pola akses metadata untuk tabel federasi Redshift, termasuk nama tabel, nama database, dan informasi ID katalog.

Lihat Juga

- Cross-account CloudTrail penebangan

Praktik, pertimbangan, dan batasan terbaik Lake Formation

Gunakan bagian ini untuk dengan cepat menemukan praktik, pertimbangan, dan batasan terbaik di dalamnya AWS Lake Formation.

Lihat Kuota layanan untuk jumlah maksimum sumber daya layanan atau operasi untuk Anda Akun AWS.

Topik

- Praktik dan pertimbangan terbaik berbagi data lintas akun
- Batasan peran terkait layanan
- Keterbatasan akses data Lintas Wilayah
- Katalog Data melihat pertimbangan dan batasan
- Batasan penyaringan data
- Pertimbangan dan batasan mode akses hibrida
- Keterbatasan untuk membawa data gudang data Amazon Redshift ke dalam AWS Glue Data Catalog
- Keterbatasan integrasi katalog Tabel S3
- Metadata sarang menyimpan pertimbangan dan batasan berbagi data
- Batasan berbagi data Amazon Redshift
- Keterbatasan integrasi Pusat Identitas IAM
- Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation
- Pertimbangan kontrol akses berbasis atribut, batasan, dan wilayah yang didukung

Praktik dan pertimbangan terbaik berbagi data lintas akun

Kemampuan lintas akun Lake Formation memungkinkan pengguna untuk berbagi data lake terdistribusi dengan aman di beberapa AWS organisasi Akun AWS, atau langsung dengan prinsipal IAM di akun lain yang menyediakan akses halus ke metadata Katalog Data dan data yang mendasarinya.

Pertimbangkan praktik terbaik berikut saat menggunakan berbagi data lintas akun Lake Formation:

- Tidak ada batasan jumlah hibah izin Lake Formation yang dapat Anda berikan kepada kepala sekolah di akun Anda sendiri. AWS Namun, Lake Formation menggunakan AWS Resource Access Manager (AWS RAM) kapasitas untuk hibah lintas akun yang dapat dibuat akun Anda dengan metode sumber daya bernama. Untuk memaksimalkan AWS RAM kapasitas, ikuti praktik terbaik berikut untuk metode sumber daya bernama:
  - Gunakan mode hibah lintas akun baru (Versi 3 ke atas di bawah pengaturan versi Cross account) untuk berbagi sumber daya dengan eksternal Akun AWS. Untuk informasi selengkapnya, lihat Memperbarui pengaturan versi berbagi data lintas akun.
  - Atur AWS akun ke dalam organisasi, dan berikan izin kepada organisasi atau unit organisasi. Hibah untuk organisasi atau unit organisasi dihitung sebagai satu hibah.

    Pemberian kepada organisasi atau unit organisasi juga menghilangkan kebutuhan untuk menerima AWS Resource Access Manager (AWS RAM) undangan pembagian sumber daya untuk hibah. Untuk informasi selengkapnya, lihat Mengakses dan melihat tabel dan database Katalog Data bersama.
  - Alih-alih memberikan izin pada banyak tabel individual dalam database, gunakan wildcard khusus Semua tabel untuk memberikan izin pada semua tabel dalam database. Pemberian pada Semua tabel dihitung sebagai hibah tunggal. Untuk informasi selengkapnya, lihat Memberikan izin pada sumber daya Katalog Data.

  > **Note**
  >
  > Untuk informasi selengkapnya tentang meminta batas yang lebih tinggi untuk jumlah pembagian sumber daya AWS RAM, lihat kuota AWS layanan di. Referensi Umum AWS

- Anda harus membuat tautan sumber daya ke database bersama agar database tersebut muncul di editor kueri Amazon Redshift Spectrum Amazon Athena dan Amazon Redshift. Demikian pula, untuk dapat menanyakan tabel bersama menggunakan Athena dan Redshift Spectrum, Anda harus membuat tautan sumber daya ke tabel. Tautan sumber daya kemudian muncul di daftar tabel editor kueri.

  Alih-alih membuat tautan sumber daya untuk banyak tabel individual untuk kueri, Anda dapat menggunakan wildcard Semua tabel untuk memberikan izin pada semua tabel dalam database. Kemudian, saat Anda membuat tautan sumber daya untuk database tersebut dan memilih tautan sumber daya basis data di editor kueri, Anda akan memiliki akses ke semua tabel di database tersebut untuk kueri Anda. Untuk informasi selengkapnya, lihat Membuat tautan sumber daya.

- Saat Anda berbagi sumber daya secara langsung dengan prinsipal di akun lain, prinsipal IAM di akun penerima mungkin tidak memiliki izin untuk membuat tautan sumber daya agar dapat menanyakan tabel bersama menggunakan Athena dan Amazon Redshift Spectrum. Alih-alih membuat tautan sumber daya untuk setiap tabel yang dibagikan, administrator data lake dapat membuat database placeholder dan memberikan `CREATE_TABLE` izin ke grup. `ALLIAMPrincipal` Kemudian, semua prinsipal IAM di akun penerima dapat membuat tautan sumber daya di database placeholder dan mulai menanyakan tabel bersama.

  Lihat contoh perintah CLI untuk memberikan izin masuk. `ALLIAMPrincipals` Memberikan izin database menggunakan metode sumber daya bernama
- Ketika izin lintas akun diberikan langsung kepada prinsipal, hanya penerima hibah yang dapat melihat izin ini. Administrator data lake di AWS akun penerima tidak dapat melihat hibah langsung ini.
- Athena dan Redshift Spectrum mendukung kontrol akses tingkat kolom, tetapi hanya untuk inklusi, bukan pengecualian. Kontrol akses tingkat kolom tidak didukung dalam pekerjaan AWS Glue ETL.
- Ketika sumber daya dibagikan dengan AWS akun Anda, Anda dapat memberikan izin pada sumber daya hanya kepada pengguna di akun Anda. Anda tidak dapat memberikan izin pada sumber daya ke AWS akun lain, ke organisasi (bahkan organisasi Anda sendiri), atau ke `IAMAllowedPrincipals` grup.
- Anda tidak dapat memberikan `DROP` atau `Super` pada database ke akun eksternal.
- Cabut izin lintas akun sebelum Anda menghapus database atau tabel. Jika tidak, Anda harus menghapus pembagian sumber daya yatim piatu di. AWS Resource Access Manager

> Lihat juga
>
> - Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation
> - `CREATE_TABLE`dalam Referensi izin Lake Formation untuk lebih banyak aturan dan batasan akses lintas akun.

Batasan peran terkait layanan

Peran terkait layanan adalah jenis peran IAM khusus yang ditautkan langsung ke. AWS Lake Formation Peran ini memiliki izin yang telah ditentukan sebelumnya yang memungkinkan Lake Formation melakukan tindakan atas nama Anda di seluruh AWS layanan.

Batasan berikut berlaku saat menggunakan peran terkait layanan (SLR) untuk mendaftarkan lokasi data dengan Lake Formation.

- Anda tidak dapat mengubah kebijakan peran terkait layanan setelah dibuat.
- Peran terkait layanan tidak mendukung pembagian sumber daya katalog terenkripsi di seluruh akun. Sumber daya terenkripsi memerlukan izin AWS KMS kunci tertentu. Peran terkait layanan memiliki izin yang telah ditentukan sebelumnya yang tidak menyertakan kemampuan untuk bekerja dengan sumber daya katalog terenkripsi di seluruh akun.
- Saat mendaftarkan beberapa lokasi Amazon S3, menggunakan peran terkait layanan dapat menyebabkan Anda melampaui batas kebijakan IAM dengan cepat. Ini terjadi karena dengan peran terkait layanan, AWS tulis kebijakan untuk Anda, dan itu bertambah sebagai satu blok besar yang mencakup semua pendaftaran Anda. Anda dapat menulis kebijakan yang dikelola pelanggan dengan lebih efisien, mendistribusikan izin di beberapa kebijakan, atau menggunakan peran berbeda untuk Wilayah yang berbeda.
- Amazon EMR di EC2 tidak dapat mengakses data yang Anda daftarkan lokasi data dengan peran terkait layanan.
- Operasi peran terkait layanan melewati kebijakan kontrol AWS layanan Anda.
- Saat Anda mendaftarkan lokasi data dengan peran terkait layanan, ia memperbarui kebijakan IAM dengan konsistensi akhirnya. Untuk informasi selengkapnya, lihat dokumentasi IAM Pemecahan Masalah di Panduan Pengguna IAM.
- Anda tidak dapat mengatur `SET_CONTEXT = TRUE` pengaturan danau data Lake Formation saat menggunakan peran terkait layanan, dan Anda menggunakan Pusat Identitas IAM. Alasannya adalah bahwa peran terkait layanan memiliki kebijakan kepercayaan yang tidak dapat diubah yang tidak sesuai dengan propagasi identitas tepercaya yang diperlukan untuk `SetContext` audit dengan kepala sekolah IAM Identity Center.

Keterbatasan akses data Lintas Wilayah

Lake Formation mendukung kueri tabel Katalog Data di seluruh Wilayah AWS. Anda dapat mengakses data di Wilayah dari Wilayah lain menggunakan Amazon Athena, Amazon EMR, dan AWS Glue ETL dengan membuat tautan sumber daya di Wilayah lain yang menunjuk ke database dan tabel sumber. Dengan akses tabel lintas wilayah, Anda dapat mengakses data di seluruh Wilayah tanpa menyalin data dasar atau metadata ke dalam Katalog Data.

Batasan berikut berlaku untuk akses tabel lintas wilayah.

- Lake Formation tidak mendukung kueri tabel Katalog Data dari Wilayah lain menggunakan Amazon Redshift Spectrum.
- Di konsol Lake Formation, tampilan database dan tabel tidak menampilkan nama database/tabel Wilayah sumber.
- Untuk melihat daftar tabel di bawah database bersama dari Wilayah lain, Anda harus terlebih dahulu membuat tautan sumber daya ke database bersama, lalu pilih tautan sumber daya, dan pilih Lihat tabel.
- Lake Formation tidak mendukung panggilan tautan sumber daya Lintas wilayah yang dilakukan oleh pengguna SAFL.
- Fitur Lintas wilayah Lake Formation tidak melibatkan biaya tambahan untuk transfer data.

Katalog Data melihat pertimbangan dan batasan

Pertimbangan dan batasan berikut berlaku untuk tampilan Katalog Data.

- Anda tidak dapat membuat tampilan Katalog Data dari konsol Lake Formation. Anda dapat membuat tampilan menggunakan AWS CLI atau SDK.
- Anda dapat membuat tampilan Katalog Data dari 10 tabel. Ini adalah batas yang sulit. Tabel referensi yang mendasari untuk tampilan dapat menjadi milik database yang sama atau database yang berbeda dalam AWS akun yang sama.
- Untuk pertimbangan dan batasan tambahan khusus untuk membuat tampilan Katalog Data menggunakan Redshift, lihat bagian Pertimbangan dan batasan tampilan Katalog Data di Panduan Pengembang Basis Data Amazon Redshift. Untuk Athena, lihat bagian Pertimbangan dan batasan tampilan Katalog Data di Panduan Pengguna Amazon Athena.
- Anda dapat membuat tampilan Katalog Data pada tabel yang terdaftar dengan Lake Formation dalam mode akses hybrid dan mode Lake Formation.

  Saat menggunakan tampilan Katalog Data dengan mode akses hibrida Lake Formation, disarankan untuk memastikan bahwa prinsipal yang menggunakan tampilan dipilih untuk izin Lake Formation untuk tabel dasar yang direferensikan dalam tampilan tanpa memberikan akses. Ini memastikan bahwa tabel dasar tidak akan diungkapkan kepada konsumen melalui izin AWS Glue IAM.
- Tidak ada batasan pada versi berbagi lintas akun untuk berbagi tampilan.
- Tampilan mendapatkan versi seperti tabel Katalog Data, saat Anda menggunakan `ALTER VIEW` pernyataan untuk dialek tampilan yang sudah dibuat. Anda tidak dapat memutar kembali ke

tampilan sebelumnya karena versi tampilan berubah dengan perubahan data yang mendasarinya. Anda dapat menghapus versi tampilan dan itu akan default ke versi terbaru berikutnya yang tersedia. Saat Anda mengubah versi tampilan, pastikan data Anda sinkron dengan skema versi tampilan yang dipilih.

- Tidak ada Katalog Data baru APIs yang diperkenalkan. Yang ada`CreateTable`,`UpdateTable`, `DeleteTable` dan `GetTable` APIs diperbarui.
- Amazon Redshift selalu membuat tampilan dengan kolom varchar dari tabel dengan string. Anda harus melemparkan kolom string ke varchar dengan panjang eksplisit saat menambahkan dialek dari mesin lain.
- Pemberian izin data lake ke `All tables` dalam database akan mengakibatkan penerima hibah memiliki izin pada semua tabel dan tampilan dalam database.
- Anda tidak dapat membuat tampilan:
  - Itu referensi pandangan lain.
  - Ketika tabel referensi adalah tautan sumber daya.
  - Ketika tabel referensi ada di akun lain.
  - Dari metastores Hive eksternal.
- Peran penentu lintas akun tidak didukung untuk tampilan Dialek Spektrum Redshift.
- Tautan sumber daya untuk dialek Athena di editor kueri Athena tidak didukung. Untuk menggunakan peran penentu lintas akun untuk dialek Athena, tambahkan akun yang menghosting tabel dasar sebagai Sumber Data di Athena.

Batasan penyaringan data

Saat Anda memberikan izin Lake Formation pada tabel Katalog Data, Anda dapat menyertakan spesifikasi pemfilteran data untuk membatasi akses ke data tertentu dalam hasil kueri dan mesin yang terintegrasi dengan Lake Formation. Lake Formation menggunakan pemfilteran data untuk mencapai keamanan tingkat kolom, keamanan tingkat baris, dan keamanan tingkat sel. Anda dapat menentukan dan menerapkan filter data pada kolom bersarang jika data sumber Anda berisi struktur bersarang.

Catatan dan batasan untuk penyaringan tingkat kolom

Ada tiga cara untuk menentukan pemfilteran kolom:

- Dengan menggunakan filter data

- Dengan menggunakan penyaringan kolom sederhana atau penyaringan kolom bersarang.
- Dengan menggunakan TAGs.

Pemfilteran kolom sederhana hanya menentukan daftar kolom untuk menyertakan atau mengecualikan. Baik konsol Lake Formation, API, dan AWS CLI mendukung pemfilteran kolom sederhana. Sebagai contoh, lihat Grant with Simple Column Filtering.

Catatan dan batasan berikut berlaku untuk pemfilteran kolom:

- AWS Glue 5.0 atau lebih tinggi mendukung kontrol akses halus melalui Lake Formation hanya untuk Apache Hive dan Apache Iceberg tabel.
- Untuk memberikan `SELECT` opsi hibah dan pemfilteran kolom, Anda harus menggunakan daftar sertakan, bukan daftar pengecualian. Tanpa opsi hibah, Anda dapat menggunakan daftar sertakan atau kecualikan.
- Untuk memberikan `SELECT` pada tabel dengan pemfilteran kolom, Anda harus diberikan `SELECT` di atas meja dengan opsi hibah dan tanpa batasan baris. Anda harus memiliki akses ke semua baris.
- Jika Anda memberikan opsi `SELECT` hibah dan pemfilteran kolom ke prinsipal di akun Anda, prinsipal tersebut harus menentukan pemfilteran kolom untuk kolom yang sama atau subset dari kolom yang diberikan saat memberikan kepada prinsipal lain. Jika Anda memberikan `SELECT` opsi hibah dan pemfilteran kolom ke akun eksternal, administrator data lake di akun eksternal dapat memberikan `SELECT` semua kolom ke prinsipal lain di akun mereka. Namun, bahkan dengan `SELECT` pada semua kolom, prinsipal itu hanya akan memiliki visibilitas pada kolom yang diberikan ke akun eksternal.
- Anda tidak dapat menerapkan pemfilteran kolom pada tombol partisi.
- Prinsipal dengan `SELECT` izin pada subset kolom dalam tabel tidak dapat diberikan`ALTER`,, `DROPDELETE`, atau `INSERT` izin pada tabel itu. Untuk kepala sekolah dengan`ALTER`,, `DROPDELETE`, atau `INSERT` izin di atas meja, jika Anda memberikan `SELECT` izin dengan pemfilteran kolom, itu tidak berpengaruh.

Catatan dan batasan berikut berlaku untuk pemfilteran kolom bersarang:

- Anda dapat menyertakan atau mengecualikan lima tingkat bidang bersarang dalam filter data.

  Example

  Col1.Col1_1.Col1_1_1.Col1_1_1_1_1.col1_1_1_1

- Anda tidak dapat menerapkan pemfilteran kolom pada bidang bersarang dalam kolom partisi.
- Jika skema tabel Anda berisi nama kolom tingkat atas (“pelanggan”.” address”) yang memiliki pola representasi bidang bersarang yang sama dalam filter data (kolom bersarang dengan nama kolom tingkat atas `customer` dan nama bidang bersarang `address` ditentukan seperti `"customer"."address"` dalam filter data), Anda tidak dapat secara eksplisit menentukan akses ke kolom tingkat atas atau bidang bersarang karena keduanya direpresentasikan menggunakan pola yang sama dalam daftar. inclusion/exclusion Ini ambigu, dan Lake Formation tidak dapat diselesaikan jika Anda menentukan kolom tingkat atas atau bidang bersarang.
- Jika kolom tingkat atas atau bidang bersarang berisi tanda kutip ganda dalam nama, Anda harus menyertakan kutipan ganda kedua saat Anda menentukan akses ke bidang bersarang dalam daftar sertakan dan kecualikan filter sel data.

  Example

  Contoh nama kolom bersarang dengan tanda kutip ganda - `a.b.double"quote`

  Example

  Contoh representasi kolom bersarang dalam filter data - `"a"."b"."double""quote"`

Batasan penyaringan tingkat sel

Ingatlah catatan dan batasan berikut untuk pemfilteran tingkat baris dan tingkat sel.

- Keamanan tingkat sel tidak didukung pada kolom bersarang, tampilan, dan tautan sumber daya.
- Semua ekspresi yang didukung pada kolom tingkat atas juga didukung pada kolom bersarang. Namun, bidang bersarang di bawah kolom partisi TIDAK boleh direferensikan saat mendefinisikan ekspresi tingkat baris bersarang.
- Keamanan tingkat sel tersedia di semua wilayah saat menggunakan mesin Athena versi 3 atau Amazon Redshift Spectrum. Untuk layanan lain, keamanan tingkat sel hanya tersedia di wilayah yang disebutkan di. Wilayah yang Didukung
- Pernyataan `SELECT INTO` tidak didukung.
- Tipe`array`, dan `map` data tidak didukung dalam ekspresi filter baris. Tipe `struct` data didukung.
- Tidak ada batasan untuk jumlah filter data yang dapat didefinisikan pada tabel, tetapi ada batas 100 filter data untuk satu prinsipal pada tabel.

- Untuk menerapkan filter data dengan ekspresi filter baris, Anda harus memiliki `SELECT` opsi hibah pada semua kolom tabel. Pembatasan ini tidak berlaku untuk administrator di akun eksternal saat hibah dibuat ke akun eksternal.
- Jika kepala sekolah adalah anggota grup dan prinsipal dan grup diberikan izin pada subset baris, izin baris efektif prinsipal adalah gabungan izin prinsipal dan izin grup.
- Nama kolom berikut dibatasi dalam tabel untuk pemfilteran tingkat baris dan tingkat sel:
  - ctid
  - oid
  - xmin
  - cmin
  - xmax
  - cmax
  - tableoid
  - insertxid
  - deletexid
  - importoid
  - redcatuniqueid
- Jika Anda menerapkan ekspresi filter semua baris pada tabel bersamaan dengan ekspresi filter lain dengan predikat, ekspresi semua baris akan menang atas semua ekspresi filter lainnya.
- Ketika izin pada subset baris diberikan ke AWS akun eksternal dan administrator data lake dari akun eksternal memberikan izin tersebut kepada prinsipal di akun tersebut, predikat filter efektif prinsipal adalah persimpangan predikat akun dan predikat apa pun yang langsung diberikan kepada prinsipal.

  Misalnya, jika akun memiliki izin baris dengan predikat `dept='hr'` dan prinsipal diberikan izin secara terpisah untuk`country='us'`, prinsipal hanya memiliki akses ke baris dengan `dept='hr'` dan. `country='us'`

Untuk informasi selengkapnya tentang penyaringan tingkat sel, lihat. Pemfilteran data dan keamanan tingkat sel di Lake Formation

Untuk pertimbangan dan batasan saat menanyakan tabel menggunakan Amazon Redshift Spectrum dengan kebijakan keamanan tingkat baris, lihat Pertimbangan dan batasan menggunakan kebijakan RLS di Panduan Pengembang Database Amazon Redshift.

Pertimbangan dan batasan mode akses hibrida

Mode akses hibrida memberikan fleksibilitas untuk mengaktifkan izin Lake Formation secara selektif untuk database dan tabel di situs Anda. AWS Glue Data Catalog
Dengan mode akses Hybrid, Anda sekarang memiliki jalur tambahan yang memungkinkan Anda mengatur izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu kebijakan izin pengguna atau beban kerja lain yang ada.

Pertimbangan dan batasan berikut berlaku untuk mode akses hybrid.

Batasan

- Memperbarui pendaftaran lokasi Amazon S3 — Anda tidak dapat mengedit parameter lokasi yang terdaftar di Lake Formation menggunakan peran terkait layanan.
- Opsi ikut serta saat menggunakan LF-tag — Bila Anda dapat memberikan izin Lake Formation menggunakan LF-tag, Anda dapat memilih prinsipal untuk menerapkan izin Lake Formation sebagai langkah berturut-turut dengan memilih database dan tabel yang memiliki LF-tag terlampir.
- Akses mode akses hibrid — Akses ke mode akses hibrida di Lake Formation terbatas pada pengguna dengan administrator danau data atau izin administrator hanya-baca.
- Memilih prinsipal — Saat ini, hanya peran administrator data lake yang dapat memilih prinsipal untuk sumber daya.
- Memilih semua tabel dalam database — Dalam hibah lintas akun, saat Anda memberikan izin, dan memilih semua tabel dalam database, Anda juga harus memilih dalam database agar izin berfungsi.

Pertimbangan-pertimbangan

- Memperbarui lokasi Amazon S3 yang terdaftar dengan Lake Formation ke mode akses hybrid — Kami tidak menyarankan mengonversi lokasi data Amazon S3 yang sudah terdaftar dengan Lake Formation ke mode akses hybrid meskipun dapat dilakukan.
- Perilaku API saat lokasi data terdaftar dalam mode akses hibrid
  - CreateTable — Lokasi dianggap terdaftar di Lake Formation terlepas dari flag mode akses hybrid dan status opt in. Dengan demikian, pengguna memerlukan izin lokasi data untuk membuat tabel.

- CreatePartition/BatchCreatePartition/UpdatePartition(saat lokasi partisi diperbarui untuk menunjuk ke lokasi yang terdaftar dengan hybrid) - Lokasi Amazon S3 dianggap terdaftar di Lake Formation terlepas dari bendera mode akses hibrida dan status pilih.
- CreateDatabase/UpdateDatabase (ketika lokasi database diperbarui untuk menunjuk ke lokasi yang terdaftar dalam mode akses hibrida) — Lokasi dianggap terdaftar di Lake Formation terlepas dari flag mode akses hybrid dan status opt in. Dengan demikian, pengguna memerlukan izin lokasi data untuk membuat atau memperbarui database.
- UpdateTable (ketika lokasi tabel diperbarui untuk menunjuk ke lokasi yang terdaftar dalam mode akses hibrida) — Lokasi dianggap terdaftar di Lake Formation terlepas dari bendera mode akses hibrida dan status pilih. Dengan demikian, pengguna memerlukan izin lokasi data untuk memperbarui tabel. Jika lokasi tabel tidak diperbarui atau menunjuk ke lokasi yang tidak terdaftar di Lake Formation, pengguna tidak memerlukan izin lokasi data untuk memperbarui tabel.

Keterbatasan untuk membawa data gudang data Amazon Redshift ke dalam AWS Glue Data Catalog

Anda dapat membuat katalog, dan mengelola akses ke data analitik di gudang data Amazon Redshift menggunakan. AWS Glue Data Catalog Batasan berikut berlaku:

- Berbagi lintas akun tidak didukung di tingkat katalog federasi. Namun, Anda dapat berbagi database dan tabel individual dari dalam katalog federasi di seluruh Akun AWS s.
- Anda harus memiliki pengaturan versi Cross account versi 4 atau lebih tinggi untuk berbagi database atau tabel dalam katalog federasi di seluruh Akun AWS s.
- Katalog Data hanya mendukung pembuatan katalog tingkat atas.
- Anda hanya dapat memperbarui deskripsi katalog di Redshift Managed Storage (RMS).
- Menyiapkan izin pada katalog federasi serta database dan tabel dalam katalog federasi ke grup tidak didukung. `IAMAllowedPrincipals`

- Operasi Data Definition Language (DDL) pada katalog dari mesin seperti Athena, Amazon EMR Spark, atau lainnya, termasuk pengaturan konfigurasi katalog, tidak didukung.
- Melakukan operasi DDL pada tabel RMS menggunakan Athena tidak didukung.
- Membuat tampilan terwujud tidak didukung, apakah itu melalui Athena, Apache Spark, AWS Glue Data Catalog the, atau konsumen Amazon Redshift.

- Athena tidak mendukung pengalaman multi-katalog. Itu hanya dapat terhubung ke satu katalog tertentu pada satu waktu. Athena tidak dapat mengakses atau melakukan kueri di beberapa katalog secara bersamaan.
- Operasi penandaan dan percabangan pada tabel Iceberg melalui Athena dan Amazon Redshift tidak didukung.
- Perjalanan Waktu pada tabel RMS tidak didukung.
- Katalog multi-level dengan tabel data lake tidak didukung. Semua data yang disimpan di Amazon S3 untuk digunakan dengan tabel data lake harus berada dalam default AWS Glue Data Catalog, dan tidak dapat diatur ke dalam katalog multi-level.
- Di Amazon Redshift, datashares tidak ditambahkan ke namespace terdaftar. Cluster dan ruang nama adalah sinonim, setelah Anda mempublikasikan cluster ke, Anda tidak dapat menambahkan AWS Glue Data Catalog data baru.
- Amazon EMR di EC2 tidak mendukung penggabungan di seluruh tabel RMS dan tabel Amazon S3. Hanya EMR Serverless yang mendukung kemampuan ini.
- Skema dan tabel eksternal tidak didukung.
- Tabel RMS hanya dapat diakses dari titik akhir ekstensi di AWS Glue Iceberg REST Catalog.
- Tabel sarang tidak dapat diakses dari mesin pihak ketiga yang terhubung ke Katalog AWS Glue Iceberg REST.
- Tingkat isolasi read_commit pada tabel RMS melalui Spark akan didukung.
- Nama database Redshift diperlakukan sebagai case-insensitive di AWS Glue Data Catalog, dibatasi hingga 128 karakter, dan dapat alfanumerik dengan tanda hubung (-) dan garis bawah (_).
- Nama katalog tidak peka huruf besar/kecil, dibatasi hingga 50 karakter, dan dapat berupa alfanumerik dengan tanda hubung (-) dan garis bawah (_).
- Amazon Redshift tidak mendukung penggunaan perintah GRANT dan REVOKE gaya Lake Formation SQL untuk mengelola izin akses pada tabel yang dipublikasikan ke. AWS Glue Data Catalog
- Keamanan tingkat baris dan kebijakan masking data dinamis yang dilampirkan ke produsen (sumber) klaster Amazon Redshift tidak akan diberlakukan. Sebagai gantinya, izin akses yang ditentukan dalam Lake Formation akan diberlakukan pada data bersama.
- Melakukan operasi Data Definition Language (DDL) dan Data Manipulation Language (DHTML) pada tautan tabel tidak didukung.
- Jika kata kunci yang dicadangkan tidak lolos dengan benar, itu akan mengakibatkan kegagalan atau kesalahan.

- Enkripsi data dalam skenario multi-katalog tidak didukung.

Keterbatasan integrasi katalog Tabel S3

Tabel Amazon S3 terintegrasi dengan AWS Glue Data Catalog (Katalog Data) dan mendaftarkan katalog sebagai lokasi data Lake Formation. Anda dapat mengatur pendaftaran ini dari konsol Lake Formation atau dengan menggunakan layanan ini APIs.

> **Note**
>
> Jika Anda memiliki kebijakan berbasis sumber daya IAM atau S3 Tables yang membatasi pengguna IAM dan peran IAM berdasarkan tag utama, Anda harus melampirkan tag utama yang sama ke peran IAM yang digunakan Lake Formation untuk mengakses data S3 Anda (misalnya,) dan memberikan peran ini izin yang diperlukan. LakeFormationDataAccessRole Ini diperlukan agar kebijakan kontrol akses berbasis tag Anda berfungsi dengan benar dengan integrasi analitik Tabel S3 Anda.

Batasan berikut berlaku untuk mengintegrasikan katalog Tabel S3 dengan Katalog Data dan Lake Formation:

- AWS Glue dan Lake Formation tidak mendukung nama kolom huruf campuran, dan mengonversi semua nama kolom menjadi huruf kecil. Anda harus memverifikasi bahwa nama kolom tabel unik saat dikonversi ke huruf kecil. Gunakan `customer_id` sebagai pengganti`customerId`. Penggunaan nama kolom kasus campuran hanya didukung selama rilis pratinjau.
- CreateCatalog API tidak dapat membuat bucket tabel di Amazon S3.
- SearchTables API tidak dapat mencari Tabel S3.

Metadata sarang menyimpan pertimbangan dan batasan berbagi data

Dengan federasi AWS Glue Data Catalog metadata (federasi Katalog Data), Anda dapat menghubungkan Katalog Data ke metastor eksternal yang menyimpan metadata untuk data Amazon S3 Anda, dan mengelola izin akses data dengan aman menggunakan. AWS Lake Formation

Pertimbangan dan batasan berikut berlaku untuk database federasi yang dibuat dari database Hive:

Pertimbangan-pertimbangan

- AWS SAM dukungan aplikasi - Anda bertanggung jawab atas ketersediaan sumber daya aplikasi yang AWS SAM menyebarkan (Amazon API Gateway dan fungsi Lambda). Pastikan koneksi antara metastore AWS Glue Data Catalog dan Hive berfungsi saat pengguna menjalankan kueri.
- Persyaratan versi metastore sarang - Anda dapat membuat database federasi hanya menggunakan Apache Hive versi 3 dan di atasnya.
- Persyaratan database yang dipetakan — Setiap database Hive harus dipetakan ke database baru di Lake Formation.
- Dukungan federasi tingkat database - Anda dapat terhubung ke Hive metastore hanya di tingkat database.
- Izin pada database federasi — Izin yang diterapkan pada database federasi atau tabel di bawah database federasi tetap ada bahkan ketika tabel sumber atau database dihapus. Saat database sumber atau tabel dibuat ulang, Anda tidak perlu memberikan izin kembali. Ketika tabel federasi dengan izin Lake Formation dihapus di sumber, izin Lake Formation masih terlihat, dan Anda dapat mencabutnya jika diperlukan.

  Jika pengguna menghapus database federasi, semua izin yang sesuai akan hilang. Membuat ulang database yang sama dengan nama yang sama, tidak akan memulihkan izin Lake Formation. Pengguna harus mengatur izin baru lagi.
- IAMAllowedIzin grup utama pada database federasi — Berdasarkan, Lake `DataLakeSettings` Formation dapat menetapkan izin ke semua database dan tabel ke grup virtual bernama. `IAMAllowedPrincipal` `IAMAllowedPrincipal`Ini mengacu pada semua kepala sekolah IAM yang memiliki akses ke sumber daya Katalog Data melalui kebijakan utama IAM dan kebijakan sumber daya. AWS Glue Jika izin ini ada pada database atau tabel, semua prinsipal diberikan akses ke database atau tabel.

  Namun, Lake Formation tidak mengizinkan `IAMAllowedPrincipal` izin pada tabel di bawah database federasi. Saat Anda membuat database federasi, pastikan Anda meneruskan `CreateTableDefaultPermissions` parameter sebagai daftar kosong.

  Untuk informasi selengkapnya, lihat Mengubah pengaturan default untuk data lake.
- Menggabungkan tabel dalam kueri — Anda dapat bergabung dengan tabel metastore Hive dengan tabel asli Katalog Data untuk menjalankan kueri.

Batasan

- Batasan sinkronisasi metadata antara metastore AWS Glue Data Catalog dan Hive — Setelah membuat koneksi metastore Hive, Anda perlu membuat database federasi untuk menyinkronkan metadata di metastore Hive dengan. AWS Glue Data Catalog Tabel di bawah database federasi disinkronkan saat runtime saat pengguna menjalankan kueri.
- Batasan membuat tabel baru di bawah database federasi — Anda tidak akan dapat membuat tabel baru di bawah database federasi.
- Batasan izin data - Dukungan untuk izin pada tampilan tabel metastore Hive tidak tersedia.

Batasan berbagi data Amazon Redshift

AWS Lake Formation memungkinkan Anda mengelola data dengan aman di datashare dari Amazon Redshift. Amazon Redshift adalah layanan gudang data skala petabyte yang dikelola sepenuhnya di Cloud. AWS Menggunakan kemampuan berbagi data, Amazon Redshift membantu Anda berbagi data. Akun AWS Untuk informasi selengkapnya tentang berbagi data Amazon Redshift, lihat Ringkasan berbagi data di Amazon Redshift.

Catatan dan batasan berikut berlaku untuk database federasi yang dibuat dari datashares Amazon Redshift:

- Persyaratan database yang dipetakan — Setiap data Amazon Redshift harus dipetakan ke database baru di Lake Formation. Ini diperlukan untuk mempertahankan nama tabel yang unik ketika representasi objek datashare diratakan dalam database Data Catalog.
- Batasan membuat tabel baru di bawah database federasi — Anda tidak akan dapat membuat tabel baru di bawah database federasi.
- Izin pada database federasi — Izin yang diterapkan pada database federasi atau tabel di bawah database federasi tetap ada bahkan ketika tabel sumber atau database dihapus. Ketika database sumber atau tabel dibuat ulang, Anda tidak perlu memberikan kembali izin. Ketika tabel federasi dengan izin Lake Formation dihapus di sumber, izin Lake Formation akan tetap terlihat dan Anda dapat mencabutnya jika diperlukan.

  Jika pengguna menghapus database federasi, semua izin yang sesuai akan hilang. Membuat ulang database yang sama dengan nama yang sama, tidak akan memulihkan izin Lake Formation. Pengguna harus mengatur izin baru lagi.
- IAMAllowedIzin grup utama pada database federasi — Berdasarkan, Lake `DataLakeSettings` Formation dapat menetapkan izin ke semua database dan tabel ke grup virtual bernama.

`IAMAllowedPrincipal` `IAMAllowedPrincipal`Ini mengacu pada semua kepala sekolah IAM yang memiliki akses ke sumber daya Katalog Data melalui kebijakan utama IAM dan kebijakan sumber daya. AWS Glue Jika izin ini ada pada database atau tabel, semua prinsipal diberikan akses ke database atau tabel.

Namun, Lake Formation tidak mengizinkan `IAMAllowedPrincipal` izin pada tabel di bawah database federasi. Saat Anda membuat database federasi, pastikan Anda meneruskan `CreateTableDefaultPermissions` parameter sebagai daftar kosong.

Untuk informasi selengkapnya, lihat Mengubah pengaturan default untuk data lake.

- Pemfilteran data — Di Lake Formation, Anda dapat memberikan izin pada tabel di bawah database federasi dengan pemfilteran tingkat kolom dan tingkat baris. Namun, Anda tidak dapat menggabungkan penyaringan tingkat kolom dan tingkat baris untuk membatasi akses pada perincian tingkat sel pada tabel di bawah database federasi.
- Pengidentifikasi sensitivitas huruf besar — Objek data Amazon Redshift yang dikelola oleh Lake Formation, akan mendukung nama tabel dan nama kolom hanya dalam huruf kecil. Jangan aktifkan pengenal sensitivitas huruf kecil untuk database, tabel, dan kolom di rangkaian data Amazon Redshift, jika mereka akan dibagikan dan dikelola menggunakan Lake Formation.
- Dukungan kueri — Anda dapat menanyakan rangkaian data Amazon Redshift yang dikelola oleh Lake Formation dengan Amazon Redshift. Athena tidak mendukung kueri datashares Amazon Redshift yang dikelola oleh Lake Formation.

Untuk informasi selengkapnya tentang batasan saat bekerja dengan rangkaian data di Amazon Redshift, lihat Batasan untuk berbagi data di Panduan Pengembang Basis Data Amazon Redshift.

Keterbatasan integrasi Pusat Identitas IAM

Dengan AWS IAM Identity Center, Anda dapat terhubung ke penyedia identitas (IdPs) dan mengelola akses secara terpusat untuk pengguna dan grup di seluruh layanan AWS analitik. Anda dapat mengonfigurasi AWS Lake Formation sebagai aplikasi yang diaktifkan di Pusat Identitas IAM, dan administrator data lake dapat memberikan izin halus kepada pengguna dan grup yang berwenang pada sumber daya. AWS Glue Data Catalog

Batasan berikut berlaku untuk integrasi Lake Formation dengan IAM Identity Center:

- Anda tidak dapat menetapkan pengguna dan grup Pusat Identitas IAM sebagai administrator data lake atau administrator hanya-baca di Lake Formation.

Pengguna dan grup IAM Identity Center dapat menanyakan sumber daya Katalog Data terenkripsi jika Anda menggunakan peran IAM yang AWS Glue dapat diambil atas nama Anda untuk mengenkripsi dan mendekripsi Katalog Data. AWS kunci terkelola tidak mendukung propagasi identitas tepercaya.

- Pengguna dan grup IAM Identity Center hanya dapat menjalankan operasi API yang tercantum dalam `AWSIAMIdentityCenterAllowListForIdentityContext` kebijakan yang disediakan oleh IAM Identity Center.
- Lake Formation mengizinkan peran IAM dari akun eksternal untuk bertindak sebagai peran operator atas nama pengguna dan grup Pusat Identitas IAM untuk mengakses sumber daya Katalog Data, tetapi izin hanya dapat diberikan pada sumber daya Katalog Data dalam akun yang dimiliki. Jika Anda mencoba memberikan izin kepada pengguna Pusat Identitas IAM dan grup pada sumber daya Katalog Data di akun eksternal, Lake Formation akan menampilkan kesalahan berikut - “Hibah lintas akun tidak didukung untuk prinsipal.”
- Saat menggunakan Lake Formation dengan IAM Identity Center, konfigurasi penetapan aplikasi diatur `false` secara default. Jika Anda memodifikasi konfigurasi ini secara langsung melalui IAM Identity Center API, Anda kemudian harus mengelola semua penetapan aplikasi secara manual menggunakan API. Lake Formation tidak secara otomatis menyinkronkan atau mengelola perubahan penugasan yang dibuat di luar alur kerja standarnya, yang dapat memengaruhi pola akses dan alur otorisasi dalam lingkungan data lake Anda.

Praktik dan pertimbangan terbaik kontrol akses berbasis tag Lake Formation

Anda dapat membuat, memelihara, dan menetapkan LF-tag untuk mengontrol akses ke database, tabel, dan kolom Katalog Data.

Pertimbangkan praktik terbaik berikut saat menggunakan kontrol akses berbasis tag Lake Formation:

- Semua LF-tag harus ditentukan sebelumnya sebelum dapat ditetapkan ke sumber daya Katalog Data atau diberikan kepada prinsipal.

  Administrator data lake dapat mendelegasikan tugas manajemen tag dengan membuat pembuat LF-tag dengan izin IAM yang diperlukan. Insinyur dan analis data memutuskan karakteristik dan hubungan untuk LF-tag. Pembuat LF-tag kemudian membuat dan memelihara LF-tag di Lake Formation.

- Anda dapat menetapkan beberapa LF-tag ke sumber daya Katalog Data. Hanya satu nilai untuk kunci tertentu yang dapat ditetapkan ke sumber daya tertentu.

  Misalnya, Anda dapat menetapkan`module=Orders`,, `region=West`dan `division=Consumer`, dan seterusnya ke database, tabel, atau kolom. Anda tidak dapat menetapkan`module=Orders,Customers`.
- Anda tidak dapat menetapkan LF-tag ke sumber daya saat membuat sumber daya. Anda hanya dapat menambahkan LF-tag ke sumber daya yang ada.
- Anda dapat memberikan ekspresi LF-tag, bukan hanya tag LF tunggal, ke prinsipal.

  Ekspresi LF-tag terlihat seperti berikut (dalam pseudo-code).

  ```
  module=sales AND division=(consumer OR commercial)
  ```

  Prinsipal yang diberikan ekspresi LF-tag ini hanya dapat mengakses sumber daya Katalog Data (database, tabel, dan kolom) yang ditetapkan `module=sales` dan salah satu atau. `division=consumer` `division=commercial` Jika Anda ingin kepala sekolah dapat mengakses sumber daya yang memiliki `module=sales` atau`division=commercial`, jangan sertakan keduanya dalam hibah yang sama. Buat dua hibah, satu untuk `module=sales` dan satu untuk`division=commercial`.

  Ekspresi LF-tag paling sederhana hanya terdiri dari satu LF-tag, seperti. `module=sales`
- Prinsipal yang diberikan izin pada LF-tag dengan beberapa nilai dapat mengakses sumber daya Katalog Data dengan salah satu dari nilai tersebut. Misalnya, jika pengguna diberikan LF-tag dengan key= `module` dan values=`orders,customers`, pengguna memiliki akses ke sumber daya yang ditetapkan salah satu atau. `module=orders` `module=customers`
- Anda harus memiliki `Grant with LF-Tag expressions` izin untuk memberikan izin data pada sumber daya Katalog Data dengan menggunakan metode LF-TBAC. Administrator data lake dan pembuat LF-tag secara implisit menerima izin ini. Prinsipal yang memiliki `Grant with LFTag expressions` izin dapat memberikan izin data pada sumber daya menggunakan:
  - metode sumber daya bernama
  - metode LF-TBAC, tetapi hanya menggunakan ekspresi LF-tag yang sama

    Misalnya, asumsikan bahwa administrator data lake membuat hibah berikut (dalam kode semu).

```
GRANT (SELECT ON TABLES) ON TAGS module=customers, region=west,south TO user1 WITH
 GRANT OPTION
```

Dalam hal ini, `user1` dapat memberikan `SELECT` tabel ke prinsipal lain dengan menggunakan metode LF-TBAC, tetapi hanya dengan ekspresi LF-tag lengkap. `module=customers, region=west,south`

- Jika prinsipal diberikan izin pada sumber daya dengan metode LF-TBAC dan metode sumber daya bernama, izin yang dimiliki prinsipal pada sumber daya adalah gabungan izin yang diberikan oleh kedua metode.
- Lake Formation mendukung pemberian `DESCRIBE` dan penggunaan LF-tag `ASSOCIATE` di seluruh akun, dan pemberian izin pada sumber daya Katalog Data di seluruh akun menggunakan metode LF-TBAC. Dalam kedua kasus, kepala sekolah adalah ID AWS akun.

> **Note**
>
> Lake Formation mendukung hibah lintas akun untuk organisasi dan unit organisasi menggunakan metode LF-TBAC. Untuk menggunakan kemampuan ini, Anda perlu memperbarui pengaturan versi akun Cross ke Versi 3.

Untuk informasi selengkapnya, lihat Cross-account berbagi data di Lake Formation.

- Sumber daya Katalog Data yang dibuat dalam satu akun hanya dapat ditandai menggunakan tag LF yang dibuat di akun yang sama. LF-tag yang dibuat dalam satu akun tidak dapat dikaitkan dengan sumber daya bersama dari akun lain.
- Menggunakan kontrol akses berbasis tag Lake Formation (LF-TBAC) untuk memberikan akses lintas akun ke sumber daya Katalog Data memerlukan penambahan kebijakan sumber daya Katalog Data untuk akun Anda. AWS Untuk informasi selengkapnya, lihat Prasyarat.
- Kunci LF-tag dan nilai LF-tag tidak boleh melebihi 50 karakter panjangnya.
- Jumlah maksimum LF-tag yang dapat ditetapkan ke sumber daya Katalog Data adalah 50.
- Batasan berikut adalah batas lunak:
  - Jumlah maksimum LF-tag yang dapat dibuat adalah 10000.
  - Jumlah maksimum nilai yang dapat didefinisikan untuk LF-tag adalah 1000.
- Kunci dan nilai tag dikonversi ke semua huruf kecil saat disimpan.
- Hanya satu nilai untuk LF-tag yang dapat ditetapkan ke sumber daya tertentu.

- Jika beberapa LF-tag diberikan kepada prinsipal dengan satu hibah, prinsipal hanya dapat mengakses sumber daya Katalog Data yang memiliki semua LF-tag.
- Jika evaluasi ekspresi LF-tag menghasilkan akses ke hanya subset kolom tabel, tetapi izin Lake Formation yang diberikan saat ada kecocokan adalah salah satu izin yang memerlukan akses kolom penuh, yaitu,,, atau `Alter` `Drop` `InsertDelete`, maka tidak ada izin tersebut yang diberikan. Sebaliknya, hanya `Describe` diberikan. Jika izin yang diberikan adalah `All` (`Super`), maka hanya `Select` dan `Describe` diberikan.
- Wildcard tidak digunakan dengan LF-tag. Untuk menetapkan LF-tag ke semua kolom tabel, Anda menetapkan LF-tag ke tabel, dan semua kolom dalam tabel mewarisi LF-tag. Untuk menetapkan LF-tag ke semua tabel dalam database, Anda menetapkan LF-tag ke database, dan semua tabel dalam database mewarisi LF-tag tersebut.
- Anda dapat membuat hingga 1000 ekspresi LF-tag di akun.
- Anda dapat menggunakan hingga 50 ekspresi LF-tag untuk memberikan izin kepada prinsipal pada sumber daya Katalog Data.
- Saat memberikan atau mencabut izin pada ekspresi tag LF sebaris, ukuran ekspresi LF-tag tidak dapat melebihi 900 byte. Untuk memberikan izin pada ekspresi LF-tag yang lebih besar, gunakan ekspresi LF-tag yang disimpan. Untuk informasi selengkapnya, lihat Membuat LF-Tag ekspresi.
- Untuk menambahkan LF-tag ke katalog federasi Redshift yang ada yang dibuat sebelum rilis ketersediaan umum dukungan LF-tag untuk katalog federasi, Anda perlu menghubungi tim dukungan untuk mendapatkan bantuan. AWS

Pertimbangan kontrol akses berbasis atribut, batasan, dan wilayah yang didukung

Pertimbangan dan batasan berikut berlaku untuk Attribute based access control (ABAC).

- ABAC tidak mendukung pemberian akses menggunakan kebijakan LF-tag.
- Izin yang dapat diberikan tidak tersedia dengan ABAC.
- ABAC tidak mendukung pemberian izin kepada pengguna IAM Identity Center.
- Saat menggunakan hibah ABAC pada tabel di Lake Formation, Lake Formation tidak memberikan `DESCRIBE` izin ke database atau katalog induk. Ini berbeda dari skenario non-ABAC, di mana Lake Formation memberikan `DESCRIBE` izin implisit ke sumber daya induk.
- Semua prinsipal dengan kunci `AmazonDataZoneProject` tag selalu diperlakukan sebagai ikut serta dalam Lake Formation untuk semua sumber daya Katalog Data.

- ABAC hanya mendukung atribut string.

Memecahkan Masalah Lake Formation

Jika Anda mengalami masalah saat bekerja dengan AWS Lake Formation, lihat topik di bagian ini.

Topik

- Pemecahan masalah umum
- Memecahkan masalah akses lintas akun
- Memecahkan masalah cetak biru dan alur kerja
- Masalah yang diketahui untuk AWS Lake Formation
- Pesan kesalahan yang diperbarui

Pemecahan masalah umum

Gunakan informasi di sini untuk membantu Anda mendiagnosis dan memperbaiki berbagai masalah Lake Formation.

Kesalahan: Izin Lake Formation tidak mencukupi <Amazon S3 location>

Upaya dilakukan untuk membuat atau mengubah sumber daya Katalog Data tanpa izin lokasi data di lokasi Amazon S3 yang ditunjukkan oleh sumber daya.

Jika database atau tabel Katalog Data menunjuk ke lokasi Amazon S3, saat Anda memberikan izin Lake Formation `CREATE_TABLE` atau`ALTER`, Anda juga harus memberikan `DATA_LOCATION_ACCESS` izin pada lokasi tersebut. Jika Anda memberikan izin ini ke akun eksternal atau organisasi, Anda harus menyertakan opsi hibah.

Setelah izin ini diberikan ke akun eksternal, administrator data lake di akun tersebut kemudian harus memberikan izin kepada prinsipal (pengguna atau peran) di akun tersebut. Saat memberikan `DATA_LOCATION_ACCESS` izin yang diterima dari akun lain, Anda harus menentukan ID katalog (ID AWS akun) dari akun pemilik. Akun pemilik adalah akun yang mendaftarkan lokasi.

Untuk informasi selengkapnya, lihat Kontrol akses data yang mendasari dan Memberikan izin lokasi data.

Kesalahan: “Izin kunci enkripsi tidak memadai untuk Glue API”

Upaya dilakukan untuk memberikan izin Lake Formation tanpa izin AWS Identity and Access Management (IAM) pada kunci AWS KMS enkripsi untuk Katalog Data terenkripsi.

Kueri saya Amazon Athena atau Amazon Redshift yang menggunakan manifes gagal

Lake Formation tidak mendukung kueri yang menggunakan manifes.

Kesalahan: “Izin Lake Formation tidak mencukupi: Wajib membuat tag di katalog”

user/role Harus menjadi administrator danau data.

Kesalahan saat menghapus administrator danau data yang tidak valid

Anda harus menghapus semua administrator danau data yang tidak valid (peran IAM yang dihapus yang didefinisikan sebagai administrator danau data) secara bersamaan. Jika Anda mencoba menghapus administrator danau data yang tidak valid secara terpisah, Lake Formation menampilkan kesalahan utama yang tidak valid.

Kesalahan: “Penelepon sementara tidak berwenang untuk melakukan operasi ini pada sumber daya...”

Upaya dilakukan untuk menanyakan tabel menggunakan tautan sumber daya yang mungkin telah dialihkan. Untuk menghindari kesalahan ini, aplikasi Anda harus mengadopsi QuerySessionContextprotokol. Lihat Lake Formation alur kerja untuk operasi API integrasi aplikasi untuk informasi selengkapnya. Jika tidak, harap tunggu 24 jam sebelum mencoba mengeksekusi kueri Anda lagi.

Memecahkan masalah akses lintas akun

Gunakan informasi di sini untuk membantu Anda mendiagnosis dan memperbaiki masalah akses lintas akun.

Topik

- Saya memberikan izin Lake Formation lintas akun tetapi penerima tidak dapat melihat sumber dayanya
- Prinsipal di akun penerima dapat melihat sumber daya Katalog Data tetapi tidak dapat mengakses data yang mendasarinya
- Kesalahan: “Asosiasi gagal karena pemanggil tidak diotorisasi” saat menerima undangan berbagi AWS RAM sumber daya
- Kesalahan: “Tidak berwenang untuk memberikan izin untuk sumber daya”
- Kesalahan: “Akses ditolak untuk mengambil informasi AWS Organisasi”
- Kesalahan: “Organisasi <organization-ID>tidak ditemukan”
- Kesalahan: “Izin Lake Formation tidak mencukupi: Kombinasi ilegal”
- ConcurrentModificationException atas grant/revoke permintaan ke akun eksternal
- Kesalahan saat menggunakan Amazon EMR untuk mengakses data yang dibagikan melalui lintas akun

Saya memberikan izin Lake Formation lintas akun tetapi penerima tidak dapat melihat sumber dayanya

- Apakah pengguna di akun penerima adalah administrator danau data? Hanya administrator danau data yang dapat melihat sumber daya pada saat berbagi.
- Apakah Anda berbagi dengan akun di luar organisasi Anda dengan menggunakan metode sumber daya bernama? Jika demikian, administrator data lake dari akun penerima harus menerima undangan berbagi sumber daya di AWS Resource Access Manager (AWS RAM).

  Untuk informasi selengkapnya, lihat the section called “Menerima undangan berbagi AWS RAM sumber daya”.
- Apakah Anda menggunakan kebijakan sumber daya tingkat akun (Katalog Data) di? AWS Glue Jika ya, maka jika Anda menggunakan metode sumber daya bernama, Anda harus menyertakan pernyataan khusus dalam kebijakan yang mengizinkan AWS RAM untuk berbagi kebijakan atas nama Anda.

  Untuk informasi selengkapnya, lihat the section called “Mengelola izin lintas akun menggunakan keduanya AWS Glue dan Formasi Danau”.
- Apakah Anda memiliki izin AWS Identity and Access Management (IAM) yang diperlukan untuk memberikan akses lintas akun?

Untuk informasi selengkapnya, lihat the section called “Prasyarat”.

- Sumber daya yang Anda berikan izin tidak boleh memiliki izin Lake Formation yang diberikan kepada grup. `IAMAllowedPrincipals`
- Apakah ada `deny` pernyataan tentang sumber daya dalam kebijakan tingkat akun?

Prinsipal di akun penerima dapat melihat sumber daya Katalog Data tetapi tidak dapat mengakses data yang mendasarinya

Prinsipal di akun penerima harus memiliki izin yang diperlukan AWS Identity and Access Management (IAM). Lihat perinciannya di Mengakses data dasar tabel bersama.

Kesalahan: “Asosiasi gagal karena pemanggil tidak diotorisasi” saat menerima undangan berbagi AWS RAM sumber daya

Setelah memberikan akses ke sumber daya ke akun yang berbeda, ketika akun penerima mencoba menerima undangan berbagi sumber daya, tindakan gagal.

```
$ 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
        }
    ]
}
```

Kesalahan terjadi karena dipanggil oleh AWS Glue ketika akun penerima menerima undangan berbagi sumber daya. `glue:PutResourcePolicy` Untuk mengatasi masalah ini, izinkan `glue:PutResourcePolicy` tindakan dengan peran yang diasumsikan yang digunakan oleh producer/grantor akun.

Kesalahan: “Tidak berwenang untuk memberikan izin untuk sumber daya”

Upaya dilakukan untuk memberikan izin lintas akun pada database atau tabel yang dimiliki oleh akun lain. Ketika database atau tabel dibagikan dengan akun Anda, sebagai administrator data lake, Anda dapat memberikan izin hanya kepada pengguna di akun Anda.

Kesalahan: “Akses ditolak untuk mengambil informasi AWS Organisasi”

Akun Anda adalah akun manajemen AWS Organizations dan Anda tidak memiliki izin yang diperlukan untuk mengambil informasi organisasi, seperti unit organisasi di akun.

Untuk informasi selengkapnya, lihat Required permissions for cross-account grants.

Kesalahan: “Organisasi <organization-ID>tidak ditemukan”

Upaya dilakukan untuk berbagi sumber daya dengan organisasi, tetapi berbagi dengan organisasi tidak diaktifkan. Aktifkan berbagi sumber daya dengan organisasi.

Untuk informasi selengkapnya, lihat Aktifkan Berbagi dengan AWS Organizations di Panduan AWS RAM Pengguna.

Kesalahan: “Izin Lake Formation tidak mencukupi: Kombinasi ilegal”

Pengguna membagikan sumber daya Katalog Data sementara izin Lake Formation diberikan kepada `IAMAllowedPrincipals` grup untuk sumber daya. Pengguna harus mencabut semua izin Lake Formation `IAMAllowedPrincipals` sebelum membagikan sumber daya.

ConcurrentModificationException atas grant/revoke permintaan ke akun eksternal

Ketika pengguna membuat beberapa hibah bersamaan and/or mencabut permintaan izin untuk prinsipal pada kebijakan LF-tag, maka Lake Formation melempar. ConcurrentModificationException Pengguna perlu menangkap pengecualian dan mencoba kembali grant/revoke permintaan yang

gagal. Menggunakan versi batch dari operasi`GrantPermissions`/`RevokePermissions`API - BatchGrantPermissionsdan BatchRevokePermissionsmeringankan masalah ini sampai batas tertentu dengan mengurangi jumlah permintaan bersamaan grant/revoke .

Kesalahan saat menggunakan Amazon EMR untuk mengakses data yang dibagikan melalui lintas akun

Saat Anda menggunakan Amazon EMR untuk mengakses data yang dibagikan dengan Anda dari akun lain, beberapa pustaka Spark akan mencoba memanggil operasi API. `Glue:GetUserDefinedFunctions` Karena izin AWS RAM terkelola versi 1 dan 2 tidak mendukung tindakan ini, Anda menerima pesan galat berikut:

```
"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"
```

Untuk mengatasi kesalahan ini, administrator data lake yang membuat pembagian sumber daya harus memperbarui izin AWS RAM terkelola yang dilampirkan ke pembagian sumber daya. Versi 3 dari izin AWS RAM terkelola memungkinkan prinsipal untuk melakukan tindakan. `glue:GetUserDefinedFunctions`

Jika Anda membuat pembagian sumber daya baru, Lake Formation menerapkan versi terbaru dari izin AWS RAM terkelola secara default, dan tidak ada tindakan yang diperlukan oleh Anda. Untuk mengaktifkan akses data lintas akun untuk pembagian sumber daya yang ada, Anda perlu memperbarui izin AWS RAM terkelola ke versi 3.

Anda dapat melihat AWS RAM izin yang ditetapkan ke sumber daya yang dibagikan dengan Anda di AWS RAM. Izin berikut disertakan dalam versi 3:

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
```

```
AWSRAMPermissionGlueAllTablesReadWriteForCatalog
AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

Untuk memperbarui versi izin AWS RAM terkelola dari pembagian sumber daya yang ada

Anda (administrator data lake) dapat memperbarui izin AWS RAM terkelola ke versi yang lebih baru dengan mengikuti petunjuk di Panduan AWS RAM Pengguna atau Anda dapat mencabut semua izin yang ada untuk jenis sumber daya dan memberikannya kembali. Jika Anda mencabut izin, AWS RAM menghapus pembagian AWS RAM sumber daya yang terkait dengan jenis sumber daya. Saat Anda memberikan kembali izin, AWS RAM buat pembagian sumber daya baru yang melampirkan versi terbaru izin terkelola. AWS RAM

Memecahkan masalah cetak biru dan alur kerja

Gunakan informasi di sini untuk membantu Anda mendiagnosis dan memperbaiki masalah cetak biru dan alur kerja.

Topik

- <role-ARN>Cetak biru saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: on resource:PassRole ”
- <role-ARN>Alur kerja saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: PassRole on resource:”
- Perayap dalam alur kerja saya gagal dengan “Sumber daya tidak ada atau pemohon tidak diizinkan untuk mengakses izin yang diminta”
- Perayap di alur kerja saya gagal dengan “Terjadi kesalahan (AccessDeniedException) saat memanggil CreateTable operasi...”

<role-ARN>Cetak biru saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: on resource:PassRole ”

Upaya dilakukan untuk membuat cetak biru oleh pengguna yang tidak memiliki izin yang cukup untuk lulus peran yang dipilih.

Perbarui kebijakan IAM pengguna agar dapat meneruskan peran, atau minta pengguna untuk memilih peran yang berbeda dengan izin peran sandi yang diperlukan.

Untuk informasi selengkapnya, lihat the section called “Referensi personas Lake Formation dan izin IAM”.

<role-ARN>Alur kerja saya gagal dengan “User: <user-ARN>is not authorized to perform: iam: PassRole on resource:”

Peran yang Anda tentukan untuk alur kerja tidak memiliki kebijakan inline yang memungkinkan peran untuk lulus sendiri.

Untuk informasi selengkapnya, lihat the section called “(Opsional) Membuat peran IAM untuk alur kerja”.

Perayap dalam alur kerja saya gagal dengan “Sumber daya tidak ada atau pemohon tidak diizinkan untuk mengakses izin yang diminta”

Salah satu kemungkinan penyebabnya adalah bahwa peran yang diteruskan tidak memiliki izin yang cukup untuk membuat tabel di database target. Berikan peran `CREATE_TABLE` izin pada database.

Perayap di alur kerja saya gagal dengan “Terjadi kesalahan (AccessDeniedException) saat memanggil CreateTable operasi...”

Salah satu kemungkinan penyebabnya adalah bahwa peran alur kerja tidak memiliki izin lokasi data pada lokasi penyimpanan target. Berikan izin lokasi data ke peran tersebut.

Untuk informasi selengkapnya, lihat the section called “`DATA_LOCATION_ACCESS`”.

Masalah yang diketahui untuk AWS Lake Formation

Tinjau masalah yang diketahui ini untuk AWS Lake Formation.

Topik

- Batasan pada penyaringan metadata tabel
- Masalah dengan mengganti nama kolom yang dikecualikan
- Masalah dengan menghapus kolom dalam tabel CSV
- Partisi tabel harus ditambahkan di bawah jalur umum
- Masalah dengan membuat database selama pembuatan alur kerja

- Masalah dengan menghapus dan kemudian membuat ulang pengguna
- Operasi API Katalog Data tidak memperbarui nilai untuk IsRegisteredWithLakeFormation parameter
- Operasi Lake Formation tidak mendukung AWS Glue Schema Registry

Batasan pada penyaringan metadata tabel

AWS Lake Formation izin tingkat kolom dapat digunakan untuk membatasi akses ke kolom tertentu dalam tabel. Saat pengguna mengambil metadata tentang tabel menggunakan konsol atau API seperti`glue:GetTable`, daftar kolom dalam objek tabel hanya berisi bidang yang dapat mereka akses. Penting untuk memahami keterbatasan penyaringan metadata ini.

Meskipun Lake Formation menyediakan metadata tentang izin kolom ke layanan terintegrasi, pemfilteran kolom yang sebenarnya dalam tanggapan kueri adalah tanggung jawab layanan terintegrasi. Klien Lake Formation yang mendukung pemfilteran tingkat kolom, termasuk Amazon Athena, Amazon Redshift Spectrum, dan Amazon EMR memfilter data berdasarkan izin kolom yang terdaftar di Lake Formation. Pengguna tidak akan dapat membaca data apa pun yang seharusnya tidak mereka akses. Saat ini, AWS Glue ETL tidak mendukung pemfilteran kolom.

> **Note**
>
> Cluster EMR tidak sepenuhnya dikelola oleh. AWS Oleh karena itu, adalah tanggung jawab administrator EMR untuk mengamankan cluster dengan benar untuk menghindari akses data yang tidak sah.

Aplikasi atau format tertentu mungkin menyimpan metadata tambahan, termasuk nama dan jenis kolom, di `Parameters` peta sebagai properti tabel. Properti ini dikembalikan tanpa dimodifikasi dan dapat diakses oleh pengguna mana pun dengan `SELECT` izin pada kolom apa pun.

Misalnya, Avro SerDe menyimpan representasi JSON dari skema tabel dalam properti tabel bernama`avro.schema.literal`, yang tersedia untuk semua pengguna dengan akses ke tabel. Kami menyarankan Anda menghindari menyimpan informasi sensitif dalam properti tabel dan menyadari bahwa pengguna dapat mempelajari skema lengkap tabel format Avro. Batasan ini khusus untuk metadata tentang tabel.

AWS Lake Formation menghapus properti tabel apa pun yang dimulai dengan `spark.sql.sources.schema` saat menanggapi permintaan `glue:GetTable` atau serupa jika

pemanggil tidak memiliki SELECT izin pada semua kolom dalam tabel. Ini mencegah pengguna mendapatkan akses ke metadata tambahan tentang tabel yang dibuat dengan Apache Spark. Saat dijalankan di Amazon EMR, aplikasi Apache Spark masih dapat membaca tabel ini, tetapi pengoptimalan tertentu mungkin tidak diterapkan, dan nama kolom peka huruf besar/kecil tidak didukung. Jika pengguna memiliki akses ke semua kolom dalam tabel, Lake Formation mengembalikan tabel yang tidak dimodifikasi dengan semua properti tabel.

Masalah dengan mengganti nama kolom yang dikecualikan

Jika Anda menggunakan izin tingkat kolom untuk mengecualikan kolom dan kemudian mengganti nama kolom, kolom tidak lagi dikecualikan dari kueri, seperti. SELECT *

Masalah dengan menghapus kolom dalam tabel CSV

Jika Anda membuat tabel Katalog Data dengan format CSV dan kemudian menghapus kolom dari skema, kueri dapat menampilkan data yang salah, dan izin tingkat kolom mungkin tidak dipatuhi.

Solusi: Buat tabel baru sebagai gantinya.

Partisi tabel harus ditambahkan di bawah jalur umum

Lake Formation mengharapkan semua partisi tabel berada di bawah jalur umum yang diatur di bidang lokasi tabel. Saat Anda menggunakan crawler untuk menambahkan partisi ke katalog, ini berfungsi dengan mulus. Tetapi jika Anda menambahkan partisi secara manual, dan partisi ini tidak berada di bawah lokasi yang diatur dalam tabel induk, akses data tidak berfungsi.

Masalah dengan membuat database selama pembuatan alur kerja

Saat membuat alur kerja dari cetak biru menggunakan konsol Lake Formation, Anda dapat membuat database target jika tidak ada. Ketika Anda melakukannya, pengguna yang masuk mendapatkan CREATE_TABLE izin pada database yang dibuat. Namun, crawler yang dihasilkan alur kerja mengasumsikan peran alur kerja saat mencoba membuat tabel. Ini gagal karena peran tidak memiliki CREATE_TABLE izin pada database.

Solusi: Jika Anda membuat database melalui konsol selama penyiapan alur kerja, sebelum menjalankan alur kerja, Anda harus memberi peran yang terkait dengan alur kerja CREATE_TABLE izin pada database yang baru saja Anda buat.

Masalah dengan menghapus dan kemudian membuat ulang pengguna

Skenario berikut menghasilkan izin Lake Formation yang salah yang dikembalikan oleh: `lakeformation:ListPermissions`

1. Buat pengguna dan berikan izin Lake Formation.
2. Hapus pengguna.
3. Buat ulang pengguna dengan nama yang sama.

`ListPermissions`mengembalikan dua entri, satu untuk pengguna lama dan satu untuk pengguna baru. Jika Anda mencoba mencabut izin yang diberikan kepada pengguna lama, izin dicabut dari pengguna baru.

Operasi API Katalog Data tidak memperbarui nilai untuk **`IsRegisteredWithLakeFormation`** parameter

Ada batasan yang diketahui bahwa operasi API Katalog Data seperti `GetTables` dan `SearchTables` tidak memperbarui nilai untuk `IsRegisteredWithLakeFormation` parameter, dan mengembalikan default, yang salah. Disarankan untuk menggunakan `GetTable` API untuk melihat nilai yang benar untuk `IsRegisteredWithLakeFormation` parameter.

Operasi Lake Formation tidak mendukung AWS Glue Schema Registry

Operasi Lake Formation tidak mendukung AWS Glue tabel yang berisi a `SchemaReference` dalam `StorageDescriptor` untuk digunakan dalam Schema Registery.

Pesan kesalahan yang diperbarui

AWS Lake Formation telah memperbarui pengecualian khusus sumber daya ke pesan `EntityNotFound` kesalahan umum untuk operasi API berikut untuk memenuhi tujuan keamanan dan kepatuhan.

- RevokePermissions
- GrantPermissions
- GetResourceLFTags
- GetTable

- GetDatabase

AWS Lake Formation API

> **Note**
>
> Referensi API yang diperbarui untuk AWS Lake Formation layanan sekarang tersedia.

Daftar Isi

- Izin APIs
  - Operasi
  - Jenis Data
- Pengaturan danau data APIs
  - Operasi
  - Jenis Data
- Integrasi Pusat Identitas IAM APIs
  - Operasi
  - Jenis Data
- Mode akses hibrid APIs
  - Operasi
  - Jenis Data
- Penjual kredensi APIs
  - Operasi
  - Jenis Data
- Menandai APIs
  - Operasi
  - Jenis Data
- Filter data APIs
  - Operasi
  - Jenis data
- Tipe data umum
  - ErrorDetail struktur

- Pola string

Izin APIs

Bagian API Izin menjelaskan operasi dan tipe data yang diperlukan untuk memberikan dan mencabut izin di. AWS Lake Formation Lihat Panduan Referensi API Lake Formation untuk semua operasi AWS Lake Formation API dan tipe data.

Operasi

- GrantPermissions
- RevokePermissions
- BatchGrantPermissions
- BatchRevokePermissions
- GetEffectivePermissionsForPath
- ListPermissions
- GetDataLakePrincipal

Jenis Data

- Sumber Daya
- DatabaseResource
- TableResource
- TableWithColumnsResource
- DataCellsFilterResourcee
- DataLocationResource
- DataLakePrincipal
- PrincipalPermissions
- PrincipalResourcePermissions
- DetailsMap
- ColumnWildcard
- BatchPermissionsRequestEntry

- BatchPermissionsFailureEntry

Pengaturan danau data APIs

Bagian ini berisi operasi API pengaturan data lake dan tipe data untuk mengelola administrator data lake.

Operasi

- GetDataLakeSettings
- PutDataLakeSettings

Jenis Data

- DataLakeSettings

Integrasi Pusat Identitas IAM APIs

Bagian ini berisi operasi untuk membuat dan mengelola integrasi Lake Formation dengan IAM Identity Center.

Operasi

- CreateLakeFormationIdentityCenterConfiguration
- DeleteLakeFormationIdentityCenterConfiguration
- DescribeLakeFormationIdentityCenterConfiguration
- UpdateLakeFormationIdentityCenterConfiguration

Jenis Data

- ExternalFilteringConfiguration

Mode akses hibrid APIs

Bagian API mode akses Hybrid menjelaskan operasi dan tipe data yang diperlukan untuk menyiapkan mode akses hibrid AWS Lake Formation. Lihat Panduan Referensi API Lake Formation untuk semua operasi AWS Lake Formation API dan tipe data.

Operasi

- CreateLakeFormationOptIn
- DeleteLakeFormationOptIn
- ListLakeFormationOptIns

Jenis Data

- Sumber Daya
- DatabaseResource
- TableResource
- Info Sumber Daya
- LakeFormationOptInsInfo
- DataLocationResource

Penjual kredensi APIs

Bagian Credential Vending API menjelaskan operasi dan tipe data yang terkait dengan bekerja dengan AWS Lake Formation layanan untuk menjual kredensi dan untuk mendaftar dan mengelola sumber daya data lake.

Operasi

- RegisterResource
- DeregisterResource
- ListResources
- GetUnfilteredTableMetadata
- GetUnfilteredPartitionsMetadata

- GetTemporaryGluePartitionCredentials
- GetTemporaryGlueTableCredentials
- UpdateResource

Jenis Data

- FilterCondition
- RowFilter
- ResourceInfo

Menandai APIs

Bagian API Tagging menjelaskan operasi dan tipe data yang terkait dengan strategi otorisasi yang mendefinisikan model izin pada atribut atau tag pasangan nilai kunci.

Operasi

- Dapatkan LFTag Ekspresi
- Daftar LFTag Ekspresi
- Hapus LFTag Ekspresi
- Perbarui LFTag Ekspresi
- Buat LFTag Ekspresi
- Menambahkan LFTags ToResource
- Hapus LFTags FromResource
- GetResourceLFTags
- Daftar LFTags
- Buat LFTag
- Dapatkan LFTag
- Perbarui LFTag
- Hapus LFTag
- SearchTablesByLFTags

- SearchDatabasesByLFTags

Jenis Data

- LFTagKeyResource
- LFTagPolicyResource
- TaggedTable
- TaggedDatabase
- LFTag
- LFTagSepasang
- LFTagKesalahan
- Kolom LFTag

Filter data APIs

Filter Data APIs menjelaskan cara mengelola filter sel data di AWS Lake Formation.

Operasi

- CreateDataCellsFilter
- DeleteDataCellsFilter
- ListDataCellsFilter
- GetDataCellsFilter
- UpdateDataCellsFilter

Jenis data

- DataCellsFilter
- RowFilter

Tipe data umum

Tipe Data Umum menjelaskan berbagai tipe data umum di. AWS Lake Formation

ErrorDetail struktur

Berisi detail tentang kesalahan.

Bidang

- `ErrorCode` — String UTF-8, sepanjang tidak kurang dari 1 atau lebih dari 255 byte, yang cocok dengan Single-line string pattern.

  Kode yang dikaitkan dengan kesalahan ini.
- `ErrorMessage` — String deskripsi, dengan panjang tidak lebih dari 2048 byte, yang cocok dengan URI address multi-line string pattern.

  Sebuah pesan yang menjelaskan kesalahan.

Pola string

API menggunakan ekspresi reguler berikut untuk menentukan konten apa yang valid untuk berbagai parameter string dan anggota:

- Pola string satu baris — `"[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*"`
- Pola string multi-baris alamat URI — `"[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*"`
- Pola string kustom #3 — `"^\w+\.\w+\.\w+$”`
- Pola string kustom #4 — `"^\w+\.\w+$”`
- Pola string kustom #5 — `"arn:aws:iam::[0-9]*:role/.*”`
- Pola string kustom #6 — `"arn:aws:iam::[0-9]*:user/.*”`
- Pola string kustom #7 — `"arn:aws:iam::[0-9]*:group/.*”`
- Pola string kustom #8 — `"arn:aws:iam::[0-9]*:saml-provider/.*”`
- Pola string kustom #9 — `"^([\p{L}\p{Z}\p{N}_.:\/=+\-@%]*)$”`
- Pola string kustom #10 — `"^([\p{L}\p{Z}\p{N}_.:\*\/=+\-@%]*)$"`
- Pola string kustom #11 — `"[\p{L}\p{N}\p{P}]*"`

Wilayah yang Didukung

Bagian ini memiliki informasi tentang dukungan Wilayah AWS dan fungsionalitas untuk Lake Formation.

Ketersediaan umum

Untuk Wilayah AWS dukungan AWS Lake Formation, lihat Daftar AWS layanan yang tersedia menurut Wilayah.

Untuk daftar titik akhir layanan Lake Formation untuk setiap Wilayah dan kuota layanan Lake Formation, lihat AWS Lake Formation titik akhir dan kuota.

AWS GovCloud (US)

Untuk ikhtisar perbedaan antara AWS GovCloud (US) Wilayah dan standar Wilayah AWS, lihat Bagaimana AWS Lake Formation perbedaannya AWS GovCloud (US).

Transaksi dan optimasi penyimpanan

Tabel yang diatur, dukungan transaksi, dan fitur pengoptimalan penyimpanan untuk Lake Formation tersedia sebagai berikut: Wilayah AWS

Nama wilayah	Parameter wilayah	Titik akhir
US East (N. Virginia)	us-east-1	lakeformation.us-east-1.amazonaws.com lakeformation-fips.us-east-1.amazonaws.com
AS Timur (Ohio)	us-east-2	lakeformation.us-east-2.amazonaws.com lakeformation-fips.us-east-2.amazonaws.com

Nama wilayah	Parameter wilayah	Titik akhir
AS Barat (Oregon)	us-west-2	lakeformation.us-west-2.amazonaws.com lakeformation-fips.us-west-2.amazonaws.com
Asia Pasifik (Mumbai)	ap-south-1	lakeformation.ap-south-1.amazonaws.com
Asia Pasifik (Seoul)	ap-northeast-2	lakeformation.ap-northeast-2.amazonaws.com
Asia Pasifik (Singapura)	ap-southeast-1	lakeformation.ap-southeast-1.amazonaws.com
Asia Pasifik (Sydney)	ap-southeast-2	lakeformation.ap-southeast-2.amazonaws.com
Asia Pasifik (Tokyo)	ap-northeast-1	lakeformation.ap-northeast-1.amazonaws.com
Eropa (Frankfurt)	eu-central-1	lakeformation.eu-central-1.amazonaws.com
Eropa (Irlandia)	eu-west-1	lakeformation.eu-west-1.amazonaws.com
Eropa (London)	eu-west-2	lakeformation.eu-west-2.amazonaws.com
Eropa (Stockholm)	eu-north-1	lakeformation.eu-north-1.amazonaws.com

Nama wilayah	Parameter wilayah	Titik akhir
Kanada (Pusat)	ca-central-1	lakeformation.ca-central-1.amazonaws.com
Amerika Selatan (Sao Paulo)	sa-east-1	lakeformation.sa-east-1.amazonaws.com

Riwayat dokumen untuk AWS Lake Formation

Tabel berikut menjelaskan perubahan penting pada dokumentasi untuk AWS Lake Formation.

Perubahan	Deskripsi	Tanggal
Dokumentasi yang diperbarui untuk kontrol akses berbasis atribut	Lake Formation memungkinkan Anda memberikan izin kepada kepala sekolah dengan atribut yang cocok pada sumber daya Katalog Data. Untuk informasi selengkapnya, lihat Kontrol akses berbasis atribut.	17 April 2025
Integrasi Amazon S3 Tabel dengan dan AWS Lake FormationAWS Glue Data Catalog	Anda sekarang dapat mengintegrasikan dan membuat katalog Tabel S3 sebagai AWS Glue Data Catalog objek dan mendaftarkan katalog sebagai lokasi data Lake Formation. Untuk informasi selengkapnya, lihat Membuat katalog Tabel Amazon S3 di. AWS Glue Data Catalog	Maret 13, 2025
`AWSLakeFormationCrossAccountManager` Kebijakan yang diperbarui Lake Formation.	Lake Formation meningkatkan AWSLakeFormationCrossAccountManagerkebijakan dengan mengganti operator `StringLike` kondisi dengan `ArnLike` operator yang memungkinkan IAM melakukan pemeriksaan format ARN.	Januari 25, 2025

Perubahan kebijakan yang diperbarui	Mendokumentasikan perubahan AWSLakeFormationDataAdminkebijakan.	Desember 3, 2024
Pembaruan multi-katalog	AWS Glue Data Catalog memungkinkan Anda membuat katalog gabungan, dan menyatukan data di seluruh danau data Amazon S3 dan gudang data Amazon Redshift serta menginteg rasikan data dari basis data operasional Anda seperti Amazon DynamoDB, dan sumber data pihak ketiga seperti Snowflake, MySQL, dan sebagainya. Untuk informasi selengkapnya, lihat Membawa data Anda ke dalam AWS Glue Data Catalog.	Desember 3, 2024
Dokumentasi diperbarui untuk ekspresi LF-tag	Anda dapat menyimpan ekspresi LF-tag dan menggunakannya kembali untuk memberikan izin pada sumber daya Katalog Data. Untuk informasi selengkapnya, lihat Mengelola ekspresi LF-tag.	November 7, 2024

Dokumentasi diperbarui untuk tampilan Data Catalog	Anda dapat membuat tampilan AWS Glue APIs dalam AWS Glue Data Catalog penggunaan selain DDLs menggunakan Amazon Athena dan Amazon Redshift. Untuk informasi selengkapnya, lihat Membangun tampilan Katalog Data.	Agustus 7, 2024
Menambahkan dokumentasi untuk penjual kredensi yang dapat diaudit	Lake Formation memungkinkan Anda untuk memasukkan konteks pengguna IAM Identity Center dalam CloudTrail acara, dan kemudian melacak pengguna yang mengakses sumber daya Anda. Untuk informasi selengkapnya, lihat Menyertakan konteks pengguna Pusat Identitas IAM di CloudTrail log.	Juli 14, 2024
Perubahan kebijakan yang diperbarui	Mendokumentasikan perubahan (menambahkan pernyataan IDs dan menghapus izin berlebihan) ke dan kebijakan. AWSLakeFormationCrossAccountManager AWSLakeFormationDataAdmin	Maret 14, 2024
Memperbarui pengaturan Lake Formation	Memperbarui langkah-langkah di AWS Lake Formation bagian Siapkan.	Februari 7, 2024

Perubahan kebijakan yang diperbarui	Menambahkan izin baru ke kebijakan sebaris peran terkait layanan. Untuk informasi selengkapnya, lihat Menggunakan peran terkait layanan untuk Lake Formation.	Februari 7, 2024
Perubahan kebijakan yang diperbarui	Mendokumentasikan perubahan LakeFormationDataAccessServiceRolePolicykebijakan.	Februari 2, 2024
Batasan Formasi Danau Konsolidasi	Membuat bagian terpadu untuk batasan dan pertimbangan Lake Formation. Untuk informasi lebih lanjut, lihat batasan Lake Formation.	15 Desember 2023
Ditambahkan dokumentasi untuk Iceberg pemadatan	Untuk kinerja baca yang lebih baik oleh layanan AWS analitik seperti Athena dan Amazon EMR, dan pekerjaan AWS Glue ETL, AWS Glue Data Catalog menyediakan pemadatan terkelola (proses yang memadatkan objek Amazon S3 kecil menjadi objek yang lebih besar) untuk tabel Iceberg di Katalog Data. Untuk informasi selengkapnya, lihat Mengoptimalkan tabel Gunung Es.	November 25, 2023

Ditambahkan dokumentasi untuk integrasi IAM Identity Center	Integrasi IAM Identity Center memungkinkan pengguna dan grup mengakses sumber daya Katalog Data yang menerapkan izin Lake Formation. Untuk informasi selengkapnya, lihat Integrasi Pusat Identitas IAM.	November 25, 2023
Ditambahkan dokumentasi untuk tampilan Data Catalog	Anda dapat membuat tampilan di AWS Glue Data Catalog yang mereferensikan hingga 10 tabel menggunakan editor SQL untuk Amazon Athena atau Amazon Redshift. Untuk informasi selengkapnya, lihat Membuat tampilan.	November 25, 2023
Memperbarui perubahan kebijakan	Mendokumentasikan perubahan AWSLakeFormationCrossAccountManagerkebijakan.	25 Oktober 2023

Ditambahkan dokumentasi untuk modus akses hybrid	Mode akses hibrida memberikan fleksibilitas untuk mengaktifkan izin Lake Formation secara selektif untuk database dan tabel di situs Anda. AWS Glue Data Catalog Dengan mode akses hybrid, Anda sekarang memiliki jalur tambahan yang memungkinkan Anda mengatur izin Lake Formation untuk kumpulan pengguna tertentu tanpa mengganggu kebijakan izin pengguna atau beban kerja lain yang ada. Untuk informasi selengkapnya, lihat Mode akses hibrid.	26 September 2023
Ditambahkan dokumentasi untuk membuat tabel Apache Iceberg	Anda sekarang dapat membuat tabel Apache Iceberg yang menggunakan format data Apache Parquet di AWS Glue Data Catalog dengan data yang berada di Amazon S3. Untuk informasi selengkapnya, lihat Membuat tabel Gunung Es.	16 Agustus 2023

Ditambahkan dokumentasi untuk akses data lintas wilayah	Lake Formation mendukung kueri tabel Katalog Data di seluruh AWS Wilayah. Anda dapat mengakses data di Wilayah dari Wilayah lain menggunakan Athena, Amazon EMR, dan menjalankan AWS Glue ETL dengan membuat tautan sumber daya di Wilayah lain yang menunjuk ke database dan tabel sumber. Anda dapat menghubungkan Katalog Data ke metastor eksternal yang menyimpan metadata untuk data Amazon S3 Anda, dan mengelola izin akses data dengan aman. AWS Lake Formation Untuk informasi selengkapnya, lihat Mengakses tabel di seluruh Wilayah.	Juni 30, 2023
Konten yang diatur ulang	Bab yang diatur ulang dalam panduan untuk mencocokkan perjalanan pengguna Lake Formation.	15 Mei 2023

Ditambahkan dokumentasi untuk federasi HMS	Anda dapat menghubungkan Katalog Data ke metastor eksternal yang menyimpan metadata untuk data Amazon S3 Anda, dan mengelola izin akses data dengan aman. AWS Lake Formation Untuk informasi selengkapnya, lihat Mengelola izin pada kumpulan data yang menggunakan metastor eksternal.	April 15, 2023
Menambahkan dokumentasi untuk berbagi data Amazon Redshift	Anda sekarang dapat mengelola data dengan aman di datashare dari Amazon Redshift menggunakan izin Lake Formation. Lake Formation mendukung akses lisensi ke data Anda melalui AWS Data Exchange. Untuk informasi selengkapnya, lihat Berbagi data di AWS Lake Formation.	30 November 2022
Support untuk berbagi data lintas akun langsung dengan kepala sekolah	Menambahkan informasi tentang berbagi data secara langsung dengan kepala sekolah IAM di akun lain. Untuk informasi selengkapnya, lihat Berbagi data lintas akun di AWS Lake Formation.	10 November 2022

Support untuk berbagi data yang AWS RAM diaktifkan menggunakan TBAC	Menambahkan informasi tentang metode LF-TBAC dalam memberikan izin Katalog Data yang digunakan untuk hibah lintas akun. AWS Resource Access Manager	10 November 2022
Menambahkan bagian tentang bekerja dengan layanan lain	Menambahkan informasi tentang bagaimana AWS layanan seperti Athena,, AWS Glue Redshift Spectrum, dan Amazon EMR dapat menggunakan Lake Formation untuk mengakses data dengan aman di lokasi Amazon S3 yang terdaftar di Lake Formation. Untuk mengetahui informasi selengkapnya, lihat Bekerja dengan AWS layanan lain.	10 November 2022
???	Menambahkan informasi tentang pemecahan masalah kesalahan saat menggunakan Amazon EMR untuk mengakses data lintas akun. Untuk informasi selengkapnya, lihat Kesalahan saat menggunakan Amazon EMR untuk mengakses data yang dibagikan melalui lintas akun.	7 November 2022

Pembaruan untuk berbagi sumber daya lintas akun	Menambahkan deskripsi tentang cara kerja pembagian sumber daya lintas akun di Lake Formation. Mendokumentasikan perubahan AWSLakeFormationCrossAccountManagerkebijakan.	6 Mei 2022
Tutorial baru	Menambahkan tutorial baru untuk membuat tabel yang diatur, mengamankan data lake, dan berbagi data lake. Untuk detail selengkapnya, lihat bagian Memulai.	20 April 2022
Halaman arahan Lake Formation baru	Memperbarui halaman arahan Lake Formation untuk menyertakan tautan untuk tutorial yang memberikan step-by-step instruksi tentang cara membangun danau data, menelan data, berbagi, dan mengamankan data lake menggunakan Lake Formation .	20 April 2022
Support untuk penjual kredenal	Menambahkan informasi tentang penjual kredenal, yang mendukung Lake Formation untuk memungkinkan layanan pihak ketiga berintegrasi dengan Lake Formation dengan menggunakan operasi API penjual kredenal. Untuk informasi lebih lanjut, lihat Cara kerja penjual kredenal di Lake Formation.	28 Februari 2022

Support untuk tabel yang diatur dan pemfilteran data tingkat lanjut	Menambahkan informasi tentang tabel yang diatur, yang mendukung transaksi ACID, pemadatan data otomatis, dan kueri perjalanan waktu. Menambahkan informasi tentang membuat filter data untuk mendukung keamanan tingkat kolom, keamanan tingkat baris, dan keamanan tingkat sel. Untuk informasi lebih lanjut, lihat Tabel yang Diatur di Lake Formation dan Data Filtering dan Keamanan Tingkat Sel di Lake Formation.	30 November 2021
Support untuk titik akhir antarmuka VPC	Menambahkan informasi tentang membuat titik akhir antarmuka virtual private cloud (VPC) untuk Lake Formation , sehingga komunikasi antara VPC dan Lake Formation Anda dilakukan sepenuhnya dan aman di dalam jaringan. AWS Untuk informasi selengkapnya, lihat Menggunakan Lake Formation with VPC Endpoints.	11 Oktober 2021

Dukungan untuk kebijakan titik akhir VPC	Menambahkan informasi tentang dukungan untuk kebijakan titik akhir Virtual Private Cloud (VPC) di Lake Formation. Untuk informasi selengkapnya, lihat Menggunakan Lake Formation with VPC Endpoints.	11 Oktober 2021
Support untuk kontrol akses berbasis tag	Kontrol akses berbasis tag Lake Formation menyediakan cara baru yang lebih skalabel untuk mengelola akses ke sumber daya Katalog Data dan data yang mendasarinya dengan menggunakan LF-tag. Untuk informasi lebih lanjut, lihat Kontrol Akses Berbasis Tag Lake Formation.	7 Mei 2021
Persyaratan keikutsertaan baru untuk pemfilteran data di Amazon EMR.	Menambahkan informasi tentang persyaratan untuk ikut serta untuk mengizink an Amazon EMR memfilter data yang dikelola oleh Lake Formation. Untuk informasi selengkapnya, lihat Mengizink an Pemfilteran Data di Amazon EMR.	9 Oktober 2020

Support untuk memberikan izin lintas akun penuh pada database Data Catalog	Menambahkan informasi tentang pemberian izin Lake Formation lengkap pada database Katalog Data di seluruh AWS akun, termasuk. `CREATE_TABLE` Untuk informasi selengkapnya, lihat Berbagi Database Katalog Data.	1 Oktober 2020
Support untuk Amazon Athena pengguna yang mengauten tikasi melalui SALL.	Menambahkan informasi tentang dukungan untuk pengguna Athena yang terhubung melalui driver JDBC atau ODBC dan mengauten tikasi melalui penyedia identitas SALL seperti Okta dan Microsoft Active Directory Federation Service (AD FS). Untuk informasi selengkapnya, lihat Integrasi AWS Layanan dengan Lake Formation.	30 September 2020
Support untuk akses lintas akun dengan Katalog Data terenkripsi	Menambahkan informasi tentang pemberian izin lintas akun saat Katalog Data dienkripsi. Untuk informasi selengkapnya, lihat Prasyarat Akses Lintas Akun.	30 Juli 2020

Support untuk akses lintas akun ke data lake	Menambahkan informasi tentang pemberian AWS Lake Formation izin pada database dan tabel Katalog Data ke AWS akun dan organisasi eksternal, dan tentang mengakses objek Katalog Data yang dibagikan dari akun eksternal. Untuk informasi selengkapnya, lihat Akses Lintas Akun.	7 Juli 2020
Integrasi dengan Amazon Quick	Menambahkan informasi tentang cara memberikan izin Lake Formation kepada pengguna Quick Enterprise Edition sehingga mereka dapat mengakses kumpulan data yang berada di lokasi Amazon S3 terdaftar. Untuk informasi selengkapnya, lihat Memberikan Izin Katalog Data.	29 Juni 2020
Pembaruan untuk menyiapkan dan Memulai Bab	Menata ulang dan meningkatkan Bab Pengaturan dan Memulai. Memperbarui izin yang direkomendasikan AWS Identity and Access Management (IAM) untuk administrator danau data.	27 Februari 2020

Support untuk AWS Key Management Service	Menambahkan informasi tentang cara dukungan Lake Formation untuk AWS Key Management Service (AWS KMS) menyederhanakan pengaturan layanan terintegrasi untuk membaca dan menulis data terenkripsi di lokasi Amazon Simple Storage Service (Amazon S3) terdaftar. Menambahkan informasi tentang cara mendaftarkan lokasi Amazon S3 yang dienkripsi. AWS KMS keys Untuk informasi selengkapnya, lihat the section called "Menambahkan lokasi Amazon S3 ke danau data Anda".	27 Februari 2020
Pembaruan untuk cetak biru dan kebijakan IAM administrator data lake	Parameter input yang diklarifikasi untuk cetak biru database inkremental. Memperbarui kebijakan IAM yang diperlukan untuk administrator data lake.	20 Desember 2019
Bab keamanan menulis ulang dan meningkatkan revisi Bab	Meningkatkan keamanan dan peningkatan chapter.	Oktober 29, 2019
Izin super menggantikan Semua izin	Memperbarui Bab Keamanan dan Peningkatan untuk mencerminkan penggantian izin `All` dengan. `Super`	10 Oktober 2019

Penambahan, koreksi, dan klarifikasi	Membuat penambahan, koreksi, dan klarifikasi berdasarkan umpan balik. Merevisi bagian keamanan. Memperbarui chapter Keamanan dan Peningkatan untuk mencerminkan penggantian grup `Everyone` dengan. `IAMAllowedPrincipals`	11 September 2019
Panduan baru	Ini adalah rilis awal dari Panduan Developer AWS Lake Formation .	8 Agustus 2019

AWS Glosarium

Untuk AWS terminologi terbaru, lihat AWS glosarium di Referensi.Glosarium AWS

Terjemahan disediakan oleh mesin penerjemah. Jika konten terjemahan yang diberikan bertentangan dengan versi bahasa Inggris aslinya, utamakan versi bahasa Inggris.