aws

Guía para desarrolladores

AWS Lake Formation

AWS Lake Formation: Guía para desarrolladores

Table of Contents

¿Qué es AWS Lake Formation?

Bienvenido a la Guía para desarrolladores de AWS SDK for JavaScript.

AWS Lake Formation le ayuda a gestionar, proteger y compartir datos a nivel mundial de forma centralizada para el análisis y el machine learning. Con Lake Formation, puede administrar el control de acceso detallado para los datos de su lago de datos en Amazon Simple Storage Service (Amazon S3) y sus metadatos en AWS Glue Data Catalog.

Lake Formation proporciona su propio modelo de permisos que aumenta el modelo de permisos de IAM. El modelo de permisos de Lake Formation permite un acceso específico a los datos almacenados en los lagos de datos mediante un sencillo mecanismo de concesión o revocación, muy similar al de un sistema de gestión de bases de datos relacionales (RDBMS). Los permisos de Lake Formation se aplican mediante controles granulares a nivel de columna, fila y celda en todos los servicios de análisis de AWS y machine learning, incluidos Amazon Athena, Amazon QuickSight, Amazon Redshift Spectrum, Amazon EMR y AWS Glue.

El modo de acceso híbrido de Lake Formation AWS Glue Data Catalog le permite proteger y acceder a los datos catalogados mediante los permisos de Lake Formation y las políticas de permisos de IAM para Amazon S3 y AWS Glue sus acciones. Con el modo de acceso híbrido, los administradores de datos pueden incorporar los permisos de Lake Formation de forma selectiva e incremental, centrándose en un caso práctico del lago de datos cada vez.

Lake Formation también le permite compartir datos interna y externamente a través de múltiples Cuentas de AWS, organizaciones de AWS o directamente con las entidades principales de IAM en otra cuenta proporcionando un acceso específico a los metadatos de AWS Glue Data Catalog y los datos subyacentes.

Temas

- Introducción a Lake Formation

Características de la Lake Formation

Lake Formation le ayuda a descomponer los silos de datos y a combinar diferentes tipos de datos estructurados y no estructurados en un repositorio centralizado. En primer lugar, identifique los almacenes de datos existentes en Amazon S3 o en bases de datos relacionales y NoSQL, y traslade los datos a su lago de datos. A continuación, rastree, catalogue y prepare los datos para su análisis. Después, proporcione a sus usuarios un acceso seguro de autoservicio a los datos a través de los servicios de análisis que elijan.

Temas

- Ingesta y administración de datos
- Administración de la seguridad
- Uso compartido de datos

Ingesta y administración de datos

Importar datos de bases de datos que ya estén en AWS

Tras especificar dónde se encuentran sus bases de datos y proporcione sus credenciales de acceso, Lake Formation lee los datos y sus metadatos (esquema) para comprender el contenido de los orígenes de datos. A continuación, importa los datos a su nuevo lago de datos y registra los metadatos en un catálogo central. Con Lake Formation, puede importar datos de bases de datos MySQL, PostgreSQL, SQL Server, MariaDB y Oracle que se ejecuten en Amazon RDS o estén alojadas en Amazon EC2. Son compatibles tanto la carga masiva de datos como la incremental.

Importar datos de otros orígenes externos

Puede usar Lake Formation para mover datos desde bases de datos en las instalaciones conectándose con Java Database Connectivity (JDBC). Identifique sus fuentes de destino y proporcione las credenciales de acceso en la consola, y Lake Formation leerá y cargará sus datos en el lago de datos. Para importar datos de bases de datos distintas de las enumeradas anteriormente, puede crear trabajos ETL personalizados con AWS Glue.

Catalogar y etiquetar sus datos

Puede utilizar rastreadores de AWS Glue para leer sus datos en Amazon S3 y extraer el esquema de la base de datos y las tablas y almacenar esos datos en un AWS Glue Data Catalog apto para búsquedas. A continuación, utilice Lake Formation Control de acceso basado en etiquetas de Lake Formation (TBAC) para administrar los permisos sobre bases de datos, tablas y columnas. Para obtener más información sobre cómo agregar tablas al Catálogo de datos, consulte Creación de tablas y bases de datos del Catálogo de datos.

Administración de la seguridad

Defina y gestione los controles de acceso

Lake Formation proporciona un único lugar para administrar los controles de acceso a los datos de su lago de datos. Puede definir políticas de seguridad que restrinjan el acceso a los datos a nivel de base de datos, tabla, columna, fila y celda. Estas políticas se aplican a usuarios y roles de IAM, y a usuarios y grupos cuando se federan a través de un proveedor de identidades externo. Puede utilizar controles detallados para acceder a los datos asegurados por Lake Formation dentro de Amazon Redshift Spectrum, Athena, AWS Glue ETL y Amazon EMR para Apache Spark. Siempre que cree identidades IAM, asegúrese de seguir las mejores prácticas IAM. Para más información, consulte las mejores prácticas de seguridad en la Guía del usuario de IAM.

Modo de acceso híbrido

El modo de acceso híbrido de Lake Formation proporciona la flexibilidad de habilitar selectivamente los permisos de Lake Formation para bases de datos y tablas en su AWS Glue Data Catalog. Con el modo de acceso híbrido, ahora tiene una ruta incremental que le permite establecer los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes. Para obtener más información, consulte Modo de acceso híbrido.

Implantar el registro de auditoría

Lake Formation proporciona registros de auditoría completos CloudTrail para monitorear el acceso y mostrar el cumplimiento de las políticas definidas centralmente. Puede auditar el historial de acceso a los datos en los servicios de análisis y de machine learning que leen los datos de su lago de datos a través de Lake Formation. Esto le permite ver qué usuarios o roles han intentado acceder a qué datos, con qué servicios y cuándo. Puede acceder a los registros de auditoría de la misma manera que accede a cualquier otro CloudTrail registro mediante las CloudTrail API y la consola. Para obtener más información sobre CloudTrail los registros, consulteRegistro de llamadas a la API de AWS Lake Formation mediante AWS CloudTrail.

Seguridad de nivel de fila y celda

Lake Formation proporciona filtros de datos que le permiten restringir el acceso a una combinación de columnas y filas. Utilice la seguridad a nivel de filas y celdas para proteger datos confidenciales como la información de identificación personal (PII). Para obtener más información sobre la seguridad a nivel de fila, consulte Información general del filtrado de datos.

Control de acceso basado en etiquetas

Utilice el control de acceso basado en etiquetas de Lake Formation para gestionar cientos o incluso miles de permisos de datos mediante la creación de etiquetas personalizadas denominadas etiquetas LF. Ahora puede definir etiquetas LF y adjuntarlas a bases de datos, tablas o columnas. A continuación, comparta el acceso controlado a través de los servicios de análisis, de machine learning (ML) y de extracción, transformación y carga (ETL) para su consumo. Las etiquetas LF garantizan que la gobernanza de los datos se pueda escalar fácilmente al reemplazar las definiciones de políticas de miles de recursos por unas pocas etiquetas lógicas. Lake Formation proporciona una búsqueda basada en texto sobre estos metadatos, para que sus usuarios puedan encontrar rápidamente los datos que necesitan analizar.

Acceso entre cuentas

Las capacidades de administración de permisos de Lake Formation simplifican la seguridad y la administración de los lagos de datos distribuidos en varias cuentas AWS a través de un enfoque centralizado, proporcionando un control de acceso específico al Catálogo de datos y a las ubicaciones de Amazon S3. Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation.

Uso compartido de datos

La capacidad de uso compartido de datos le permite establecer permisos sobre conjuntos de datos almacenados en diferentes orígenes de datos como Amazon Redshift sin necesidad de migrar datos o metadatos a Amazon S3 o AWS Glue Data Catalog. Puede utilizar cualquiera de los métodos siguientes para compartir datos en Lake Formation:

Para obtener más información, consulte Uso compartido de datos en Lake Formation.

- Integración de Lake Formation con el uso compartido de datos de Amazon Redshift. Utilice Lake Formation para administrar de forma centralizada los permisos de acceso a nivel de base de datos, tabla, columna y fila de los recursos compartidos de datos de Amazon Redshift y restringir el acceso de los usuarios a los objetos dentro de un recurso compartido de datos.

- Conexión de AWS Glue Data Catalog a metaalmacenes externos. Conecte AWS Glue Data Catalog a metaalmacenes externos para gestionar los permisos de acceso a los conjuntos de datos de Amazon S3 mediante Lake Formation. No es necesaria la migración de los metadatos a AWS Glue Data Catalog.

  Para obtener más información, consulte Administración de los permisos de los conjuntos de datos que utilizan metaalmacenes externos
- Integración de Lake Formation con el intercambio de datos de AWS. Lake Formation admite la concesión de licencias de acceso a sus datos mediante AWS Data Exchange. Si está interesado en licenciar sus datos de Lake Formation, consulte Qué es AWS Data Exchange en la Guía del usuario de AWS Data Exchange.

AWS Lake Formation: Cómo funciona

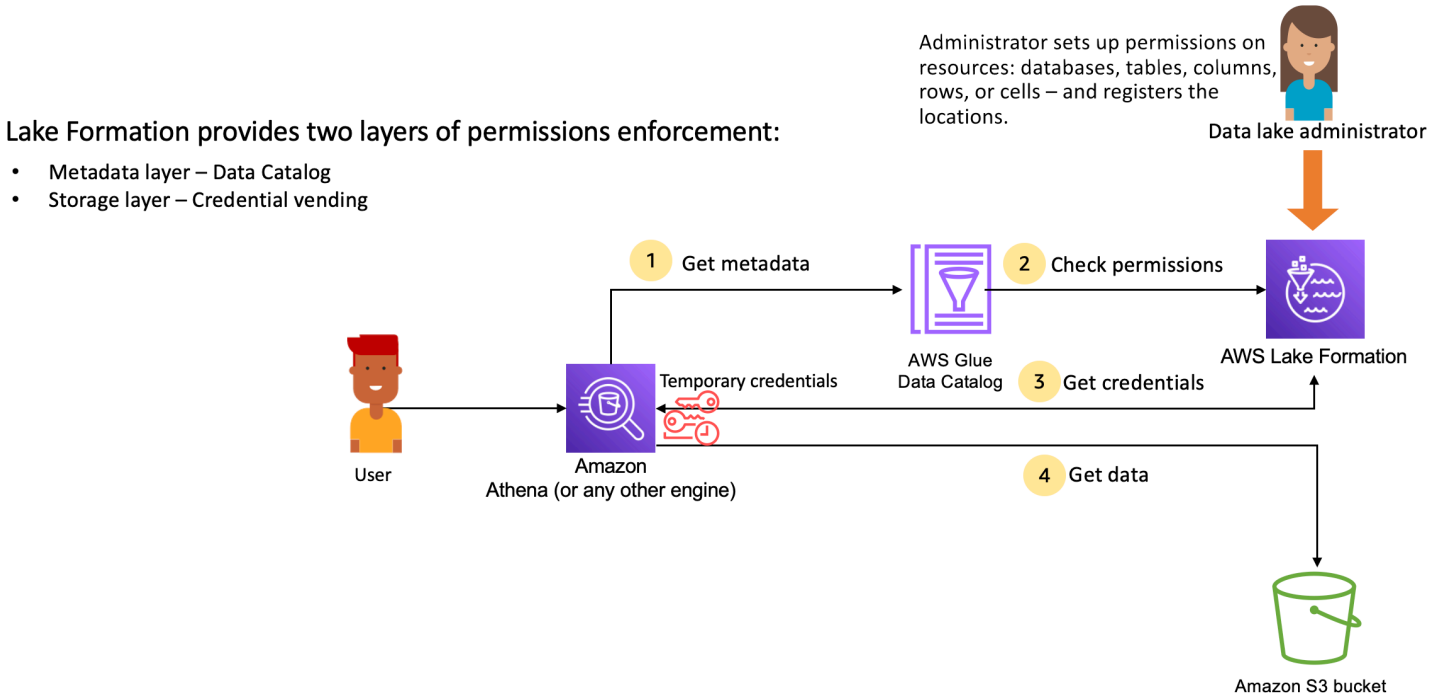
AWS Lake Formation proporciona un modelo de permisos de sistema de administración de bases de datos relacionales (RDBMS) para conceder o revocar el acceso a recursos del catálogo de datos como bases de datos, tablas y columnas con datos subyacentes en Amazon S3. Los permisos de Lake Formation, fáciles de gestionar, sustituyen a las complejas políticas de bucket de Amazon S3 y a las correspondientes políticas de IAM.

En Lake Formation, puede implementar permisos en dos niveles:

- Aplicación de permisos a nivel de metadatos en los recursos del catálogo de datos, como bases de datos y tablas
- Administración de los permisos de acceso al almacenamiento en los datos subyacentes guardados en Amazon S3 en nombre de los motores integrados

Flujo de trabajo de administración de permisos de Lake Formation

Lake Formation se integra con motores de análisis para consultar almacenes de datos de Amazon S3 y objetos de metadatos que están registrados en Lake Formation. El diagrama siguiente ilustra cómo funciona la administración de permisos en Lake Formation.



Pasos generales para la gestión de permisos de Lake Formation

Antes de que Lake Formation pueda proporcionar controles de acceso para los datos de su lago de datos, un administrador del lago de datos o un usuario con permisos administrativos configura las políticas de usuario individuales de las tablas del catálogo de datos para permitir o denegar el acceso a las tablas del catálogo de datos utilizando los permisos de Lake Formation.

A continuación, el administrador del lago de datos o un usuario delegado por el administrador concede permisos de Lake Formation a los usuarios en las bases de datos y tablas del catálogo de datos y registra la ubicación de Amazon S3 de la tabla en Lake Formation.

1. Obtener metadatos. Una entidad principal (usuario) envía una consulta o un script ETL a un motor analítico integrado como Amazon Athena, AWS Glue, Amazon EMR o Amazon Redshift Spectrum. El motor analítico integrado identifica la tabla que se solicita y envía una petición de metadatos al catálogo de datos.
2. Comprobar permisos. El catálogo de datos comprueba los permisos del usuario con Lake Formation, y si el usuario está autorizado a acceder a la tabla, devuelve al motor los metadatos que el usuario está autorizado a ver.
3. Obtener credenciales. El catálogo de datos permite al motor saber si la tabla está administrada por Lake Formation o no. Si los datos subyacentes están registrados en Lake Formation, el motor analítico solicita a Lake Formation que proporcione acceso a los datos mediante la concesión de un acceso temporal.

4. Obtener datos. Si el usuario está autorizado a acceder a la tabla, Lake Formation proporciona acceso temporal al motor analítico integrado. Mediante el acceso temporal, el motor analítico obtiene los datos de Amazon S3 y aplica el filtrado necesario, como el de columnas, filas o celdas. Cuando el motor termina de ejecutar el trabajo, devuelve los resultados al usuario. Este proceso se denomina expedición de credenciales.

   Si la tabla no está administrada por Lake Formation, la segunda llamada del motor analítico se hace directamente a Amazon S3. Para el acceso a los datos se evalúan la política de buckets de Amazon S3 y la política de usuarios de IAM correspondientes.

   Siempre que utilice políticas de IAM, compruebe que sigue las mejores prácticas IAM. Para más información, consulte Prácticas recomendadas de seguridad en IAM en la Guía del usuario de IAM.

Temas

Permisos de metadatos

Lake Formation proporciona autorización y control de acceso para el catálogo de datos. Cuando un rol de IAM efectúa una llamada a la API del catálogo de datos desde cualquier sistema, el catálogo de datos verifica los permisos de datos del usuario y solo devuelve los metadatos a los que el usuario tiene permisos para acceder. Por ejemplo, si un rol de IAM tiene acceso solo a una tabla dentro de una base de datos, y un servicio o un usuario que asume el rol ejecuta la operación `GetTables`, la respuesta contendrá solo esa tabla, con independencia del número de tablas de la base de datos.

Configuración predeterminada: permisos de grupo **`IAMAllowedPrincipal`**

De forma predeterminada, AWS Lake Formation asigna los permisos de todas las bases de datos y tablas a un grupo virtual denominado `IAMAllowedPrincipal`. Este grupo es único y visible solo dentro de Lake Formation. El grupo `IAMAllowedPrincipal` incluye a todas las entidades principales de IAM que tienen acceso a los recursos del catálogo de datos a través de las políticas principales de IAM y las políticas de recursos de AWS Glue. Si este permiso existe en una base de datos o tabla, se concederá a todas las entidades principales acceso a la base de datos o tabla.

Si desea proporcionar permisos más específicos en una base de datos o tabla, elimine el permiso `IAMAllowedPrincipal` y Lake Formation aplicará todas las demás políticas asociadas con esa base de datos o tabla. Por ejemplo, si existe una política que permite al Usuario A acceder a la base de datos A con permisos `DESCRIBE`, y el `IAMAllowedPrincipal` existe con todos los permisos, el Usuario A seguirá ejecutando todas las demás acciones, hasta que se revoque el permiso `IAMAllowedPrincipal`.

Además, de forma predeterminada, el grupo `IAMAllowedPrincipal` tiene permisos sobre todas las bases de datos y tablas nuevas cuando se crean. Hay dos configuraciones que controlan este comportamiento. La primera es a nivel de cuenta y de región, que lo habilita para las bases de datos recién creadas, y la segunda es a nivel de base de datos. Para modificar la configuración predeterminada, consulte Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido.

Concesión de permisos

Los administradores del lago de datos pueden conceder permisos del catálogo de datos a las entidades principales para que estas puedan crear y gestionar bases de datos y tablas, y puedan acceder a los datos subyacentes.

Permisos a nivel de bases de datos y tablas

Cuando concede permisos dentro de Lake Formation, el concedente debe especificar la entidad principal a la que conceder permisos, los recursos sobre los que concederlos y las acciones que el concedente debe tener acceso para ejecutar. Para la mayoría de los recursos dentro de Lake Formation, la lista de entidades principales y los recursos para conceder permisos son similares, pero las acciones que puede ejecutar una entidad principal difieren en función del tipo de recurso. Por ejemplo, los permisos `SELECT` están disponibles en las tablas para leerlas, pero los permisos `SELECT` no están permitidos en las bases de datos. El `CREATE_TABLE` permiso está válido en las bases de datos, pero no en las tablas.

Puede conceder permisos de AWS Lake Formation siguiendo dos métodos:

- Método de recurso con nombre. Para elegir los nombres de las bases de datos y las tablas al conceder permisos a los usuarios.
- Control de acceso basado en etiquetas LF (LF-TBAC). Los usuarios crean etiquetas LF, las asocian a los recursos del catálogo de datos, conceden permisos `Describe` sobre las etiquetas LF, asocian permisos a usuarios individuales y escriben políticas de permisos LF con etiquetas LF

para distintos usuarios. Dichas políticas basadas en etiquetas LF se aplican a todos los recursos del catálogo de datos que estén asociados a esos valores de etiquetas LF.

> **Note**
>
> Las etiquetas LF son exclusivas de Lake Formation. Solo son visibles en Lake Formation y no deben confundirse con las etiquetas de recursos AWS.

LF-TBAC es una característica que permite a los usuarios agrupar recursos en categorías de etiquetas LF definidas por el usuario y aplicar permisos sobre esos grupos de recursos. Por lo tanto, es la mejor manera de escalar los permisos a través de un gran número de recursos del catálogo de datos.

Para obtener más información, consulte Control de acceso basado en etiquetas de Lake Formation.

Al conceder permisos a una entidad principal, Lake Formation evalúa los permisos como una unión de todas las políticas para ese usuario. Por ejemplo, si tiene dos políticas sobre una tabla para una entidad principal en la que una política concede permisos a las columnas col1, col2 y col3 según el método de recursos con nombre, y la otra política concede permisos a la misma tabla y entidad principal a col5 y col6 mediante etiquetas LF, los permisos efectivos serán una unión de los permisos, que serían col1, col2, col3, col5 y col6. Esto también incluye filas y filtros de datos.

Permisos de ubicación de datos

Los permisos de ubicación de datos ofrecen a los usuarios no administrativos la posibilidad de crear bases de datos y tablas en ubicaciones específicas de Amazon S3. Si un usuario intenta crear una base de datos o una tabla en una ubicación para la que no tiene permisos, la tarea de creación falla. De este modo, se evita que los usuarios creen tablas en ubicaciones arbitrarias dentro del lago de datos y se controla dónde pueden leer y escribir datos dichos usuarios. Existe un permiso implícito al crear tablas en la ubicación de Amazon S3 dentro de la base de datos en la que se está creando. Para obtener más información, consulte Conceder permisos de ubicación de datos.

Crear permisos para tablas y bases de datos

De forma predeterminada, los usuarios no administrativos no tienen permisos para crear bases de datos o tablas dentro de una base de datos. La creación de bases de datos se controla a nivel de cuenta mediante los ajustes de Lake Formation, de modo que solo las entidades principales

autorizadas pueden crear bases de datos. Para obtener más información, consulte Creación de una base de datos. Para crear una tabla, una entidad principal necesita permiso `CREATE_TABLE` en la base de datos donde se está creando la tabla. Para obtener más información, consulte Creación de tablas.

Permisos implícitos y explícitos

Lake Formation proporciona permisos implícitos en función de la persona y de las acciones que esta lleve a cabo. Por ejemplo, los administradores del lago de datos obtienen automáticamente permisos `DESCRIBE` para todos los recursos del catálogo de datos, permisos de ubicación de datos para todas las ubicaciones, permisos para crear bases de datos y tablas en todas las ubicaciones, así como y permisos `Grant` y `Revoke` sobre cualquier recurso. Los creadores de bases de datos obtienen automáticamente todos los permisos sobre las bases de datos que crean, y los creadores de tablas obtienen todos los permisos sobre las tablas que crean. Para obtener más información, consulte Permisos implícitos de Lake Formation.

Permisos concedibles

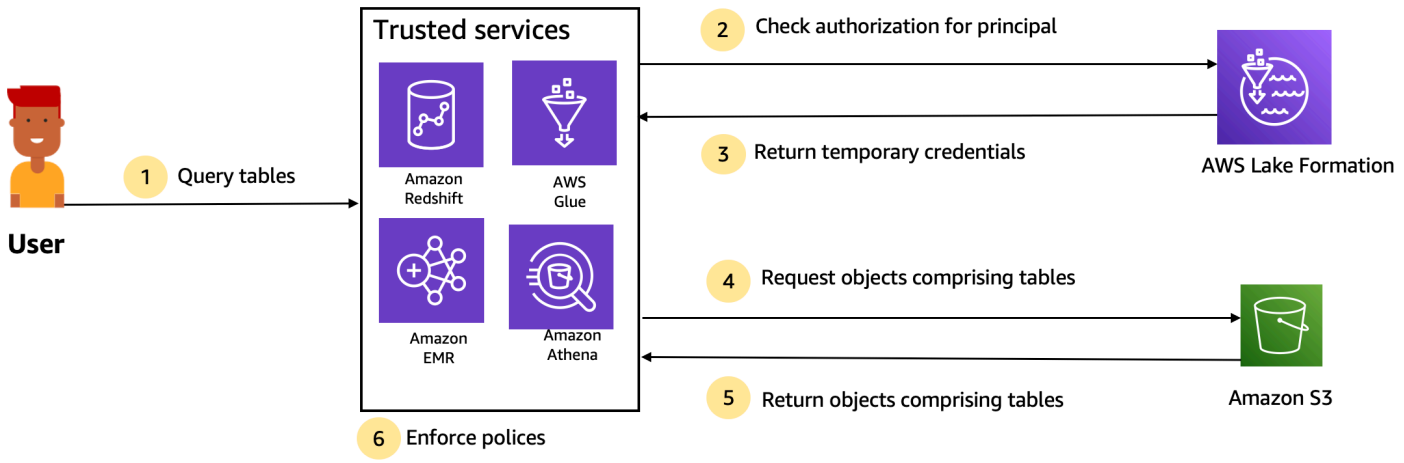
Los administradores del lago de datos tienen la posibilidad de delegar la gestión de los permisos a usuarios no administrativos proporcionándoles permisos concedibles. Cuando a una entidad principal se le proporcionan permisos concedibles sobre un recurso y un conjunto de permisos, esa entidad principal adquiere la capacidad de conceder permisos a otras entidades principales sobre ese recurso.

Administración del acceso al almacenamiento

Lake Formation utiliza la funcionalidad de expendición de credenciales para proporcionar acceso temporal a los datos de Amazon S3. La expedición de credenciales, o expedición de tokens, es un patrón común que proporciona credenciales temporales a usuarios, servicios o alguna otra entidad con el fin de conceder acceso a corto plazo a un recurso.

Lake Formation aprovecha este patrón para proporcionar acceso a corto plazo a servicios de análisis AWS como Athena para acceder a los datos en nombre de la entidad principal que llama. Al conceder los permisos, los usuarios no necesitan actualizar sus políticas de bucket de Amazon S3 ni sus políticas de IAM, y no necesitan tener acceso directo a Amazon S3.

El diagrama siguiente muestra cómo Lake Formation proporciona acceso temporal a las ubicaciones registradas:



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

1. Una entidad principal (usuario) introduce una consulta o solicitud de datos para una tabla a través de un servicio integrado de confianza como Athena, Amazon EMR, Redshift Spectrum o AWS Glue.
2. El servicio integrado comprueba la autorización de Lake Formation para la tabla y las columnas solicitadas y evalúa la autorización. Si el usuario no está autorizado, Lake Formation deniega el acceso a los datos y la consulta falla.
3. En cuanto la autorización tiene éxito y se activa la autorización de almacenamiento para la tabla y el usuario, el servicio integrado recupera las credenciales temporales de Lake Formation para acceder a los datos.
4. El servicio integrado utiliza las credenciales temporales de Lake Formation para solicitar objetos de Amazon S3.
5. Amazon S3 proporciona los objetos de Amazon S3 al servicio integrado. Los objetos de Amazon S3 contienen todos los datos de la tabla.
6. El servicio integrado efectúa la aplicación necesaria de las políticas de Lake Formation, como el filtrado a nivel de columnas, filas o celdas. El servicio integrado procesa las consultas y devuelve los resultados al usuario.

Habilitar la aplicación de permisos a nivel de almacenamiento para las tablas del catálogo de datos

De forma predeterminada, la aplicación a nivel de almacenamiento no está activada para las tablas del catálogo de datos. Para habilitar la aplicación a nivel de almacenamiento, debe registrar la ubicación de Amazon S3 de sus datos de origen con Lake Formation y proporcionar un rol de IAM.

Los permisos a nivel de almacenamiento se habilitarán para todas las tablas con la misma ruta de ubicación de la tabla o prefijo de la ubicación de Amazon S3.

Cuando un servicio integrado solicita acceso a la ubicación de los datos en nombre de un usuario, el servicio Lake Formation asume este papel y devuelve las credenciales al servicio solicitado con permisos de alcance reducido al recurso para que pueda producirse el acceso a los datos. El rol de IAM registrado debe tener todo el acceso necesario a la ubicación de Amazon S3, incluidas las claves AWS KMS.

Para obtener más información, consulte Registro de una ubicación de Amazon S3.

Servicios de AWS compatibles

AWS servicios analíticos como Athena, Redshift Spectrum, Amazon EMR, AWS Glue, Amazon QuickSight, y Amazon SageMaker se integran con AWS Lake Formation utilizando las operaciones API de expedición de credenciales de Lake Formation. Para ver una lista completa de los servicios de AWS que se integran con Lake Formation, así como el nivel de detalle y los formatos de tabla que compatibilizan, consulte Trabajar con otros AWS servicios.

Uso compartido de datos entre cuentas en Lake Formation

Con Lake Formation, puede compartir recursos del catálogo de datos (bases de datos y tablas) dentro de una cuenta AWS y entre cuentas en una sencilla configuración utilizando el método de recursos con nombre o etiquetas LF. Puede compartir una base de datos completa o seleccionar tablas de una base de datos con cualquier entidad principal de IAM (Roles de IAM y usuarios) de una cuenta, con otras cuentas AWS a nivel de cuenta o directamente con entidades principales de IAM de otra cuenta.

También puede compartir tablas del catálogo de datos con filtros de datos para restringir el acceso a los detalles a nivel de fila y a nivel de celda. Lake Formation utiliza AWS Resource Access Manager (AWS RAM) para facilitar la concesión de permisos entre cuentas. Cuando un recurso se comparte entre dos cuentas, AWS RAM envía invitaciones a la cuenta de destinatario. Cuando un usuario acepta una invitación para compartir de AWS RAM, AWS RAM proporciona los permisos necesarios a Lake Formation para disponer de los recursos del catálogo de datos, así como la habilitación de la aplicación del nivel de almacenamiento. Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation.

Cuando el administrador del lago de datos de la cuenta de destinatario acepta el recurso compartido AWS RAM, los recursos compartidos están disponibles en la cuenta de destinatario. El administrador

del lago de datos concede más permisos de Lake Formation sobre el recurso compartido a entidades principales de IAM adicionales en la cuenta de destinatario, si el administrador tiene permisos `GRANTABLE` sobre el recurso compartido.

Sin embargo, las entidades principales no pueden consultar los recursos compartidos utilizando Athena o Redshift Spectrum sin un enlace de recursos. Un enlace de recursos es una entidad del catálogo de datos y es similar al concepto de Linux-Symlink.

El administrador del lago de datos de la cuenta de destinatario crea un enlace de recurso en el recurso compartido. El administrador concede permisos de `Describe` sobre el enlace de recursos con los permisos requeridos sobre el recurso compartido original a los usuarios adicionales. A continuación, un usuario de la cuenta de destinatario puede utilizar el enlace de recursos para consultar el recurso compartido utilizando Athena y Redshift Spectrum. Para obtener más información sobre los enlaces de recursos, consulte Creación de enlaces de recursos.

Componentes de Lake Formation

AWS Lake Formation se apoya en la interacción de varios componentes para crear y administrar su lago de datos.

Consola de Lake Formation

La consola de Lake Formation se utiliza para definir y gestionar el lago de datos y conceder y revocar los permisos de Lake Formation. Puede usar los esquemas de la consola para descubrir, limpiar, transformar e ingerir datos. También puede habilitar o deshabilitar el acceso a la consola para usuarios individuales de Lake Formation.

API e interfaz de línea de comandos de Lake Formation

Lake Formation proporciona operaciones de API a través de varios SDK específicos de lenguaje y la AWS Command Line Interface (AWS CLI). La API de Lake Formation funciona en combinación con la API de AWS Glue. La API de Lake Formation se centra principalmente en la gestión de los permisos de Lake Formation, mientras que la API de AWS Glue proporciona una API de catálogo de datos y una infraestructura administrada para definir, programar y ejecutar operaciones de ETL en sus datos.

Para obtener más información acerca de la API de AWS Glue, consulte la Guía para desarrolladores de AWS Glue. Para obtener más información sobre el uso de la AWS CLI, consulte Referencia de comandos de la AWS CLI.

Otros servicios de AWS

Lake Formation utiliza los siguientes servicios:

- AWS Glue para orquestar las tareas y los rastreadores en la transformación de datos mediante las transformaciones de AWS Glue.
- IAM para conceder políticas de permisos a las entidades principales de Lake Formation. El modelo de permisos de Lake Formation amplía el modelo de permisos de IAM para proteger su lago de datos.

Terminología de Lake Formation

A continuación se indican algunos términos importantes que encontrará en esta guía.

Lago de datos

El lago de datos son sus datos persistentes almacenados en Amazon S3 y administrados por Lake Formation mediante un catálogo de datos. En general, un lago de datos almacena lo siguiente:

- Datos estructurados y no estructurados
- Datos sin procesar y datos transformados

Para que una ruta de Amazon S3 esté dentro de un lago de datos, debe estar registrada en Lake Formation.

Acceso a los datos

Lake Formation proporciona un acceso seguro y específico a los datos a través de un nuevo modelo de concesión/revocación de permisos que aumenta las políticas de AWS Identity and Access Management (IAM).

Los analistas y científicos de datos pueden utilizar toda la cartera de servicios de análisis y machine learning de AWS, como Amazon Athena, para acceder a los datos. Las políticas de seguridad configuradas de Lake Formation ayudan a garantizar que los usuarios solo puedan acceder a los datos para los que están autorizados.

Modo de acceso híbrido

Gracias al modo de acceso Hyrbid, podrá proteger y acceder a los datos catalogados utilizando tanto los permisos de Lake Formation como los de IAM y Amazon S3. El modo de acceso híbrido permite a los administradores de datos incorporar los permisos de Lake Formation de forma selectiva e incremental, centrándose en un caso práctico de lago de datos cada vez.

Esquema

Un esquema es una plantilla de administración de datos que permite incorporar datos fácilmente a un lago de datos. Lake Formation proporciona varios esquemas, cada uno para un tipo de fuente predefinido, como una base de datos relacional o registros de AWS CloudTrail. A partir de un esquema, puede crear un flujo de trabajo. Los flujos de trabajo constan de rastreadores de AWS Glue, trabajos y desencadenadores que se generan para orquestar la carga y actualización de datos. Los esquemas toman como entrada el origen de datos, el destino de estos y la programación para configurar el flujo de trabajo.

Flujo de trabajo

Un flujo de trabajo es un contenedor para un conjunto de trabajos de AWS Glue, rastreadores y desencadenantes relacionados. El flujo de trabajo se crea en Lake Formation y se ejecuta en el servicio AWS Glue. Lake Formation puede seguir el estado de un flujo de trabajo como una entidad única.

Cuando define un flujo de trabajo, selecciona el esquema en el que se basa. A continuación, puede ejecutar flujos de trabajo a petición o según un calendario.

Los flujos de trabajo que cree en Lake Formation son visibles en la consola AWS Glue como un gráfico acíclico dirigido (DAG). Utilizando el DAG, puede seguir el progreso del flujo de trabajo y solucionar problemas.

Catálogo de datos

El catálogo de datos es su almacén persistente de metadatos. Se trata de un servicio administrado para almacenar, anotar y compartir metadatos en la nube de AWS del mismo modo que lo haría en un metaalmacén de Apache Hive. Proporciona un repositorio uniforme donde los sistemas dispares pueden almacenar y encontrar metadatos para rastrear los datos en silos de datos, y luego utilizar esos metadatos para consultar y transformar los datos. Lake Formation utiliza el catálogo de datos

AWS Glue para almacenar metadatos sobre lagos de datos, orígenes de datos, transformaciones y objetivos.

Los metadatos sobre orígenes de datos y objetivos se presentan en forma de bases de datos y tablas. Las tablas almacenan información sobre el esquema, la ubicación, etc. Las bases de datos son colecciones de tablas. Lake Formation proporciona una jerarquía de permisos para controlar el acceso a las bases de datos y tablas del catálogo de datos.

Cada cuenta AWS dispone de un catálogo de datos por región de AWS.

Datos subyacentes

Los datos subyacentes se refieren a los datos de origen o datos dentro de los lagos de datos a los que apuntan las tablas del catálogo de datos.

Entidad principal

Una entidad principal es un usuario o rol de AWS Identity and Access Management (IAM) o un usuario de Active Directory.

Administrador de lago de datos

Un administrador de un lago de datos es una entidad principal que puede conceder a cualquier entidad principal (incluida la propia) permisos sobre cualquier recurso o ubicación de datos del catálogo de datos. Designe a un administrador del lago de datos como primer usuario del catálogo de datos. Este usuario puede entonces conceder permisos más específicos de recursos a otras entidades principales.

> **Note**
>
> Los usuarios administrativos de IAM —usuarios con la política administrada `AdministratorAccess` de AWS— no son automáticamente administradores del lago de datos. Por ejemplo, no pueden conceder permisos de Lake Formation sobre objetos del catálogo a menos que se les hayan concedido permisos para hacerlo. Sin embargo, pueden utilizar la consola de Lake Formation o la API para designarse como administradores del lago de datos.

Para obtener información sobre las capacidades de un administrador de lago de datos, consulte Permisos implícitos de Lake Formation. Para obtener información sobre la designación de un usuario como administrador del lago de datos, consulte Crear un administrador de lago de datos.

Integraciones de servicios de AWS con Lake Formation

Puede utilizar Lake Formation para administrar los permisos de acceso en el nivel de columnas, bases de datos y tablas de los datos almacenados en Amazon S3. Después de que sus datos se registren en Lake Formation, puede utilizar servicios de análisis AWS como AWS Glue, Amazon Athena, Amazon Redshift Spectrum o Amazon EMR para consultarlos. Los siguientes servicios AWS Lake Formation se integran con AWS y respetan los permisos de Lake Formation.

Servicio de AWS	Detalles de la integración
AWS Glue	Tema de referencia: Utilizándolo con AWS Lake FormationAWS Glue AWS Glue y Lake Formation comparten el mismo Catálogo de datos. Para las operaciones de la consola (como ver una lista de tablas) y todas las operaciones de la API, los usuarios de AWS Glue solo pueden acceder a las bases de datos y tablas en las que tienen permisos de Lake Formation.
Amazon Athena	Tema de referencia: Uso AWS Lake Formation con Amazon Athena Utilice Lake Formation para conceder o denegar permisos para leer datos en Amazon S3. Cuando los usuarios Amazon Athena seleccion an el catálogo AWS Glue en el editor de consultas, solo pueden consultar las bases de datos, tablas y columnas en las que tienen permisos de Lake Formation. No se admiten las consultas mediante el uso de manifiestos. Actualmente, Lake Formation no admite la administración de permisos en operaciones de escritura como `VACUUM`, `MERGE`, `UPDATE` y `OPTIMIZE` en tablas en formatos de tabla abierta. Además de las entidades principales que se autentican con Athena a través de AWS Identity and Access Management (IAM), Lake Formation es compatible con los usuarios de Athena que se

Servicio de AWS	Detalles de la integración
	conectan a través del controlador JDBC u ODBC y se autentican a través de SAML. Entre los proveedores SAML compatibles se encuentran Okta y Microsoft Active Directory Federation Service (AD FS).
Amazon Redshift Spectrum	Tema de referencia: Uso AWS Lake Formation con Amazon Redshift Spectrum Cuando los usuarios de Amazon Redshift crean un esquema externo en una base de datos en el AWS Glue Data Catalog, pueden consultar solo las tablas y columnas de ese esquema sobre las que tienen permisos de Lake Formation.
Edición Amazon QuickSight Enterprise	Referencia: Uso AWS Lake Formation con Amazon QuickSight Cuando un usuario de Amazon QuickSight Enterprise Edition consulta un conjunto de datos en una ubicación de Amazon S3, el usuario debe tener el `SELECT` permiso de Lake Formation sobre los datos.
Amazon EMR	Referencia: Uso AWS Lake Formation con Amazon EMR Puede integrar los permisos de Lake Formation al crear un clúster de Amazon EMR con un rol de tiempo de ejecución. Un rol en tiempo de ejecución es un rol de IAM que se asocia a los trabajos o consultas de Amazon EMR, y luego Amazon EMR utiliza este rol para acceder a los recursos.

Lake Formation también colabora con AWS Key Management Service (AWS KMS) para permitirle configurar más fácilmente otros servicios integrados para cifrar y descifrar datos en ubicaciones de Amazon Simple Storage Service (Amazon S3).

Recursos adicionales de Lake Formation

Temas

Blogs

- Resumen del año 2022 de AWS Lake Formation
- Arquitectura de datos moderna multirregional altamente resiliente
- Uso compartido entre cuentas mediante etiquetas LF para dirigir a las entidades principales de IAM
- Panel de inventario de permisos de Lake Formation
- Malla de datos impulsada por eventos

Charlas técnicas y seminarios web

- re:Invent 2020: Lagos de datos: construya, proteja y comparta fácilmente con AWS Lake Formation
- re:Invent 2022: Creación y funcionamiento de un lago de datos en Amazon S3
- AWS Summit SF 2022: Comprender y conseguir una arquitectura de datos moderna
- AWS Summit ATL 2022: Lagos de datos modernos con AWS Lake Formation, Amazon Redshift y AWS Glue
- AWS Summit ANZ 2022: Lagos de datos, lake houses y malla de datos: ¿qué, por qué y cómo?
- AWS Charlas técnicas en línea: Simplificación de los permisos y la gobernanza del lago de datos

Arquitectura moderna

- Patrones de arquitectura modernos

Recursos de la malla de datos

- Crear una arquitectura de datos moderna y un patrón de malla de datos a escala mediante el control de acceso a AWS Lake Formation basado en etiquetas

- Cómo JPMorgan Chase creó una arquitectura de malla de datos para generar un valor significativo y mejorar su plataforma de datos empresariales
- Crear una malla de datos en AWS

Guías de prácticas recomendadas

- Guías de prácticas recomendadas de AWS Lake Formation

Introducción a Lake Formation

Le recomendamos que lea las siguientes secciones:

- AWS Lake Formation: Cómo funciona. Conozca la terminología esencial y cómo interactúan los distintos componentes.
- Introducción a Lake Formation. Obtenga información sobre los requisitos previos y complete las tareas de configuración importantes.
- Tutoriales— Sigue step-by-step los tutoriales para aprender a usar Lake Formation.
- Seguridad en AWS Lake Formation. Comprenda cómo puede ayudar a proteger el acceso a los datos en Lake Formation.

Introducción a Lake Formation

Si aún no está registrado en AWS o necesita ayuda para comenzar, asegúrese de realizar las tareas siguientes.

Temas

Completar las tareas iniciales de configuración de AWS

Para utilizar AWS Lake Formation, primero debe completar las siguientes tareas:

Temas

Registro para obtener una Cuenta de AWS

Si no dispone de una Cuenta de AWS, siga estos pasos para crear una.

Creación de una Cuenta de AWS

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

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

   Al registrarse en una Cuenta de AWS, se crea un Usuario raíz de la cuenta de AWS. El usuario raíz tendrá acceso a todos los Servicios de AWS y recursos de esa cuenta. Como práctica recomendada de seguridad, asigne acceso administrativo a un usuario administrativo y utilice únicamente el usuario raíz para realizar tareas que requieran acceso de usuario raíz.

AWS le enviará un correo electrónico de confirmación después de completar el proceso de registro. Puede ver la actividad de la cuenta y administrar la cuenta en cualquier momento entrando en https://aws.amazon.com/ y seleccionando Mi cuenta.

Crear un usuario administrativo

Después de registrarse para obtener una Cuenta de AWS, proteja su Usuario raíz de la cuenta de AWS, habilite AWS IAM Identity Center y cree un usuario administrativo para no utilizar el usuario raíz en las tareas cotidianas.

Protección de su Usuario raíz de la cuenta de AWS

1. Inicie sesión en AWS Management Console como propietario de cuenta eligiendo Usuario raíz e introduzca el correo electrónico de su Cuenta de AWS. En la siguiente página, escriba su contraseña.

   Para obtener ayuda para iniciar sesión con el usuario raíz, consulte Iniciar sesión como usuario raíz en la Guía del usuario de AWS Sign-In.
2. Active la autenticación multifactor (MFA) para el usuario raíz.

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

Creación de un usuario administrativo

1. Activar IAM Identity Center

   Para obtener instrucciones, consulte Activación de AWS IAM Identity Center en la Guía del usuario de AWS IAM Identity Center.
2. En el Centro de identidades de IAM, conceda acceso administrativo a un usuario administrativo.

   Para ver un tutorial sobre cómo utilizar Directorio de IAM Identity Center como origen de identidad, consulte Configuración del acceso de los usuarios con el Directorio de IAM Identity Center predeterminado en la Guía del usuario de AWS IAM Identity Center.

Cómo iniciar sesión como usuario administrativo

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

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

Conceder acceso programático

Los usuarios necesitan acceso programático si desean interactuar con AWS fuera de la AWS Management Console. La forma de conceder el acceso programático depende del tipo de usuario que acceda a AWS.

Para conceder acceso programático a los usuarios, seleccione una de las siguientes opciones.

¿Qué usuario necesita acceso programático?	Para	Por
Identidad del personal (Usuarios administrados en el Centro de identidades de IAM)	Utilice credenciales temporales para firmar las solicitudes programáticas a la AWS CLI, los SDK de AWS o las API de AWS.	Siga las instrucciones de la interfaz que desea utilizar: • Para la AWS CLI, consulte Configuración del AWS CLI para su usoAWS IAM Identity Center en la Guía del usuario de AWS Command Line Interface. • Para ver los SDK de AWS, las herramientas y las API de AWS, consulte la Autenticación del Centro de identidades de IAM en la Guía de referencia y herramientas de SDK de AWS.

¿Qué usuario necesita acceso programático?	Para	Por
IAM	Utilice credenciales temporales para firmar las solicitudes programáticas a la AWS CLI, los SDK de AWS o las API de AWS.	Siga las instrucciones que aparecen en Usar credenciales temporales con recursos de AWS de la Guía del usuario de IAM.
IAM	(No recomendado) Utilice credenciales a largo plazo para firmar las solicitudes programáticas a la AWS CLI, los SDK de AWS o las API AWS.	Siga las instrucciones de la interfaz que desea utilizar: • Para la AWS CLI, consulte Autenticación mediante credenciales de usuario de IAM en la Guía del usuario de AWS Command Line Interface. • Para ver los SDK y las herramientas de AWS, consulte Autenticar mediante credenciales a largo plazo en la Guía de referencia de SDK y herramientas de AWS. • Para las API de AWS, consulte Administrar claves de acceso para usuarios de IAM en la Guía del usuario de IAM.

Configuración de AWS Lake Formation

Las siguientes secciones proporcionan información sobre cómo configurar Lake Formation por primera vez. No todos los temas de esta sección son necesarios para empezar a utilizar Lake Formation. Puede utilizar las instrucciones para configurar el modelo de permisos de Lake Formation

para administrar sus objetos AWS Glue Data Catalog y ubicaciones de datos existentes en Amazon Simple Storage Service (Amazon S3).

1. Crear un administrador de lago de datos
2. Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido
3. the section called “Configurar una ubicación de Amazon S3 para el lago de datos”
4. the section called “Asignar permisos a usuarios de Lake Formation”
5. the section called “Integración de IAM Identity Center”
6. the section called “(Opcional) Configuración de filtrado de datos externo”
7. the section called “(Opcional) Conceda acceso a la clave de cifrado del Catálogo de datos”
8. (Opcional) Cree un rol de IAM para los flujos de trabajo

Esta sección muestra cómo configurar los recursos de Lake Formation de dos maneras diferentes:

- Uso de una plantilla AWS CloudFormation
- Uso de la consola de Lake Formation

Para configurar Lake Formation mediante la consola de AWS, vaya a Crear un administrador de lago de datos.

Configurar los recursos de Lake Formation usando una plantilla de AWS CloudFormation

> **Note**
>
> La AWS CloudFormation pila realiza los pasos 1 a 6 de los anteriores, excepto los pasos 2 y 5. Realice Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido y the section called “Integración de IAM Identity Center” manualmente desde la consola de Lake Formation.

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation como administrador de IAM en la región Este de EE. UU. (Norte de Virginia).
2. Seleccione Lanzar pila.
3. Seleccione Siguiente en la pantalla Crear pila.

4. Introduzca un Nombre de pila.
5. Para DatalakeAdminNamey DatalakeAdminPassword, introduzca su nombre de usuario y contraseña como usuario administrador de Data Lake.
6. Para DatalakeUser1Name y DatalakeUser1Password, introduce tu nombre de usuario y contraseña para el usuario de data lake analyst.
7. Para ello DataLakeBucketName, introduce el nombre del nuevo depósito que se va a crear.
8. Elija Siguiente.
9. En la página siguiente, seleccione Siguiente.
10. Revise los detalles en la última página y acepte que AWS CloudFormation pueda crear recursos de IAM.
11. Seleccione Crear.

    La creación de la pila puede llevar hasta dos minutos.

Eliminar recursos

Si quiere limpiar los recursos de la pila de AWS CloudFormation:

1. Anule el registro del bucket de Amazon S3 creado por su pila, registrado como ubicación del lago de datos.
2. Elimine la pila de AWS CloudFormation. Esto eliminará todos los recursos creados por la pila.

Crear un administrador de lago de datos

Los administradores del lago de datos son inicialmente los únicos usuarios o roles de AWS Identity and Access Management (IAM) que pueden conceder permisos de Lake Formation sobre ubicaciones de datos y recursos del Catálogo de datos a cualquier entidad principal (incluida la propia). Para obtener más información acerca de las capacidades de administrador de lago de datos, consulte Permisos implícitos de Lake Formation. De forma predeterminada, con Lake Formation puede crear hasta 30 administradores de lago de datos.

Puede crear un administrador de lagos de datos mediante la consola de Lake Formation o la operación `PutDataLakeSettings` de la API de Lake Formation.

Para crear un administrador de lago de datos, se requieren los permisos siguientes. El usuario `Administrator` tiene estos permisos de forma implícita.

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

Si concede a un usuario la política `AWSLakeFormationDataAdmin`, dicho usuario no podrá crear usuarios administradores adicionales de Lake Formation.

Para crear un administrador de lago de datos (consola)

1. Si el usuario que va a ser administrador del lago de datos aún no existe, utilice la consola de IAM para crearlo. De lo contrario, elija un usuario existente para ser administrador del lago de datos.

   > **Note**
   >
   > Le recomendamos que no seleccione un usuario administrativo de IAM (usuario con la política administrada `AdministratorAccess` de AWS) como administrador del lago de datos.

   Vincule las políticas administradas de AWS siguientes al usuario:

Políticas	¿Obligatoria?	Notas
`AWSLakeFormationDataAdmin`	Obligatorio	Permisos básicos de administrador del lago de datos. Esta política administrada de AWS contiene una denegación explícita de la operación de la API de Lake Formation, `PutDataLakeSetting` que impide a los usuarios crear nuevos administradores del lago de datos.
`AWSGlueConsoleFullAccess` , `CloudWatchLogsReadOnlyAccess`	Opcional	Vincule estas políticas si el administrador del lago de datos va a solucionar los flujos de trabajo creados a partir de esquemas de Lake Formation. Estas políticas permiten al administrador del

Políticas	¿Obligatoria?	Notas
		lago de datos consultar la informaci ón de solución de problemas en la consola de AWS Glue y en la consola de Amazon CloudWatch Logs. Para obtener más información acerca de los flujos de trabajo, consulte the section called “Importación de datos mediante flujos de trabajo”.
`AWSLakeFormationCrossAccountManager`	Opcional	Vincule esta política para que el administrador del lago de datos pueda conceder y revocar permisos entre cuentas sobre los recursos del Catálogo de datos. Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation.
`AmazonAthenaFullAccess`	Opcional	Adjunte esta política si el administr ador del lago de datos va a ejecutar consultas en Amazon Athena.

2. Adjunte la siguiente política en línea, que otorga al administrador del lago de datos permiso para crear el rol vinculado al servicio de Lake Formation. Un posible nombre para la política es `LakeFormationSLR`.

   El rol vinculado al servicio permite al administrador del lago de datos registrar más fácilmente las ubicaciones de Amazon S3 con Lake Formation. Para más información sobre el rol vinculado al servicio Lake Formation, consulte the section called “Uso de roles vinculados a servicios”.

   > **Important**
   >
   > En toda la política siguiente, sustituya *<account-id>* por un número de cuenta de AWS válido.

```
{
```

```
    "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. (Opcional) Vincule la siguiente política `PassRole` insertada al usuario. Esta política permite al administrador del lago de datos crear y ejecutar flujos de trabajo. El permiso `iam:PassRole` capacita al flujo de trabajo asumir el rol `LakeFormationWorkflowRole` para crear rastreadores y trabajos, y vincular el rol a los rastreadores y trabajos creados. Un posible nombre para la política es `UserPassRole`.

> **Important**
>
> Sustituya *<account-id>* por un número de cuenta AWS válido.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "PassRolePermissions",
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
```

```
            ],
            "Resource": [
                "arn:aws:iam::<account-id>:role/LakeFormationWorkflowRole"
            ]
        }
    ]
}
```

4. (Opcional) Adjunte esta política adicional insertada si su cuenta va a conceder o recibir permisos entre cuentas de Lake Formation. Esta política permite al administrador del lago de datos ver y aceptar invitaciones para compartir recursos de AWS Resource Access Manager (AWS RAM). Además, para los administradores de lago de datos de la cuenta de administración de AWS Organizations, la política incluye un permiso para permitir las concesiones entre cuentas a las organizaciones. Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation.

   Un posible nombre para la política es `RAMAccess`.

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

5. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/ e inicie sesión como el usuario administrador que creó en Crear un usuario administrativo o como usuario con una política administrada de usuario `AdministratorAccess` de AWS.
6. Si aparece la ventana de bienvenida a Lake Formation, elija el usuario de IAM que creó o seleccionó en el Paso 1 y, a continuación, elija Comenzar.
7. Si no aparece la ventana de bienvenida a Lake Formation, siga estos pasos para configurar un administrador de Lake Formation.

a. En el panel de navegación, en Administradores, seleccione Roles y tareas administrativas. En la sección de administradores del lago de datos de la página de la consola, seleccione Agregar.

b. En el cuadro de diálogo Agregar administradores, como tipo de acceso seleccione Administrador del lago de datos.

c. Para los usuarios y roles de IAM, elija el usuario de IAM que creó o seleccionó en el Paso 1 y, a continuación, seleccione Guardar.

Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido

Lake Formation comienza con la configuración «Usar solo el control de acceso de IAM» habilitada para garantizar la compatibilidad con el comportamiento de AWS Glue Data Catalog existente. Este ajuste le permite administrar el acceso a sus datos en el lago de datos y sus metadatos mediante políticas de IAM y políticas de bucket de Amazon S3.

Para facilitar la transición de los permisos del lago de datos de un modelo IAM y Amazon S3 a los permisos de Lake Formation, le recomendamos que utilice el modo de acceso híbrido para el Catálogo de datos. Con el modo de acceso híbrido, dispone de una vía incremental en la que puede habilitar los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir a otros usuarios o cargas de trabajo existentes.

Para obtener más información, consulte Modo de acceso híbrido.

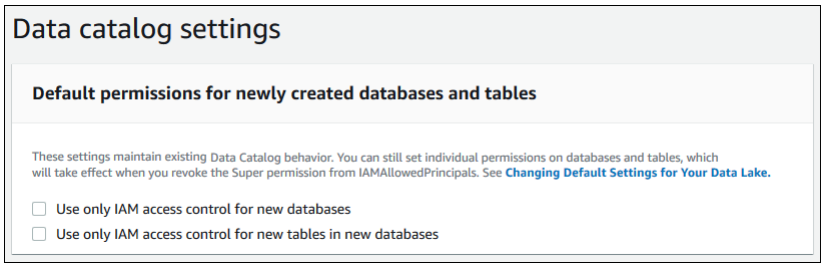
Desactive la configuración predeterminada para mover a todos los usuarios existentes de una tabla a Lake Formation en un solo paso.

> **Important**
>
> Si ya hay bases de datos y tablas de AWS Glue Data Catalog, no siga las instrucciones indicadas en esta sección. En su lugar, siga las instrucciones en the section called “Actualización de los permisos de datos AWS Glue al modelo de Lake Formation”.

> **Warning**
>
> Si cuenta con una automatización que crea bases de datos y tablas en el Catálogo de datos, los pasos siguientes podrían provocar un error en los trabajos de automatización y extracción, transformación y carga (ETL) posteriores. Continúe solo después de haber modificado sus procesos existentes o de haber concedido permisos explícitos de Lake Formation a las entidades principales requeridas. Para obtener información sobre los permisos de Lake Formation, consulte the section called “Referencia de permisos de Lake Formation”.

Para cambiar la configuración predeterminada del Catálogo de datos

1. Continúe en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. Verifique que ha iniciado sesión como el usuario administrador que había creado en Crear un usuario administrativo o como un usuario con la política administrada `AdministratorAccess` de AWS.
2. Modifique la configuración del Catálogo de datos:
   a. En el panel de navegación, Administración, seleccione Configuración del Catálogo de datos.
   b. Desmarque ambas casillas y seleccione Guardar.



3. Revoque el permiso de `IAMAllowedPrincipals` a los creadores de bases de datos.
   a. En el panel de navegación, en Administración, seleccione Roles y tareas administrativas.
   b. En la página de la consola de Roles y tareas administrativas, en la sección Creadores de bases de datos, seleccione el grupo `IAMAllowedPrincipals` y elija Revocar.

      Aparece el cuadro de diálogo para revocar permisos, en el que se muestra que `IAMAllowedPrincipals` tiene el permiso para crear base de datos.
   c. Seleccione Revocar.

Asignar permisos a usuarios de Lake Formation

Cree un usuario para tener acceso al lago de datos en AWS Lake Formation. Este usuario tiene el privilegio mínimo de permisos para consultar el lago de datos.

Para más información sobre la creación de usuarios o grupos, consulte Identidades IAM en la Guía del usuario de IAM.

Para adjuntar permisos a un usuario no administrador con el fin de acceder a los datos de Lake Formation

1. Abra la consola de IAM en https://console.aws.amazon.com/iam e inicie sesión como un usuario administrador creado en Crear un usuario administrativo o como usuario con la política administrada `AdministratorAccess` de AWS.
2. Elija Usuarios o Grupos de usuarios.
3. En la lista, seleccione el nombre del usuario o del grupo en el que integrará una política.

   Elija Permisos.
4. Elija Agregar permisos y seleccione Adjuntar políticas directamente. Introduzca `Athena` en el campo de texto Filtrar políticas. En la lista de resultados, marque la casilla de `AmazonAthenaFullAccess`.
5. Pulse el botón Crear política. En la página Crear política, elija la pestaña JSON. Copie la política siguiente y péguela en el editor de políticas.

```
{
    "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. Seleccione el botón Siguiente en la parte inferior hasta que vea la página Revisar la política. Escriba un nombre para la política, por ejemplo `DatalakeUserBasic`. Seleccione Crear política y, a continuación, cierre la pestaña Políticas o la ventana del navegador.

Configurar una ubicación de Amazon S3 para el lago de datos

Para utilizar Lake Formation con el fin de administrar y proteger los datos de su lago de datos, primero debe registrar una ubicación de Amazon S3. Al registrar una ubicación, se registran esa ruta de Amazon S3 y todas las carpetas incluidas en esa ruta, lo que permite a Lake Formation aplicar los permisos de nivel de almacenamiento. Cuando el usuario solicita datos de un motor integrado como Amazon Athena, Lake Formation proporciona acceso a los datos en lugar de utilizar los permisos del usuario.

Cuando registra una ubicación, especifica un rol de IAM que concede permisos de lectura y escritura en esa ubicación. Lake Formation asume ese rol al proporcionar credenciales temporales a los servicios de AWS integrados que solicitan acceso a los datos en la ubicación registrada de Amazon S3. Puede especificar el rol vinculado al servicio (SLR) de Lake Formation o crear el suyo propio.

Utilice un rol personalizado en las siguientes situaciones:

- Planeas publicar métricas en Amazon CloudWatch Logs. La función definida por el usuario debe incluir una política para añadir registros a los CloudWatch registros y publicar métricas, además de los permisos de SLR. Para ver un ejemplo de política en línea que concede los CloudWatch permisos necesarios, consulte. Requisitos de los roles utilizados para registrar ubicaciones
- La ubicación de Amazon S3 está en una cuenta diferente. Para obtener más detalles, consulte the section called “Registrar una ubicación de Amazon S3 en otra cuenta AWS”.
- La ubicación de Amazon S3 contiene datos cifrados con una Clave administrada de AWS. Para más detalles, consulte Registro de una ubicación cifrada de Amazon S3 y Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS.

- Tiene previsto acceder a la ubicación de Amazon S3 mediante Amazon EMR. Para obtener más información sobre los requisitos de los roles, consulte Roles de IAM para Lake Formation en la Guía de administración de Amazon EMR.

El rol que elija debe tener los permisos necesarios, tal y como se describe en Requisitos de los roles utilizados para registrar ubicaciones. Para obtener instrucciones sobre cómo registrar una ubicación de Amazon S3, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

(Opcional) Configuración de filtrado de datos externo

Si tiene intención de analizar y procesar los datos de su lago de datos utilizando motores de consulta de terceros, debe optar por permitir que los motores externos accedan a los datos gestionados por Lake Formation. Si no lo hace, los motores externos no podrán acceder a los datos de las ubicaciones de Amazon S3 que estén registradas en Lake Formation.

Lake Formation es compatible con los permisos a nivel de columna para restringir el acceso a columnas específicas de una tabla. Los servicios analíticos integrados como Amazon Athena, Amazon Redshift Spectrum y Amazon EMR recuperan metadatos de tablas no filtrados del AWS Glue Data Catalog. El filtrado real de las columnas en las respuestas a las consultas es responsabilidad del servicio integrado. Es responsabilidad de los administradores externos gestionar adecuadamente los permisos para evitar el acceso no autorizado a los datos.

Para permitir que motores de terceros accedan a los datos y los filtren (consola)

1. Continúe en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. Compruebe que ha iniciado sesión como una entidad principal con el permiso de IAM en la operación de la API `PutDataLakeSettings` de Lake Formation. El usuario administrador de IAM que ha creado en Registro para obtener una Cuenta de AWS tiene este permiso.
2. En el panel de navegación, Administración, seleccione Configuración de integración de aplicaciones.
3. En la página Configuración de integración de aplicaciones, haga lo siguiente:

    a. Marque la casilla Permitir que motores externos filtren datos en las ubicaciones de Amazon S3 registradas en Lake Formation.

    b. Introduzca los valores de las etiquetas de sesión definidos para motores de terceros.

c. Para los ID de cuenta de AWS, introduzca los ID de cuenta desde los que los motores de terceros pueden acceder a las ubicaciones registradas en Lake Formation. Pulse Intro después de cada ID de cuenta.

d. Seleccione Guardar.

Para permitir que los motores externos accedan a los datos sin validar las etiquetas de sesión, consulte Integración de aplicaciones para un acceso completo a la tabla

(Opcional) Conceda acceso a la clave de cifrado del Catálogo de datos

Si el AWS Glue Data Catalog está cifrado, conceda permisos AWS Identity and Access Management de (IAM) sobre la clave AWS KMS a las entidades principales que necesiten conceder permisos de Lake Formation en las bases de datos y tablas del Catálogo de datos.

Para obtener más información, consulte la Guía para desarrolladores de AWS Key Management Service.

(Opcional) Cree un rol de IAM para los flujos de trabajo

Con AWS Lake Formation, puede importar sus datos utilizando flujos de trabajo ejecutados por rastreadores de AWS Glue. Un flujo de trabajo define el origen de datos y la programación para importar los datos a su lago de datos. Puede definir fácilmente los flujos de trabajo mediante los esquemas o las plantillas proporcionados por Lake Formation.

Al crear un flujo de trabajo, debe asignarle un rol de AWS Identity and Access Management (IAM) que conceda a Lake Formation los permisos necesarios para ingerir los datos.

En el procedimiento siguiente se presupone que está familiarizado con IAM.

Crear un rol de IAM para flujos de trabajo

1. Abra la consola de IAM en https://console.aws.amazon.com/iam e inicie sesión como el usuario administrador que creó en Crear un usuario administrativo o como usuario con la política administrada `AdministratorAccess` de AWS.
2. En el panel de navegación, seleccione Roles, Crear rol.
3. En la página Crear rol, elija el servicio AWS y, a continuación, Glue. Elija Siguiente.
4. En la página Añadir permisos, busque la política AWSGlueServiceRolegestionada y active la casilla de verificación situada junto al nombre de la política en la lista. A continuación, complete

el asistente de creación de roles, asignando al rol el nombre `LakeFormationWorkflowRole`. Para terminar, seleccione Crear rol.

5. De regreso en la página Roles, busque `LakeFormationWorkflowRole` y elija el nombre del rol.
6. En la página Resumen del rol, en la pestaña Permisos, elija Crear política insertada. En la pantalla Crear política, vaya a la pestaña JSON y añada la siguiente política insertada. Un posible nombre para la política es `LakeFormationWorkflow`.

> **Important**
>
> En la siguiente política, sustituya *<account-id>* por un número de Cuenta de AWS válido.

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

A continuación se describen brevemente los permisos de esta política:

- `lakeformation:GetDataAccess` permite que los trabajos creados por el flujo de trabajo escriban en la ubicación de destino.

- `lakeformation:GrantPermissions` permite que el flujo de trabajo conceda el permiso de `SELECT` sobre las tablas de destino.
- `iam:PassRole` permite al servicio asumir el rol de `LakeFormationWorkflowRole` para crear rastreadores y trabajos (instancias de flujos de trabajo) y asociar el rol a los rastreadores y trabajos creados.

7. Compruebe que el rol `LakeFormationWorkflowRole` tenga dos políticas asociadas.
8. Si va a ingerir datos de fuera de la ubicación del lago de datos, añada una política insertada que conceda permisos para leer los datos de origen.

Actualización de los permisos de datos AWS Glue al modelo AWS Lake Formation

Los permisos AWS Lake Formation permiten un control de acceso específico para los datos de su lago de datos. Puede utilizar el modelo de permisos de Lake Formation para administrar sus objetos AWS Glue Data Catalog y ubicaciones de datos existentes en Amazon Simple Storage Service (Amazon S3).

El modelo de permisos de Lake Formation utiliza permisos básicos de AWS Identity and Access Management (IAM) para el acceso al servicio de API. Restringe los datos a los que sus usuarios y esos servicios pueden acceder a través de la funcionalidad de Lake Formation. En comparación, el modelo AWS Glue otorga el acceso a los datos mediante permisos de IAM de control de acceso específico. Para hacer el cambio, siga los pasos de esta guía.

Para obtener más información, consulte Descripción general de los permisos de Lake Formation .

Temas

- Paso 7: Limpiar las políticas de IAM existentes

Acerca de la actualización al modelo de permisos de Lake Formation

Para mantener la compatibilidad con AWS Glue, de forma predeterminada, AWS Lake Formation concede el permiso `Super` al grupo `IAMAllowedPrincipals` en todos los recursos existentes del catálogo de datos AWS Glue, y concede el permiso `Super` en los nuevos recursos del catálogo de datos si está activada la configuración de Utilizar solo control de acceso de IAM. Esto hace que el acceso a los recursos del catálogo de datos y a las ubicaciones de Amazon S3 esté controlado únicamente por las políticas de AWS Identity and Access Management (IAM). El grupo `IAMAllowedPrincipals` incluye a todos los Usuarios y roles de IAM a los que sus políticas de IAM permiten acceder a los objetos de su catálogo de datos. El permiso `Super` permite a una entidad principal efectuar todas las operaciones compatibles con Lake Formation en la base de datos o tabla sobre la que se concede.

Puede empezar a utilizar Lake Formation para administrar el acceso a sus datos registrando las ubicaciones de los recursos existentes del catálogo de datos en Lake Formation o utilizando el modo de acceso híbrido. Cuando registre la ubicación de Amazon S3 en modo de acceso híbrido, puede habilitar los permisos de Lake Formation optando por entidades principales para las bases de datos y las tablas bajo esa ubicación.

Para facilitar la transición de los permisos del lago de datos de un modelo IAM y Amazon S3 a los permisos de Lake Formation, le recomendamos que utilice el modo de acceso híbrido para el catálogo de datos. Con el modo de acceso híbrido, dispone de una vía incremental en la que puede habilitar los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir a otros usuarios o cargas de trabajo existentes.

Para obtener más información, consulte Modo de acceso híbrido.

Inhabilite la configuración predeterminada del catálogo de datos para mover todos los usuarios existentes de una tabla a Lake Formation en un solo paso.

Para empezar a utilizar los permisos de Lake Formation con sus bases de datos y tablas existentes del catálogo de datos AWS Glue, debe hacer lo siguiente:

1. Determine los permisos IAM existentes de sus usuarios para cada base de datos y tabla.
2. Reproduzca estos permisos en Lake Formation.
3. Para cada ubicación de Amazon S3 que contenga datos:

a. Revoque el permiso `Super` del grupo `IAMAllowedPrincipals` en cada recurso del catálogo de datos que haga referencia a esa ubicación.

b. Registre la ubicación en Lake Formation.

4. Elimine las políticas de IAM de control de acceso preciso existentes.

> **Important**
>
> Para añadir nuevos usuarios mientras está en el proceso de transición de su catálogo de datos, debe configurar permisos AWS Glue granulares en IAM como antes. También debe replicar esos permisos en Lake Formation como se describe en esta sección. Si los nuevos usuarios disponen de las políticas de IAM básicas descritas en esta guía, pueden enumerar las bases de datos o tablas que tengan concedido el permiso `Super` a `IAMAllowedPrincipals`. También pueden ver los metadatos de esos recursos.

Siga los pasos de esta sección para actualizar al modelo de permisos Lake Formation. Comience por the section called “Paso 1: Enumerar los permisos existentes”.

Paso 1: Enumerar los permisos existentes de usuarios y roles

Para empezar a utilizar los permisos AWS Lake Formation con sus bases de datos y tablas AWS Glue existentes, primero debe determinar los permisos existentes de sus usuarios.

> **Important**
>
> Antes de empezar, asegúrese de haber completado las tareas de Introducción.

Temas

Uso de la operación de la API

Utilice la operación de la API ListPoliciesGrantingServiceAccess de AWS Identity and Access Management (IAM) para determinar las políticas de IAM vinculadas a cada entidad principal (usuario o rol). A partir de las políticas que aparecen en los resultados, puede determinar los permisos IAM que se conceden a la entidad principal. Debe invocar la API para cada entidad principal por separado.

Example

El siguiente ejemplo de AWS CLI devuelve las políticas adjuntas al usuario `glue_user1`.

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

El comando devuelve resultados similares al siguiente:

```
{
    "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
}
```

Utilización de la AWS Management Console

También puede ver esta información en la consola de AWS Identity and Access Management (IAM), en la pestaña Access Advisor de la página de Resumen del usuario o rol:

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. En el panel de navegación, seleccione Users (Usuarios) o Roles.
3. Elija un nombre de la lista para abrir su página Resumen y elija la pestaña Access Advisor.
4. Inspeccione cada una de las políticas para determinar la combinación de bases de datos, tablas y acciones para las que cada usuario tiene permisos.

   Recuerde inspeccionar los roles además de los usuarios durante este proceso porque sus trabajos de tratamiento de datos podrían estar asumiendo roles para acceder a los datos.

Uso AWS CloudTrail

Otra forma de determinar los permisos existentes es buscar en AWS CloudTrail las llamadas a la API AWS Glue en las que el campo `additionaleventdata` de los registros contenga una entrada `insufficientLakeFormationPermissions`. Esta entrada enumera la base de datos y la tabla sobre las que el usuario necesita permisos de Lake Formation para llevar a cabo la misma acción.

Se trata de registros de acceso a datos, por lo que no se garantiza que generen una lista completa de los usuarios y sus permisos. Recomendamos elegir un intervalo de tiempo amplio para capturar la mayoría de los patrones de acceso a los datos de sus usuarios, por ejemplo, varias semanas o meses.

Para obtener más información, consulte Ver eventos con el historial de eventos de CloudTrail en la Guía del usuario de AWS CloudTrail.

A continuación, puede configurar los permisos de Lake Formation para que coincidan con los AWS Glue permisos. Consulte Paso 2: Configurar permisos equivalentes de Lake Formation.

Paso 2: Configurar permisos equivalentes de Lake Formation

Utilizando la información recopilada en AWS Glue, conceda permisos AWS Lake Formation para que coincidan con los permisos Paso 1: Enumerar los permisos existentes de usuarios y roles. Utilice cualquiera de los siguientes métodos para las concesiones:

- Utilizar la consola de Lake Formation o la AWS CLI.

Consulte the section called “Concesión y revocación de permisos del catálogo de datos”.

- Utilice las operaciones `GrantPermissions` o `BatchGrantPermissions` de la API.

  Consulte API de permisos.

Para obtener más información, consulte Descripción general de los permisos de Lake Formation .

Después de configurar los permisos de Lake Formation, continúe con Paso 3: Conceder a los usuarios permisos de IAM para usar Lake Formation.

Paso 3: Conceder a los usuarios permisos de IAM para usar Lake Formation

Para usar el modelo de permisos AWS Lake Formation, las entidades principales deben tener permisos de AWS Identity and Access Management (IAM) sobre las API de Lake Formation.

Cree la siguiente política en IAM y vincúlela a cada usuario que necesite acceder a su lago de datos. Llame `LakeFormationDataAccess` a la política.

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

A continuación, actualice los permisos de Lake Formation una ubicación de datos cada vez. Consulte Paso 4: Cambiar sus almacenes de datos al modelo de permisos de Lake Formation.

Paso 4: Cambiar sus almacenes de datos al modelo de permisos de Lake Formation

Actualice los permisos de Lake Formation una ubicación de datos cada vez. Para ello, repita toda esta sección hasta registrar todas las rutas de Amazon Simple Storage Service (Amazon S3) a las que hace referencia el catálogo de datos.

Temas

Verificar los permisos de Lake Formation

Antes de registrar una ubicación, complete una etapa de verificación para asegurarse de que las entidades principales correctas tienen los permisos de Lake Formation necesarios y de que no se conceden permisos de Lake Formation a entidades principales que no deberían tenerlos. Mediante la operación de la API `GetEffectivePermissionsForPath` de Lake Formation, identifique los recursos del catálogo de datos que hacen referencia a la ubicación de Amazon S3, junto con las entidades principales que tienen permisos sobre dichos recursos.

El siguiente ejemplo de AWS CLI devuelve las bases de datos y las tablas del catálogo de datos que hacen referencia al bucket de Amazon S3 `products`.

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

Observe la opción `profile`. Se recomienda ejecutar el comando como administrador de un lago de datos.

Lo siguiente es un extracto de los resultados obtenidos.

```
{
        "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**
>
> Si su catálogo de datos AWS Glue está cifrado, `GetEffectivePermissionsForPath` devuelve solo las bases de datos y tablas que se crearon o modificaron después de la disponibilidad general de Lake Formation.

Proteja los recursos del catálogo de datos existentes

A continuación, revoque el permiso `Super` de cada tabla `IAMAllowedPrincipals` y base de datos que haya identificado para la ubicación.

> **Warning**
>
> Si cuenta con una automatización que crea bases de datos y tablas en el catálogo de datos, los pasos siguientes podrían provocar un error en los trabajos de automatización y extracción, transformación y carga (ETL) posteriores. Continúe solo después de haber modificado sus procesos existentes o de haber concedido permisos explícitos de Lake Formation a las entidades principales requeridas. Para obtener información sobre los permisos de Lake Formation, consulte the section called “Referencia de permisos de Lake Formation”.

Para revocar **`Super`** de **`IAMAllowedPrincipals`** sobre una tabla

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, elija Tablas.
3. En la página Tablas, seleccione el botón que hay junto a la tabla que desee.
4. En el menú Acciones, seleccione Restaurar.
5. En el cuadro de diálogo Revocar permisos, en la lista Usuarios y roles de IAM, desplácese hacia abajo hasta el encabezado Grupo y seleccione IAMAllowedPrincipals.
6. En Permisos de tabla, asegúrese de que la opción Super esté seleccionada y, a continuación, elija Revocar.

Para revocar **`Super`** desde **`IAMAllowedPrincipals`** sobre una base de datos

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, seleccione Databases (Bases de datos).
3. En la página Bases de datos, seleccione el botón que hay junto a la base de datos que desee.
4. En el menú Actions, seleccione Editar.
5. En la página Editar base de datos, desactive Utilizar solo el control de acceso IAM para las nuevas tablas de esta base de datos y, a continuación, seleccione Guardar.
6. De vuelta en la página Bases de datos, asegúrese de que la base de datos sigue seleccionada y, a continuación, en el menú Acciones, elija Revocar.
7. En el cuadro de diálogo Revocar permisos, en la lista Usuarios y roles de IAM, desplácese hacia abajo hasta el encabezado Grupo y seleccione IAMAllowedPrincipals.
8. En Permisos de base de datos, asegúrese de que la opción Super está seleccionada y, a continuación, elija Revocar.

Active los permisos de Lake Formation para su ubicación de Amazon S3

A continuación, registre la ubicación de Amazon S3 con Lake Formation. Para ello, puede utilizar el proceso descrito en Añadir una ubicación de Amazon S3 a su lago de datos. O bien, utilice la operación de la API `RegisterResource`, descrita en API de expedición de credenciales.

> **Note**
>
> Si se registra una ubicación principal, no es necesario registrar las ubicaciones secundarias.

Tras finalizar estos pasos y comprobar que sus usuarios pueden acceder a sus datos, habrá actualizado con éxito los permisos de Lake Formation. Continúe en el paso siguiente, Paso 5: Proteger los nuevos recursos del catálogo de datos.

Paso 5: Proteger los nuevos recursos del catálogo de datos

A continuación, proteja todos los nuevos recursos del catálogo de datos cambiando la configuración predeterminada del catálogo de datos. Desactive las opciones para utilizar solo el control de acceso AWS Identity and Access Management de (IAM) para las nuevas bases de datos y tablas.

> **Warning**
>
> Si cuenta con una automatización que crea bases de datos y tablas en el catálogo de datos, los pasos siguientes podrían provocar un error en los trabajos de automatización y extracción, transformación y carga (ETL) posteriores. Continúe solo después de haber modificado sus procesos existentes o de haber concedido permisos explícitos de Lake Formation a las entidades principales requeridas. Para obtener información sobre los permisos de Lake Formation, consulte the section called “Referencia de permisos de Lake Formation”.

Para cambiar la configuración predeterminada del catálogo de datos

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como usuario administrativo de IAM (el usuario `Administrator` u otro usuario con la política administrada `AdministratorAccess` de AWS).
2. En el panel de navegación, seleccione Settings (Configuración).
3. En la página Configuración del catálogo de datos, desactive ambas casillas de verificación y, a continuación, seleccione Guardar.

El siguiente paso es conceder a los usuarios acceso a bases de datos o tablas adicionales en el futuro. Consulte Paso 6: Proporcionar a los usuarios una nueva política de IAM para el futuro acceso al lago de datos.

Paso 6: Proporcionar a los usuarios una nueva política de IAM para el futuro acceso al lago de datos

Para conceder a sus usuarios acceso a bases de datos o tablas adicionales del catálogo de datos en el futuro, debe darles la política insertada básica AWS Identity and Access Management (IAM) que se indica a continuación. Llame `GlueFullReadAccess` a la política.

> **Important**
>
> Si adjunta esta política a un usuario antes de revocar `Super` desde `IAMAllowedPrincipals` en cada base de datos y tabla de su catálogo de datos, ese usuario podrá ver todos los metadatos de cualquier recurso sobre el que se conceda `Super` a `IAMAllowedPrincipals`.

```
{
    "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**
>
> Las políticas integradas designadas en este paso y en los anteriores contienen permisos IAM mínimos. Para conocer las políticas sugeridas para los administradores del lago de datos, los analistas de datos y otros personajes, consulte the section called “Personas de Lake Formation y referencia de permisos IAM”.

A continuación, proceda a Paso 7: Limpiar las políticas de IAM existentes.

Paso 7: Limpiar las políticas de IAM existentes

Después de configurar los permisos de AWS Lake Formation y de crear y adjuntar las políticas básicas de control de acceso de AWS Identity and Access Management (IAM), complete el siguiente paso final:

- Elimine de los usuarios, grupos y roles las antiguas políticas de IAM de control de acceso específico que replicó en Lake Formation.

Al hacerlo, se asegura de que esas entidades principales ya no tengan acceso directo a los datos de Amazon Simple Storage Service (Amazon S3). A continuación, puede administrar el acceso al lago de datos para esas entidades principales totalmente a través de Lake Formation.

AWS Lake Formation y puntos de conexión de VPC de interfaz (AWS PrivateLink)

Amazon VPC es un servicio de AWS que puede utilizar para lanzar recursos de AWS en una red virtual que defina. Con una VPC, puede controlar la configuración de la red, como el rango de direcciones IP, las subredes, las tablas de ruteo y las gateways de red.

Si utiliza Amazon Virtual Private Cloud (Amazon VPC) para alojar sus recursos AWS, puede establecer una conexión privada entre su VPC y Lake Formation. Esta conexión se utiliza para que Lake Formation pueda comunicarse con los recursos de su VPC sin pasar por la Internet pública.

Puede establecer una conexión privada entre la VPC y AWS Lake Formation mediante la creación de un punto de conexión de VPC de interfaz. Los puntos de conexión de la interfaz están impulsados por AWS PrivateLink, una tecnología con la que puede acceder de forma privada a las API de Lake

Formation sin necesidad de una puerta de enlace de Internet, un dispositivo NAT, una conexión VPN o una conexión AWS Direct Connect. Las instancias de su VPC no necesitan direcciones IP públicas para comunicarse con las API de Lake Formation. El tráfico entre su VPC y Lake Formation no sale de la red de Amazon.

Cada punto de conexión de la interfaz está representado por una o más interfaces de red elásticas en las subredes.

Para obtener más información, consulte Puntos de conexión de VPC de interfaz (AWS PrivateLink) en la Guía del usuario de Amazon VPC.

Consideraciones para los puntos de conexión VPC de Lake Formation

Antes de configurar un punto de conexión de VPC de interfaz para Lake Formation, asegúrese de revisar las propiedades y limitaciones del punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Lake Formation es compatible con hacer llamadas a todas sus acciones API desde su VPC. Puede utilizar Lake Formation con puntos de conexión VPC en todos los Regiones de AWS que sean compatibles tanto con Lake Formation como con los puntos de conexión VPC de Amazon.

Creación del punto de conexión de VPC de la interfaz para Lake Formation

Puede crear un punto de conexión de VPC para el servicio Lake Formation utilizando la consola de Amazon VPC o la AWS Command Line Interface (AWS CLI). Para obtener más información, consulte Creación de un punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Cree un punto de conexión de VPC para Lake Formation utilizando el siguiente nombre de servicio:

- com.amazonaws.*region*.lakeformation

Si habilita el DNS privado para el punto de conexión, puede hacer solicitudes de API a Lake Formation utilizando su nombre DNS predeterminado para la Región, por ejemplo, `lakeformation.us-east-1.amazonaws.com`.

Para obtener más información, consulte Acceso a un servicio a través de un punto de conexión de interfaz en la Guía del usuario de Amazon VPC.

Creación de una política de punto de conexión de VPC para Lake Formation

Lake Formation es compatible con las políticas de punto de conexión de VPC. Una política de punto de conexión de VPC es una política de recursos AWS Identity and Access Management (IAM) que se adjunta a un punto de conexión cuando se crea o modifica el punto de conexión.

Puede adjuntar una política a su punto de conexión de VPC que controle el acceso a Lake Formation. La política especifica la siguiente información:

- La entidad principal que puede realizar acciones.
- Las acciones que se pueden realizar.
- Los recursos en los que se pueden llevar a cabo las acciones.

Para obtener más información, consulte Control del acceso a los servicios con puntos de conexión de VPC en la guía del usuario de Amazon VPC.

Ejemplo: política de punto de conexión de VPC para acciones de Lake Formation

El ejemplo siguiente de política de punto de conexión de VPC para Lake Formation permite la expedición de credenciales utilizando los permisos de Lake Formation. Puede utilizar esta política para ejecutar consultas con permisos de Lake Formation desde un clúster de Amazon Redshift o un clúster Amazon EMR situado en una subred privada.

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

> **Note**
>
> Si no adjunta una política al crear un punto de conexión, se adjunta una política predeterminada que permite el acceso completo al servicio.

Para obtener más información, consulte estos temas en la documentación de Amazon VPC:

- ¿Qué es Amazon VPC?
- Crear un punto de conexión de interfaz
- Utilizar políticas de punto de conexión de VPC

Tutoriales

Los siguientes tutoriales están organizados en tres temas y proporcionan instrucciones paso a paso sobre cómo construir un lago de datos, ingestar datos, compartir y asegurar lagos de datos utilizando AWS Lake Formation:

1. Crear un lago de datos e ingerir datos: aprenda a crear un lago de datos y utilice esquemas para mover, almacenar, catalogar, limpiar y organizar sus datos. También aprenderá a configurar tablas gobernadas. Una tabla gobernada es un nuevo tipo de tabla de Amazon S3 compatible con transacciones atómicas, coherentes, aisladas y duraderas (ACID).

   Antes de comenzar, asegúrese de que ha completado los pasos detallados en Introducción a Lake Formation.

   - Creación de un lago de datos a partir de una AWS CloudTrail fuente

     Cree y cargue su primer lago de datos utilizando sus propios registros de CloudTrail como origen de datos.

   - Creación de un lago de datos a partir de un origen JDBC en Lake Formation

     Cree un lago de datos utilizando uno de sus almacenes de datos accesibles desde JDBC, como una base de datos relacional, como origen de datos.

2. Proteger los lagos de datos: aprenda a utilizar controles de acceso basados en etiquetas y a nivel de fila para proteger y gestionar eficazmente el acceso a sus lagos de datos.

   - Configuración de permisos para formatos de almacenamiento de tablas abiertas en Lake Formation

     Este tutorial muestra cómo configurar los permisos para los formatos de tablas transaccionales de código abierto (tablas Iceberg de Apache, Hudi de Apache y Linux Foundation Delta Lake) en Lake Formation.

   - Gestión de un lago de datos mediante el control de acceso basado en etiquetas de Lake Formation

     Aprenda a administrar el acceso a los datos dentro de un lago de datos mediante el control de acceso basado en etiquetas en Lake Formation.

   - Protección de los lagos de datos con control de acceso a nivel de fila

Los permisos a nivel de fila de le permiten proporcionar acceso a filas específicas de una tabla en función de las políticas de control y cumplimiento de los datos.

3. Compartir datos: aprenda a compartir sus datos de forma segura entre Cuentas de AWS mediante el control de acceso basado en etiquetas (TBAC) y gestione los permisos específicos en los conjuntos de datos compartidos entre Cuentas de AWS.
   - Compartir un lago de datos utilizando el control de acceso basado en etiquetas de Lake Formation y recursos con nombre

     En este tutorial, aprenderá a compartir sus datos de forma segura entre Cuentas de AWS con Lake Formation.
   - Compartir un lago de datos utilizando el control de acceso específico de Lake Formation

     En este tutorial, aprenderá a compartir rápida y fácilmente conjuntos de datos utilizando Lake Formation cuando gestione varios Cuentas de AWS con AWS Organizations.

Temas

Creación de un lago de datos a partir de una AWS CloudTrail fuente

Este tutorial le guía por las acciones que debe realizar en la consola de Lake Formation para crear y cargar su primer lago de datos desde una AWS CloudTrail fuente.

Pasos generales para crear un lago de datos

1. Registre una ruta de Amazon Simple Storage Service (Amazon S3) como lago de datos.
2. Conceda a Lake Formation permisos para escribir en el Catálogo de datos y en las ubicaciones de Amazon S3 del lago de datos.
3. Cree una base de datos para organizar las tablas de metadatos en el Catálogo de datos.
4. Utilice un esquema para crear un flujo de trabajo. Ejecute el flujo de trabajo para incorporar datos de un origen de datos.
5. Configure sus permisos de Lake Formation para permitir que otros administren los datos del Catálogo de datos y del lago de datos.
6. Configure Amazon Athena para consultar los datos que haya importado en su lago de datos de Amazon S3.
7. Para algunos tipos de almacenes de datos, configure Amazon Redshift Spectrum para consultar los datos que importó a su lago de datos de Amazon S3.

Temas

Destinatarios previstos

En la siguiente tabla se enumeran los roles utilizadas en este tutorial para crear un lago de datos.

Destinatarios previstos

Rol	Descripción
Administrador de IAM	Tiene la política AWS gestionada:`AdministratorAccess` . Puede crear roles de IAM y buckets de Amazon S3.
Administrador de lagos de datos	Usuario que puede acceder al Catálogo de datos, crear bases de datos y conceder permisos de Lake Formation a otros usuarios. Tiene menos permisos de IAM que el administrador de IAM, pero suficientes para administrar el lago de datos.
Analista de datos	Usuario que puede ejecutar consultas en el lago de datos. Solo tiene permisos suficientes para ejecutar consultas.
Rol de flujo de trabajo	Rol con las políticas de IAM necesarias para ejecutar un flujo de trabajo. Para obtener más información, consulte (Opcional) Cree un rol de IAM para los flujos de trabajo.

Requisitos previos

Antes de empezar

- Asegúrese de haber completado las tareas de Configuración de AWS Lake Formation.
- Conozca la ubicación de sus CloudTrail registros.
- Athena requiere que el analista de datos cree un bucket de Amazon S3 para almacenar los resultados de las consultas antes de usar Athena.

Se supone que está familiarizado con el AWS Identity and Access Management (IAM). Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Paso 1: Crear un usuario de análisis de datos

Este usuario tiene el conjunto mínimo de permisos para consultar el lago de datos.

1. Abra la consola de IAM en https://console.aws.amazon.com/iam. Inicie sesión como el usuario administrador que creó en la política gestionada Crear un usuario administrativo o como usuario con la política `AdministratorAccess` AWS gestionada.
2. Cree una tabla con nombre `datalake_user` con la configuración siguiente:
   - Habilite AWS Management Console el acceso.
   - Defina una contraseña y no solicite restablecerla.
   - Adjunte la política `AmazonAthenaFullAccess` AWS gestionada.
   - Vincule la siguiente política insertada. Llame a la política `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": "*"
        }
    ]
}
```

Paso 2: Añadir permisos para leer los AWS CloudTrail registros a la función de flujo de trabajo

1. Asocie la siguiente política insertada al rol `LakeFormationWorkflowRole`. La política otorga permiso para leer tus AWS CloudTrail registros. Llame a la política `DatalakeGetCloudTrail`.

   Para crear la función de `LakeFormationWorkflowRole`, consulte (Opcional) Cree un rol de IAM para los flujos de trabajo.

   > **Important**
   >
   > <your-s3-cloudtrail-bucket>Sustitúyala por la ubicación de tus CloudTrail datos en Amazon S3.

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

2. Compruebe que haya tres políticas adjuntadas al rol.

Paso 3: Crear un bucket de Amazon S3 para el lago de datos

Cree el bucket de Amazon S3 que será la ubicación raíz de su lago de datos.

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3/ e inicie sesión como el usuario administrador que creó en Crear un usuario administrativo.
2. Elija Crear bucket y siga el asistente para crear un bucket llamado *<yourName>*`-datalake-cloudtrail`, donde *<yourName>* es su inicial y apellido. Por ejemplo: `jdoe-datalake-cloudtrail`.

Para obtener instrucciones detalladas sobre la creación de un bucket de Amazon S3, consulte Crear un bucket.

Paso 4: Registrar una ruta de Amazon S3

Registre una ruta de Amazon S3 como ubicación raíz de su lago de datos.

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, en Registrar e ingerir, elija las Ubicaciones de los lagos de datos.
3. Seleccione Registrar ubicación y, a continuación, Examinar.
4. Seleccione el bucket *<yourName>*`-datalake-cloudtrail` que creó anteriormente, acepte el rol de IAM `AWSServiceRoleForLakeFormationDataAccess` predeterminado y, a continuación, elija Registrar ubicación.

   Para obtener más información sobre cómo registrar ubicaciones, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

Paso 5: Conceder permisos de ubicación de datos

Las entidades principales deben tener permisos de ubicación de datos en una ubicación de lago de datos para crear tablas o bases de datos del Catálogo de datos que apunten a esa ubicación. Debe conceder permisos de ubicación de datos al rol de IAM para los flujos de trabajo, de modo que el flujo de trabajo pueda escribir en el destino de la ingesta de datos.

1. En el panel de navegación, bajo Permisos, seleccione Ubicaciones de datos.
2. Elija Conceder y, en el cuadro de diálogo Conceder permisos, seleccione lo siguiente:

   a. En Usuario de IAM y roles, elija `LakeFormationWorkflowRole`.

   b. En Ubicaciones de almacenamiento, elija su bucket de *<yourName>*`-datalake-cloudtrail`.
3. Elija Conceder.

Para obtener más información sobre permisos de ubicación de datos, consulte Underlying data access control.

Paso 6: Crear una base de datos en Data Catalog

Las tablas de metadatos del Catálogo de datos de Lake Formation se almacenan en una base de datos.

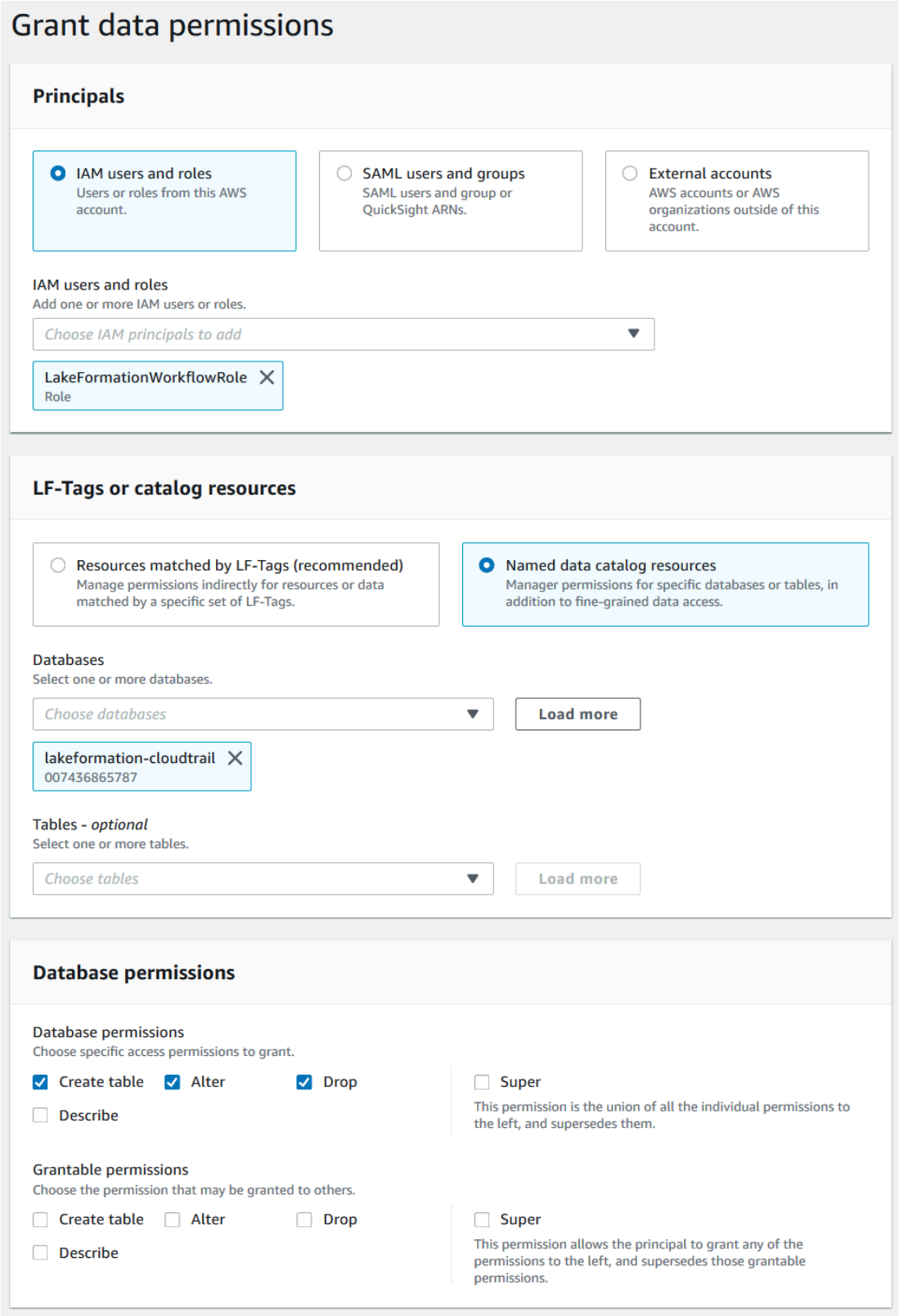
1. En el panel de navegación, Catálogo de datos, elija Tablas.
2. Seleccione Crear base de datos y, en Detalles de la base de datos, introduzca el nombre `lakeformation_cloudtrail`.
3. Deje los demás campos en blanco y elija Crear base de datos.

Paso 7: Conceder permisos de datos

Debe conceder permisos para crear tablas de metadatos en el Catálogo de datos. Como el flujo de trabajo se ejecutará con el rol `LakeFormationWorkflowRole`, debe conceder estos permisos al rol.

1. En la consola de Lake Formation, en el panel de navegación, en Catálogo de datos, seleccione Bases de datos.
2. Elija la base de datos `lakeformation_cloudtrail` y, a continuación, en la lista desplegable Acciones, seleccione Conceder bajo el encabezado Permisos.
3. En el cuadro de diálogo Conceder permisos de datos, seleccione lo siguiente:
    a. En Entidades principales, Usuario de IAM y roles, elija `LakeFormationWorkflowRole`.
    b. En Etiquetas LF o recursos del catálogo, elija Recursos de Catálogo de datos con nombre.
    c. En Bases de datos, debería ver que la base de datos `lakeformation_cloudtrail` ya está agregada.
    d. En Permisos de base de datos, seleccione Crear tabla, Modificar y Borrar, y desactive Super si está seleccionada.

    El cuadro de diálogo Conceder permisos de datos ahora tendrá el aspecto que se muestra en esta captura de pantalla.



4. Elija Conceder.

Para obtener más información sobre permisos de Lake Formation, consulte Administrar los permisos de Lake Formation.

Paso 8: Utilizar un esquema para crear un flujo de trabajo.

Para leer los CloudTrail registros, entender su estructura y crear las tablas adecuadas en el catálogo de datos, necesitamos configurar un flujo de trabajo que consista en AWS Glue rastreadores, tareas, activadores y flujos de trabajo. Los esquemas de Lake Formation simplifican este proceso.

El flujo de trabajo genera las tareas, los rastreadores y los activadores que descubren e ingieren datos en su lago de datos. Cree un flujo de trabajo basado en uno de los esquemas predefinidos de Lake Formation.

1. En la consola de Lake Formation, en el panel de navegación, elija Esquemas y, a continuación, Usar esquema.
2. En la página Usar un esquema, en Tipo de esquema, elija. AWS CloudTrail
3. En Importar fuente, selecciona una CloudTrail fuente y una fecha de inicio.
4. En Destino de importación, especifique estos parámetros:

Bases de datos de destino	`lakeformation_cloudtrail`
Ubicación de almacenamiento de destino	`s3://<yourName> -datalake-cloudtrail`
Formato de los datos	Parquet

5. Para ver la frecuencia de la importación, seleccione Ejecutar bajo demanda.
6. En Opciones de importación, especifique estos parámetros:

Nombre del flujo de trabajo	`lakeformationcloudtrailtest`
Rol de IAM	`LakeFormationWorkflowRole`
Prefijo de tabla	`cloudtrailtest`

> **Note**
>
> Debe estar en minúscula.

7. Seleccione Crear y espere a que la consola informe de que el flujo de trabajo se ha creado correctamente.

> **Tip**
>
> ¿Ha recibido este mensaje de error?
> `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...`
> Si es así, compruebe que ha sustituido <account-id>en la política integrada para el usuario administrador del lago de datos por un número de AWS cuenta válido.

Paso 9: Ejecutar el flujo de trabajo

Como especificó que el flujo de trabajo es run-on-demand, debe iniciarlo manualmente.

- En la página Esquemas, seleccione el flujo de trabajo `lakeformationcloudtrailtest` y, en el menú Acciones, elija Iniciar.

  A medida que se ejecuta el flujo de trabajo, puede ver su progreso en la columna Estado de la última ejecución. Pulse el botón de actualización de vez en cuando.

  El estado pasa de EN EJECUCIÓN a Detectando, Importando y FINALIZADO.

  Cuando se complete el flujo de trabajo:

  - El Catálogo de datos tendrá nuevas tablas de metadatos.
  - Sus CloudTrail registros se incorporarán al lago de datos.

  Si se produce un error en el flujo de trabajo, haga lo siguiente:

a. Seleccione el flujo de trabajo y, en el menú Acciones, elija Ver gráfico.

   El flujo de trabajo se abre en la consola de AWS Glue.

b. Asegúrese de que se seleccione el flujo de trabajo y elija la pestaña Historial.

c. En Historial, seleccione la ejecución más reciente y seleccione Ver detalles de la ejecución.

d. Seleccione un trabajo o un rastreador fallidos en el gráfico dinámico (tiempo de ejecución) y revise el mensaje de error. Los nodos con errores aparecen en rojo o amarillo.

Paso 10: Conceder SELECT en las tablas

Debe conceder el permiso `SELECT` a las nuevas tablas del Catálogo de datos para que el analista de datos pueda consultar los datos a los que apuntan las tablas.

> **Note**
>
> Un flujo de trabajo concede automáticamente el permiso `SELECT` sobre las tablas que crea al usuario que lo ejecutó. Dado que el administrador del lago de datos ejecutó este flujo de trabajo, debe conceder `SELECT` al analista de datos.

1. En la consola de Lake Formation, en el panel de navegación, en Catálogo de datos, seleccione Bases de datos.
2. Elija la base de datos `lakeformation_cloudtrail` y, a continuación, en la lista desplegable Acciones, seleccione Conceder bajo el encabezado Permisos.
3. En el cuadro de diálogo Conceder permisos de datos, seleccione lo siguiente:

   a. En Entidades principales, Usuario de IAM y roles, elija `datalake_user`.

   b. En Etiquetas LF o recursos del catálogo, elija Recursos de Catálogo de datos con nombre.

   c. En Bases de datos, la base de datos `lakeformation_cloudtrail` ya debería estar seleccionada.

   d. Para Tablas, elija `cloudtrailtest-cloudtrail`.

   e. En Permisos de tabla y columna, elija Seleccionar.

4. Elija Conceder.

El paso siguiente se efectúa como analista de datos.

Paso 11: Consultar el lago de datos mediante Amazon Athena

Utilice la Amazon Athena consola para consultar los CloudTrail datos de su lago de datos.

1. Abra la consola de Athena en https://console.aws.amazon.com/athena/ e inicie sesión como analista de datos, usuario `datalake_user`.
2. Si es necesario, elija Comenzar para continuar con el editor de consultas de Athena.
3. Para Origen de datos, elija AwsDataCatalog.
4. En Database (Base de datos), elija `lakeformation_cloudtrail`.

   Se rellena la lista de Tablas.
5. En el menú desplegable (3 puntos en horizontal) situado junto a la tabla `cloudtrailtest-cloudtrail`, seleccione Vista previa de la tabla y, a continuación, seleccione Ejecutar.

   La consulta se ejecuta y muestra 10 filas de datos.

   Si no ha utilizado Athena antes, primero debe configurar una ubicación de Amazon S3 en la consola de Athena para almacenar los resultados de las consultas. El `datalake_user` debe disponer de los permisos necesarios para acceder al bucket de Amazon S3 que elija.

> **Note**
>
> Ahora que ha completado el tutorial, conceda permisos de datos y permisos de ubicación de datos a las entidades principales de su organización.

Creación de un lago de datos a partir de un origen JDBC en Lake Formation

Este tutorial le guía a través de los pasos a seguir en la consola de AWS Lake Formation para crear y cargar su primer lago de datos desde un origen JDBC utilizando Lake Formation.

Temas

- Destinatarios previstos
- Requisitos previos del tutorial

- Paso 1: Crear un usuario analista de datos
- Paso 2: Crear una conexión en AWS Glue
- Paso 3: Crear un bucket de Amazon S3 para el lago de datos
- Paso 4: Registrar una ruta de Amazon S3
- Paso 5: Conceder permisos de ubicación de datos
- Paso 6: Crear una base de datos en el catálogo de datos
- Paso 7: Conceder permisos de datos
- Paso 8: Utilizar un esquema para crear un flujo de trabajo.
- Paso 9: Ejecutar el flujo de trabajo
- Paso 10: Conceder SELECCIONAR en las tablas
- Paso 11: Consultar el lago de datos mediante Amazon Athena
- Paso 12: Consultar los datos del lago de datos mediante Amazon Redshift Spectrum
- Paso 13: Conceder o revocar los permisos de Lake Formation mediante Amazon Redshift Spectrum

Destinatarios previstos

La siguiente tabla enumera los roles utilizados en este tutorial de JDBC de AWS Lake Formation.

Rol	Descripción
Administrador de IAM	Un usuario que puede crear usuarios y roles de AWS Identity and Access Management (IAM) y buckets de Amazon Simple Storage Service (Amazon S3). Tiene la política administrada `AdministratorAccess` de AWS.
Administrador de lago de datos	Usuario que puede acceder al catálogo de datos, crear bases de datos y conceder permisos de Lake Formation a otros usuarios. Tiene menos permisos de IAM que el administrador de IAM, pero suficientes para administrar el lago de datos.

Rol	Descripción
Analista de datos	Usuario que puede ejecutar consultas en el lago de datos. Solo tiene permisos suficientes para ejecutar consultas.
Rol de flujo de trabajo	Rol con las políticas de IAM necesarias para ejecutar un flujo de trabajo.

Para obtener información sobre los requisitos previos para completar el tutorial, consulte Requisitos previos del tutorial.

Requisitos previos del tutorial

Antes de comenzar eltutorial JDBC de AWS Lake Formation, asegúrese de:

- Completar las tareas de Introducción a Lake Formation.
- Decidir qué almacén de datos accesible mediante JDBC desea utilizar para el tutorial.
- Reunir la información necesaria para crear una conexión AWS Glue de tipo JDBC. Este objeto del catálogo de datos incluye la dirección URL del almacén de datos, las credenciales de inicio de sesión y, si el almacén de datos se creó en una Amazon Virtual Private Cloud (Amazon VPC), información adicional de configuración específica de la VPC. Para más información, consulte Definición de conexiones en el catálogo de datos de AWS Glue en la Guía del programador de AWS Glue.

El tutorial presupone que conoce AWS Identity and Access Management (IAM). Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Para empezar, vaya al the section called “Paso 1: Crear un usuario analista de datos”.

Paso 1: Crear un usuario analista de datos

En este paso, creará un usuario AWS Identity and Access Management de (IAM) para que sea el analista de datos de su lago de datos en AWS Lake Formation.

Este usuario tiene el conjunto mínimo de permisos para consultar el lago de datos.

1. Abra la consola de IAM en https://console.aws.amazon.com/iam. Inicie sesión como el usuario administrador que creó en Crear un usuario administrativo o como usuario con la política administrada `AdministratorAccess` de AWS.
2. Cree una tabla con nombre `datalake_user` con la configuración siguiente:
   - Habilite el acceso a AWS Management Console.
   - Defina una contraseña y no solicite restablecerla.
   - Vincule la política administrada `AmazonAthenaFullAccess` de AWS.
   - Vincule la siguiente política insertada. Llame a la política `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": "*"
        }
    ]
}
```

Paso 2: Crear una conexión en AWS Glue

> **Note**
>
> Omita este paso si ya tiene una conexión en AWS Glue a su origen de datos JDBC.

AWS Lake Formation accede a orígenes de datos JDBC a través de una conexión AWS Glue. Una conexión es un objeto del catálogo de datos que contiene toda la información necesaria para conectarse al origen de datos. Puede crear una conexión utilizando la consola de AWS Glue.

Para crear una conexión

1. Abra la consola de AWS Glue en https://console.aws.amazon.com/glue/, e inicie sesión como el usuario administrador que creó en Crear un usuario administrativo.
2. En el panel de navegación, en Data catalog (Catálogo de datos), elija Connections (Conexiones).
3. En la página Connectors (Conectores), seleccione Create custom connector (Crear conector personalizado).
4. En la página de propiedades del conector, introduzca **`datalake-tutorial`** como nombre de conexión y elija JDBC como tipo de conexión. A continuación, elija Next.
5. Continúe con el asistente de conexión y guarde la conexión.

   Para obtener información sobre cómo crear una conexión, consulte Propiedades de la conexión JDBC de AWS Glue en la Guía para desarrolladores de AWS Glue.

Paso 3: Crear un bucket de Amazon S3 para el lago de datos

En este paso, creará el bucket de Amazon Simple Storage Service (Amazon S3) que será la ubicación raíz de su lago de datos.

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3/ e inicie sesión como el usuario administrador que creó en Crear un usuario administrativo.
2. Elija Crear bucket y siga el asistente para crear un bucket llamado *`<yourName>`*`-datalake-tutorial`, donde *`<yourName>`* es su inicial y apellido. Por ejemplo: `jdoe-datalake-tutorial`.

Para obtener instrucciones detalladas sobre la creación de un bucket de Amazon S3, consulte ¿Cómo se crea un bucket de S3? en la Guía del usuario de Amazon Simple Storage Service.

Paso 4: Registrar una ruta de Amazon S3

En este paso, registrará una ruta de Amazon Simple Storage Service (Amazon S3) como ubicación raíz de su lago de datos.

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, en Registrar e ingerir, elija las Ubicaciones del lago de datos.
3. Seleccione Registrar ubicación y, a continuación, Examinar.
4. Seleccione el bucket de *<yourName>*`-datalake-tutorial` que creó anteriormente, acepte el rol de IAM `AWSServiceRoleForLakeFormationDataAccess` predeterminado y, a continuación, elija Registrar ubicación.

   Para obtener más información sobre cómo registrar ubicaciones, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

Paso 5: Conceder permisos de ubicación de datos

Las entidades principales deben tener permisos de ubicación de datos en una ubicación de lago de datos para crear tablas o bases de datos del catálogo de datos que apunten a esa ubicación. Debe conceder permisos de ubicación de datos al rol de IAM para los flujos de trabajo, de modo que el flujo de trabajo pueda escribir en el destino de la ingesta de datos.

1. En la consola de Lake Formation, desde panel de navegación, Permisos, seleccione Ubicaciones de datos.
2. Elija Conceder y, en el cuadro de diálogo Conceder permisos:
   a. En Usuario de IAM y roles, seleccione `LakeFormationWorkflowRole`.
   b. En Ubicaciones de almacenamiento, elija su bucket de *<yourName>*`-datalake-tutorial`.
3. Elija Grant (Conceder).

Para obtener más información sobre permisos de ubicación de datos, consulte Underlying data access control.

Paso 6: Crear una base de datos en el catálogo de datos

Las tablas de metadatos del catálogo de datos de Lake Formation se almacenan en una base de datos.

1. En la consola de Lake Formation, en el panel de navegación, Catálogo de datos, seleccione Bases de datos.
2. Seleccione Crear base de datos y, en Detalles de la base de datos, introduzca el nombre `lakeformation_tutorial`.
3. Deje los demás campos en blanco y elija Crear base de datos.

Paso 7: Conceder permisos de datos

Debe conceder permisos para crear tablas de metadatos en el catálogo de datos. Como el flujo de trabajo se ejecutará con el rol `LakeFormationWorkflowRole`, debe conceder estos permisos al rol.

1. En la consola de Lake Formation, en el panel de navegación, en Permisos, seleccione Permisos de lago de datos.
2. Seleccione Conceder y, en el cuadro de diálogo Conceder permisos de datos:
    a. En Entidades principales, Usuario de IAM y roles, seleccione `LakeFormationWorkflowRole`.
    b. En Etiquetas LF o recursos del catálogo, elija Recursos de catálogo de datos con nombre.
    c. En Bases de datos, elija la base de datos que creó anteriormente, `lakeformation_tutorial`.
    d. En Permisos de base de datos, seleccione Crear tabla, Modificar y Borrar, y desactive Super si está marcado.
3. Elija Grant (Conceder).

Para obtener más información sobre permisos de Lake Formation, consulte Descripción general de los permisos de Lake Formation .

Paso 8: Utilizar un esquema para crear un flujo de trabajo.

El flujo de trabajo de AWS Lake Formation genera los trabajos de AWS Glue, rastreadores y desencadenadores que descubren e ingieren datos en su lago de datos. Cree un flujo de trabajo basado en uno de los esquemas predefinidos de Lake Formation.

1. En la consola de Lake Formation, en el panel de navegación, elija Esquemas y, a continuación, Usar esquema.
2. En la página Usar un esquema, en Tipo de esquema, elija Instantánea de base de datos.
3. En Importar origen, para Conexión de base de datos, elija la conexión que acaba de crear, `datalake-tutorial`, o elija una conexión existente para sus orígenes de datos.
4. En Ruta de datos de origen, introduzca la ruta desde la que se van a ingerir los datos en el formulario *<database>*/*<schema>*/*<table>*.

   Puede sustituir el carácter de porcentaje (%) por esquema o tabla. En las bases de datos que admiten esquemas, introduzca *<base de datos>*/*<esquema>*/% para hacer coincidir todas las tablas de *<esquema>* dentro de *<base de datos>*. Oracle Database y MySQL no admiten esquemas en la ruta; en su lugar, introduzca *<base de datos>*/%. En el caso de Oracle Database, *<base de datos>* es el identificador del sistema (SID).

   Por ejemplo, si una base de datos Oracle tiene `orcl` como SID, introduzca `orcl/%` para que coincidan con todas las tablas a las que tiene acceso el usuario especificado en la conexión JDCB.

   > **Important**
   >
   > Este campo distingue entre mayúsculas y minúsculas.

5. En Destino de importación, especifique estos parámetros:

Bases de datos de destino	`lakeformation_tutorial`
Ubicación de almacenamiento de destino	`s3://`*<yourName>* `-datalake-tutorial`
Formato de los datos	(Elija Parquet o CSV)

6. Para ver la frecuencia de la importación, seleccione Ejecutar bajo demanda.

7. En Opciones de importación, especifique estos parámetros:

Nombre del flujo de trabajo	`lakeformationjdbctest`
Rol de IAM	`LakeFormationWorkflowRole`
Prefijo de tabla	`jdbctest` **Note** Debe estar en minúscula.

8. Seleccione Crear y espere a que la consola informe de que el flujo de trabajo se ha creado correctamente.

> **Tip**
>
> ¿Ha recibido este mensaje de error?
> `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...`
> Si es así, compruebe que ha sustituido *<id de cuenta>* en la política integrada para el usuario administrador del lago de datos por un número de cuenta AWS válido.

Paso 9: Ejecutar el flujo de trabajo

Dado que ha especificado que el flujo de trabajo se ejecute bajo demanda, debe iniciar manualmente el flujo de trabajo en AWS Lake Formation.

1. En la consola de Lake Formation, en la página Esquemas, seleccione el flujo de trabajo `lakeformationjdbctest`.
2. Elija Acciones y, a continuación, Pegar.
3. Conforme se ejecuta el flujo de trabajo, vea su progreso en la columna Estado de la última ejecución. Pulse el botón de actualización de vez en cuando.

El estado pasa de EN EJECUCIÓN a Detectando, Importando y FINALIZADO.

Cuando finalice el flujo de trabajo:

- El catálogo de datos tendrá nuevas tablas de metadatos.
- Sus datos se ingestan al lago de datos.

Si se produce un error en el flujo de trabajo:

a. Seleccione el flujo de trabajo. Elija Acciones y, a continuación, elija Ver gráfico.

   El flujo de trabajo se abre en la consola de AWS Glue.

b. Seleccione un flujo de trabajo y, a continuación, la pestaña Historial.

c. Seleccione la ejecución más reciente y elija Ver detalles de la ejecución.

d. Seleccione un trabajo o un rastreador fallidos en el gráfico dinámico (tiempo de ejecución) y revise el mensaje de error. Los nodos con errores aparecen en rojo o amarillo.

Paso 10: Conceder SELECCIONAR en las tablas

Debe conceder el permiso `SELECT` en las nuevas tablas del catálogo de datos en AWS Lake Formation para que el analista de datos pueda consultar los datos a los que apuntan las tablas.

> **Note**
>
> Un flujo de trabajo concede automáticamente el permiso `SELECT` sobre las tablas que crea al usuario que lo ejecutó. Como el administrador del lago de datos ejecutó este flujo de trabajo, debe conceder `SELECT` al analista de datos.

1. En la consola de Lake Formation, en el panel de navegación, en Permisos, seleccione Permisos de lago de datos.
2. Seleccione Conceder y, en el cuadro de diálogo Conceder permisos de datos:

   a. En Entidades principales, Usuario de IAM y roles, seleccione `datalake_user`.

   b. En Etiquetas LF o recursos del catálogo, elija Recursos de catálogo de datos con nombre.

   c. En Bases de datos, seleccione `lakeformation_tutorial`.

Se rellena la lista Tablas.

d. En Tablas, elija una o más tablas como su origen de datos.

e. En Permisos de tabla y columna, elija Seleccionar.

3. Elija Grant (Conceder).

El paso siguiente se efectúa como analista de datos.

Paso 11: Consultar el lago de datos mediante Amazon Athena

Utilice la consola de Amazon Athena para consultar los datos en su lago de datos.

1. Abra la consola de Athena en https://console.aws.amazon.com/athena/ e inicie sesión como analista de datos, usuario `datalake_user`.
2. Si es necesario, elija Comenzar para continuar con el editor de consultas de Athena.
3. En Origen de datos, elija AWSDataCatalog.
4. En Database (Base de datos), elija `lakeformation_tutorial`.

   Se rellena la lista Tablas.
5. En el menú emergente situado junto a una de las tablas, seleccione Vista previa de la tabla.

   La consulta se ejecuta y muestra 10 filas de datos.

Paso 12: Consultar los datos del lago de datos mediante Amazon Redshift Spectrum

Puede configurar Amazon Redshift Spectrum para consultar los datos que importó a su lago de datos de Amazon Simple Storage Service (Amazon S3). En primer lugar, cree un rol AWS Identity and Access Management de (IAM) que se utilice para lanzar el clúster de Amazon Redshift y para consultar los datos de Amazon S3. A continuación, conceda a este rol los permisos `Select` en las tablas que desee consultar. Después, conceda permisos al usuario para utilizar el editor de consultas de Amazon Redshift. Por último, cree un clúster de Amazon Redshift y ejecute consultas.

Creará el clúster como administrador y lo consultará como analista de datos.

Para obtener más información sobre Amazon Redshift Spectrum, consulte Uso de Amazon Redshift Spectrum para consultar datos externos en la Guía para desarrolladores de bases de datos de Amazon Redshift.

Para configurar los permisos para ejecutar consultas de Amazon Redshift

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/. Inicie sesión como el usuario administrador que creó en Crear un usuario administrativo (nombre de usuario `Administrator`) o como un usuario con la política administrada `AdministratorAccess` de AWS.
2. En el panel de navegación, seleccione Policies (Políticas).

   Si es la primera vez que elige Policies (Políticas), aparecerá la página Welcome to Managed Policies (Bienvenido a políticas administradas). Elija Get Started (Comenzar).
3. Elija Create Policy (Crear política).
4. Seleccione la pestaña JSON.
5. Pegue el siguiente documento de política 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. Cuando haya terminado, seleccione Review (Revisar) para revisar la política. El validador de políticas notifica los errores de sintaxis.
7. En la página Revisar política, introduzca el Nombre **`RedshiftLakeFormationPolicy`** para la política que está creando. (Opcional) Introduzca una descripción. Revise el Summary (Resumen) de la política para ver los permisos concedidos por su política. A continuación, elija Create policy (Crear política) para guardar su trabajo.
8. En el panel de navegación de la consola de IAM, seleccione Roles y, a continuación, seleccione Create role (Crear rol).
9. En Select trusted entity (Seleccionar entidad de confianza), elija AWS service (Servicio de).
10. Elija el servicio Amazon Redshift para asumir este rol.
11. Elija el caso de uso Redshift Customizable (Redshift personalizable) para su servicio. A continuación, elija Next: Permissions (Siguiente: permisos).
12. Busque la política de permisos que ha creado, `RedshiftLakeFormationPolicy`, y marque la casilla situada junto al nombre de la política en la lista.
13. Elija Next: Tags (Siguiente: etiquetas).
14. Elija Next: Review (Siguiente: revisar).
15. En Role name (Nombre del rol), escriba el nombre **`RedshiftLakeFormationRole`**.
16. (Opcional) En Role description (Descripción del rol), escriba una descripción para el nuevo rol.
17. Revise el rol y, a continuación, seleccione Create role (Crear rol).

Para conceder permisos **`Select`** en la tabla para hacer consultas en la base de datos de Lake Formation

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, en Permisos, elija Permisos de lago de datos y seleccione Conceder.
3. Proporcione la siguiente información:

   - Para los usuarios y roles de IAM, elija el rol de IAM que creó, `RedshiftLakeFormationRole`. Cuando se ejecuta el editor de consultas de Amazon Redshift, este utiliza el rol de IAM para dar permiso a los datos.
   - En Database (Base de datos), elija `lakeformation_tutorial`.

Se rellena la lista de tablas.

- En Tabla, elija una tabla dentro del origen de datos que desee consultar.
- Elija el permiso Seleccionar.

4. Elija Grant (Conceder).

Para configurar Amazon Redshift Spectrum y ejecutar consultas

1. Abra la consola de Amazon Redshift en https://console.aws.amazon.com/redshift. Inicie sesión como usuario `Administrator`.
2. Elija Create cluster.
3. En la página Crear clúster, introduzca `redshift-lakeformation-demo` el identificador del clúster.
4. Para el tipo de nodo, seleccione dc2.large.
5. Desplácese hacia abajo y, en Configuraciones de base de datos, introduzca o acepte estos parámetros:

   - Nombre de usuario administrador: `awsuser`
   - Contraseña del usuario administrador: (*`Choose a password`*)

6. Amplíe los permisos del clúster y, para ver los roles de IAM disponibles, elija RedshiftLakeFormationRole. A continuación, seleccione Add IAM role (Añadir rol de IAM).
7. Si debe utilizar un puerto diferente al valor predeterminado de 5439, junto a Configuraciones adicionales, desactive la opción Utilizar valores predeterminados. Amplíe la sección de configuraciones de base de datos e introduzca un nuevo número de puerto de base de datos.
8. Elija Create cluster.

   Se carga la página Clústeres.

9. Espere hasta que el estado del clúster pase a ser Disponible. Seleccione el icono de actualización periódicamente.
10. Conceda permiso al analista de datos para ejecutar consultas en el clúster. Para ello, complete los siguientes pasos.

    a. Abra la consola de IAM en https://console.aws.amazon.com/iam/, e Inicie sesión como usuario `Administrator`.

b. En el panel de navegación, seleccione Usuarios y vincule las siguientes políticas administradas al usuario `datalake_user`.

- `AmazonRedshiftQueryEditor`
- `AmazonRedshiftReadOnlyAccess`

11. Cierre sesión en la consola de Amazon Redshift y vuelva a iniciarla como usuario `datalake_user`.
12. En la barra de herramientas vertical izquierda, elija el icono EDITOR para abrir el editor de consultas y conectarse al clúster. Si aparece el cuadro de diálogo Conectar a la base de datos, elija el nombre del clúster `redshift-lakeformation-demo` e introduzca el nombre de la base de datos **`dev`**, el nombre de usuario **`awsuser`** y la contraseña que creó. Seleccione Connect to database (Conectar a base de datos).

> **Note**
>
> Si no se le solicitan los parámetros de conexión y ya hay otro clúster seleccionado en el editor de consultas, elija Cambiar conexión para abrir el cuadro de diálogo Conectar a la base de datos.

13. En el cuadro de texto Nueva consulta 1, introduzca y ejecute la siguiente instrucción para asignar la base de datos `lakeformation_tutorial` en Lake Formation al nombre del esquema de Amazon Redshift `redshift_jdbc`:

> **Important**
>
> Sustituya *<account-id>* por un número de cuenta AWS válido y *<región>* por un nombre de región AWS válido (por ejemplo, `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. En la lista de esquemas, bajo Seleccionar esquema, elija redshift_jdbc.

Se rellena la lista de tablas. El editor de consultas muestra solo las tablas en las que se le concedieron permisos para el lago de datos de Lake Formation.

15. En el menú emergente situado junto al nombre de una tabla, seleccione Vista previa de los datos.

    Amazon Redshift devuelve las 10 primeras filas.

    Ahora puede ejecutar consultas en las tablas y columnas para las que tiene permisos.

Paso 13: Conceder o revocar los permisos de Lake Formation mediante Amazon Redshift Spectrum

Amazon Redshift es compatible con la capacidad de conceder y revocar permisos de Lake Formation en bases de datos y tablas mediante instrucciones SQL modificadas. Estas instrucciones son similares a las existentes de Amazon Redshift. Para obtener más información, consulte CONCEDER y REVOCAR en la Guía para desarrolladores de base de datos de Amazon Redshift.

Configuración de permisos para formatos de almacenamiento de tablas abiertas en Lake Formation

En AWS Lake Formation es posible administrar los permisos de acceso para Formatos de tabla abiertos (OTF), como Apache Iceberg, Apache Hudi y Linux Foundation Delta Lake. En este tutorial, aprenderá a crear Iceberg, Hudi y Delta Lake con tablas de manifiesto de enlaces simbólicos en el AWS Glue Data Catalog usando AWS Glue, configurar permisos específicos con Lake Formation y a consultar datos con Amazon Athena.

> **Note**
>
> Los servicios de análisis de AWS no admiten todos los formatos de tablas transaccionales. Para obtener más información, consulte Trabajar con otros AWS servicios. Este tutorial describe manualmente la creación de una nueva base de datos y una tabla en el catálogo de datos utilizando únicamente AWS Glue trabajos.

Este tutorial incluye una AWS CloudFormation plantilla para una configuración rápida. Puede revisarla y personalizarla para adaptarla a sus necesidades.

Temas

- Destinatarios previstos

Destinatarios previstos

Este tutorial está dirigido a administradores de IAM, administradores de lago de datos y analistas empresariales. En la tabla siguiente se enumeran los roles utilizados en este tutorial para crear una tabla gobernada con Lake Formation.

Rol	Descripción
Administrador de IAM	Usuario que puede crear usuarios de IAM, roles y buckets de Amazon S3. Tiene la política administrada `AdministratorAccess` de AWS.
Administrador de lago de datos	Usuario que puede acceder al Catálogo de datos, crear bases de datos y conceder permisos de Lake Formation a otros usuarios. Tiene menos permisos de IAM que el administrador de IAM, pero suficientes para administrar el lago de datos.
Analista de negocios	Usuario que puede ejecutar consultas en el lago de datos. Tiene permisos para ejecutar consultas.

Requisitos previos

Antes de empezar este tutorial, debe disponer de una en la Cuenta de AWS que pueda iniciar sesión como usuario con los permisos correctos. Para obtener más información, consulte Registro para obtener una Cuenta de AWS y Crear un usuario administrativo.

En el tutorial, se asume que conoce las políticas y los roles de IAM. Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Para completar este tutorial, debe configurar los recursos de AWS siguientes:

- Administrador de lago de datos
- Configuración del lago de datos de Lake Formation
- Motor Amazon Athena, versión 3

Para crear un administrador de lago de datos

1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ como usuario administrador. Creará recursos en la región EE.UU. Este (Norte de Virginia) para este tutorial.
2. Desde la consola de Lake Formation, en Permisos del panel de navegación, seleccione Roles y tareas administrativas.
3. Seleccione Elegir administradores en Administradores del lago de datos.
4. En la ventana emergente para administrar gestores de lagos de datos, en Usuarios y roles de IAM, seleccione Usuario de IAM administrador.
5. Seleccione Guardar.

Para habilitar la configuración del lago de datos

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/. En el panel de navegación, en Catálogo de datos, seleccione Configuración. Desmarque lo siguiente:
   - Usar solo el control de acceso de IAM para nuevas bases de datos.
   - Usar solo el control de acceso de IAM para nuevas tablas en bases de datos nuevas.
2. En Configuración de la versión entre cuentas, seleccione Versión 3 como versión entre cuentas.
3. Seleccione Guardar.

Para actualizar el motor Amazon Athena a la versión 3

1. Abra la consola de Athena en https://console.aws.amazon.com/athena/.
2. Seleccione el grupo de trabajo y elija el principal.

3. Asegúrese de que el grupo de trabajo tenga una versión mínima de 3. Si no es así, edite el grupo de trabajo, elija Manual en Actualizar el motor de consultas y seleccione la versión 3.
4. Seleccione Guardar cambios.

Paso 1: Aprovisionar recursos

En esta sección se explica cómo configurar los recursos de AWS mediante una plantilla de AWS CloudFormation.

Para cree recursos mediante una plantilla de AWS CloudFormation.

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation como administrador de IAM en la región Este de EE. UU. (Norte de Virginia).
2. Seleccione Lanzar pila.
3. Seleccione Siguiente en la pantalla Crear pila.
4. Introduzca un Nombre de pila.
5. Elija Siguiente.
6. En la página siguiente, seleccione Siguiente.
7. Revise los detalles en la última página y acepte que AWS CloudFormation pueda crear recursos de IAM.
8. Seleccione Crear.

   La creación de la pila puede llevar hasta dos minutos.

El lanzamiento de la pila de formación de nube crea los recursos siguientes:

- lf-otf-datalake-123456789012 — Depósito de Amazon S3 para almacenar datos

> **Note**
>
> El identificador de cuenta adjunto al nombre del bucket de Amazon S3 se sustituye por el identificador de su cuenta.

- lf-otf-tutorial-123456789012 — Depósito de Amazon S3 para almacenar los resultados de las consultas y los scripts de trabajo AWS Glue
- lficebergdb: base de datos AWS Glue Iceberg

- lfhudidb: base de datos AWS Glue Hudi
- lfdeltadb: base de datos AWS Glue Delta
- native-iceberg-create — AWS Glue trabajo que crea una tabla de iceberg en el catálogo de datos
- native-hudi-create — AWS Glue trabajo que crea una tabla Hudi en el catálogo de datos
- native-delta-create — AWS Glue trabajo que crea una tabla Delta en el catálogo de datos
- LF-OTF GlueServiceRole -: función de IAM que se transfiere AWS Glue para ejecutar los trabajos. Este rol incluye las políticas necesarias para acceder a los recursos, como el Catálogo de datos, el bucket de Amazon S3, etc.
- LF-OTF RegisterRole -: función de IAM para registrar la ubicación de Amazon S3 en Lake Formation. Este rol lleva `LF-Data-Lake-Storage-Policy` asociada.
- lf-consumer-analystuser — Usuario de IAM para consultar los datos mediante Athena
- lf-consumer-analystuser-credentials — Contraseña del usuario analista de datos almacenada en AWS Secrets Manager

Una vez finalizada la creación de la pila, vaya a la pestaña de resultados y anote los valores de:

- AthenaQueryResultLocation — Ubicación de Amazon S3 para el resultado de la consulta de Athena
- BusinessAnalystUserCredentials — Contraseña para el usuario analista de datos

  Para recuperar el valor de la contraseña:

  1. Elija el valor de `lf-consumer-analystuser-credentials` accediendo a la consola de Secrets Manager.
  2. En la sección Valor secreto, seleccione Recuperar valor secreto.
  3. Anote el valor secreto de la contraseña.

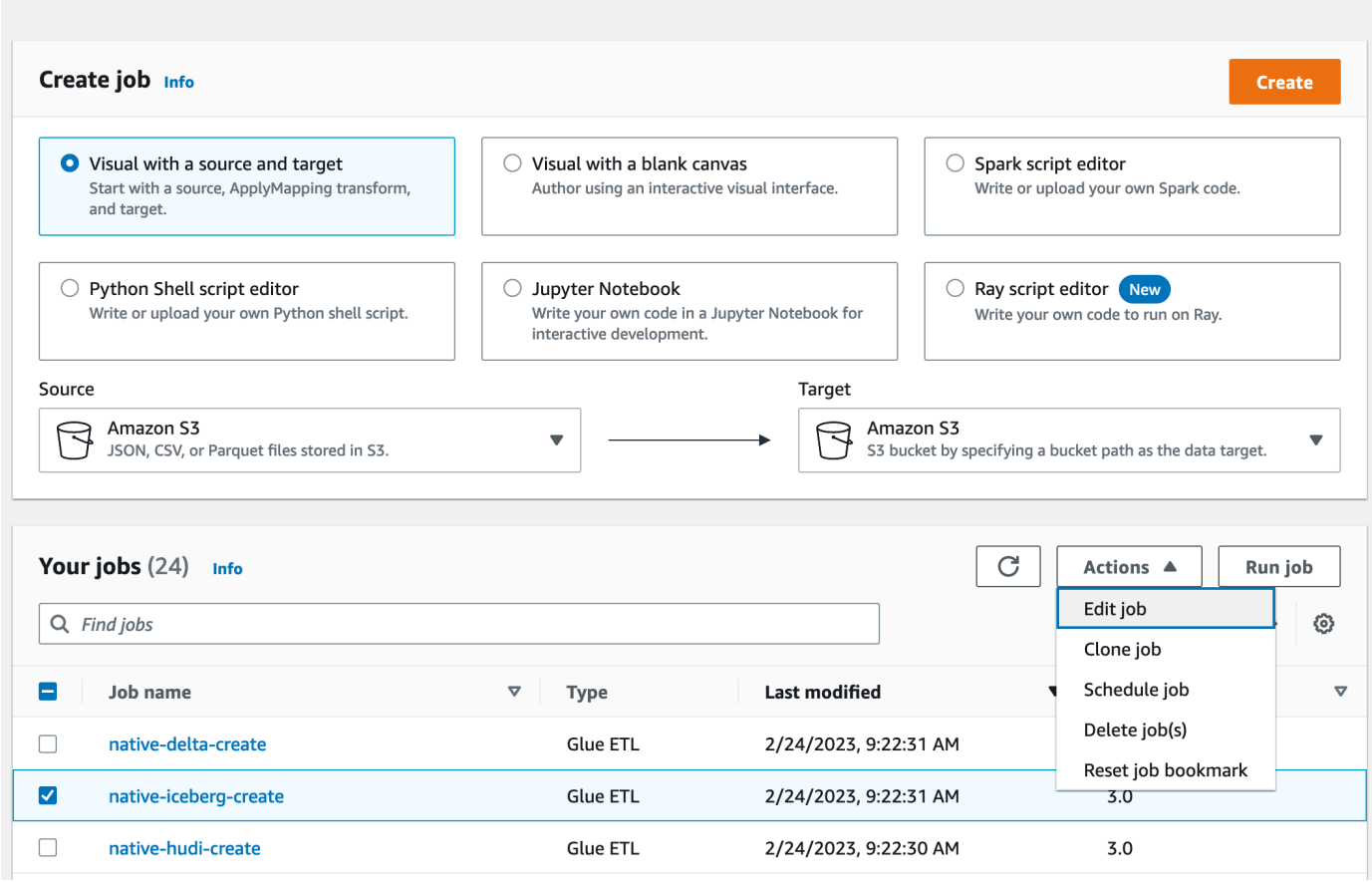
Paso 2: Configurar los permisos para una tabla de Iceberg

En esta sección, aprenderá a crear una tabla de Iceberg en AWS Glue Data Catalog, configurar permisos de datos en AWS Lake Formation y consultar datos con Amazon Athena.

Para crear una tabla de Iceberg

En este paso, ejecutará un trabajo AWS Glue que creará una tabla de transacciones de Iceberg en el Catálogo de datos.

1. Abra la consola de AWS Glue en https://console.aws.amazon.com/glue/ en la región Este de EE. UU. (Norte de Virginia) como usuario administrador del lago de datos.
2. En el panel de navegación izquierdo, seleccione trabajos.
3. Seleccione `native-iceberg-create`.



4. En Acciones, seleccione Editar trabajo.
5. En Detalles del trabajo, expanda Propiedades avanzadas y marque la casilla junto a Usar AWS Glue Data Catalog como metaalmacén de Hive para añadir los metadatos de la tabla en AWS Glue Data Catalog. Esto especifica AWS Glue Data Catalog como metaalmacén de los recursos del Catálogo de datos utilizados en el trabajo y facilita que los permisos de Lake Formation se apliquen después sobre los recursos del catálogo.
6. Seleccione Guardar.
7. Elija Ejecutar. Puede ver el estado del trabajo mientras se ejecuta.

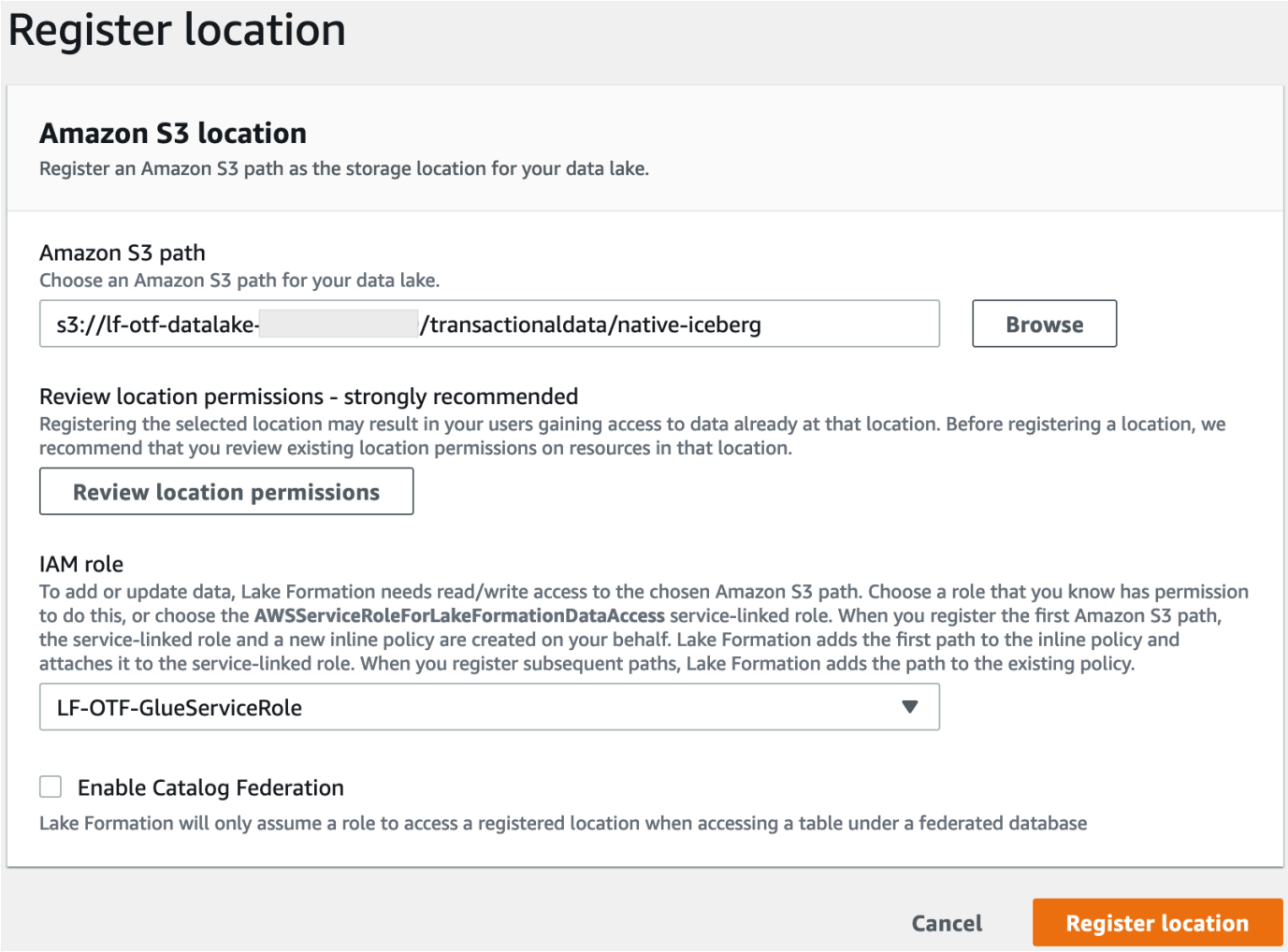
   Para obtener más información acerca de trabajos de AWS Glue, consulte el apartado sobre operar con trabajos en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue.

Este trabajo crea una tabla de Iceberg de nombre `product` en la base de datos `lficebergdb`. Verifique la tabla de productos en la consola de Lake Formation.

Para registrar la ubicación de los datos con Lake Formation

A continuación, registre una ruta de Amazon S3 como ubicación de su lago de datos.

1. Abra la consola de Lake Formation como administrador del lago de datos en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, en Registrar e ingerir, seleccione Ubicación de los datos.
3. En la parte superior derecha de la consola, seleccione Registrar ubicación.
4. En la página Registrar ubicación, introduzca:
   - Ruta de Amazon S3: seleccione Examinar y elija `lf-otf-datalake-123456789012`. Haga clic en la flecha derecha (>) junto a la ubicación raíz de Amazon S3 para ir a la ubicación `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-iceberg`.
   - Rol de IAM: seleccione `LF-OTF-RegisterRole` como rol de IAM.
   - Seleccione Registrar ubicación.

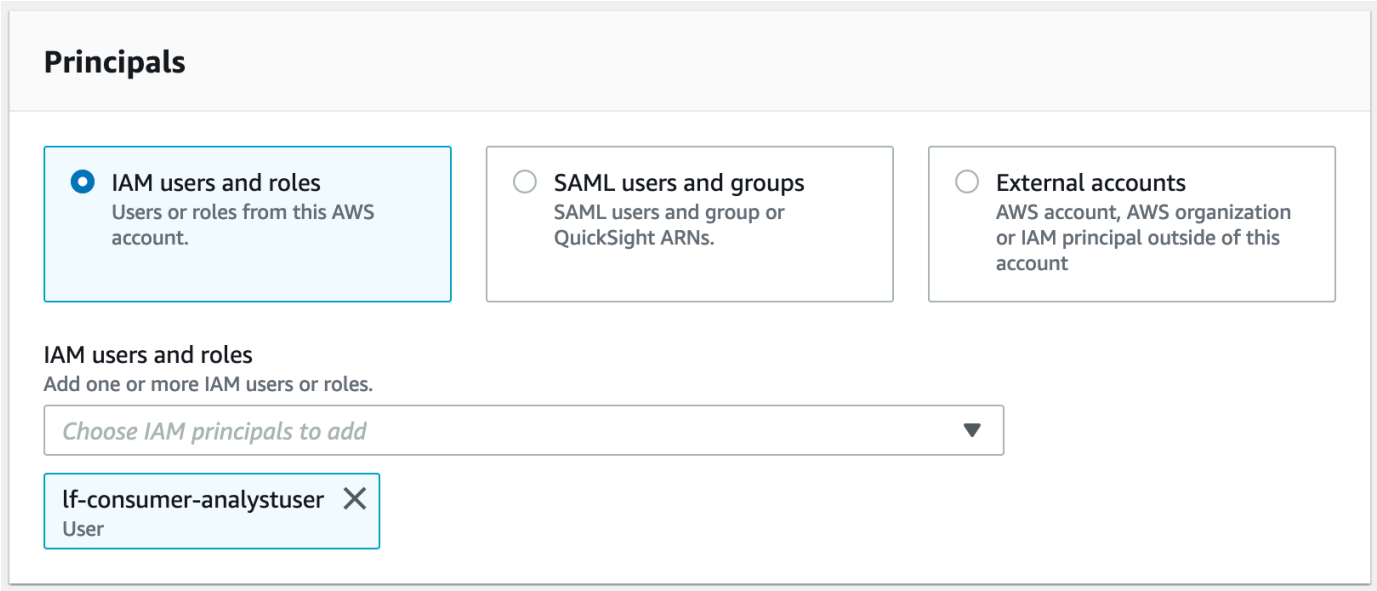


Para obtener más información sobre el registro de una ubicación de datos en Lake Formation, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

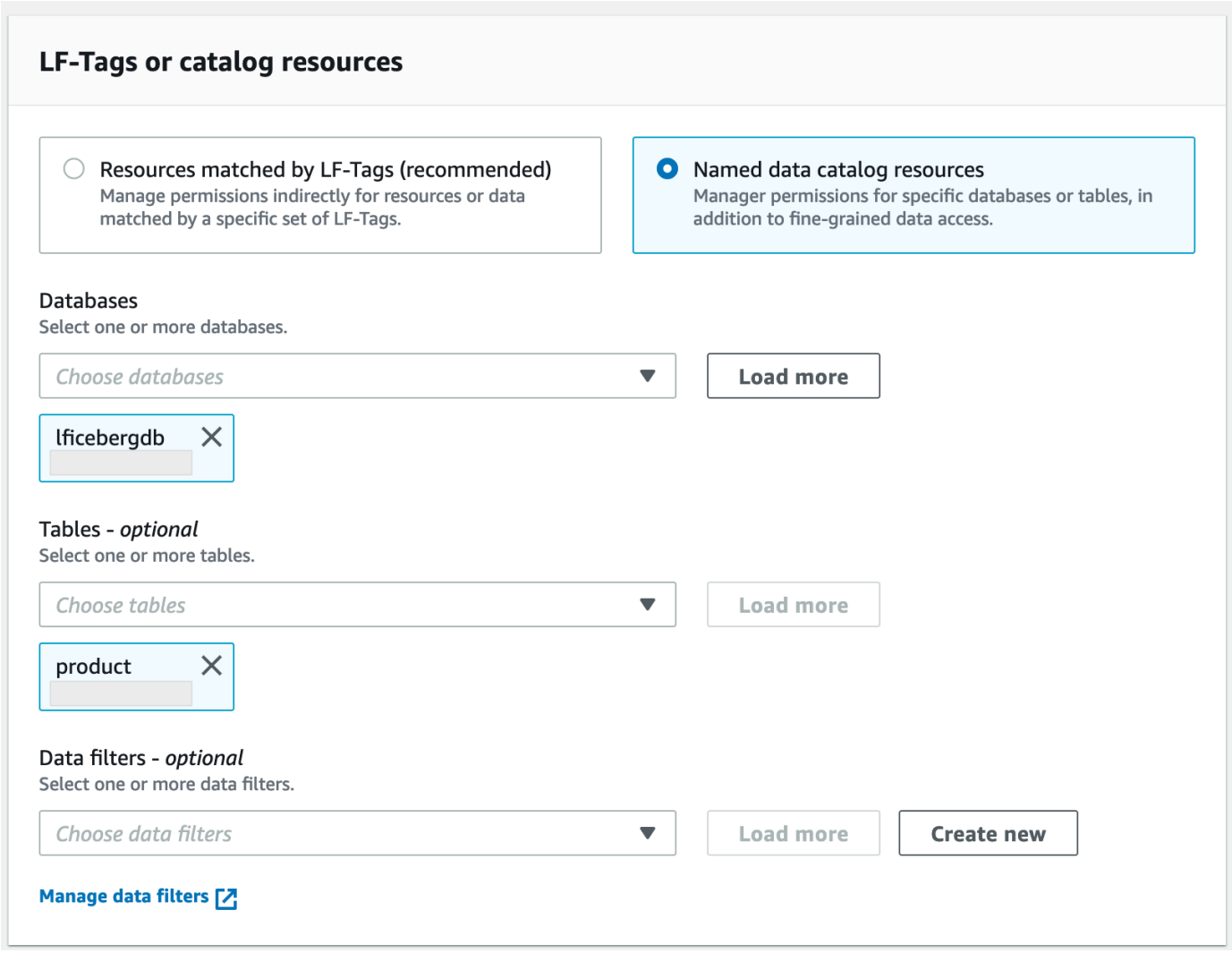
Para conceder permisos a Lake Formation en la tabla de Iceberg

En este paso, concederemos permisos de lago de datos al usuario analista de negocios.

1. En Permisos del lago de datos, seleccione Conceder.
2. En la pantalla Conceder permisos de datos, seleccione Usuarios y roles de IAM.
3. Seleccione `lf-consumer-analystuser` en la lista desplegable.



4. Seleccione un recurso de Catálogo de datos con nombre.
5. En Bases de datos, seleccione `lficebergdb`.
6. En Tablas, seleccione `product`.



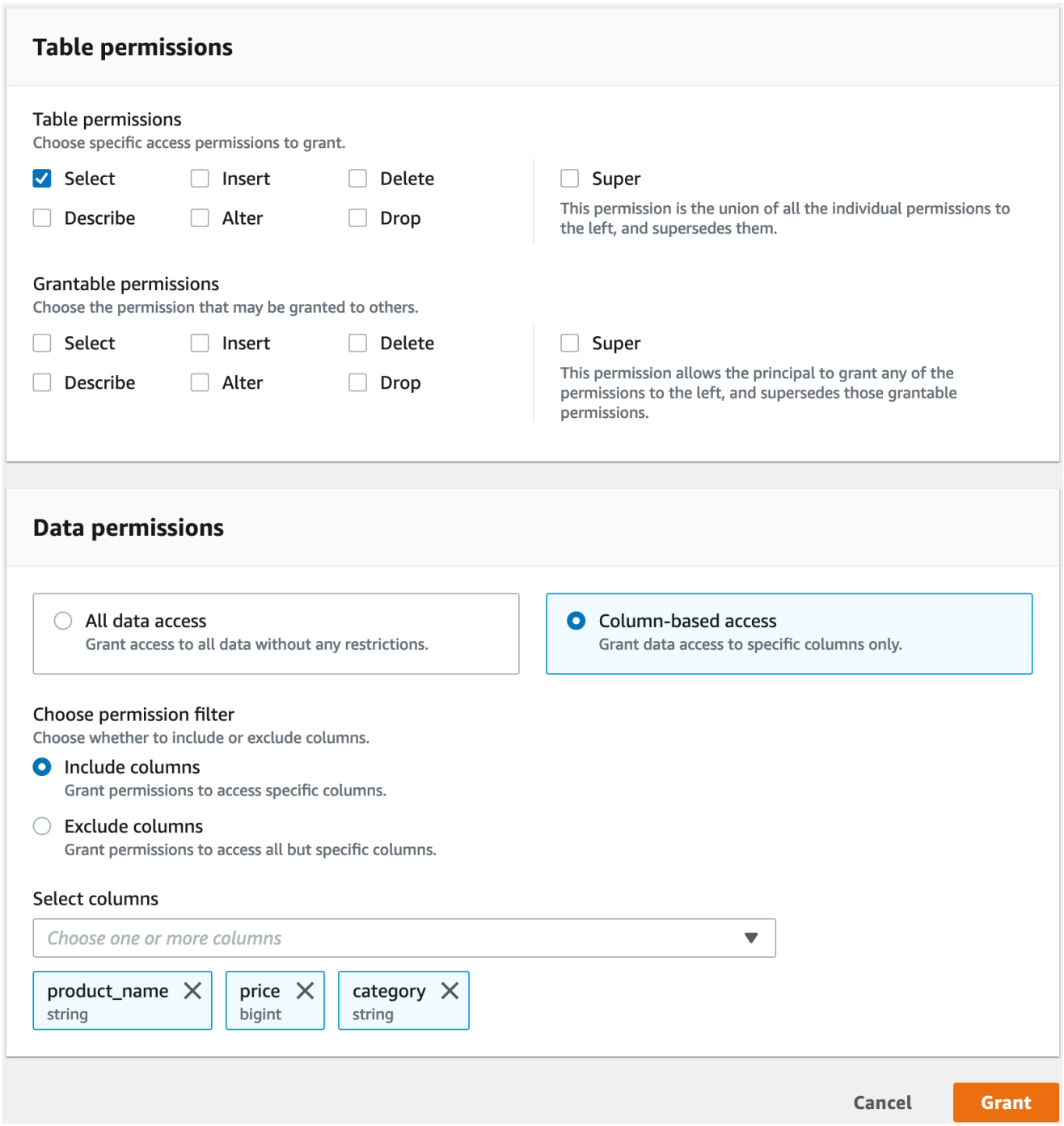
7. A continuación, puede conceder el acceso basado en columnas especificando las columnas.

   a. En Permisos de tabla, marque Seleccionar.

   b. En Permisos de datos, seleccione Acceso basado en columnas e Incluir columnas.

   c. Seleccione las columnas `product_name`, `price` y `category`.

   d. Elija Conceder.



Para consultar la tabla de Iceberg con Athena

Ahora puede empezar a consultar la tabla de Iceberg que creó con Athena. Si es la primera vez que ejecuta consultas en Athena, debe configurar una ubicación de resultados de consulta. Para obtener más información, consulte Especificación de una ubicación de resultados de consulta.

1. Cierre sesión como usuario administrador del lago de datos e inicie sesión como `lf-consumer-analystuser` en la región EE.UU. Este (Virginia del Norte) con la contraseña indicada anteriormente en el AWS CloudFormation resultado.
2. Abra la consola de Athena en https://console.aws.amazon.com/athena/.
3. Elija Configuración y seleccione Administrar.
4. En el cuadro Ubicación del resultado de la consulta, indique la ruta al bucket que creó en los resultados de AWS CloudFormation. Copie el valor de **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial-123456789012/athena-results/) y seleccione Guardar.
5. Ejecute la consulta siguiente para obtener una vista previa de los 10 registros almacenados en la tabla de Iceberg:

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

   Para obtener más información sobre la consulta de tablas de Iceberg con Athena, vaya a Consulta de tablas de Iceberg en la Guía del usuario de Amazon Athena.

Paso 3: Configurar los permisos para una tabla de Hudi

En esta sección, aprenderá a crear una tabla de Hudi en AWS Glue Data Catalog, configurar permisos de datos en AWS Lake Formation y consultar datos con Amazon Athena.

Para crear una tabla de Hudi

En este paso, ejecutará un trabajo AWS Glue que creará una tabla transaccional de Hudi en el Catálogo de datos.

1. Inicie sesión en la AWS Glue consola en https://console.aws.amazon.com/glue/ en la región EE.UU. Este (Norte de Virginia)

   como usuario administrador del lago de datos.
2. En el panel de navegación izquierdo, seleccione trabajos.
3. Seleccione `native-hudi-create`.
4. En Acciones, seleccione Editar trabajo.
5. En Detalles del trabajo, expanda Propiedades avanzadas y marque la casilla junto a Usar AWS Glue Data Catalog como metaalmacén de Hive para añadir los metadatos de la tabla en AWS Glue Data Catalog. Esto especifica AWS Glue Data Catalog como metaalmacén de los recursos

del Catálogo de datos utilizados en el trabajo y facilita que los permisos de Lake Formation se apliquen después sobre los recursos del catálogo.

6. Seleccione Guardar.
7. Elija Ejecutar. Puede ver el estado del trabajo mientras se ejecuta.

   Para obtener más información acerca de trabajos de AWS Glue, consulte el apartado sobre operar con trabajos en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue.

   Este trabajo crea una tabla de Hudi (vaca) en la base de datos: lfhudidb. Verifique la tabla `product` en la consola de Lake Formation.

Para registrar la ubicación de los datos con Lake Formation

A continuación, registre una ruta de Amazon S3 como ubicación raíz de su lago de datos.

1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ como usuario administrador del lago de datos.
2. En el panel de navegación, en Registrar e ingerir, seleccione Ubicación de los datos.
3. En la parte superior derecha de la consola, seleccione Registrar ubicación.
4. En la página Registrar ubicación, introduzca:

   - Ruta de Amazon S3: seleccione Examinar y elija `lf-otf-datalake-123456789012`. Haga clic en la flecha derecha (>) junto a la ubicación raíz de Amazon S3 para ir a la ubicación `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-hudi`.
   - Rol de IAM: seleccione `LF-OTF-RegisterRole` como rol de IAM.
   - Seleccione Registrar ubicación.

Para conceder permisos de lago de datos en la tabla de Hudi

En este paso, concederemos permisos de lago de datos al usuario analista de negocios.

1. En Permisos del lago de datos, seleccione Conceder.
2. En la pantalla Conceder permisos de datos, seleccione Usuarios y roles de IAM.
3. `lf-consumer-analystuser` en el menú desplegable.
4. Seleccione un recurso de Catálogo de datos con nombre.
5. En Bases de datos, seleccione `lfhudidb`.

6. En Tablas, seleccione `product`.
7. A continuación, puede conceder el acceso basado en columnas especificando las columnas.
   a. En Permisos de tabla, marque Seleccionar.
   b. En Permisos de datos, seleccione Acceso basado en columnas e Incluir columnas.
   c. Seleccione las columnas `product_name`, `price` y `category`.
   d. Elija Conceder.

Para consultar la tabla de Hudi con Athena

Ahora puede empezar a consultar la tabla de Hudi que creó con Athena. Si es la primera vez que ejecuta consultas en Athena, debe configurar una ubicación de resultados de consulta. Para obtener más información, consulte Especificación de una ubicación de resultados de consulta.

1. Cierre sesión como usuario administrador del lago de datos e inicie sesión como `lf-consumer-analystuser` en la región EE.UU. Este (Virginia del Norte) con la contraseña indicada anteriormente en el AWS CloudFormation resultado.
2. Abra la consola de Athena en https://console.aws.amazon.com/athena/.
3. Elija Configuración y seleccione Administrar.
4. En el cuadro Ubicación del resultado de la consulta, indique la ruta al bucket que creó en los resultados de AWS CloudFormation. Copie el valor de **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial-123456789012/athena-results/) y guárdelo.
5. Ejecute la consulta siguiente para obtener una vista previa de los 10 registros almacenados en la tabla de Hudi:

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

   Para obtener más información sobre la consulta de tablas de Hudi, vaya a Consulta de tablas de Hudi en la Guía del usuario de Amazon Athena.

Paso 4: Configurar los permisos para una tabla de Hudi

En esta sección, aprenderá a crear una tabla Delta Lake con un archivo de manifiesto de enlaces simbólicos en AWS Glue Data Catalog, a configurar los permisos de datos en AWS Lake Formation y a consultar los datos mediante Amazon Athena.

Para crear una tabla de Delta Lake.

En este paso, ejecutará un trabajo de AWS Glue para crear una tabla transaccional de Delta Lake en el Catálogo de datos.

1. Inicie sesión en la AWS Glue consola en https://console.aws.amazon.com/glue/ en la región EE.UU. Este (Norte de Virginia)

   como usuario administrador del lago de datos.
2. En el panel de navegación izquierdo, seleccione trabajos.
3. Seleccione `native-delta-create`.
4. En Acciones, seleccione Editar trabajo.
5. En Detalles del trabajo, expanda Propiedades avanzadas y marque la casilla junto a Usar AWS Glue Data Catalog como metaalmacén de Hive para añadir los metadatos de la tabla en AWS Glue Data Catalog. Esto especifica AWS Glue Data Catalog como metaalmacén de los recursos del Catálogo de datos utilizados en el trabajo y facilita que los permisos de Lake Formation se apliquen después sobre los recursos del catálogo.
6. Seleccione Guardar.
7. Seleccione Ejecutar en Acciones.

   Este trabajo crea una tabla de Delta Lake de nombre `product` en la base de datos `lfdeltadb`. Verifique la tabla `product` en la consola de Lake Formation.

Para registrar la ubicación de los datos con Lake Formation

A continuación, registre una ruta de Amazon S3 como ubicación raíz de su lago de datos.

1. Abra la consola de Lake Formation como administrador del lago de datos en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, en Registrar e ingerir, seleccione Ubicación de los datos.
3. En la parte superior derecha de la consola, seleccione Registrar ubicación.
4. En la página Registrar ubicación, introduzca:

   - Ruta de Amazon S3: seleccione Examinar y elija `lf-otf-datalake-123456789012`. Haga clic en la flecha derecha (>) junto a la ubicación raíz de Amazon S3 para ir a la ubicación `s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-delta`.
   - Rol de IAM: seleccione `LF-OTF-RegisterRole` como rol de IAM.

- Seleccione Registrar ubicación.

Para conceder permisos de lago de datos en la tabla de Delta Lake

En este paso, concederemos permisos de lago de datos al usuario analista de negocios.

1. En Permisos del lago de datos, seleccione Conceder.
2. En la pantalla Conceder permisos de datos, seleccione Usuarios y roles de IAM.
3. `lf-consumer-analystuser` en el menú desplegable.
4. Seleccione un recurso de Catálogo de datos con nombre.
5. En Bases de datos, seleccione `lfdeltadb`.
6. En Tablas, seleccione `product`.
7. A continuación, puede conceder el acceso basado en columnas especificando las columnas.

   a. En Permisos de tabla, marque Seleccionar.
   b. En Permisos de datos, seleccione Acceso basado en columnas e Incluir columnas.
   c. Seleccione las columnas `product_name`, `price` y `category`.
   d. Elija Conceder.

Para consultar la tabla de Delta Lake con Athena

Ahora puede empezar a consultar la tabla de Delta Lake que creó con Athena. Si es la primera vez que ejecuta consultas en Athena, debe configurar una ubicación de resultados de consulta. Para obtener más información, consulte Especificación de una ubicación de resultados de consulta.

1. Cierre sesión como usuario administrador del lago de datos e inicie sesión como `BusinessAnalystUser` en la región Este de EE. UU. (Norte de Virginia) con la contraseña indicada anteriormente en el resultado de AWS CloudFormation.
2. Abra la consola de Athena en https://console.aws.amazon.com/athena/.
3. Elija Configuración y seleccione Administrar.
4. En el cuadro Ubicación del resultado de la consulta, indique la ruta al bucket que creó en los resultados de AWS CloudFormation. Copie el valor de **`AthenaQueryResultLocation`** (s3://lf-otf-tutorial-123456789012/athena-results/) y guarde.
5. Ejecute la consulta siguiente para obtener una vista previa de los 10 registros almacenados en la tabla de Delta Lake:

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

Para obtener más información sobre la consulta de tablas de Delta Lake, vaya a Consulta de tablas de Delta Lake en la Guía del usuario de Amazon Athena.

Paso 5: Limpiar los recursos de AWS

Para limpiar los recursos:

Para evitar que se le cobren cargos no deseados, elimine los recursos que utilizó para este tutorial. Cuenta de AWS AWS

1. Inicie sesión en la AWS CloudFormation consola en https://console.aws.amazon.com/cloudformation como administrador de IAM.
2. Elimine la pila de formación de nube. Las tablas que ha creado se eliminan automáticamente con la pila.

Gestión de un lago de datos mediante el control de acceso basado en etiquetas de Lake Formation

Miles de clientes crean lagos de datos a escala de petabytes en AWS. Muchos de estos clientes utilizan AWS Lake Formation para crear y compartir fácilmente sus lagos de datos en toda la organización. A medida que aumenta el número de tablas y usuarios, los responsables y administradores de datos buscan formas de gestionar fácilmente los permisos en los lagos de datos a escala. El control de acceso basado en etiquetas de Lake Formation (LF-TBAC) resuelve este problema al permitir que los administradores de datos creen etiquetas LF (en función de su clasificación y ontología de datos) que luego se pueden adjuntar a los recursos.

LF-TBAC es una estrategia de autorización que define permisos basados en atributos. En Lake Formation, estos atributos se denominan etiquetas LF. Puede adjuntar etiquetas LF a los recursos del Catálogo de datos y a las entidades principales de Lake Formation. Los administradores de lagos de datos pueden asignar y revocar permisos en los recursos de Lake Formation mediante etiquetas LF. Para obtener más información, consulte Control de acceso basado en etiquetas de Lake Formation.

Este tutorial muestra cómo crear una política de control de acceso basada en etiquetas de Lake Formation utilizando un conjunto de datos AWS público. Además, muestra cómo consultar tablas, bases de datos y columnas que tienen asociadas políticas de acceso basadas en etiquetas de Lake Formation.

Puede utilizar LF-TBAC para los siguientes casos de uso:

- Tiene un gran número de tablas y entidades principales a los que el administrador del lago de datos debe conceder acceso
- Desea clasificar sus datos en función de una ontología y conceder permisos en función de la clasificación
- El administrador del lago de datos desea asignar los permisos de forma dinámica, con acoplamiento flexible

A continuación, se indican los pasos generales para configurar permisos mediante LF-TBAC:

1. El administrador de datos define la ontología de las etiquetas con dos etiquetas LF: `Confidential` y `Sensitive`. Los datos `Confidential=True` tienen controles de acceso más estrictos. Los datos `Sensitive=True` requieren un análisis específico por parte del analista.
2. El administrador de datos asigna diferentes niveles de permisos al ingeniero de datos para crear tablas con diferentes etiquetas LF.
3. El ingeniero de datos crea dos bases de datos: `tag_database` y `col_tag_database`. Todas las tablas `tag_database` están configuradas con `Confidential=True`. Todas las tablas de `col_tag_database` están configuradas con `Confidential=False`. Algunas columnas de la tabla en `col_tag_database` tienen etiquetas `Sensitive=True` para necesidades de análisis específicas.
4. El ingeniero de datos concede permiso de lectura al analista para las tablas con una condición de expresión específica `Confidential=True` y `Confidential=False`, `Sensitive=True`.
5. Con esta configuración, el analista de datos puede centrarse en hacer el análisis con los datos correctos.

Temas

- Paso 2: Registre la ubicación de sus datos, cree una ontología de etiquetas LF y conceda permisos
- Paso 3: Crear bases de datos de Lake Formation
- Paso 4: Conceder permisos de tabla
- Paso 5: Ejecutar una consulta en Amazon Athena para verificar los permisos
- Paso 6: Limpiar los recursos de AWS

Destinatarios previstos

Este tutorial está dirigido a administradores, ingenieros y analistas de datos. Cuando se trata de administrar AWS Glue Data Catalog y organizar los permisos en Lake Formation, los administradores de datos de las cuentas productoras tienen una propiedad funcional de acuerdo con las funciones que respaldan y pueden conceder acceso a varios consumidores, organizaciones externas y cuentas.

En la siguiente tabla se enumeran los roles utilizados en este tutorial:

Rol	Descripción
Administrador de datos	El usuario `lf-data-steward` tiene el acceso siguiente: • Acceso de lectura a todos los recursos en el Catálogo de datos • Puede crear etiquetas LF y asociarlas al rol de ingeniero de datos para conceder permisos a otras entidades principales
Ingeniero de datos	El usuario `lf-data-engineer` tiene el acceso siguiente: • Acceso completo de lectura, escritura y actualización a todos los recursos del Catálogo de datos • Permisos de localización de datos en el lago de datos

Rol	Descripción
	• Puede asociar etiquetas LF y vincularlas al Catálogo de datos • Puede adjuntar etiquetas LF a los recursos, lo que proporciona acceso a las entidades principales de acuerdo con cualquier política creada por los administradores de datos
Analista de datos	El usuario `lf-data-analyst` tiene el siguiente acceso: • Acceso específico a los recursos compartidos por las políticas de acceso basadas en etiquetas de Lake Formation

Requisitos previos

Antes de empezar este tutorial, debe tener una Cuenta de AWS en la que pueda iniciar sesión como usuario administrativo con los permisos correctos. Para obtener más información, consulte Completar las tareas iniciales de configuración de AWS.

En este tutorial, se supone que está familiarizado con IAM. Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Paso 1: Aprovisionar recursos

Este tutorial incluye una plantilla de AWS CloudFormation para una configuración rápida. Puede revisarla y personalizarla para adaptarla a sus necesidades. La plantilla crea tres funciones diferentes (enumeradas enDestinatarios previstos) para realizar este ejercicio y copia el nyc-taxi-data conjunto de datos en su bucket local de Amazon S3.

- Un bucket de Amazon S3.
- Los escenarios apropiados de Lake Formation
- Los recursos adecuados de Amazon EC2
- Tres roles de IAM con credenciales

Crear recursos

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation en la región Este de EE. UU. (Norte de Virginia).
2. Seleccione Lanzar pila.
3. Elija Siguiente.
4. En la sección Configuración de usuario, introduzca la contraseña para tres roles: `DataStewardUserPassword`, `DataEngineerUserPassword` y `DataAnalystUserPassword`.
5. Revise los detalles en la última página y seleccione Acepto que AWS CloudFormation podría crear recursos IAM.
6. Seleccione Crear.

   La creación de la pila puede tardar hasta cinco minutos.

> **Note**
>
> Después de completar el tutorial, puede eliminar la pila en AWS CloudFormation para evitar seguir incurriendo en cargos. Compruebe que los recursos se hayan eliminado correctamente en el estado de evento de la pila.

Paso 2: Registre la ubicación de sus datos, cree una ontología de etiquetas LF y conceda permisos

En este paso, el usuario administrador de datos define la ontología de etiquetas con dos etiquetas LF `Confidential` y`Sensitive`, además, permite a los directores de IAM específicos adjuntar etiquetas LF recién creadas a los recursos.

Registre una ubicación de datos y defina la ontología de las etiquetas LF

1. Siga el primer paso como usuario administrador de datos (`lf-data-steward`) para verificar los datos en Amazon S3 y el Catálogo de datos en Lake Formation.

   a. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ `lf-data-steward` con la contraseña utilizada al implementar la AWS CloudFormation pila.

b. En el panel de navegación, en Permisos, elija Roles y tareas administrativas.

c. Elija Agregar en la sección de administradores de Data Lake.

d. En la página Añadir administrador, para los usuarios y roles de IAM, elija el usuario`lf-data-steward`.

e. Seleccione Guardar para añadir `lf-data-steward` como administrador de Lake Formation.

2. A continuación, actualice la configuración del Catálogo de datos para usar el permiso de Lake Formation para controlar los recursos del catálogo en lugar del control de acceso basado en IAM.

a. En el panel de navegación, Administración, seleccione Configuración del Catálogo de datos.

b. Desmarque Usar solo el control de acceso de IAM para las nuevas bases de datos.

c. Desmarque Usar solo el control de acceso de IAM para las nuevas bases de datos.

d. Haga clic en Guardar.

3. A continuación, registre la ubicación de los datos para el lago de datos.

a. En el panel de navegación, bajo Administración, seleccione Ubicaciones de los lagos de datos.

b. Seleccione Registrar ubicación.

c. En la página Registrar ubicación, para la ruta de Amazon S3, introduzcas`3://lf-tagbased-demo-`*`Account-ID`*.

d. En rol de IAM, deje el valor predeterminado `AWSServiceRoleForLakeFormationDataAccess` como está.

e. Elija Lake Formation como modo de permiso.

f. Elija Registrar ubicación.

4. A continuación, cree la ontología definiendo una etiqueta LF.

a. En Permisos, en el panel de navegación, elija etiquetas L y permisos. .

b. Seleccione Añadir etiqueta LF.

c. En Clave, escriba `Confidential`.

d. En Valores, añada `True` y `False`.

e. Elija Añadir etiquetas LF.

f. Repita los pasos para crear la etiqueta LF `Sensitive` con el valor. `True`

Ha creado todas las etiquetas L necesarias para este ejercicio.

Conceder permisos a los usuarios de IAM

1. A continuación, otorgue a entidades principales específicas de IAM la capacidad de adjuntar etiquetas LF recién creadas a los recursos.

   a. En Permisos, en el panel de navegación, elija etiquetas L y permisos.

   b. En la sección de permisos con etiquetas LF, selecciona Otorgar permisos.

   c. En el tipo de permiso, selecciona permisos de par clave-valor con etiqueta LF.

   d. Seleccione Roles y usuarios de IAM.

   e. En Roles y usuarios de IAM, busque y elija el rol `lf-data-engineer`.

   f. En la sección Etiquetas LF, añada la clave `Confidential` con valores `True` y y `False` la clave con valor. `key Sensitive True`

   g. En Permisos, selecciona Describir y asociar para ver los permisos y los permisos concedibles.

   h. Elija Conceder.

2. A continuación, conceda permisos a `lf-data-engineer` para crear bases de datos en nuestro Catálogo de datos y en el bucket de Amazon S3 subyacente creado por AWS CloudFormation.

   a. En Administración, en el panel de navegación, selecciona Funciones y tareas administrativas.

   b. En la sección Creadores de bases de datos, elija Conceder.

   c. En Usuarios de AIM y roles, elija el rol `lf-data-engineer`.

   d. En Permisos del catálogo, seleccione Crear base de datos.

   e. Elija Conceder.

3. A continuación, conceda permisos en el bucket (`s3://lf-tagbased-demo-`*`Account-ID`*`)` de Amazon S3 al usuario `lf-data-engineer`.

   a. En el panel de navegación, bajo Permisos, seleccione Ubicaciones de datos.

   b. Elija Conceder.

   c. Seleccione Mi cuenta.

   d. En Usuarios de AIM y roles, elija el rol `lf-data-engineer`.

e. En Ubicaciones de almacenamiento, introduzca el bucket de Amazon S3 creado por la plantilla (`s3://lf-tagbased-demo-`*`Account-ID`*) de AWS CloudFormation.

f. Elija Conceder.

4. A continuación, conceda permisos `lf-data-engineer` concedibles sobre los recursos asociados a la expresión LF-tag. `Confidential=True`

   a. En el panel de navegación, en Permisos, seleccione Permisos de lago de datos.

   b. Elija Conceder.

   c. Seleccione Roles y usuarios de IAM.

   d. Elija el rol `lf-data-engineer`.

   e. En la sección de etiquetas LF o recursos del catálogo, seleccione los recursos que coincidan con las etiquetas LF.

   f. Selecciona Añadir par clave-valor de etiqueta LF.

   g. Añada la clave `Confidential` con los valores `True`.

   h. En la sección Permisos de base de datos, seleccione Describir en Permisos de bases de datos y Permisos concedibles.

   i. En la sección Permisos de tabla, selecciona Describir, Seleccionar y Modificar tanto para los permisos de tabla como para los permisos concedibles.

   j. Elija Conceder.

5. A continuación, conceda permisos `lf-data-engineer` concedibles sobre los recursos asociados a la expresión de etiqueta LF. `Confidential=False`

   a. En el panel de navegación, en Permisos, seleccione Permisos de lago de datos.

   b. Elija Conceder.

   c. Seleccione Roles y usuarios de IAM.

   d. Elija el rol `lf-data-engineer`.

   e. Seleccione Recursos que coincidan con las etiquetas LF.

   f. Elija Añadir etiquetas LF.

   g. Añada la clave `Confidential` con los valores `False`.

   h. En la sección Permisos de base de datos, seleccione Describir en Permisos de bases de datos y Permisos concedibles.

   i. En la sección Permisos de tabla y columna, no selecciones nada.

   j. Elija Conceder.

6. A continuación, concedemos permisos **`lf-data-engineer`** concedibles sobre los recursos asociados a los pares clave-valor de la etiqueta LF y. `Confidential=False Sensitive=True`

   a. En el panel de navegación, en Permisos, seleccione Permisos de datos.

   b. Elija Conceder.

   c. Seleccione Roles y usuarios de IAM.

   d. Elija el rol `lf-data-engineer`.

   e. En la sección de etiquetas LF o recursos del catálogo, selecciona Recursos que coincidan con etiquetas LF.

   f. Seleccione Añadir etiqueta LF.

   g. Añada la clave `Confidential` con los valores `False`.

   h. Selecciona Añadir par clave-valor de etiqueta LF.

   i. Añada la clave `Sensitive` con los valores `True`.

   j. En la sección Permisos de base de datos, seleccione Describir en Permisos de bases de datos y Permisos concedibles.

   k. En la sección Permisos de tabla, selecciona Describir, Seleccionar y Modificar tanto para los permisos de tabla como para los permisos concedibles.

   l. Elija Conceder.

Paso 3: Crear bases de datos de Lake Formation

En este paso, creará dos bases de datos y adjuntará etiquetas LF a las bases de datos y columnas específicas con fines de prueba.

Cree sus bases de datos y su tabla para el acceso a nivel de base de datos

1. En primer lugar, cree la base de datos`tag_database`, la tabla `source_data` y adjunte las etiquetas LF adecuadas.

   a. En la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), en Catálogo de datos, elija Databases.

   b. Elija Crear base de datos.

   c. En Nombre, ingrese `tag_database`.

d. En Ubicación, introduzca la ubicación de Amazon S3 creada por la plantilla (`s3://lf-tagbased-demo-Account-ID/tag_database/`) de AWS CloudFormation.

e. Desmarque Usar solo el control de acceso de IAM para las nuevas tablas de esta base de datos.

f. Elija Crear base de datos.

2. A continuación, cree una nueva tabla dentro de `tag_database`.

a. En la página Bases de datos, seleccione la base de datos `tag_database`.

b. Seleccione Ver tablas y haga clic en Crear tabla.

c. En Nombre, ingrese `source_data`.

d. En Base de datos elija la base de datos `tag_database`.

e. En Formato de tabla, elija AWS GlueTabla estándar.

f. En la sección Los datos se encuentran en, elija Ruta especificada en otra cuenta.

g. En Incluir ruta, introduzca la ruta a `tag_database` creada por la plantilla (`s3://lf-tagbased-demoAccount-ID/tag_database/`) de AWS CloudFormation.

h. En Formato de datos, seleccione CSV.

i. En Esquema de carga, introduzca la matriz JSON de estructura de columnas siguiente para crear un esquema:

```
[
              {
                   "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. Seleccione Cargar. Tras cargar el esquema, el esquema de la tabla será similar a la siguiente captura de pantalla:

#	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. Elija Enviar.

3. A continuación, adjunte etiquetas LF a nivel de base de datos.

   a. En la página Bases de datos, busque y seleccione `tag_database`.
   b. En el menú Acciones, seleccione Editar etiquetas LF.
   c. Seleccione Asignar una nueva etiqueta LF.
   d. En Llaves asignadas, elige la `Confidential` etiqueta LF que creaste anteriormente.
   e. En Valores, elija `True`.
   f. Seleccione Guardar.

   Esto completa la asignación de la etiqueta LF a la base de datos tag_database.

Cree su base de datos y su tabla para el acceso a nivel de base de datos

Repita los pasos siguientes para crear la base de datos y la tabla`source_data_col_lvl`, `col_tag_database` y adjunte las etiquetas LF a nivel de columna.

1. En la página Bases de datos, seleccione Crear base de datos.
2. En Nombre, ingrese `col_tag_database`.
3. En Ubicación, introduzca la ubicación de Amazon S3 creada por la plantilla (`s3://lf-tagbased-demo-`*`Account-ID`*`/col_tag_database/`) de AWS CloudFormation.
4. Desmarque Usar solo el control de acceso de IAM para las nuevas tablas de esta base de datos.
5. Elija Crear base de datos.
6. En la página Bases de datos, seleccione la nueva base de datos (`col_tag_database`).
7. Seleccione Ver tablas y haga clic en Crear tabla.
8. En Nombre, ingrese `source_data_col_lvl`.
9. En Base de datos, elija su nueva base de datos (`col_tag_database`).
10. En Formato de tabla, elija AWS GlueTabla estándar.
11. En Los datos se encuentran en, elija Ruta especificada en otra cuenta.
12. Introduzca la ruta de Amazon S3 para `col_tag_database` (`s3://lf-tagbased-demo-`*`Account-ID`*`/col_tag_database/`).
13. En Formato de datos, seleccione `CSV`.
14. En `Upload schema`, introduzca el siguiente esquema JSON:

```
[
                {
                     "Name": "vendorid",
                     "Type": "string"


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


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


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


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


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


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


                },
```

```
                {
                  "Name": "passenger_count",
                  "Type": "string"


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


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


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


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


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


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


               },
               {
                  "Name": "ehail_fee",
```

```
                    "Type": "string"


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


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


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


               }
]
```

15. Elija `Upload`. Tras cargar el esquema, el esquema de la tabla será similar a la siguiente captura de pantalla:

#	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. Seleccione Enviar para completar la creación de la tabla.
17. Ahora, asocie la `Sensitive=True` etiqueta LF a las columnas `vendorid` y. `fare_amount`

    a. En la página Tablas, seleccione la tabla que ha creado (`source_data_col_lvl`).

    b. En el menú Acciones, selecciona Esquema.

    c. Seleccione la columna `vendorid` y elija Editar etiquetas LF.

    d. En Teclas asignadas, seleccione Sensible.

    e. En Valores, elija Verdadero.

    f. Seleccione Guardar.
18. A continuación, asocie la etiqueta `Confidential=False` LF a. `col_tag_database` Esto es necesario `lf-data-analyst` para poder describir la base de datos `col_tag_database` cuando se inicia sesión desde. Amazon Athena

    a. En la página Bases de datos, busque y seleccione `col_tag_database`.

    b. En el menú Acciones, seleccione Editar etiquetas LF.

    c. Seleccione Asignar una nueva etiqueta LF.

    d. En Llaves asignadas, elige la `Confidential` etiqueta LF que creaste anteriormente.

    e. En Valores, elija `False`.

    f. Seleccione Guardar.

Paso 4: Conceder permisos de tabla

Conceda permisos a los analistas de datos para el uso de las bases de datos `tag_database` y la tabla `col_tag_database` utilizando etiquetas LF `Confidential` y `Sensitive`.

1. Siga estos pasos para conceder permisos al `lf-data-analyst` usuario sobre los objetos asociados a la etiqueta LF `Confidential=True` (Database:TAG_Database) para tener la base de datos y permisos sobre las tablas. `Describe` `Select`

    a. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ as`lf-data-engineer`.

    b. En Permisos, selecciona Permisos de Data Lake.

    c. Elija Conceder.

    d. En Entidades principales, seleccione Roles y usuarios de IAM.

e. En Roles y usuarios de IAM, elija `lf-data-analyst`.

f. En Etiquetas LF o recursos del catálogo, selecciona Recursos que coincidan con etiquetas LF.

g. Elija Añadir etiquetas LF.

h. En Clave, elija `Confidential`.

i. En Valores, elija `True`.

j. En Permisos de base de datos, seleccione `Describe`.

k. En Permisos de tabla, elija Seleccionar y Describir.

l. Elija Conceder.

2. A continuación, repita los pasos para conceder permisos a los analistas de datos para la expresión de etiquetas LF. `Confidential=False` Esta Etiqueta LF se utiliza para describir la tabla `col_tag_database` y la tabla `source_data_col_lvl` cuando se inicia sesión como `lf-data-analyst` desde Amazon Athena.

   a. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ as`lf-data-engineer`.

   b. En la página Bases de datos, seleccione la base de datos `col_tag_database`.

   c. Elija Acciones y Conceder.

   d. En Entidades principales, seleccione Roles y usuarios de IAM.

   e. En Roles y usuarios de IAM, elija `lf-data-analyst`.

   f. Seleccione los recursos que coincidan con las etiquetas LF.

   g. Seleccione Añadir etiqueta LF.

   h. En Clave, elija `Confidential`.

   i. En Valores, elija `False`.

   j. En Permisos de base de datos, seleccione `Describe`.

   k. En Permisos de tabla, no seleccione nada.

   l. Elija Conceder.

3. A continuación, repita los pasos para conceder permisos a los analistas de datos para la expresión de etiquetas LF para y. `Confidential=False Sensitive=True` Esta Etiqueta LF se utiliza para describir la tabla `col_tag_database` y la tabla `source_data_col_lvl` (nivel de columna) cuando se inicia sesión como `lf-data-analyst` desde Amazon Athena.

a. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ como `lf-data-engineer`.

b. En la página Bases de datos, seleccione la base de datos `col_tag_database`.

c. Elija Acciones y Conceder.

d. En Entidades principales, seleccione Roles y usuarios de IAM.

e. En Roles y usuarios de IAM, elija `lf-data-analyst`.

f. Seleccione los recursos que coincidan con las etiquetas LF.

g. Elija Añadir etiquetas LF.

h. En Clave, elija `Confidential`.

i. En Valores, elija `False`.

j. Elija Añadir etiquetas LF.

k. En Clave, elija `Sensitive`.

l. En Valores, elija `True`.

m. En Permisos de base de datos, seleccione `Describe`.

n. En Permisos de tabla, seleccione `Select` y `Describe`.

o. Elija Conceder.

Paso 5: Ejecutar una consulta en Amazon Athena para verificar los permisos

En este paso, utilice Amazon Athena para ejecutar consultas `SELECT` en las dos tablas (`source_data and source_data_col_lvl`). Utilice la ruta de Amazon S3 como ubicación de los resultados de la consulta (`s3://lf-tagbased-demo-`*`Account-ID`*`/athena-results/`).

1. Abra la consola de Athena en https://console.aws.amazon.com/athena/ como `lf-data-analyst`.
2. En el editor de consultas de Athena, selecciona `tag_database` en el panel izquierdo.
3. Seleccione el icono de opciones de menú adicionales (tres puntos verticales) situado junto a `source_data` y elija Vista previa de la tabla.
4. Elija Ejecutar consulta.

La consulta tardará unos minutos en ejecutarse. La consulta muestra todas las columnas de la salida porque la etiqueta LF está asociada a nivel de base de datos y la tabla `source_data` heredó automáticamente la `LF-tag` de la base de datos `tag_database`.

5. Ejecute otra consulta con `col_tag_database` y `source_data_col_lvl`.

   La segunda consulta devuelve las dos columnas que estaban etiquetadas como `Non-Confidential` y `Sensitive`.

6. También puede comprobar el comportamiento de la política de acceso basada en etiquetas de Lake Formation en las columnas para las que no tiene concesiones de política. Cuando se selecciona una columna sin etiquetar de la tabla `source_data_col_lvl`, Athena devuelve un error. Por ejemplo, puede ejecutar la siguiente consulta para elegir columnas `geolocationid` sin etiquetar:

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

Paso 6: Limpiar los recursos de AWS

Para evitar cargos no deseados en su Cuenta de AWS, elimine los recursos de AWS que ha usado en este tutorial.

1. Inicie sesión en la consola de Lake Formation como `lf-data-engineer` y borre las bases de datos `tag_database` y `col_tag_database`.
2. Luego, inicie sesión como `lf-data-steward` y borre todos los Permisos de etiqueta L, los Permisos de datos y los Permisos de ubicación de datos concedidos antes `lf-data-engineer` y `lf-data-analyst`..
3. Inicie sesión en la consola de Amazon S3 como propietario de la cuenta con las credenciales de IAM que utilizó para implementar la pila AWS CloudFormation.
4. Elimine los buckets siguientes:

   - lf-tagbased-demo-accesslogs*- identificador de cuenta*
   - lf-tagbased-demo*- identificador de cuenta*

5. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation y borre la pila que había creado. Espera a que el estado de la pila cambie a. `DELETE_COMPLETE`

Protección de los lagos de datos con control de acceso a nivel de fila

Los permisos a nivel de fila de AWS Lake Formation le permiten proporcionar acceso a filas específicas de una tabla en función de las políticas de control y cumplimiento de los datos. Si tiene tablas grandes que almacenan miles de millones de registros, necesita una forma de permitir que los diferentes usuarios y equipos accedan solo a los datos que pueden ver. El control de acceso a nivel de fila es una forma sencilla y eficaz de proteger los datos, a la vez que permite a los usuarios acceder a los datos que necesitan para desarrollar su trabajo. Lake Formation proporciona informes centralizados de auditoría y cumplimiento al identificar qué entidades principales accedieron a qué datos, cuándo lo hicieron y a través de qué servicios.

En este tutorial aprenderá cómo funcionan los controles de acceso a nivel de fila en Lake Formation y cómo configurarlos.

Este tutorial incluye una plantilla AWS CloudFormation para configurar rápidamente los recursos necesarios. Puede revisarla y personalizarla para adaptarla a sus necesidades.

Temas

Destinatarios previstos

Este tutorial está dirigido a administradores, ingenieros y analistas de datos. En la siguiente tabla se enumeran los roles y responsabilidades de un propietario y un consumidor de datos.

Rol	Descripción
Administrador de IAM	Usuario que puede crear usuarios, roles y buckets de Amazon Simple Storage Service

Rol	Descripción
	(Amazon S3). Tiene la política administrada `AdministratorAccess` de AWS.
Administrador de lagos de datos	Usuario responsable de configurar el lago de datos, crear filtros de datos y conceder permisos a los analistas de datos.
Analista de datos	Usuario que puede ejecutar consultas en el lago de datos. Los analistas de datos que residen en diferentes países (en nuestro caso, EE. UU. y Japón) solo pueden analizar las reseñas de productos de clientes ubicados en su propio país y, por motivos de cumplimiento, no deberían poder ver los datos de clientes ubicados en otros países.

Requisitos previos

Antes de empezar este tutorial, debe tener una Cuenta de AWS en la que pueda iniciar sesión como usuario administrativo con los permisos correctos. Para obtener más información, consulte Completar las tareas iniciales de configuración de AWS.

En este tutorial, se asume que conoce IAM. Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Cambiar la configuración de Lake Formation

> **Important**
>
> Antes de lanzar la plantilla de AWS CloudFormation, desactive la opción Usar solo el control de acceso de IAM para nuevas bases de datos o tablas en Lake Formation siguiendo los pasos indicados a continuación:

1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ en la región Este de EE. UU. (Norte de Virginia) o la región Oeste de EE. UU. (Oregón).

2. En Catálogo de datos, seleccione Configuración.
3. Deseleccione Usar solo el control de acceso de IAM para nuevas bases de datos y Usar solo el control de acceso de IAM para nuevas tablas en las nuevas bases de datos.
4. Seleccione Guardar.

Paso 1: Aprovisionar recursos

Este tutorial incluye una plantilla de AWS CloudFormation para una configuración rápida. Puede revisarla y personalizarla para adaptarla a sus necesidades. La plantilla AWS CloudFormation incluye los recursos siguientes:

- Usuarios y políticas para:
  - DataLakeAdmin
  - DataAnalystUS
  - DataAnalystJP
- Configuración y permisos del lago de datos de Lake Formation
- Una función de Lambda (para recursos personalizados AWS CloudFormation respaldados por Lambda) que se utiliza para copiar archivos de datos de muestra del bucket público de Amazon S3 al bucket de Amazon S3
- Un bucket de Amazon S3 que sirva como nuestro lago de datos
- Una base de datos AWS Glue Data Catalog, tabla y partición

Crear recursos

Siga estos pasos para crear sus recursos con la plantilla AWS CloudFormation.

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation en la región Este de EE. UU. (Norte de Virginia).
2. Elija Lanzar pila.
3. Elija Siguiente en la pantalla Crear pila.
4. Introduzca Nombre de pila.
5. Para DataLakeAdminUserName y DataLakeAdminUserPassword, introduzca su nombre de usuario de IAM y la contraseña del usuario administrador del lago de datos.

6. Para DataAnalystusUsUserName y DataAnalystUsUserPassword, introduzca el nombre de usuario y la contraseña del nombre de usuario y la contraseña que desee para el usuario analista de datos responsable del mercado estadounidense.
7. Para DataAnalystusJpUserName y DataAnalystJpUserPassword, introduzca el nombre de usuario y la contraseña del nombre de usuario y la contraseña que desee para el usuario analista de datos responsable del mercado japonés.
8. En DataLakeBucketName, escriba el nombre de su bucket de datos.
9. Para DatabaseName y TableName, deje el valor predeterminado.
10. Elija Siguiente.
11. En la siguiente página, elija Siguiente.
12. Revise los detalles en la última página y seleccione Acepto que AWS CloudFormation podría crear recursos IAM.
13. Seleccione Create (Crear).

    La creación de la pila puede tardar un minuto en completarse.

Paso 2: Consultar sin filtros de datos

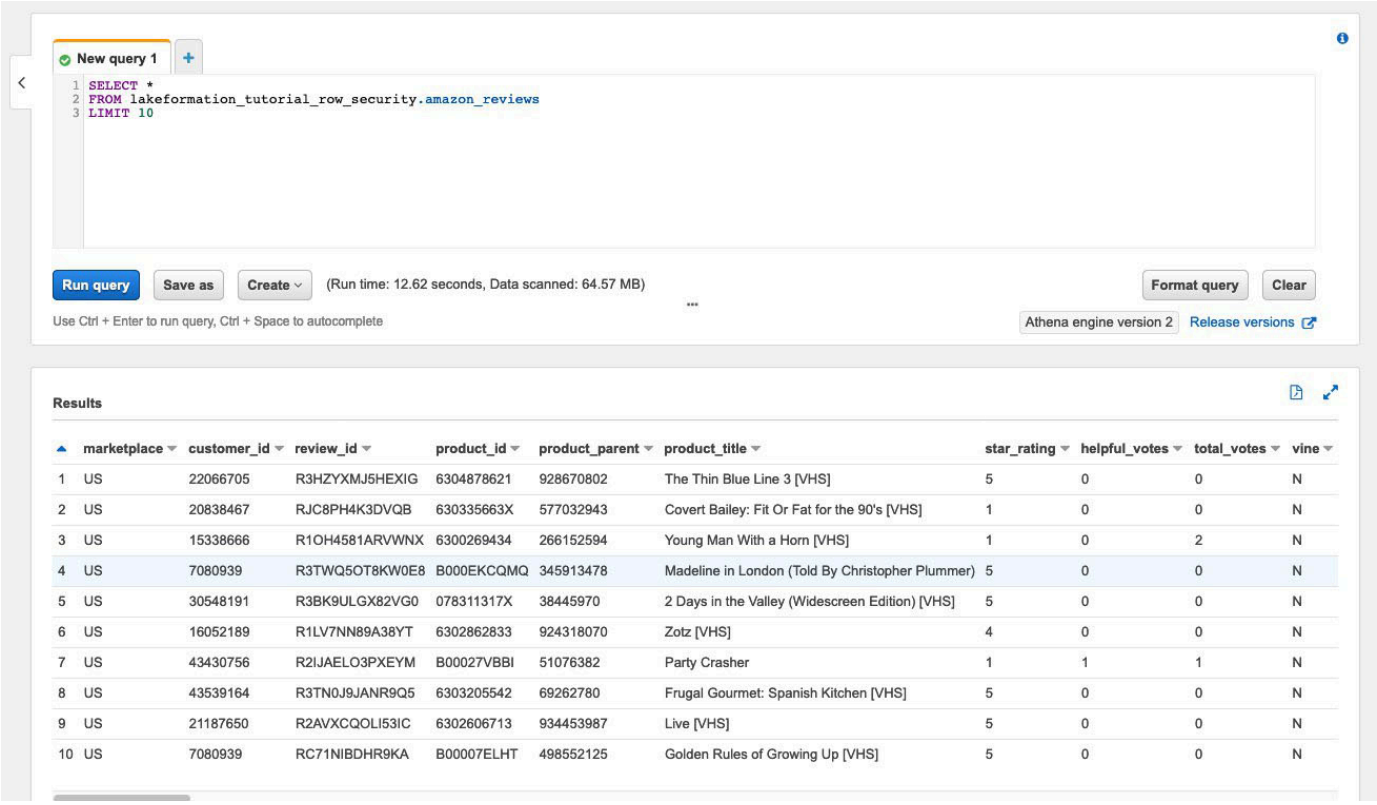
Una vez configurado el entorno, puede consultar la tabla de reseñas de productos. Primero consulte la tabla sin controles de acceso a nivel de fila para asegurarse de que puede ver los datos. Si es la primera vez que ejecuta consultas en Amazon Athena, debe configurar la ubicación de los resultados de la consulta.

Consulte la tabla sin control de acceso a nivel de fila

1. Inicie sesión en la consola de Athena en https://console.aws.amazon.com/athena/ como usuario `DatalakeAdmin` y ejecute la consulta siguiente:

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

   La captura de pantalla siguiente muestra el resultado de la consulta. Esta tabla solo tiene una partición, `product_category=Video`, por lo que cada registro es un comentario de revisión de un producto de vídeo.

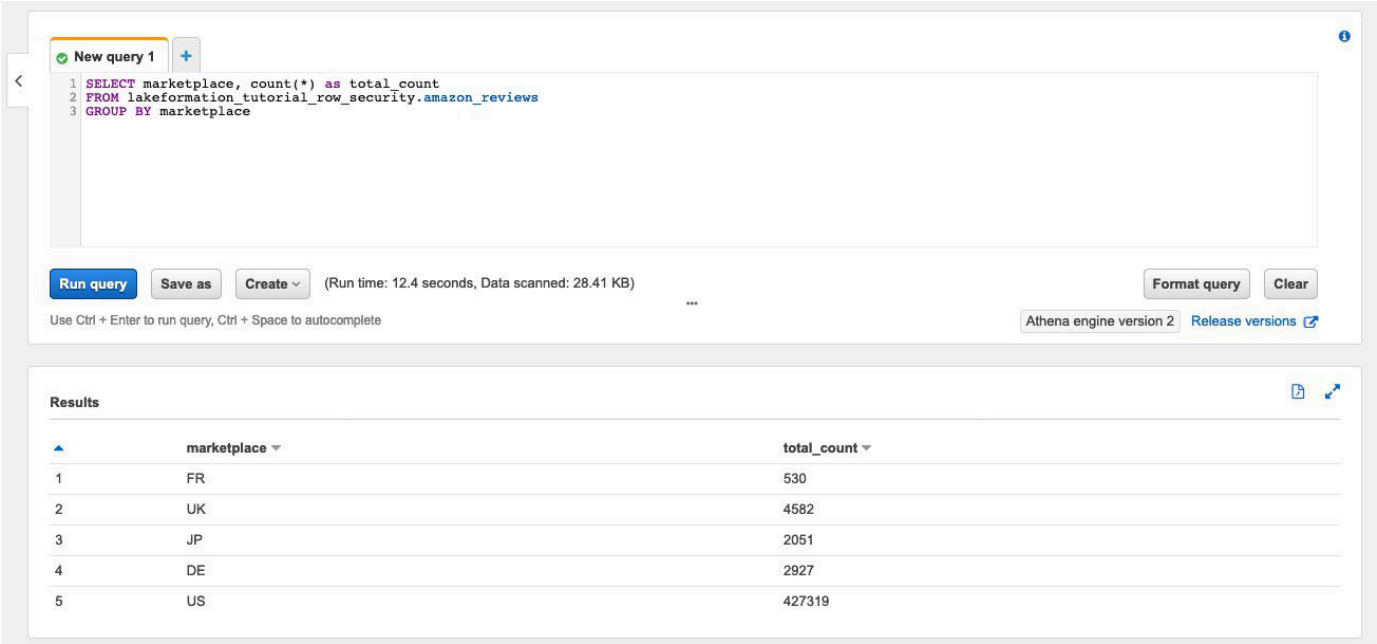


	marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine
1	US	22066705	R3HZYXMJ5HEXIG	6304878621	928670802	The Thin Blue Line 3 [VHS]	5	0	0	N
2	US	20838467	RJC8PH4K3DVQB	630335663X	577032943	Covert Bailey: Fit Or Fat for the 90's [VHS]	1	0	0	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. A continuación, ejecute una consulta de agregación para recuperar el número total de registros por cada `marketplace`.

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

La captura de pantalla siguiente muestra el resultado de la consulta. La columna `marketplace` tiene cinco valores diferentes. En los pasos siguientes, configurará los filtros basados en filas utilizando la columna `marketplace`.



Paso 3: Configurar los filtros de datos y conceder permisos

En este tutorial se utilizan dos analistas de datos, uno responsable del mercado estadounidense y otro del mercado japonés. Cada analista utiliza Athena para analizar las opiniones de los clientes únicamente para su mercado específico. Cree dos filtros de datos diferentes, uno para el analista responsable del mercado estadounidense y otro para el responsable del mercado japonés. A continuación, conceda a los analistas sus correspondiente permisos.

Cree filtros de datos y conceda permisos

1. Cree un filtro para restringir el acceso a los datos del `marketplace` US.

   a. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ en la región Este de EE. UU. (Norte de Virginia) como usuario `DatalakeAdmin`.

   b. Elija Filtros de datos.

   c. Elija Crear filtro.

   d. En Nombre del filtro de datos, introduzca `amazon_reviews_US`.

   e. En Base de datos de destino elija la base de datos `lakeformation_tutorial_row_security`.

   f. En Tabla de objetivos, elija la tabla `amazon_reviews`.

g. En Acceso a nivel de columna, deje el valor predeterminado.

h. En Expresión de filtro de filas, introduzca `marketplace='US'`.

i. Elija Crear filtro.

2. Cree un filtro para restringir el acceso a los datos del `marketplace` japonés.

a. En la página Filtros de datos, seleccione Crear filtro nuevo.

b. En Nombre del filtro de datos, introduzca `amazon_reviews_JP`.

c. En Base de datos de destino elija la base de datos `lakeformation_tutorial_row_security`.

d. En Tabla de destino, elija la tabla `table amazon_reviews`.

e. En Acceso a nivel de columna, deje el valor predeterminado.

f. En Expresión de filtro de filas, introduzca `marketplace='JP'`.

g. Elija Crear filtro.

3. A continuación, conceda permisos a los analistas de datos que utilizan estos filtros de datos. Siga estos pasos para conceder permisos al analista de datos estadounidense (`DataAnalystUS`):

a. En Permisos, elija Permisos de lago de datos.

b. En Permiso de datos, seleccione Conceder.

c. En el Entidades principales, elija Usuarios y roles de IAM y seleccione el rol `DataAnalystUS`.

d. En Etiquetas LF o recursos del catálogo, elija Recursos de catálogo de datos con nombre.

e. En Database (Base de datos), elija `lakeformation_tutorial_row_security`.

f. Para Tablas (opcional), elija `amazon_reviews`.

g. Para Filtros de datos: opcional, seleccione `amazon_reviews_US`.

h. Para Permisos de filtro de datos, elija Seleccionar.

i. Elija Grant (Conceder).

4. Siga estos pasos para conceder permisos al analista de datos japonés (`DataAnalystJP`):

a. En Permisos, elija Permisos de lago de datos.

b. En Permiso de datos, seleccione Conceder.

c. En el Entidades principales, elija Usuarios y roles de IAM y seleccione el rol `DataAnalystJP`.

d. En Etiquetas LF o recursos del catálogo, elija Recursos de catálogo de datos con nombre.

e. En Database (Base de datos), elija `lakeformation_tutorial_row_security`.

f. Para Tablas (opcional), elija `amazon_reviews`.

g. Para Filtros de datos: opcional, seleccione `amazon_reviews_JP`.

h. Para Permisos de filtro de datos, elija Seleccionar.

i. Elija Grant (Conceder).

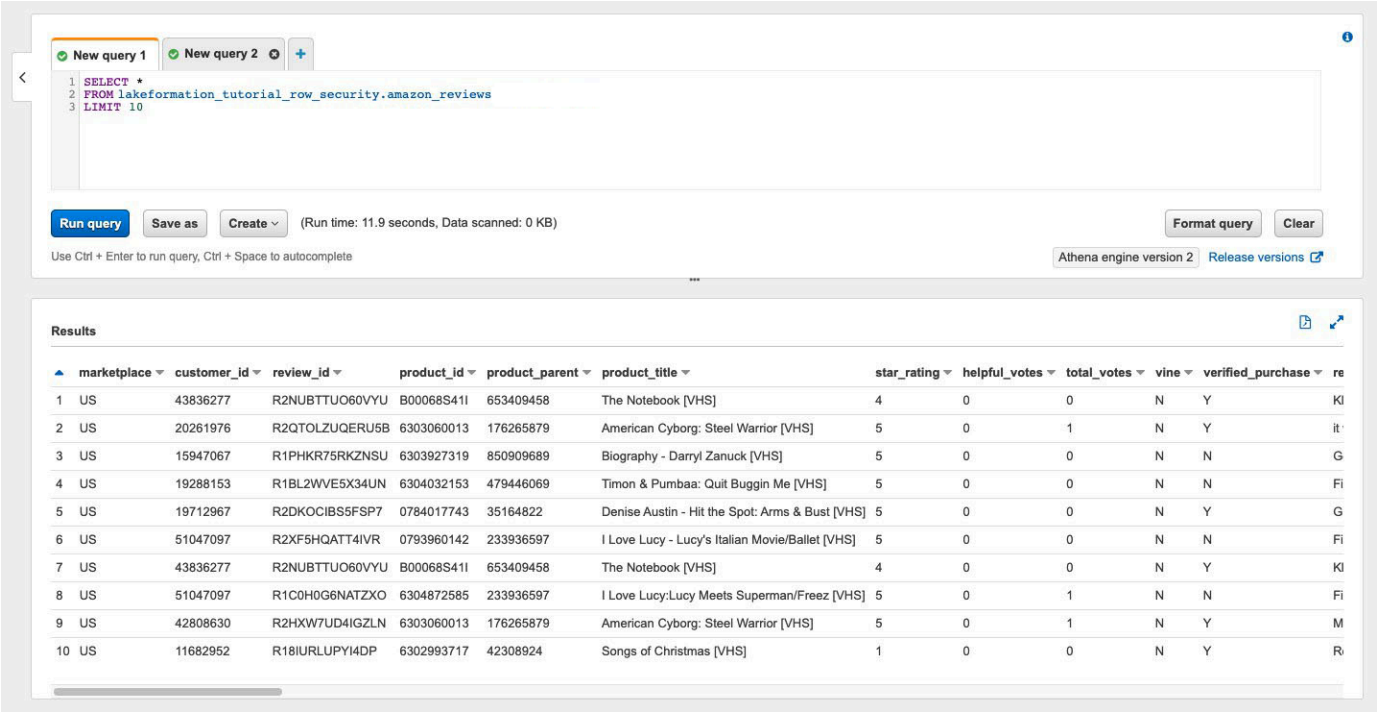
Paso 4: Consultar con filtros de datos

Con los filtros de datos adjuntos a la tabla de reseñas de productos, ejecute algunas consultas y compruebe cómo Lake Formation aplica los permisos.

1. Abra la consola de Athena en https://console.aws.amazon.com/athena/ como usuario `DataAnalystUS`.
2. Ejecute la siguiente consulta para recuperar algunos registros, que se filtran en función de los permisos de nivel de fila que definimos:

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

La captura de pantalla siguiente muestra el resultado de la consulta.



	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. Del mismo modo, ejecute una consulta para contar el número total de registros por mercado.

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

El resultado de la consulta solo muestra el US `marketplace` en los resultados. Esto se debe a que el usuario solo puede ver las filas en las que el valor de la columna `marketplace` sea igual a US.

4. Cambie al usuario `DataAnalystJP` y ejecute la misma consulta.

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

El resultado de la consulta solo muestra el `marketplace JP` en los resultados.

5. Ejecute la consulta para contar el número total de registros por `marketplace`.

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

El resultado de la consulta solo muestra la fila que pertenece al `marketplace JP` en los resultados.

Paso 5: Limpiar los recursos de AWS

Limpiar recursos

Para evitar cargos no deseados en su Cuenta de AWS, elimine los recursos de AWS que ha usado en este tutorial.

- Elimine la pila de formación de nube.

Compartir un lago de datos utilizando el control de acceso basado en etiquetas de Lake Formation y recursos con nombre

Este tutorial muestra cómo puede configurar AWS Lake Formation para compartir de forma segura los datos almacenados en un lago de datos con varias empresas, organizaciones o unidades de negocio, sin tener que copiar toda la base de datos. Hay dos opciones para compartir sus bases de datos y tablas con otra Cuenta de AWS mediante el control de acceso entre cuentas de Lake Formation:

- Control de acceso basado en etiquetas de Lake Formation (recomendado)

  El control de acceso basado en etiquetas de Lake Formation es una estrategia de autorización que define permisos basados en atributos. En Lake Formation, estos atributos se denominan etiquetas LF. Para obtener más detalles, consulte Gestión de un lago de datos mediante el control de acceso basado en etiquetas de Lake Formation.
- Recursos con nombre de Lake Formation

  El método de recursos con nombre de Lake Formation es una estrategia de autorización que define los permisos para los recursos. Los recursos incluyen bases de datos, tablas y columnas. Los administradores del lago de datos pueden asignar y revocar permisos sobre los recursos de Lake Formation. Para obtener más detalles, consulte Compartir datos entre cuentas en Lake Formation.

  Recomendamos utilizar recursos con nombre si el administrador del lago de datos prefiere conceder permisos explícitamente a recursos individuales. Cuando utiliza el método de recurso

con nombre para conceder permisos de Lake Formation sobre un recurso del catálogo de datos a una cuenta externa, Lake Formation utiliza el AWS Resource Access Manager (AWS RAM) para compartir el recurso.

Temas

Destinatarios previstos

Este tutorial está dirigido a administradores de datos, ingenieros de datos y analistas de datos. Cuando se trata de compartir tablas del catálogo de datos de AWS Glue y administrar permisos en Lake Formation, los administradores de datos de las cuentas productoras tienen la propiedad funcional basada en sus funciones, y pueden conceder acceso a varios consumidores, organizaciones externas y cuentas. La tabla siguiente enumera los roles que se utilizan en este tutorial:

Rol	Descripción
DataLakeAdminProducer	El usuario de IAM administrador del lago de datos tiene el acceso siguiente: • Acceso completo de lectura, escritura y actualización a todos los recursos del catálogo de datos • Posibilidad de conceder permisos a los recursos

Rol	Descripción
	• Puede crear enlaces a recursos para la tabla compartida • Puede adjuntar etiquetas LF a los recursos, lo que proporciona acceso a las entidades principales en función de cualquier política creada por los administradores de datos.
DataLakeAdminConsumer	El usuario de IAM administrador del lago de datos tiene el acceso siguiente: • Acceso completo de lectura, escritura y actualización a todos los recursos del catálogo de datos • Posibilidad de conceder permisos a los recursos • Puede crear enlaces a recursos para la tabla compartida • Puede adjuntar etiquetas LF a los recursos, lo que proporciona acceso a las entidades principales en función de cualquier política creada por los administradores de datos.
DataAnalyst	El usuario DataAnalyst tiene los siguientes accesos: • Acceso específico a los recursos compartid os por las políticas de acceso basadas en etiquetas de Lake Formation o mediante el método de recursos con nombre

Configurar los ajustes del catálogo de datos de Lake Formation en la cuenta del productor

Antes de comenzar este tutorial, debe disponer de una Cuenta de AWS que pueda utilizar para iniciar sesión como usuario administrativo con los permisos correctos. Para obtener más información, consulte Completar las tareas iniciales de configuración de AWS.

En este tutorial, se supone que está familiarizado con IAM. Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Configurar los ajustes del catálogo de datos de Lake Formation en la cuenta del productor

> **Note**
>
> En este tutorial, la cuenta que tiene la tabla de origen se denomina cuenta del productora, y la que necesita acceder a la tabla de origen se denomina cuenta del consumidor.

Lake Formation proporciona su propio modelo de gestión de permisos. Para mantener la compatibilidad con el modelo de permisos de IAM, el permiso de `Super` se concede al grupo `IAMAllowedPrincipals` en todos los recursos de AWS Glue Data Catalog existentes de forma predeterminada. Además, la configuración de control de acceso para utilizar solo IAM está activada para los nuevos recursos del catálogo de datos. Este tutorial utiliza el control de acceso específico utilizando permisos de Lake Formation y políticas de IAM para el control de acceso menos específico. Para obtener más información, consulte Métodos para el control de acceso específico. En consecuencia, antes de utilizar una plantilla AWS CloudFormation para una configuración rápida, deberá cambiar los ajustes del catálogo de datos de Lake Formation en la cuenta del productor.

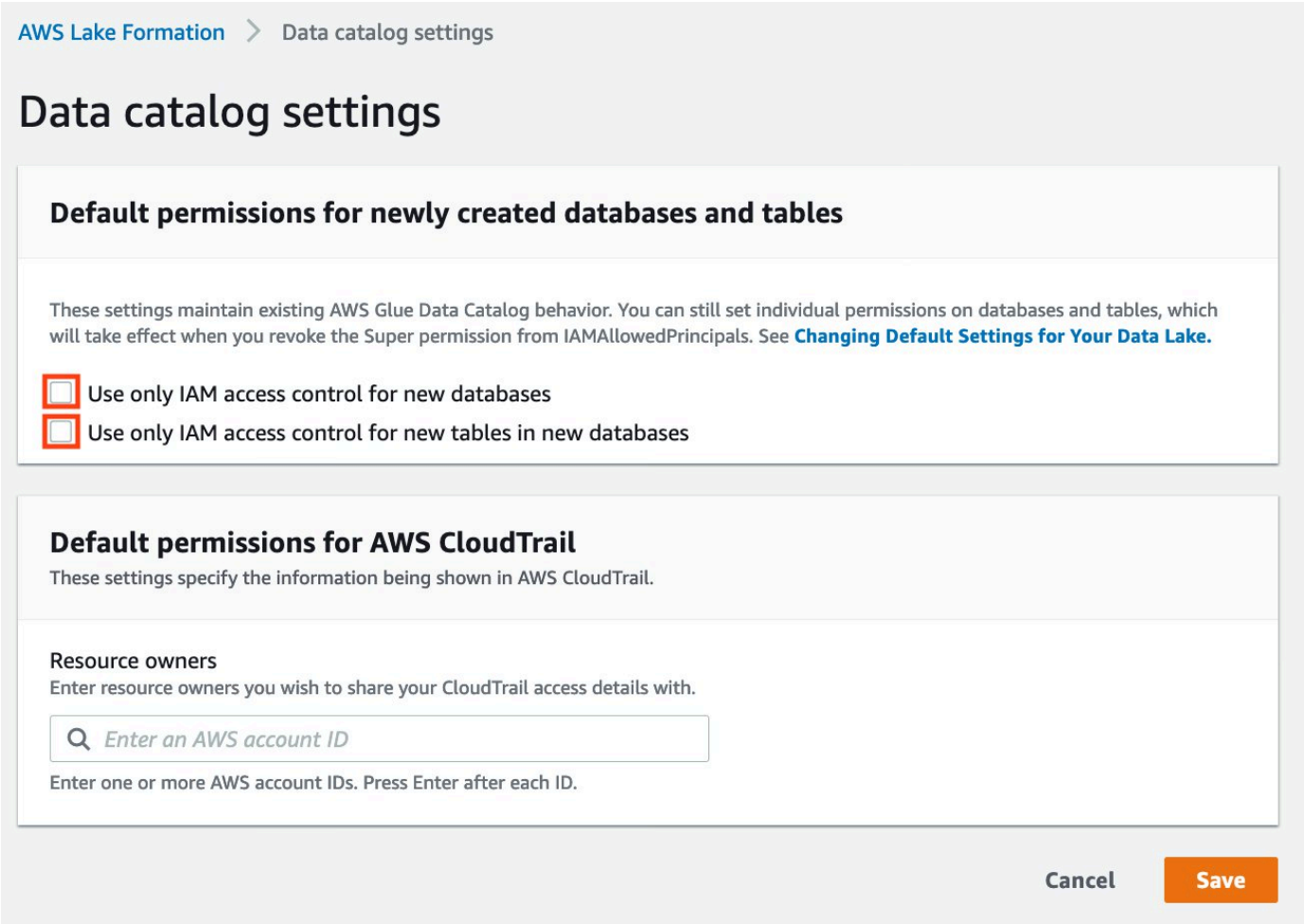
> **Important**
>
> Esta configuración afecta a todas las bases de datos y tablas recién creadas, por lo que se recomienda completar este tutorial en una cuenta que no sea de producción o en una cuenta nueva. Además, si utiliza una cuenta compartida (como la cuenta de desarrollo de su empresa), asegúrese de que no afecta a otros recursos. Si prefiere mantener la configuración de seguridad predeterminada, deberá completar un paso adicional cuando comparta recursos con otras cuentas, en el que revocará el permiso Super predeterminado

de `IAMAllowedPrincipals` en la base de datos o tabla. Trataremos los detalles más adelante en este tutorial.

Para configurar los ajustes del catálogo de datos de Lake Formation en la cuenta del productor:

1. Acceda a la AWS Management Console utilizando la cuenta de productor como usuario administrador, o como usuario con permiso para la API `PutDataLakeSettings` de Lake Formation.
2. En la consola de Lake Formation, desde panel de navegación, Catálogo de datos, seleccione Configuración.
3. Desmarque Usar solo control de acceso IAM para nuevas bases de datos y Usar solo control de acceso IAM para nuevas tablas en nuevas bases de datos.

   Seleccione Guardar.



También puede retirar los permisos de `CREATE_DATABASE` para `IAMAllowedPrincipals` en la sección de roles y tareas administrativas, creadores de bases de datos. Solo entonces podrá decidir quién puede crear una nueva base de datos mediante los permisos de Lake Formation.

Paso 1: Aprovisionar sus recursos mediante plantillas AWS CloudFormation

La plantilla CloudFormation para la cuenta del productor genera los siguientes recursos:

- Un bucket de Amazon S3 que sirva como lago de datos
- Una función de Lambda (para recursos personalizados respaldados por Lambda AWS CloudFormation). Esta función sirve para copiar archivos de datos de muestra del bucket público de Amazon S3 al bucket de Amazon S3.
- Usuarios de IAM y políticas: DataLakeAdminProducer.
- La configuración y los permisos apropiados de Lake Formation, como:
  - Definición del administrador del lago de datos de Lake Formation en la cuenta del productor
  - Registro de un bucket de Amazon S3 como ubicación del lago de datos de Lake Formation (cuenta de productor)
- Una base de datos, tabla y partición de AWS Glue Data Catalog. Como hay dos opciones para compartir recursos entre Cuentas de AWS, esta plantilla crea dos conjuntos independientes de base de datos y tablas.

La plantilla de AWS CloudFormation para la cuenta del consumidor genera los recursos siguientes:

- Usuarios de IAM y políticas:
  - DataLakeAdminConsumer
  - DataAnalyst
- Una base de datos de AWS Glue Data Catalog. Esta base de datos sirve para crear enlaces de recursos a recursos compartidos.

Crear sus recursos en la cuenta del productor

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation en la región Este de EE. UU. (Norte de Virginia).
2. Seleccione Lanzar pila.

3. Elija Next (Siguiente).
4. En Nombre de la pila, escriba un nombre para la pila, como `stack-producer`.
5. En la sección Configuración de usuario, introduzca la contraseña para `ProducerDatalakeAdminUserName` y `ProducerDatalakeAdminUserPassword`.
6. En DataLakeBucketName, escriba el nombre de su bucket de lago de datos. El nombre debe ser único de forma global.
7. Para DatabaseName y TableName, deje los valores predeterminados.
8. Elija Next (Siguiente).
9. En la página siguiente, seleccione Siguiente.
10. Revise los detalles en la última página y acepte que AWS CloudFormation pueda crear recursos de IAM.
11. Seleccione Create (Crear).

    La creación de la pila puede llevar hasta un minuto.

Crear sus recursos en la cuenta del consumidor

1. Inicie sesión en la consola de AWS CloudFormation en https://console.aws.amazon.com/cloudformation en la región Este de EE. UU. (Norte de Virginia).
2. Seleccione Lanzar pila.
3. Elija Next (Siguiente).
4. En Nombre de la pila, escriba un nombre para la pila, como `stack-consumer`.
5. En la sección Configuración de usuario, introduzca la contraseña para `ConsumerDatalakeAdminUserName` y `ConsumerDatalakeAdminUserPassword`.
6. Para `DataAnalystUserName` y `DataAnalystUserPassword`, introduzca el nombre de usuario y la contraseña que desee para el usuario de IAM del analista de datos.
7. En DataLakeBucketName, escriba el nombre de su bucket de lago de datos. El nombre debe ser único de forma global.
8. En DatabaseName, deje los valores predeterminados.
9. Para `AthenaQueryResultS3BucketName`, introduzca el nombre del bucket de Amazon S3 que almacena los resultados de consultas de Amazon Athena. Si no dispone de un bucket de Amazon S3, cree uno nuevo.
10. Elija Next (Siguiente).

11. En la página siguiente, seleccione Siguiente.
12. Revise los detalles en la última página y acepte que AWS CloudFormation pueda crear recursos de IAM.
13. Seleccione Create (Crear).

    La creación de la pila puede llevar hasta un minuto.

> **Note**
>
> Después de completar el tutorial, elimine la pila de AWS CloudFormation para evitar incurrir en cargos. Compruebe que los recursos se hayan eliminado correctamente en el estado de evento de la pila.

Paso 2: Requisitos previos para compartir entre cuentas de Lake Formation

Antes de compartir recursos con Lake Formation, hay requisitos previos tanto para el método de control de acceso basado en etiquetas como para el método de recurso con nombre.

Completar los requisitos previos del control de acceso basado en etiquetas para el intercambio de datos entre cuentas

- Para más información sobre los requisitos para compartir datos entre cuentas, consulte la sección Requisitos previos del capítulo sobre compartir datos entre cuentas.

  Para compartir recursos del catálogo de datos con la versión 3 o superior de la configuración de versión entre cuentas, el concedente debe disponer de los permisos de IAM definidos en la política administrada `AWSLakeFormationCrossAccountManager` de AWS en su cuenta.

  Si utiliza la versión 1 o la versión 2 de la configuración de la versión entre cuentas, antes de poder utilizar el método de control de acceso basado en etiquetas para conceder acceso entre cuentas a los recursos, deberá añadir el siguiente objeto de permisos `JSON` entre cuentas a la política de recursos del catálogo de datos en la cuenta de productor. De este modo, la cuenta del consumidor permite a la cuenta del consumidor acceder al catálogo de datos cuando `glue:EvaluatedByLakeFormationTags` es cierto. Además, esta condición se cumple para los recursos sobre los que haya concedido permiso mediante etiquetas de permiso de Lake Formation a la cuenta del consumidor. Esta política es necesaria para cada Cuenta de AWS a la que conceda permisos.

La política siguiente debe estar dentro de un elemento `Statement`. En la sección siguiente, analizamos toda la política de IAM.

```
{
    "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
        }
    }
}
```

Completar los requisitos previos del método de recursos con nombre para compartir entre cuentas

1. Si en su cuenta no existe una política de recursos del catálogo de datos, las concesiones que efectúe entre cuentas procederán como de costumbre. Sin embargo, si existe una política de recursos del catálogo de datos, debe añadirle la siguiente instrucción para permitir que sus concesiones entre cuentas tengan éxito si se efectúan con el método de recursos con nombre. Si piensa utilizar solo el método de recursos con nombre, o solo el método de control de acceso basado en etiquetas, puede omitir este paso. En este tutorial, evaluamos ambos métodos, y necesitamos añadir la política siguiente.

   La política siguiente debe estar dentro de un elemento `Statement`. En la sección siguiente, analizamos toda la política de IAM.

```
{
          "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. Ahora, añada la política de recursos AWS Glue Data Catalog con la AWS Command Line Interface (AWS CLI).

   Si concede permisos entre cuentas utilizando tanto el método de control de acceso basado en etiquetas como el método de recursos con nombre, debe establecer el argumento `EnableHybrid` en "true" cuando añada las políticas anteriores. Debido a que esta opción no es actualmente Compatibilidad en la consola, y debe utilizar la API `glue:PutResourcePolicy` y la AWS CLI.

   En primer lugar, cree un documento de políticas (como policy.json) y añada las dos políticas anteriores. Sustituya *`consumer-account-id`* por el *`ID de cuenta`* de la Cuenta de AWS que recibe la concesión, *`region`* por la región del catálogo de datos que contiene las bases de datos y las tablas sobre las que está concediendo permisos, y *`account-id`* por el ID de Cuenta de AWS del productor.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "ram.amazonaws.com"
            },
            "Action": "glue:ShareResource",
            "Resource": [
```

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

Escriba el comando AWS CLI siguiente. Sustituya *glue-resource-policy* por los valores correctos (como file://policy.json).

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

Para obtener más información, consulte put-resource-policy.

Paso 3: Implementar el uso compartido entre cuentas mediante el método de control de acceso basado en etiquetas

En esta sección, encontrará los siguientes pasos generales:

1. Definir una etiqueta LF.
2. Asignar la etiqueta LF al recurso de destino.

3. Conceder permisos de etiqueta LF a la cuenta de consumidor.
4. Conceder permisos sobre datos a la cuenta de consumidor.
5. Opcionalmente, revocar los permisos para `IAMAllowedPrincipals` sobre las bases de datos, las tablas y las columnas.
6. Crear un enlace de recursos en la base de datos compartida.
7. Crear una etiqueta LF y asignarla a la base de datos de destino.
8. Conceder permisos de datos de etiqueta LF a la cuenta de consumidor.

Definir una etiqueta LF.

> **Note**
>
> Si ha iniciado sesión en su cuenta de productor, ciérrela antes de completar los pasos siguientes.

1. Inicie sesión en la cuenta del productor como administrador del lago de datos en https://console.aws.amazon.com/lakeformation/. Utilice el número de cuenta del productor, el nombre de usuario de IAM (el predeterminado es `DatalakeAdminProducer`) y la contraseña que especificó durante la creación de la pila AWS CloudFormation.
2. En la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), en el panel de navegación, en Permisos y en Roles y tareas administrativas, elija Etiquetas LF.
3. Seleccione Añadir etiqueta LF.

Asignar la etiqueta LF al recurso de destino.

Asigne la etiqueta LF al recurso de destino y conceda permisos de datos a otra cuenta

Como administrador de un lago de datos, puede adjuntar etiquetas a los recursos. Si tiene previsto utilizar un rol independiente, es posible que tenga que conceder permisos de descripción y vinculación a ese rol.

1. En el panel de navegación, en Catálogo de datos, elija Bases de datos.
2. Seleccione la base de datos `(lakeformation_tutorial_cross_account_database_tbac)` de destino y, en el menú Acciones, elija Editar etiquetas LF.

En este tutorial, asignará una etiqueta LF a una base de datos, pero también podrá asignar etiquetas LF a tablas y columnas.

3. Seleccione Asignar una nueva etiqueta LF.
4. Añada la clave `Confidentiality` y el valor `public`.
5. Seleccione Guardar.

Conceder permisos de etiqueta LF a la cuenta de consumidor.

Si aún está en la cuenta del productor, conceda permisos a la cuenta del consumidor para acceder a la etiqueta LF.

1. En el panel de navegación, en Permisos, Roles y tareas administrativas, Permisos de etiquetas LF, seleccione Conceder.
2. En Entidades principales, seleccione Cuentas externas.
3. Introduzca el ID Cuenta de AWS de destino.

   Las Cuentas de AWS de la misma organización aparecen automáticamente. De lo contrario, tendrá que introducir el ID de Cuenta de AWS manualmente. En el momento de redactar este documento, el control de acceso basado en etiquetas de Lake Formation no es compatible con la concesión de permisos a organizaciones o unidades de organización.
4. Para las etiquetas LF, elija la clave y los valores de la etiqueta LF que se va a compartir con la cuenta del consumidor (clave `Confidentiality` y valor `public`).
5. En Permisos, seleccione Describir en Permisos de etiqueta LF.

   Los permisos de etiquetas LF son permisos otorgados a la cuenta del consumidor. Los permisos concedibles son permisos que la cuenta de consumidor puede conceder a otras entidades principales.
6. Elija Grant (Conceder).

   En este punto, el administrador del lago de datos del consumidor debería ser capaz de encontrar la etiqueta de la política que se comparte a través de la cuenta del consumidor de la consola Lake Formation, en Permisos, Roles y tareas administrativas, etiquetas LF.

Conceder permiso de datos a la cuenta del consumidor

Ahora proporcionaremos acceso a los datos a la cuenta del consumidor especificando una expresión de etiquetas LF y concediendo a la cuenta del consumidor acceso a cualquier tabla o base de datos que coincida con la expresión.

1. En el panel de navegación, en Permisos, Permisos de lago de datos, seleccione Conceder.
2. En Entidades principales, seleccione Cuentas externas e introduzca el ID Cuenta de AWS de destino.
3. Para las etiquetas LF o los recursos del catálogo, elija la clave y los valores de la etiqueta LF que se va a compartir con la cuenta del consumidor (clave `Confidentiality` y valor `public`).
4. En Permisos, bajo Recursos que coinciden con las etiquetas LF (recomendado) elija Agregar etiqueta LF.
5. Seleccione la clave y el valor de la etiqueta que se va a compartir con la cuenta del consumidor (clave `Confidentiality` y valor `public`).
6. Para Permisos de base de datos, seleccione Describir en Permisos de bases de datos para conceder permisos de acceso a nivel de base de datos.
7. El administrador del lago de datos del consumidor debería poder encontrar la etiqueta de política que se comparte a través de la cuenta del consumidor en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/, en Permisos, Roles y tareas administrativas, Etiquetas LF.
8. Seleccione Describir en Permisos concedibles para que la cuenta de consumidor pueda conceder permisos a nivel de base de datos a sus usuarios.
9. En Permisos de tabla y columna, elija Seleccionar y Describir en Permisos de tabla.
10. Elija Seleccionar y Describir en Permisos concedibles.
11. Elija Grant (Conceder).

Revocar los permisos para **`IAMAllowedPrincipals`** sobre las bases de datos, las tablas y las columnas (opcional).

Al principio de este tutorial, cambió la configuración del catálogo de datos de Lake Formation. Si omitió esa parte, este paso es obligatorio. Si ha cambiado la configuración del catálogo de datos de Lake Formation, puede omitir este paso.

En este paso, necesitamos revocar el permiso Super por defecto de `IAMAllowedPrincipals` en la base de datos o tabla. Para obtener más información, consulte Paso 4: Cambiar sus almacenes de datos al modelo de permisos de Lake Formation.

Antes de revocar el permiso para `IAMAllowedPrincipals`, asegúrese de que ha concedido a las entidades principales de IAM existentes el permiso necesario a través de Lake Formation. Esto conlleva tres pasos:

1. Añadir el permiso de IAM al rol o usuario de IAM de destino con la acción `GetDataAccess` de Lake Formation (con la política de IAM).
2. Conceder al rol o usuario de IAM de destino permisos de datos de Lake Formation (alterar, seleccionar, etc.).
3. A continuación, revocar los permisos para `IAMAllowedPrincipals`. De lo contrario, después de revocar los permisos para `IAMAllowedPrincipals`, es posible que las entidades principales de IAM existentes ya no puedan acceder a la base de datos de destino o al catálogo de datos.

   La revocación del permiso Super para `IAMAllowedPrincipals` es necesaria cuando desea aplicar el modelo de permisos Lake Formation (en lugar del modelo de políticas de IAM) para gestionar el acceso de los usuarios dentro de una misma cuenta o entre varias cuentas utilizando el modelo de permisos de Lake Formation. No tiene que revocar el permiso de `IAMAllowedPrincipals` para otras tablas en las que desee mantener el modelo de política de IAM tradicional.

   En este punto, el administrador del lago de datos de la cuenta de consumidor debería poder encontrar la base de datos y la tabla que se comparten a través de la cuenta de consumidor en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/, en Catálogo de datos, bases de datos. Si no es así, confirme si lo siguiente está bien configurado:

   1. Están asignados la etiqueta y los valores de política correctos a las bases de datos y tablas de destino.
   2. Están asignados el permiso de etiqueta y el permiso de datos correctos a la cuenta del consumidor.
   3. Está revocado el permiso super predeterminado de `IAMAllowedPrincipals` en la base de datos o tabla.

Crear un enlace de recursos a la tabla compartida

Cuando un recurso se comparte entre cuentas, y los recursos compartidos no se ubican en el catálogo de datos de las cuentas consumidoras. Para que estén disponibles, y consultar los datos subyacentes de una tabla compartida utilizando servicios como Athena, necesitamos crear un enlace de recursos a la tabla compartida. Un enlace de recursos es un objeto del catálogo de datos que es un enlace a una base de datos o tabla local o compartida. Para obtener más información, consulte Creación de enlaces de recursos. Al crear un enlace de recursos, puede:

- Asignar un nombre diferente a una base de datos o tabla que se ajuste a sus políticas de nombres de recursos del catálogo de datos.
- Utilizar servicios como Athena y Redshift Spectrum para consultar bases de datos o tablas compartidas.

Para crear un enlace de recursos:

1. Si ha iniciado sesión en su cuenta de consumidor, cierre la sesión.
2. Inicie sesión como administrador del lago de datos de la cuenta del consumidor. Utilice el ID de cuenta del consumidor, el nombre de usuario de IAM (predeterminado DatalakeAdminConsumer) y la contraseña que especificó durante la creación de la pila AWS CloudFormation.
3. En la consola Lake Formation (https://console.aws.amazon.com/lakeformation/), en el panel de navegación, en Catálogo de datos, Bases de datos, elija la base de datos compartida `lakeformation_tutorial_cross_account_database_tbac`.

   Si no ve la base de datos, vuelva a revisar los pasos anteriores para ver si todo está configurado correctamente.
4. Elija Ver las tablas.
5. Elija la tabla compartida `amazon_reviews_table_tbac`.
6. En el menú Acciones, elija Crear recurso.
7. En Nombre del enlace de recursos, introduzca un nombre (para este tutorial, `amazon_reviews_table_tbac_resource_link`).
8. En Base de datos, seleccione la base de datos en la que se creó el enlace al recurso (para esta publicación, la pila AWS CloudFormation creó la base de datos `lakeformation_tutorial_cross_account_database_consumer`).
9. Seleccione Create (Crear).

El enlace de recursos aparece en Catálogo de datos, Tablas.

Crear una etiqueta LF y asignarla a la base de datos de destino

Las etiquetas de Lake Formation se encuentran en el mismo catálogo de datos que los recursos. Esto significa que las etiquetas creadas en la cuenta del productor no están disponibles para su uso al conceder acceso a los enlaces de recursos en la cuenta del consumidor. Es necesario crear un conjunto separado de etiquetas LF en la cuenta del consumidor para utilizar el control de acceso basado en etiquetas LF al compartir los enlaces de recursos en la cuenta del consumidor.

1. Defina la etiqueta LF en la cuenta del consumidor. Para este tutorial, utilizamos la clave `Division` y los valores `sales`, `marketing` y `analyst`.
2. Asigne la clave de etiquetas LF `Division` y el valor `analyst` a la base de datos `lakeformation_tutorial_cross_account_database_consumer`, donde se crea el enlace de recursos.

Conceder permiso de datos de etiquetas LF al consumidor

Como paso final, conceda permiso de datos de etiquetas LF al consumidor.

1. En el panel de navegación, en Permisos, Permisos de lago de datos, seleccione Conceder.
2. En Entidades principales, seleccione los usuarios y roles de IAM y elija el usuario `DataAnalyst`.
3. En las etiquetas LF o los recursos del catálogo, elija los recursos que coincidan con etiquetas LF (recomendado).
4. Elija la clave división y el valor analista.
5. Para los permisos de la base de datos, seleccione Describir en Permisos de la base de datos.
6. En Permisos de tabla y columna, elija Seleccionar y Describir en Permisos de tabla.
7. Elija Grant (Conceder).
8. Repita estos pasos para el usuario `DataAnalyst`, donde la clave de la etiqueta LF es `Confidentiality` y el valor es `public`.

   En este punto, el usuario analista de datos de la cuenta de consumidor debería poder encontrar la base de datos y el enlace de recursos, y consultar la tabla compartida a través de la consola

Athena en https://console.aws.amazon.com/athena/. Si no es así, confirme si lo siguiente está bien configurado:

- Está creado el enlace de recursos para la tabla compartida
- Ha concedido al usuario acceso a las etiquetas LF compartidas por la cuenta de productor
- Ha concedido al usuario acceso a las etiquetas LF asociadas al enlace de recursos y a la base de datos en la que se crea el enlace de recursos.
- Compruebe si ha asignado las etiquetas LF correctas al enlace de recursos y a la base de datos en la que se ha creado el enlace de recursos.

Paso 4: Implementar el método de recurso con nombre

Para utilizar el método de recursos con nombre, le guiaremos a través de los siguientes pasos básicos:

1. De manera opcional, revocar el permiso para `IAMAllowedPrincipals` sobre la base de datos, las tablas y las columnas.
2. Conceder permiso de datos a la cuenta del consumidor.
3. Aceptar un recurso compartido de AWS Resource Access Manager.
4. Crear un enlace de recursos para la tabla compartida.
5. Concede permiso de datos para la tabla compartida al consumidor.
6. Concede permiso de datos para el enlace de recursos al consumidor.

Revocar permiso para **`IAMAllowedPrincipals`** sobre las bases de datos, las tablas y las columnas (opcional).

- Al principio de este tutorial, cambiamos la configuración del catálogo de datos de Lake Formation. Si omitió esa parte, este paso es obligatorio. Para obtener instrucciones, consulte el paso opcional de la sección anterior.

Conceder permiso de datos a la cuenta del consumidor

1. > **Note**
   >
   > Si ha iniciado sesión en la cuenta de productor como otro usuario, cierre primero la sesión.

   Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ con la cuenta de productor administrador del lago de datos utilizando el ID Cuenta de AWS, el nombre de usuario de IAM (por defecto es `DatalakeAdminProducer`) y la contraseña especificados durante la creación de la pila AWS CloudFormation.
2. En la página Permisos, en Permisos del lago de datos elija Conceder.
3. En Entidades principales, seleccione Cuentas externas e introduzca uno o varios ID de Cuenta de AWS o ID AWS de organizaciones. Para más información, consulte: Organizaciones AWS.

   Las organizaciones a las que pertenece la cuenta de productor y las Cuentas de AWS dentro de la misma organización aparecen automáticamente. De lo contrario, introduzca manualmente el ID de la cuenta o el ID de la organización.
4. Para las etiquetas LF o los recursos del catálogo, elija `Named data catalog resources`.
5. En Bases de datos elija la base de datos `lakeformation_tutorial_cross_account_database_named_resource`.
6. Seleccione Añadir etiqueta LF.
7. En Tablas, seleccione Todas las tablas.
8. En Permisos de columna de tabla, elija Seleccionar y Describir en Permisos de tabla.
9. Elija Seleccionar y Describir en Permisos concedibles.
10. De forma opcional, para Permisos de datos, elija Acceso simple basado en columnas si se requiere una gestión de permisos a nivel de columna.
11. Elija Grant (Conceder).

Si no ha revocado el permiso para `IAMAllowedPrincipals`, aparecerá un error de concesión de permisos. En este punto, debería ver la tabla de destino que se comparte a través de AWS RAM con la cuenta del consumidor en Permisos, Permisos de datos.

Aceptar un recurso compartido de AWS RAM.

> **Note**
>
> Este paso es necesario solo para el uso compartido basado en Cuenta de AWS, no para el basado en la organización.

1. Inicie sesión en la consola de AWS en https://console.aws.amazon.com/connect/ con la cuenta de consumidor administrador del lago de datos utilizando el nombre de usuario de IAM (predeterminado DatalakeAdminConsumer) y la contraseña especificados durante la creación de la pila AWS CloudFormation.
2. En la consola de AWS RAM, en el panel de navegación, Compartido conmigo, Recursos compartidos, elija el recurso compartido de Lake Formation. El estado debe ser Pendiente.
3. Elija Acción y Conceder.
4. Confirme los detalles del recurso y seleccione Aceptar recurso compartido.

   En este punto, el administrador del lago de datos de la cuenta del consumidor debería poder encontrar el recurso compartido en la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/) en Catálogo de datos, Bases de datos.

Crear un enlace de recursos para la tabla compartida

- Siga las instrucciones de Paso 3: Implementar el uso compartido entre cuentas mediante el método de control de acceso basado en etiquetas (paso 6) para crear un enlace de recursos para una tabla compartida. Asigne un nombre al enlace de recursos `amazon_reviews_table_named_resource_resource_link`. Cree el enlace de recursos en la base de datos `lakeformation_tutorial_cross_account_database_consumer`.

Conceder permiso de datos al consumidor sobre la tabla compartida

Para conceder permiso de datos sobre la tabla compartida al consumidor:

1. En la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), en Permisos, Permisos de lago de datos, elija Conceder.
2. En Entidades principales, seleccione los usuarios y roles de IAM y elija el usuario `DataAnalyst`.

3. Para las etiquetas LF o los recursos del catálogo, elija recursos del catálogo de datos con nombre.
4. En Bases de datos elija la base de datos `lakeformation_tutorial_cross_account_database_named_resource`. Si no ve la base de datos en la lista desplegable, elija Cargar más.
5. En Tablas, seleccione la tabla `amazon_reviews_table_named_resource`.
6. En Permisos de tabla y columna, elija Seleccionar y Describir en Permisos de tabla.
7. Elija Grant (Conceder).

Conceder permiso de datos al consumidor sobre el enlace de recursos

Además de conceder al usuario del lago de datos permiso para acceder a la tabla compartida, también debe concederle permiso para acceder al enlace de recursos.

1. En la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), en Permisos, Permisos de lago de datos, elija Conceder.
2. En Entidades principales, seleccione los usuarios y roles de IAM y elija el usuario `DataAnalyst`.
3. Para las etiquetas LF o los recursos del catálogo, elija recursos del catálogo de datos con nombre.
4. En Bases de datos elija la base de datos `lakeformation_tutorial_cross_account_database_consumer`. Si no ve la base de datos en la lista desplegable, elija Cargar más.
5. En Tablas, seleccione la tabla `amazon_reviews_table_named_resource_resource_link`.
6. Para ver los permisos de enlace de recursos, seleccione Describir en Permisos de enlace de recursos.
7. Elija Grant (Conceder).

   En este punto, el usuario analista de datos de la cuenta de consumidor debería poder encontrar la base de datos y el enlace de recursos, y consultar la tabla compartida mediante la consola de Athena.

   Si no es así, confirme si lo siguiente está bien configurado:

   - Está creado el enlace de recursos para la tabla compartida

- Ha concedido al usuario acceso a la tabla compartida por la cuenta de productor
- Ha concedido al usuario acceso al enlace de recursos y a la base de datos para la que se ha creado el enlace de recursos

Paso 5: Limpiar recursos de AWS

Para evitar cargos no deseados en su Cuenta de AWS, elimine los recursos de AWS que ha usado en este tutorial.

1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ utilizando la cuenta de productor y elimine o cambie lo siguiente:
   - Compartir recursos de AWS Resource Access Manager
   - Etiquetas de Lake Formation
   - Pila de AWS CloudFormation
   - Ajustes de Lake Formation
   - AWS Glue Data Catalog
2. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ utilizando la cuenta de consumidor y elimine o cambie lo siguiente:
   - Etiquetas de Lake Formation
   - Pila de AWS CloudFormation

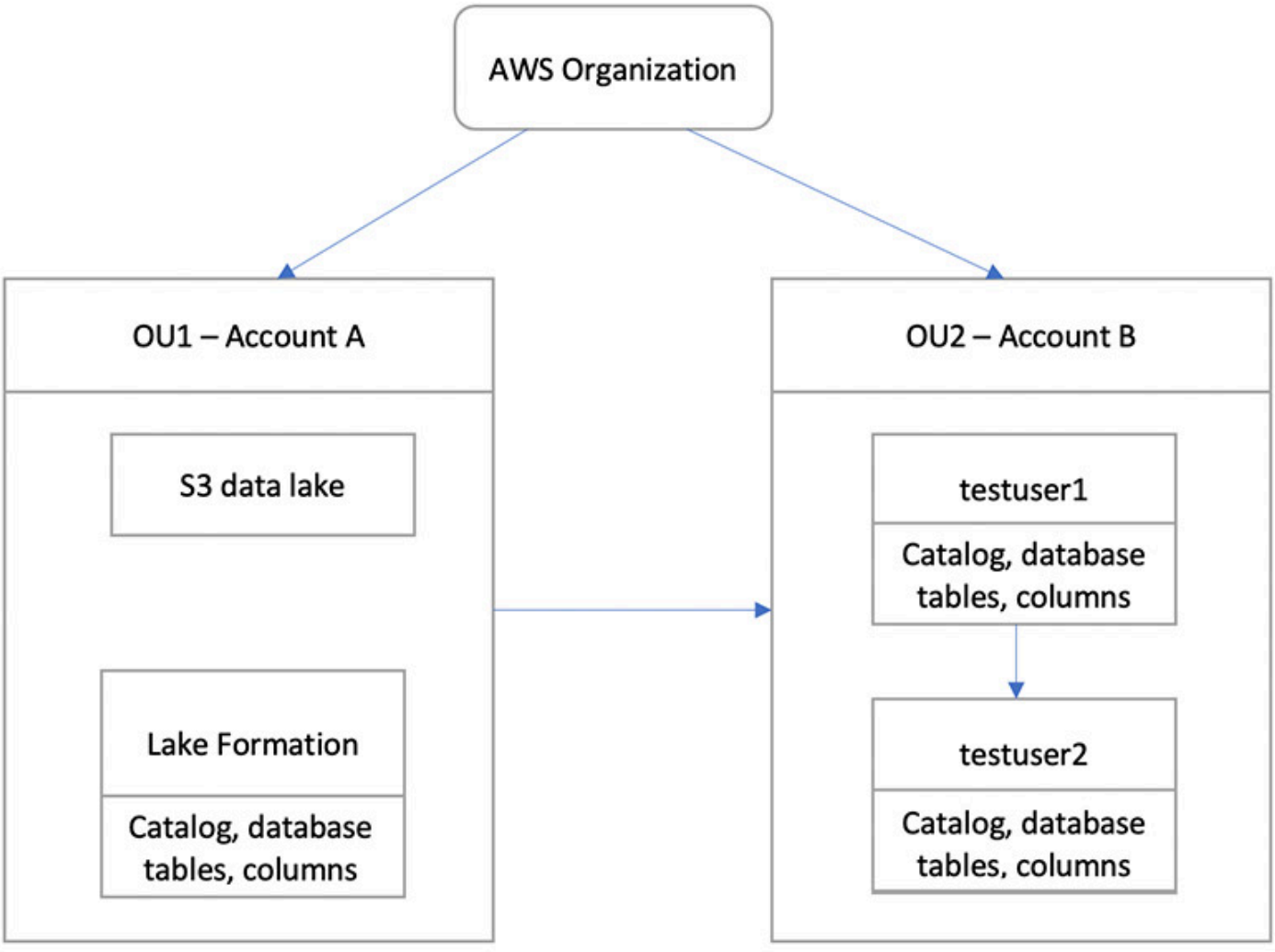
Compartir un lago de datos utilizando el control de acceso específico de Lake Formation

Este tutorial proporciona instrucciones detalladas sobre cómo compartir rápida y fácilmente conjuntos de datos utilizando Lake Formation cuando gestiona varios Cuentas de AWS con AWS Organizations. La definición de permisos específicos sirve para controlar el acceso a datos confidenciales.

Los siguientes procedimientos también muestran cómo un administrador del lago de datos de la Cuenta A puede proporcionar acceso específico para la Cuenta B, y cómo un usuario de la Cuenta B, actuando como administrador de datos, puede conceder acceso específico a la tabla compartida para otros usuarios de su cuenta. Los administradores de datos de cada cuenta pueden delegar

independientemente el acceso a sus propios usuarios, lo que proporciona autonomía a cada equipo o línea de negocio (LOB).

El caso práctico supone que está utilizando AWS Organizations para gestionar su Cuentas de AWS. El usuario de la Cuenta A en una unidad organizativa (UO1) concede acceso a los usuarios de la Cuenta B en la UO2. Puede utilizar el mismo enfoque cuando no utilice organizaciones, como cuando solo tiene unas pocas cuentas. El siguiente diagrama ilustra el control de acceso específico de los conjuntos de datos en un lago de datos. El lago de datos está disponible en la Cuenta A. El administrador del lago de datos de la Cuenta A proporciona acceso específico para la Cuenta B. El diagrama también muestra que un usuario de la Cuenta B proporciona acceso a nivel de columna de la tabla del lago de datos de la Cuenta A a otro usuario de la Cuenta B.



Temas

- Destinatarios previstos

Destinatarios previstos

Este tutorial está dirigido a administradores de datos, ingenieros de datos y analistas de datos. La tabla siguiente enumera los roles que se utilizan en este tutorial:

Rol	Descripción
Administrador de IAM	Usuario que tiene la política administrada de AWS: `AdministratorAccess` .
Administrador de lago de datos	Usuario que tiene la política administrada de AWS: `AWSLakeFormationDataAdmin` vinculada al rol.
Analista de datos	Usuario que tiene la política administrada de AWS: `AmazonAthenaFullAccess` vinculada.

Requisitos previos

Antes de comenzar este tutorial, debe disponer de una Cuenta de AWS que pueda utilizar para iniciar sesión como usuario administrativo con los permisos correctos. Para obtener más información, consulte Completar las tareas iniciales de configuración de AWS.

En este tutorial, se supone que está familiarizado con IAM. Para obtener más información acerca de IAM, consulte la Guía del usuario de IAM.

Para este tutorial, necesita los siguientes recursos:

- Dos unidades organizativas:
  - UO1: contiene la Cuenta A
  - UO2: contiene la Cuenta B

- Una ubicación (bucket) del lago de datos de Amazon S3 en la Cuenta A.
- Un usuario administrador del lago de datos en la Cuenta A. Puede crear un administrador del lago de datos utilizando la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/) o la operación `PutDataLakeSettings` de la API de Lake Formation.
- Lake Formation configurado en la Cuenta A, y la ubicación del lago de datos de Amazon S3 registrada con Lake Formation en la Cuenta A.
- Dos usuarios en la Cuenta B con las siguientes políticas administradas de IAM:
  - testuser1: tiene vinculadas las políticas administradas de AWS `AWSLakeFormationDataAdmin`.
  - testuser2 - Tiene vinculada la política administrada `AmazonAthenaFullAccess` de AWS.
- Una base de datos testdb en la base de datos Lake Formation para la Cuenta B.

Paso 1: Proporcionar acceso específico a otra cuenta

Conozca cómo un administrador del lago de datos de la Cuenta A proporciona acceso específico a la Cuenta B.

Conceder acceso específico a otra cuenta

1. Inicie sesión en la AWS Management Console en https://console.aws.amazon.com/connect/ en la Cuenta A como administrador del lago de datos.
2. Abra la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), y elija Empezar.
3. En el panel de navegación, seleccione Bases de datos.
4. Elija Create database (Crear base de datos).
5. En la sección de detalles de la base de datos, seleccione Base de datos.
6. En Nombre, introduzca un nombre (para este tutorial, utilizamos `sampledb01`).
7. Asegúrese de que no esté marcada la opción para utilizar solo el control de acceso IAM para las nuevas tablas de esta base de datos. Si se deja sin seleccionar, podremos controlar el acceso desde Lake Formation.
8. Elija Create database (Crear base de datos).
9. En la página Bases de datos, elija su base de datos `sampledb01`.
10. En el menú Acciones, elija Conceder.
11. En la sección Conceder permisos, seleccione Cuenta externa.

12. Para el ID de Cuenta de AWS o el ID de organización AWS, introduzca el ID de la Cuenta B en la UO2.
13. Para Tabla, elija la tabla a la que desea que tenga acceso la Cuenta B (para este post, utilizamos la tabla `acc_a_area`). De forma opcional, puede conceder acceso a columnas dentro de la tabla, como hacemos en esta publicación.
14. En Incluir columnas¸ elija las columnas a las que desea que tenga acceso la Cuenta B (para este post, concedemos permisos a tipo, nombre e identificadores).
15. En Columnas, seleccione Incluir columnas.
16. En Permisos de tabla, elija Seleccionar.
17. En Permisos concedibles, elija Seleccionar. Los permisos concedibles son necesarios para que los usuarios administradores de la Cuenta B puedan conceder permisos a otros usuarios de la Cuenta B.
18. Elija Grant (Conceder).
19. En el panel de navegación, elija Tablas.
20. Podrá consultar una conexión activa en la sección de Cuentas de AWS y organizaciones de AWS con acceso.

Crear un recurso compartido

Los servicios integrados como Amazon Athena no pueden acceder directamente a las bases de datos o tablas de las cuentas. Por consiguiente, debe crear un enlace de recursos para que Athena pueda acceder a los enlaces de recursos de su cuenta a bases de datos y tablas de otras cuentas. Cree un enlace de recursos a la tabla (`acc_a_area`) para que los usuarios de la Cuenta B puedan consultar sus datos con Athena.

1. Inicie sesión en la consola de AWS en https://console.aws.amazon.com/connect/ en la Cuenta B como `testuser1`.
2. En la consola de Lake Formation (https://console.aws.amazon.com/lakeformation/), en el panel de navegación, seleccione Tablas. Debería ver las tablas a las que la Cuenta A ha proporcionado acceso.
3. Elija la tabla `acc_a_area`.
4. En el menú Acciones, elija Crear enlace de recursos.
5. En Nombre del enlace de recursos, introduzca un nombre (para este tutorial, `acc_a_area_rl`).
6. En Base de datos, elija su base de datos (`testdb`).

7. Seleccione Create (Crear).
8. En el panel de navegación, elija Tablas.
9. Elija la tabla `acc_b_area_rl`.
10. En el menú Acciones, seleccione Ver datos.

    Se le redirige a la consola de Athena, donde debería ver la base de datos y la tabla.

    Ahora puede ejecutar una consulta en la tabla para ver el valor de la columna para la que se proporcionó acceso a testuser1 desde la Cuenta B.

Paso 2: Proporcionar acceso específico a un usuario de la misma cuenta

Esta sección muestra cómo un usuario de la cuenta B (`testuser1`), que actúa como administrador de datos, proporciona acceso específico a otro usuario de la misma cuenta (`testuser2`) al nombre de la columna de la tabla compartida `aac_b_area_rl`.

Conceder acceso específico a un usuario de la misma cuenta

1. Inicie sesión en la consola de AWS en https://console.aws.amazon.com/connect/ en la Cuenta B como `testuser1`.
2. En la consola de Lake Formation, en el panel de navegación, seleccione Tablas.

   Puede conceder permisos sobre una tabla a través de su enlace de recursos. Para ello, en la página Tablas, seleccione el enlace del recurso `acc_b_area_rl`, y en el menú Acciones, elija Conceder en destino.
3. En la sección Conceder permisos, seleccione Mi cuenta.
4. En Roles y usuarios de IAM, elija el usuario `testuser2`.
5. En Columna, elija el nombre de la columna.
6. En Permisos de tabla, elija Seleccionar.
7. Elija Grant (Conceder).

   Cuando cree un enlace de recursos, solo podrá verlo y acceder a él el usuario. Para permitir que otros usuarios de su cuenta accedan al enlace de recursos, debe conceder permisos sobre el propio enlace de recursos. Debe conceder los permisos DESCRIBIR o ELIMINAR. En la página Tablas, vuelva a seleccionar la tabla y, en el menú Acciones, Conceder.
8. En la sección Conceder permisos, seleccione Mi cuenta.

9. En Roles y usuarios de IAM, elija el usuario `testuser2`.
10. En Permisos de enlace de recursos, seleccione Describir.
11. Elija Grant (Conceder).
12. Inicie sesión en la consola de AWS en la cuenta B como `testuser2`.

    En la consola de Athena (https://console.aws.amazon.com/athena/), debería ver la base de datos y la tabla `acc_b_area_rl`. Ahora puede ejecutar una consulta en la tabla para consultar el valor de la columna a la que `testuser2` tiene acceso.

Permisos de incorporación a Lake Formation

AWS Lake Formation utiliza el AWS Glue Data Catalog para almacenar metadatos de los datos de Amazon S3 en forma de bases de datos y tablas. Las tablas almacenan información sobre los datos subyacentes, incluida la información sobre esquemas, particiones y ubicaciones de datos. Las bases de datos son colecciones de tablas. El catálogo de datos también contiene enlaces a recursos, que son enlaces a bases de datos y tablas compartidas en cuentas externas, y se utilizan para el acceso entre cuentas a los datos del lago de datos. Cada cuenta AWS dispone de un catálogo de datos por región AWS.

Lake Formation proporciona un modelo de permisos de sistema de administración de bases de datos relacionales (RDBMS) para conceder o revocar el acceso a bases de datos, tablas y columnas en el catálogo de datos con datos subyacentes en Amazon S3.

Antes de conocer los detalles del modelo de permisos de Lake Formation, es útil repasar los siguientes antecedentes:

- Los lagos de datos administrados por Lake Formation residen en ubicaciones designadas en Amazon Simple Storage Service (Amazon S3).
- Lake Formation mantiene un catálogo de datos que contiene metadatos sobre los orígenes de datos que se importarán en sus lagos de datos, como los datos de registros y bases de datos relacionales, y sobre los datos de sus lagos de datos en Amazon S3. Los metadatos se organizan en forma de bases de datos y tablas. Las tablas de metadatos contienen el esquema, la ubicación, la partición y otra información sobre los datos que representan. Las bases de datos son colecciones de tablas.
- El catálogo de datos de Lake Formation es el mismo catálogo de datos utilizado por AWS Glue. Puede utilizar rastreadores AWS Glue para crear tablas del catálogo de datos, y trabajos de extracción, transformación y carga (ETL) AWS Glue para poblar los datos subyacentes en sus lagos de datos.
- Las bases de datos y las tablas del catálogo de datos se denominan recursos del catálogo de datos. Las tablas del catálogo de datos se denominan tablas de metadatos para distinguirlas de las tablas de los orígenes de datos o de los datos tabulares de Amazon S3. Los datos a los que apuntan las tablas de metadatos en Amazon S3 o en los orígenes de datos se denominan datos subyacentes.

- Una entidad principal es un usuario o rol, un usuario o grupo de Amazon QuickSight, un usuario o grupo que se autentifica con Lake Formation a través de un proveedor SAML, o para el control de acceso entre cuentas, un ID de cuenta AWS, ID de organización o ID de unidad organizativa.
- Los rastreadores AWS Glue crean tablas de metadatos, pero también puede crear manualmente tablas de metadatos con la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI). Al crear una tabla de metadatos, debe especificar una ubicación. Al crear una base de datos, la ubicación es opcional. Las ubicaciones de las tablas pueden ser ubicaciones de Amazon S3 o ubicaciones de orígenes de datos como una base de datos de Amazon Relational Database Service (Amazon RDS). Las ubicaciones de las bases de datos son siempre ubicaciones de Amazon S3.
- Los servicios que se integran con Lake Formation, como Amazon Athena y Amazon Redshift, pueden acceder al catálogo de datos para obtener metadatos y comprobar la autorización para ejecutar consultas. Para obtener una lista completa de los servicios integrados, consulte Integraciones de servicios de AWS con Lake Formation.

Temas

Descripción general de los permisos de Lake Formation

Hay dos tipos principales de permisos en AWS Lake Formation:

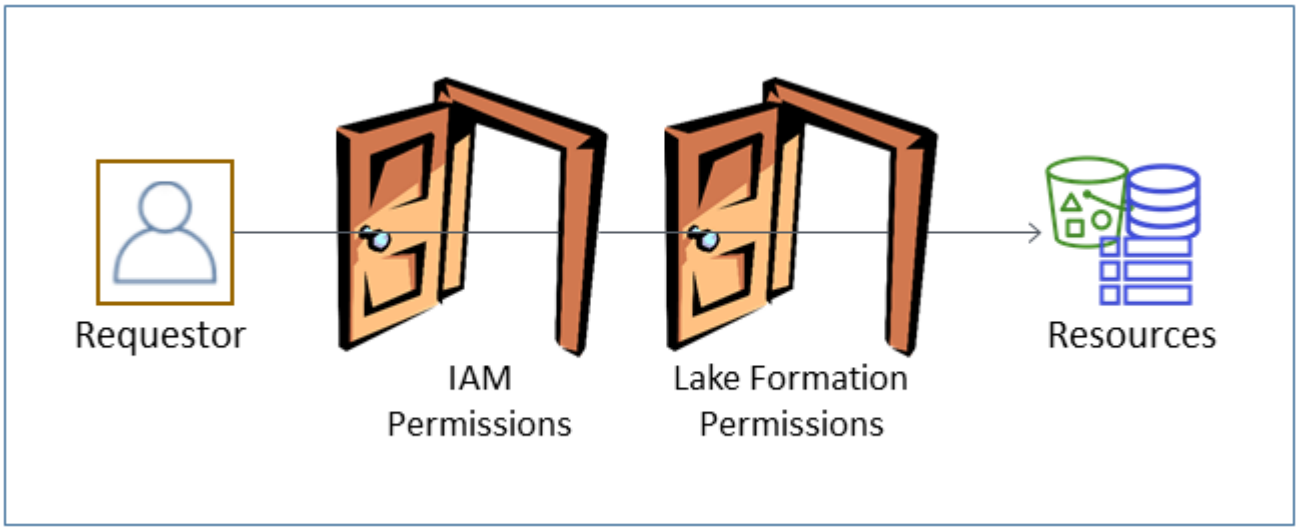
- Acceso a los metadatos. Permisos sobre los recursos del catálogo de datos (Permisos sobre el catálogo de datos).

Con estos permisos las entidades principales pueden crear, leer, actualizar y eliminar bases de datos de metadatos y tablas del catálogo de datos.

- Acceso a datos subyacentes. Permisos sobre ubicaciones en Amazon Simple Storage Service (Amazon S3) (permisos de acceso a datos y permisos de ubicación de datos).
  - Los permisos del lago de datos permiten a las entidades principales leer y escribir datos en las ubicaciones subyacentes de Amazon S3, datos a los que apuntan los recursos del catálogo de datos.
  - Los permisos de ubicación de datos permiten a las entidades principales crear y modificar bases de datos y tablas de metadatos que apunten a ubicaciones específicas de Amazon S3.

Para ambas áreas, Lake Formation utiliza una combinación de permisos de Lake Formation y permisos AWS Identity and Access Management (IAM). El modelo de permisos IAM se compone de políticas de IAM. El modelo de permisos de Lake Formation se implementa como comandos CONCEDER/REVOCAR de estilo DBMS, como `Grant SELECT on` *`tableName`* `to` *`userName`*.

Cuando una entidad principal efectúa una solicitud para acceder a los recursos del catálogo de datos o a los datos subyacentes, para que la solicitud tenga éxito, debe pasar las comprobaciones de permisos tanto de IAM como de Lake Formation.



Los permisos de Lake Formation controlan el acceso a los recursos del catálogo de datos, a las ubicaciones de Amazon S3 y a los datos subyacentes en dichas ubicaciones. Los permisos de IAM controlan el acceso a la Lake Formation y a los recursos y las API de AWS Glue. Así, aunque tenga el permiso Lake Formation para crear una tabla de metadatos en el catálogo de datos

(CREATE_TABLE), su operación fallará si no tiene el permiso IAM en la API `glue:CreateTable`. (¿Por qué un permiso `glue:`? porque Lake Formation usa el catálogo de datos AWS Glue).

> **Note**
>
> Los permisos de Lake Formation se aplican solo en la Región en la que fueron concedidos.

AWS Lake Formation requiere que cada entidad principal (usuario o rol) esté autorizada a ejecutar acciones sobre los recursos administrados por Lake Formation. El administrador del lago de datos u otra entidad principal con permisos para conceder permisos de Lake Formation concede a la entidad principal las autorizaciones necesarias.

Cuando concede un permiso de Lake Formation a una entidad principal, puede conceder de forma opcional la capacidad de pasar ese permiso a otra entidad principal.

Puede utilizar la API de Lake Formation, la AWS Command Line Interface (AWS CLI) o las páginas Permisos de datos y Ubicaciones de datos de la consola de Lake Formation para conceder y revocar permisos de Lake Formation.

Métodos para el control de acceso específico

Con un lago de datos, el objetivo es tener un control de acceso detallado a los datos. En Lake Formation, esto significa un control de acceso específico a los recursos del catálogo de datos y a las ubicaciones de Amazon S3. Puede lograr un control de acceso específico con uno de los siguientes métodos.

Método	Permisos de Lake Formation	Permisos de IAM	Comentarios
Método 1	Open	Específicos	Este es el método predeterminado por compatibilidad inversa con AWS Glue. • Abierto significa que el permiso especial `Super` se concede al grupo `IAMAllowedPrincipals` , donde se crea automáticamente `IAMAllowedPrincipals` e incluye a todos

Método	Permisos de Lake Formation	Permisos de IAM	Comentarios
			los usuarios y roles de IAM que tienen permitido el acceso a los recursos de su catálogo de datos por sus políticas de IAM, y el permiso `Super` permite a una entidad principal todas las operaciones de Lake Formation compatibles en la base de datos o tabla en la que se concede. Esto hace que el acceso a los recursos del catálogo de datos y a las ubicaciones de Amazon S3 esté controlado únicamente por las políticas de IAM. Para obtener más información, consulte Cambiar la configuración predeterminada de su lago de datos y Actualización de los permisos de datos AWS Glue al modelo AWS Lake Formation. • Específico significa que las políticas de IAM controlan todo el acceso a los recursos del catálogo de datos y a los buckets individuales de Amazon S3. En la consola de Lake Formation, este método aparece como Utilizar solo control de acceso IAM.

Método	Permisos de Lake Formation	Permisos de IAM	Comentarios
Método 2	Específicos	Básicos	Esta es la opción recomendada. • El acceso específico implica la concesión de permisos limitados de Lake Formation a entidades principal es individuales sobre los recursos del catálogo de datos, las ubicaciones de Amazon S3 y los datos subyacentes en dichas ubicaciones. • Básicos significa permisos más amplios sobre operaciones individuales y sobre el acceso a las ubicaciones de Amazon S3. Por ejemplo, una política de IAM básica podría incluir `"glue:*"` o `"glue:Create*"` en lugar de `"glue:CreateTables"`, dejando que los permisos de Lake Formation controlen si una entidad principal puede o no crear objetos de catálogo. También significa dar a las entidades principal es acceso a las API que necesitan para hacer su trabajo, pero bloqueando otras API y recursos. Por ejemplo, puede crear una política de IAM que permita a una entidad principal crear recursos del catálogo de datos y crear y ejecutar flujos de trabajo, pero que no permita la creación de conexiones AWS Glue o funciones definidas por el usuario. Vea los ejemplos más adelante en esta sección.

> **Important**
>
> Tenga en cuenta lo siguiente:
>
> - De forma predeterminada, Lake Formation tiene activada la configuración de control de acceso Utilizar solo IAM por compatibilidad con el comportamiento existente del catálogo de datos AWS Glue. Le recomendamos que desactive esta configuración después de pasar a usar los permisos de Lake Formation. Para obtener más información, consulte Cambiar la configuración predeterminada de su lago de datos.
> - Los administradores de lagos de datos y los creadores de bases de datos tienen permisos implícitos de Lake Formation que debe comprender. Para obtener más información, consulte Permisos implícitos de Lake Formation.

Control de acceso a los metadatos

Para el control de acceso a los recursos del catálogo de datos, en el siguiente análisis se asume un control de acceso detallado con los permisos de Lake Formation y un control de acceso más detallado con las políticas de IAM.

Existen dos métodos distintos para la concesión de permisos de Lake Formation sobre los recursos del catálogo de datos:

- Control de acceso a recursos con nombre. Con este método, el usuario concede permisos sobre bases de datos o tablas específicas especificando nombres de bases de datos o tablas. Las concesiones tienen esta forma:

  Conceda permisos a las entidades principales sobre los recursos [con opción de concesión].

  Con la opción de concesión, puede permitir que el beneficiario conceda los permisos a otras entidades principales.
- Control de acceso basado en etiquetas. Con este método, asigna una o varias etiquetas LF a las bases de datos, tablas y columnas del catálogo de datos, y concede permisos sobre una o varias etiquetas LF a las entidades principales. Cada etiqueta LF es un par clave-valor, como `department=sales`. Una entidad principal con etiquetas LF que coincidan con las etiquetas LF de un recurso puede acceder a ese recurso. Este método se recomienda para los lagos de datos con un gran número de bases de datos y tablas. Se explica en detalle en Control de acceso basado en etiquetas de Lake Formation.

Los permisos que una entidad principal tiene sobre un recurso son la unión de los permisos concedidos por ambos métodos.

La siguiente tabla resume los permisos disponibles de Lake Formation en los recursos del catálogo de datos. Los títulos de las columnas indican el recurso sobre el que se concede el permiso.

Catálogo	Base de datos	Tabla
CREATE_DATABASE	CREATE_TABLE	ALTER
	ALTER	DROP
	DROP	DESCRIBE
	DESCRIBE	SELECT*
		INSERT*
		DELETE*

Por ejemplo, el permiso CREATE_TABLE se concede sobre una base de datos. Esto significa que la entidad principal está autorizada a crear tablas en esa base de datos.

Los permisos con un asterisco (*) se conceden sobre los recursos del catálogo de datos, pero se aplican a los datos subyacentes. Por ejemplo, el permiso DROP sobre una tabla de metadatos le permite eliminar la tabla del catálogo de datos. Sin embargo, el permiso DELETE concedido sobre la misma tabla le permite eliminar los datos subyacentes de la tabla en Amazon S3, utilizando, por ejemplo, una instrucción SQL DELETE. Con estos permisos, también puede ver la tabla en la consola de Lake Formation y recuperar información sobre la tabla con la API AWS Glue. Así, SELECT, INSERT y DELETE son tanto permisos del catálogo de datos como permisos de acceso a los datos.

Al conceder SELECT en una tabla, puede añadir un filtro que incluya o excluya una o más columnas. Esto permite un control de acceso específico sobre las columnas de la tabla de metadatos, limitando las columnas que los usuarios de los servicios integrados pueden ver al ejecutar consultas. Esta capacidad no está disponible solo con políticas de IAM.

También hay un permiso especial denominado Super. El permiso Super permite a una entidad principal efectuar todas las operaciones compatibles con Lake Formation en la base de datos o tabla sobre la que se concede. Este permiso puede coexistir con los demás permisos de Lake

Formation. Por ejemplo, puede conceder `Super`, `SELECT` y `INSERT` sobre una tabla de metadatos. La entidad principal puede efectuar todas las acciones compatibles en la tabla, y cuando revoca `Super`, permanecen los permisos `SELECT` y `INSERT`.

Para obtener más información sobre cada permiso, consulte Referencia de permisos de Lake Formation.

> **Important**
>
> Para consultar una tabla del catálogo de datos creada por otro usuario, debe tener al menos un permiso de Lake Formation sobre la tabla. Si se le concede al menos un permiso sobre la tabla, también podrá ver la base de datos contenedora de la tabla.

Puede conceder o revocar los permisos del catálogo de datos utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI). A continuación se muestra un ejemplo de comando AWS CLI que concede al usuario permiso `datalake_user1` para crear tablas en la base de datos `retail`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

A continuación, se muestra un ejemplo de política de IAM de control de acceso básico que complementa el control de acceso específico con los permisos de Lake Formation. Permite efectuar todas las operaciones sobre cualquier base de metadatos o tabla.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "glue:*Database*",
                "glue:*Table*",
                "glue:*Partition*"
            ],
            "Resource": "*"
        }
    ]
```

```
}
```

El siguiente ejemplo también es básico, pero algo más restrictivo. Permite operaciones de solo lectura en todas las bases de datos de metadatos y tablas del catálogo de datos de la cuenta y región designadas.

```
{
    "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:*"
        }
    ]
}
```

Compare estas políticas con la siguiente, que implementa un control de acceso específico basado en IAM. Concede permisos solo sobre un subconjunto de tablas de la base de datos de metadatos de gestión de relaciones con los clientes (CRM) en la cuenta y la región designadas.

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

Para más ejemplos de políticas de control de acceso básicas, consulte Personas de Lake Formation y referencia de permisos IAM.

Control de acceso a los datos subyacentes

Cuando un servicio integrado AWS solicita acceso a datos en una ubicación de Amazon S3 cuyo acceso está controlado por AWS Lake Formation, Lake Formation suministra credenciales temporales para acceder a los datos.

Para permitir que Lake Formation controle el acceso a los datos subyacentes en una ubicación de Amazon S3, el usuario registra dicha ubicación con Lake Formation.

Después de registrar una ubicación de Amazon S3, puede comenzar a conceder los siguientes permisos de Lake Formation:

- Permisos de acceso a los datos (`SELECT`, `INSERT`, y `DELETE`)) en las tablas del catálogo de datos que apuntan a esa ubicación.
- Permisos de ubicación de datos en esa ubicación.

Los permisos de ubicación de datos de Lake Formation controlan la capacidad de crear recursos del catálogo de datos que apunten a determinadas ubicaciones de Amazon S3. Los permisos de ubicación de datos proporcionan una capa adicional de seguridad a las ubicaciones dentro del lago de datos. Cuando concede el permiso `CREATE_TABLE` o `ALTER` a una entidad principal, también concede permisos de ubicación de datos para limitar las ubicaciones en las que la entidad principal puede crear o modificar tablas de metadatos.

Las ubicaciones de Amazon S3 son buckets o prefijos bajo un bucket, pero no objetos individuales de Amazon S3.

Puede conceder permisos de ubicación de datos a una entidad principal utilizando la consola de Lake Formation, la API o la AWS CLI. El formato general de una concesión es el siguiente:

```
grant DATA_LOCATION_ACCESS to principal on S3 location [with grant option]
```

Si incluye `with grant option`, el beneficiario puede conceder los permisos a otras entidades principales.

Recuerde que los permisos de Lake Formation siempre funcionan en combinación con los permisos de AWS Identity and Access Management (IAM) para un control de acceso específico. Para los permisos de lectura/escritura en los datos subyacentes de Amazon S3, los permisos IAM se conceden de la siguiente manera:

Cuando registra una ubicación, especifica un rol de IAM que concede permisos de lectura y escritura en esa ubicación. Lake Formation asume ese rol al proporcionar credenciales temporales a los servicios de AWS integrados. Un rol normal puede tener adjunta la siguiente política, en la que la ubicación registrada es el bucket `awsexamplebucket`.

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

Lake Formation proporciona un rol vinculado al servicio que puede utilizar durante el registro para crear automáticamente políticas como esta. Para obtener más información, consulte Uso de roles vinculados a servicios para Lake Formation.

Por lo tanto, el registro de una ubicación de Amazon S3 concede los permisos `s3:` IAM necesarios sobre dicha ubicación, donde los permisos vienen especificados por el rol utilizado para registrar la ubicación.

> **Important**
>
> Evite registrar un bucket de Amazon S3 que tenga activada la opción El solicitante paga. Para los buckets registrados en Lake Formation, el rol utilizado para registrar el bucket se considera siempre como el solicitante. Si otra cuenta de AWS accede al bucket, se cobrará al propietario del bucket por el acceso a los datos si el rol pertenece a la misma cuenta que el propietario del bucket.

Para el acceso de lectura/escritura a los datos subyacentes, además de los permisos de Lake Formation, las entidades principales también necesitan el siguiente permiso de IAM:

```
lakeformation:GetDataAccess
```

Con este permiso, Lake Formation concede la solicitud de credenciales temporales para acceder a los datos.

> **Note**
>
> Amazon Athena requiere que el usuario cuente con el permiso `lakeformation:GetDataAccess`. Otros servicios integrados requieren que su rol de ejecución subyacente cuente con el permiso `lakeformation:GetDataAccess`.

Este permiso está incluido en las políticas sugeridas en el Personas de Lake Formation y referencia de permisos IAM.

En resumen, para permitir a las entidades principales de Lake Formation leer y escribir datos subyacentes con acceso controlado por los permisos de Lake Formation:

- Registre las ubicaciones de Amazon S3 que contienen los datos con Lake Formation.
- Las entidades principales creadoras de tablas del catálogo de datos que apunten a ubicaciones de datos subyacentes deben tener permisos de ubicación de datos.

- Las entidades principales que leen y escriben datos subyacentes deben tener permisos de acceso a los datos de Lake Formation en las tablas del catálogo de datos que apuntan a las ubicaciones de datos subyacentes.
- Las entidades principales que lean y escriban datos subyacentes deben tener el permiso `lakeformation:GetDataAccess` IAM cuando la ubicación de datos subyacente esté registrada en Lake Formation.

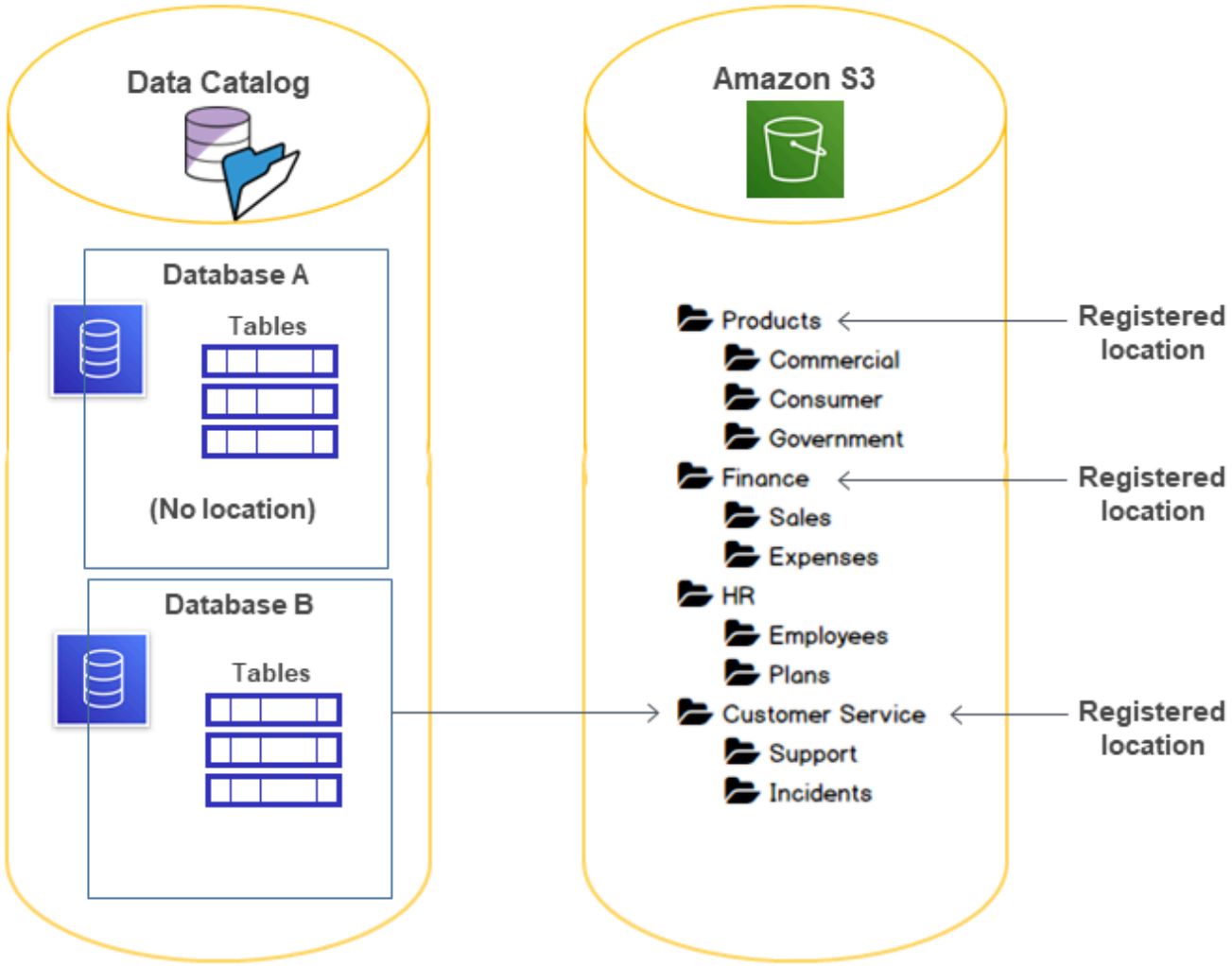
> **Note**
>
> El modelo de permisos de Lake Formation no impide el acceso a las ubicaciones de Amazon S3 a través de la API o la consola de Amazon S3 si tiene acceso a ellas a través de las políticas de IAM o Amazon S3. Puede adjuntar políticas de IAM a las entidades principales para bloquear este acceso.

Más información sobre los permisos de ubicación de datos

Los permisos de ubicación de datos rigen el resultado de las operaciones de creación y actualización en las bases de datos y tablas del catálogo de datos. Las normas son las siguientes:

- Una entidad principal debe tener permisos explícitos o implícitos de ubicación de datos en una ubicación de Amazon S3 para crear o actualizar una base de datos o tabla que especifique dicha ubicación.
- El permiso explícito `DATA_LOCATION_ACCESS` se concede utilizando la consola, la API o AWS CLI.
- Los permisos implícitos se conceden cuando una base de datos tiene una propiedad de ubicación que apunta a una ubicación registrada, la entidad principal tiene el permiso `CREATE_TABLE` en la base de datos y la entidad principal intenta crear una tabla en esa ubicación o en una ubicación secundaria.
- Si a una entidad principal se le conceden permisos de ubicación de datos en una ubicación, la entidad principal tendrá permisos de ubicación de datos en todas las ubicaciones secundarias.
- Una entidad principal no necesita permisos de ubicación de datos para efectuar las operaciones de lectura/escritura de los datos subyacentes. Basta con tener los o permisos de acceso a los datos `SELECT` o `INSERT`. Los permisos de ubicación de datos se aplican solo a la creación de recursos del catálogo de datos que apunten a la ubicación.

Piense en el escenario que se muestra en el siguiente diagrama.



En este diagrama:

- Los buckets de Amazon S3 `Products`, `Finance` y `Customer Service` están registrados en Lake Formation.
- `Database A` no tiene propiedad de ubicación, y `Database B` tiene una propiedad de ubicación que apunta al bucket `Customer Service`.
- El usuario `datalake_user` tiene `CREATE_TABLE` en ambas bases de datos.
- Se han concedido al usuario `datalake_user` permisos de ubicación de datos solo en el bucket `Products`.

A continuación se muestran los resultados cuando el usuario `datalake_user` intenta crear una tabla de catálogo en una base de datos concreta en una ubicación determinada.

Ubicación donde **`datalake_user`** intenta crear una tabla

Base de datos y ubicación	Éxito o fracaso	Motivo
Base de datos A en `Finance/Sales`	Fracaso	Sin permiso de ubicación de datos
Base de datos A en `Products`	Correcto	Tiene permiso de ubicación de datos
Base de datos A en `HR/Plans`	Correcto	La ubicación no está registrada
Base de datos B en `Customer Service/Incidents`	Correcto	La base de datos tiene la propiedad de ubicación en `Customer Service`

Para obtener más información, consulte lo siguiente:

- Añadir una ubicación de Amazon S3 a su lago de datos
- Referencia de permisos de Lake Formation
- Personas de Lake Formation y referencia de permisos IAM

Personas de Lake Formation y referencia de permisos IAM

Esta sección enumera algunas personas sugeridas de Lake Formation y sus permisos AWS Identity and Access Management (IAM) sugeridos. Para obtener información sobre los permisos de Lake Formation, consulte the section called “Referencia de permisos de Lake Formation”.

AWS Lake Formation personas

En la siguiente tabla se muestran las AWS Lake Formation personas sugeridas.

Personas de Lake Formation

Persona	Descripción
Administrador de IAM (superusuario)	(Obligatorio) Usuario que puede crear usuarios y roles de IAM. Tiene la política `AdministratorAccess` AWS gestionad

Persona	Descripción
	a. Tiene todos los permisos sobre todos los recursos de Lake Formation. Puede añadir administradores del lago de datos. No puede conceder permisos de Lake Formation si no ha sido designado también administrador del lago de datos.
Administrador del lago de datos	(Obligatorio) Usuario que puede registrar las ubicaciones de Amazon S3, acceder al catálogo de datos, crear bases de datos, crear y ejecutar flujos de trabajo, conceder permisos de Lake Formation a otros usuarios y ver los AWS CloudTrail registros . Tiene menos permisos IAM que el administrador IAM, pero suficientes para administrar el lago de datos. No se pueden añadir otros administradores del lago de datos.
Administrador de solo lectura	(Opcional) Usuario que puede ver entidades principales, recursos del Catálogo de datos, permisos y registros de AWS CloudTrail , sin permisos para las actualizaciones.
Ingeniero de datos	(Opcional) Usuario que puede crear bases de datos, crear y ejecutar rastreadores y flujos de trabajo, y conceder permisos de Lake Formation sobre las tablas del Catálogo de datos que crean los rastreadores y los flujos de trabajo. Le recomenda mos que convierta a todos los ingenieros de datos en creadores de bases de datos. Para obtener más información, consulte Creación de una base de datos.
Analista de datos	(Opcional) Usuario que puede ejecutar consultas en el lago de datos utilizando, por ejemplo, Amazon Athena. Solo tiene permisos suficientes para ejecutar consultas.
Rol de flujo de trabajo	(Obligatorio) Rol que ejecuta un flujo de trabajo en nombre de un usuario. Este rol se especifica cuando se crea un flujo de trabajo a partir de un esquema.

AWS políticas gestionadas para Lake Formation

Puede conceder los permisos AWS Identity and Access Management (de IAM) necesarios para trabajar con ellas AWS Lake Formation mediante políticas AWS gestionadas y políticas integradas. Las siguientes políticas AWS gestionadas están disponibles para Lake Formation.

AWS política gestionada: AWSLakeFormationDataAdmin

AWSLakeFormationDataAdminla política otorga acceso administrativo AWS Lake Formation y servicios relacionados, como la administración AWS Glue de lagos de datos.

Puede asociar `AWSLakeFormationDataAdmin` a los usuarios, grupos y roles.

Detalles del permiso

- `CloudTrail`— Permite a los directores ver los AWS CloudTrail registros. Esto es necesario para revisar cualquier error en la configuración del lago de datos.
- `Glue`. Permite a las entidades principales ver, crear y actualizar tablas de metadatos y bases de datos en el Catálogo de datos. Esto incluye las operaciones de la API que comienzan con `Get`, `List`, `Create`, `Update`, `Delete` y `Search`. Esto es necesario para administrar los metadatos de las tablas del lago de datos.
- `IAM`. Permite a las entidades principales recuperar información sobre los usuarios y roles de IAM y las políticas asociadas a los roles. Esto es necesario para que el administrador de datos revise y enumere los usuarios y roles de IAM para conceder los permisos de Lake Formation.
- `Lake Formation`. Concede a los administradores de los lagos de datos los permisos necesarios de Lake Formation para administrar los lagos de datos.
- `S3`. Permite a las entidades principales recuperar información sobre los buckets de Amazon S3 y sus ubicaciones con el fin de establecer la ubicación de datos para los lagos de datos.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:*",
                "cloudtrail:DescribeTrails",
                "cloudtrail:LookupEvents",
                "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:GetWorkflow",
                "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": "*"
        },
        {
            "Effect": "Deny",
            "Action": [
                "lakeformation:PutDataLakeSettings"
            ],
            "Resource": "*"
        }
    ]
}
```

> **Note**
>
> La política `AWSLakeFormationDataAdmin` no concede todos los permisos necesarios a los administradores del lago de datos. Se necesitan permisos adicionales para crear y ejecutar flujos de trabajo y registrar ubicaciones con el rol vinculado al servicio `AWSServiceRoleForLakeFormationDataAccess`. Para obtener más información, consulte Crear un administrador de lago de datos y Uso de roles vinculados a servicios para Lake Formation.

AWS política gestionada: AWSLakeFormationCrossAccountManager

AWSLakeFormationCrossAccountManagerla política proporciona acceso entre cuentas a AWS Glue los recursos a través de Lake Formation y otorga acceso de lectura a otros servicios requeridos, como AWS Organizations y AWS RAM.

Puede asociar `AWSLakeFormationCrossAccountManager` a los usuarios, grupos y roles.

Detalles del permiso

Esta política incluye los siguientes permisos.

- `Glue`. Permite a las entidades principales establecer o eliminar la política de recursos del Catálogo de datos para el control de acceso.
- `Organizations`. Permite a las entidades principales recuperar información sobre cuentas y unidades organizativas (UO) de una organización.
- `ram:CreateResourceShare`. Permite a las entidades principales crear un recurso compartido.
- `ram:UpdateResourceShare`. Permite a las entidades principales modificar algunas propiedades del recurso compartido especificado.
- `ram:DeleteResourceShare`. Permite a las entidades principales eliminar el recurso compartido especificado.
- `ram:AssociateResourceShare`. Permite a las entidades principales añadir la lista especificada de entidades principales y la lista de recursos a un recurso compartido.
- `ram:DisassociateResourceShare`. Permite a las entidades principales eliminar a las entidades principales o recursos especificados de la participación en el recurso compartido especificado.

- `ram:GetResourceShares`. Permite a las entidades principales recuperar detalles sobre los recursos compartidos que le pertenecen o que se comparten.
- `ram:RequestedResourceType`. Permite a las entidades principales recuperar el tipo de recurso (base de datos, tabla o catálogo).
- `AssociateResourceSharePermission`— Permite a los directores añadir o reemplazar el AWS RAM permiso para un tipo de recurso incluido en un recurso compartido. Puede tener exactamente un permiso asociado a cada tipo de recurso en el recurso compartido.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ram:CreateResourceShare"
            ],
            "Resource": "*",
            "Condition": {
                "StringLikeIfExists": {
                    "ram:RequestedResourceType": [
                        "glue:Table",
                        "glue:Database",
                        "glue:Catalog"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "ram:UpdateResourceShare",
                "ram:DeleteResourceShare",
                "ram:AssociateResourceShare",
                "ram:DisassociateResourceShare",
                "ram:GetResourceShares"
            ],
            "Resource": "*",
            "Condition": {
                "StringLike": {
                    "ram:ResourceShareName": [
                        "LakeFormation*"
```

```
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "ram:AssociateResourceSharePermission"
            ],
            "Resource": "*",
            "Condition": {
                "StringLike": {
                    "ram:PermissionArn": [
                        "arn:aws:ram::aws:permission/AWSRAMLFEnabled*"
                    ]
                }
            }
       },
        {
            "Effect": "Allow",
            "Action": [
                "glue:PutResourcePolicy",
                "glue:DeleteResourcePolicy",
                "organizations:DescribeOrganization",
                "organizations:DescribeAccount",
                "ram:Get*",
                "ram:List*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "organizations:ListRoots",
                "organizations:ListAccountsForParent",
                "organizations:ListOrganizationalUnitsForParent"
            ],
            "Resource": "*"
        }
    ]
}
```

AWS política gestionada: AWSGlueConsoleFullAccess

AWSGlueConsoleFullAccessla política otorga acceso total a AWS Glue los recursos cuando una identidad a la que está asociada la política utiliza la AWS Management Console. Si sigue la convención de nomenclatura para los recursos especificados en esta política, los usuarios dispondrán de todas las funciones de la consola. Esta política suele estar asociada a los usuarios de la AWS Glue consola.

Además, AWS Glue Lake Formation asume la función de servicio `AWSGlueServiceRole` para permitir el acceso a servicios relacionados, incluidos Amazon Elastic Compute Cloud (Amazon EC2), Amazon Simple Storage Service (Amazon S3) y Amazon. CloudWatch

AWS managed policy:LakeFormationDataAccessServiceRolePolicy

Esta política está asociada a un rol vinculado al servicio denominado `ServiceRoleForLakeFormationDataAccess` que permite al servicio realizar acciones sobre los recursos a petición suya. No puedes adjuntar esta política a tus identidades de IAM.

Esta política permite a los AWS servicios integrados de Lake Formation, como Amazon Athena Amazon Redshift, utilizar la función vinculada al servicio para descubrir los recursos de Amazon S3.

Para obtener más información, consulte Uso de roles vinculados a servicios para Lake Formation.

Detalles del permiso

Esta política incluye el siguiente permiso.

- `s3:ListAllMyBuckets`— Devuelve una lista de todos los depósitos propiedad del remitente autenticado de la solicitud.

```
{
 "Version": "2012-10-17",
 "Statement": [
  {
   "Sid": "LakeFormationDataAccessServiceRolePolicy",
   "Effect": "Allow",
   "Action": [
    "s3:ListAllMyBuckets"
   ],
   "Resource": [
    "arn:aws:s3:::*"
```

```
    ]
   }
  ]
}
```

Lake Formation actualiza sus políticas AWS gestionadas

Consulta los detalles sobre las actualizaciones de las políticas AWS gestionadas de Lake Formation desde que este servicio comenzó a rastrear estos cambios.

Cambio	Descripción	Fecha
Política `LakeFormationDataAccessServiceRolePolicy` actualizada de Lake Formation.	Lake Formation mejoró la LakeFormationDataAccessServiceRolePolicypolítica al agregar un elemento Sid a la declaración de política.	Febrero de 2024
Política `AWSLakeFormationCrossAccountManager` actualizada de Lake Formation.	Lake Formation mejoró la AWSLakeFormationCrossAccountManagerpolítica al agregar un nuevo permiso para permitir el intercambio de datos entre cuentas en modo de acceso híbrido.	Octubre de 2023
Política `AWSLakeFormationCrossAccountManager` actualizada de Lake Formation.	Lake Formation mejoró la AWSLakeFormationCrossAccountManagerpolítica para crear solo un recurso compartido por cuenta de destinatario cuando el recurso se comparte por primera vez. Todos los recursos compartidos a partir de entonces con la misma cuenta se adjuntan al mismo recurso compartido.	6 de mayo de 2022
Lake Formation inició el seguimiento de los cambios.	Lake Formation comenzó a rastrear los cambios en sus políticas AWS gestionadas.	6 de mayo de 2022

Permisos sugeridos para las personas

Los siguientes son los permisos sugeridos para cada persona. El administrador de IAM no está incluido porque ese usuario tiene todos los permisos sobre todos los recursos.

Temas

Permisos de administrador del lago de datos

> **Important**
>
> En las siguientes políticas, <account-id>sustitúyalo por un número de AWS cuenta válido y <workflow_role>sustitúyelo por el nombre de un rol que tenga permisos para ejecutar un flujo de trabajo, tal y como se define enPermisos del rol de flujo de trabajo.

Tipo de política	Política
AWS políticas gestionadas	• `AWSLakeFormationDataAdmin` • `LakeFormationDataAccessServiceRolePolicy` (política de funciones vinculadas al servicio) • `AWSGlueConsoleFullAccess` (opcional) • `CloudWatchLogsReadOnlyAccess` (Optional) • `AWSLakeFormationCrossAccountManager` (Optional) • `AmazonAthenaFullAccess` (opcional)

Tipo de política	Política
	Para obtener información sobre las políticas AWS administradas opcionales, consulte. the section called “Crear un administrador de lago de datos”

Política insertada (para crear el rol vinculado al servicio de Lake Formation)	{
	"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"
	}
	]
	}

Tipo de política	Política

(Opcional) Política insertada (política de passrole para el rol de flujo de trabajo). Esto solo es necesario si el administrador del lago de datos crea y ejecuta flujos de trabajo.	<pre>{
	    "Version": "2012-10-17",
	    "Statement": [
	        {
	            "Sid": "PassRolePermissions",
	            "Effect": "Allow",
	            "Action": [
	                "iam:PassRole"
	            ],
	            "Resource": [
	                "arn:aws:iam:: <account-id> :role/<workflow_role> "
	            ]
	        }
	    ]
	}</pre>

(Opcional) Política en línea (si su cuenta concede o recibe permisos entre cuentas de Lake Formation). Esta política sirve para aceptar o rechazar las invitaciones para compartir AWS RAM recursos y para permitir la concesión de permisos entre cuentas a las organizaciones. `ram:EnableSharingWithAwsOrganization` solo es obligatorio para los administradores del lago de datos de la cuenta de AWS Organizations administración.	<pre>{
	    "Version": "2012-10-17",
	    "Statement": [
	        {
	            "Effect": "Allow",
	            "Action": [
	                "ram:AcceptResourceShareInvitation",
	                "ram:RejectResourceShareInvitation",
	                "ec2:DescribeAvailabilityZones",
	                "ram:EnableSharingWithAwsOrganization"
	            ],
	            "Resource": "*"
	        }
	    ]
	}</pre>

Permisos de administrador de solo lectura

Tipo de política	Política
Política insertada (básica)	{ "Version":"2012-10-17", "Statement":[{ "Effect":"Allow", "Action":["lakeformation:GetEffectivePermissionsForPath", "lakeformation:ListPermissions", "lakeformation:ListDataCellsFilter", "lakeformation:GetDataCellsFilter", "lakeformation:SearchDatabasesByLFTags", "lakeformation:SearchTablesByLFTags", "lakeformation:GetLFTag", "lakeformation:ListLFTags", "lakeformation:GetResourceLFTags", "lakeformation:ListLakeFormationOption", "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",

Tipo de política	Política

	"iam:ListRoles",
	"iam:GetRole",
	"iam:GetRolePolicy"
	],
	"Resource":"*"
	},
	{
	"Effect":"Deny",
	"Action":[
	"lakeformation:PutDataLakeSettings"
	],
	"Resource":"*"
	}
	]
	}

Permisos de ingeniero de datos

> **Important**
>
> En las siguientes políticas, <account-id>sustitúyalo por un número de AWS cuenta válido y <workflow_role>sustitúyalo por el nombre del rol del flujo de trabajo.

Tipo de política	Política
AWS política gestionada	`AWSGlueConsoleFullAccess`
Política insertada (básica)	{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", "lakeformation:GrantPermissions", "lakeformation:RevokePermissions", "lakeformation:BatchGrantPermissions",

Tipo de política	Política
	"lakeformation:BatchRevokePermissions", "lakeformation:ListPermissions", "lakeformation:AddLFTagsToResource", "lakeformation:RemoveLFTagsFromResou rce", "lakeformation:GetResourceLFTags", "lakeformation:ListLFTags", "lakeformation:GetLFTag", "lakeformation:SearchTablesByLFTags", "lakeformation:SearchDatabasesByLFTa gs", "lakeformation:GetWorkUnits", "lakeformation:GetWorkUnitResults", "lakeformation:StartQueryPlanning", "lakeformation:GetQueryState", "lakeformation:GetQueryStatistics"], "Resource": "*" }] }

Tipo de política	Política

Política insertada (para las operaciones en tablas gobernadas, incluidas las operaciones dentro de las transacciones)	{
	"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": "*"
	}
	]
	}

Tipo de política	Política

Política insertada (para el control de acceso a metadatos mediante el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC))	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"lakeformation:AddLFTagsToResource",
	"lakeformation:RemoveLFTagsFromResou
	rce",
	"lakeformation:GetResourceLFTags",
	"lakeformation:ListLFTags",
	"lakeformation:GetLFTag",
	"lakeformation:SearchTablesByLFTags",
	"lakeformation:SearchDatabasesByLFTags"
	],
	"Resource": "*"
	}
	]
	}

Política insertada (política de passrole para el rol de flujo de trabajo)	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *<account-id>* :role/*<workflow
	_role>* "
	]
	}
	]
	}

Permisos de analista de datos

Tipo de política	Política
AWS política gestionada	`AmazonAthenaFullAccess`
Política insertada (básica)	`{ "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": "*" }] }`
(Opcional) Política insertada (para las operaciones en tablas gobernadas, incluidas las operaciones dentro de las transacciones)	`{ "Version": "2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [ "lakeformation:StartTransaction", "lakeformation:CommitTransaction", "lakeformation:CancelTransaction", "lakeformation:ExtendTransaction", "lakeformation:DescribeTransaction",`

Tipo de política	Política
	"lakeformation:ListTransactions", "lakeformation:GetTableObjects", "lakeformation:UpdateTableObjects", "lakeformation:DeleteObjectsOnCancel"], "Resource": "*" }] }

Permisos del rol de flujo de trabajo

Este rol cuenta con los permisos necesarios para ejecutar un flujo de trabajo. Se especifica un rol con estos permisos cuando se crea un flujo de trabajo.

> **Important**
>
> En las siguientes políticas, <region>sustitúyalo por un identificador de AWS región válido (por ejemplous-east-1), por <account-id>un número de AWS cuenta válido, por <workflow_role>el nombre del rol del flujo de trabajo y por <your-s3-cloudtrail-bucket>la ruta de Amazon S3 a tus AWS CloudTrail registros.

Tipo de política	Política
AWS política gestionada	`AWSGlueServiceRole`
Política insertada (acceso a datos)	{ "Version": "2012-10-17", "Statement": [{ "Sid": "Lakeformation", "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", "lakeformation:GrantPermissions"],

Tipo de política	Política

	<pre>"Resource": "*"
	}
	]
	}</pre>

Política insertada (política de passrole para el rol de flujo de trabajo)	<pre>{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: <account-id> :role/<workflow_role> "
	]
	}
	]
	}</pre>

Política en línea (para ingerir datos fuera del lago de datos, por ejemplo, AWS CloudTrail registros)	<pre>{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": ["s3:GetObject", "s3:ListBucket"],
	"Resource": ["arn:aws:s3::: <your-s3-cloudtrail-bucket> /*"]
	}
	]
	}</pre>

Cambiar la configuración predeterminada de su lago de datos

Para mantener la compatibilidad con versiones anteriores de AWS Glue, AWS Lake Formation tiene la siguiente configuración de seguridad inicial:

- El permiso `Super` se concede al grupo `IAMAllowedPrincipals` sobre todos los recursos existentes del catálogo de datos de AWS Glue.
- La configuración "Usar solo control de acceso de IAM" está activada para los nuevos recursos del catálogo de datos.

Estos ajustes hacen que el acceso a los recursos del catálogo de datos y a las ubicaciones de Amazon S3 esté controlado únicamente por las políticas de AWS Identity and Access Management (IAM). Los permisos individuales de Lake Formation no están en vigor.

El grupo `IAMAllowedPrincipals` incluye a todos los usuarios y roles de IAM a los que sus políticas de IAM permiten el acceso a los recursos de su catálogo de datos. El permiso `Super` permite a una entidad principal efectuar todas las operaciones compatibles con Lake Formation en la base de datos o tabla sobre la que se ha concedido.

Para cambiar la configuración de seguridad de modo que el acceso a los recursos del catálogo de datos (bases de datos y tablas) sea administrado por los permisos de Lake Formation, haga lo siguiente:

1. Cambie la configuración de seguridad predeterminada para los nuevos recursos. Para obtener más información, consulte Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido.
2. Cambiar la configuración de los recursos existentes del catálogo de datos. Para obtener más información, consulte Actualización de los permisos de datos AWS Glue al modelo AWS Lake Formation.

Cambiar la configuración de seguridad predeterminada mediante la operación de la API **`PutDataLakeSettings`** de Lake Formation

También puede cambiar la configuración de seguridad predeterminada mediante la operación de la API PutDataLakeSettings de Lake Formation. Esta acción toma como argumentos un ID de catálogo opcional y una estructura DataLakeSettings.

Para imponer el control de acceso a los metadatos y a los datos subyacentes por parte de Lake Formation en las nuevas bases de datos y tablas, codifique la estructura `DataLakeSettings` de la manera siguiente.

> **Note**
>
> Sustituya *<AccountID>* por un ID de cuenta AWS válido y *<Username>* por un nombre de usuario de IAM válido. Puede especificar más de un usuario como administrador del lago de datos.

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ],
        "CreateDatabaseDefaultPermissions": [],
        "CreateTableDefaultPermissions": []
    }
}
```

También puede codificar la estructura de la siguiente manera. Omitir el parámetro `CreateDatabaseDefaultPermissions` o `CreateTableDefaultPermissions` equivale a pasar una lista vacía.

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ]
    }
}
```

Esta acción revoca efectivamente todos los permisos de Lake Formation del grupo `IAMAllowedPrincipals` en las nuevas bases de datos y tablas. Cuando cree una base de datos, puede anular esta configuración.

Para imponer el control de acceso a los metadatos y a los datos subyacentes solo mediante IAM en las nuevas bases de datos y tablas, codifique la estructura de `DataLakeSettings` de la manera siguiente.

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

Esto concede al grupo `IAMAllowedPrincipals` el permiso `Super` de Lake Formation sobre nuevas bases de datos y tablas. Cuando cree una base de datos, puede anular esta configuración.

> **Note**
>
> En la estructura `DataLakeSettings` anterior, el único valor permitido para `DataLakePrincipalIdentifier` es `IAM_ALLOWED_PRINCIPALS`, y el único valor permitido para `Permissions` es `ALL`.

Permisos implícitos de Lake Formation

AWS Lake Formation concede las siguientes concesiones implícitas a los administradores del lago de datos, a los creadores de bases de datos y a los creadores de tablas.

Administrador de lago de datos

- Tener acceso de `Describe` a todos los recursos del catálogo de datos excepto a los recursos compartidos desde otra cuenta directamente a otra entidad principal. Este acceso no puede ser revocado por un administrador.
- Disponga de permisos de ubicación de datos en todas las partes del lago de datos.
- Puede conceder o revocar el acceso a cualquier recurso del catálogo de datos a cualquier entidad principal (incluida la propia). Este acceso no puede ser revocado por un administrador.
- Puede crear bases de datos en el catálogo de datos.
- Puede conceder el permiso para crear una base de datos a otro usuario.

> **Note**
>
> Los administradores del lago de datos pueden registrar ubicaciones de Amazon S3 solo si disponen de permisos de IAM para hacerlo. Las políticas del administrador del lago de datos sugeridas en esta guía conceden esos permisos. Además, los administradores del lago de datos no tienen permisos implícitos para eliminar bases de datos o alterar/eliminar tablas creadas por otros. Sin embargo, pueden concederse a sí mismos permisos para hacerlo.

Para obtener más información sobre los administradores del lago de datos, consulte Crear un administrador de lago de datos.

Creadores de bases de datos

- Tienen todos los permisos sobre las bases de datos que crean, tienen permisos sobre las tablas que crean en la base de datos y pueden conceder a otras entidades principales de la misma cuenta AWS permiso para crear tablas en la base de datos. Un creador de bases de datos que también disponga de la política administrada `AWSLakeFormationCrossAccountManager` de AWS puede conceder permisos sobre la base de datos a otras cuentas AWS u organizaciones.

  Los administradores del lago de datos pueden utilizar la consola de Lake Formation o la API para designar a los creadores de las bases de datos.

  > **Note**
  >
  > Los creadores de bases de datos no tienen permisos implícitos sobre las tablas que otros crean en la base de datos.

Para obtener más información, consulte Creación de una base de datos.

Creadores de tablas

- Tienen todos los permisos sobre las tablas que crean.
- Pueden conceder permisos sobre todas las tablas que creen a las entidades principales de la misma cuenta AWS.
- Pueden conceder permisos sobre todas las tablas que creen a otras cuentas de AWS u organizaciones si disponen de la política administrada `AWSLakeFormationCrossAccountManager` de AWS.
- Pueden ver las bases de datos que contienen las tablas que crean.

Referencia de permisos de Lake Formation

Para realizar AWS Lake Formation las operaciones, los directores necesitan tanto los permisos de Lake Formation como los permisos AWS Identity and Access Management (IAM). En general, los permisos IAM se conceden mediante políticas de control de acceso poco específicas, como se describe en the section called “Descripción general de los permisos de Lake Formation ”. Puede conceder permisos a Lake Formation mediante la consola, la API o AWS Command Line Interface (AWS CLI).

Para saber cómo conceder o revocar permisos de Lake Formation, consulte the section called "Concesión y revocación de permisos del catálogo de datos" y the section called "Conceder permisos de ubicación de datos".

> **Note**
>
> Los ejemplos de esta sección muestran cómo conceder permisos a entidades principales en la misma cuenta AWS . Para ver ejemplos de concesiones entre cuentas, consulte the section called "Cómo compartir datos entre cuentas".

Permisos de Lake Formation por tipo de recurso

A continuación se indican los permisos válidos de Lake Formation disponibles para cada tipo de recurso:

Recurso	Permiso
`Database`	`ALL (Super)`
	`ALTER`
	`CREATE_TABLE`
	`DESCRIBE`
	`DROP`
`Table`	`ALL (Super)`
	`ALTER`
	`DELETE`
	`DESCRIBE`
	`DROP`
	`INSERT`

Recurso	Permiso
	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
LF-Tag policy - Database	ALL (Super)
	ALTER
	CREATE_TABLE
	DESCRIBE
	DROP
LF-Tag policy - Table	ALL (Super)
	ALTER
	DESCRIBE

Recurso	Permiso
	DELETE
	DROP
	INSERT
	SELECT
Resource link - Database or Table	DESCRIBE
	DROP
Table with data filters	DESCRIBE
	DROP
	SELECT
Table with column filter	SELECT

Temas

- Lake Formation otorga y revoca órdenes AWS CLI
- Permisos de Lake Formation

Lake Formation otorga y revoca órdenes AWS CLI

Cada descripción de permisos de esta sección incluye ejemplos de cómo se concede el permiso mediante un AWS CLI comando. Las siguientes son las sinopsis de Lake Formation grant-permissions y revoke-permissions AWS CLI sus comandos.

```
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>]
```

Para obtener descripciones detalladas de estos comandos, consulte grant-permissions y revoke-permissions en la Referencia de comandos de AWS CLI . En esta sección se proporciona información adicional sobre la opción `--principal`.

El valor de la opción `--principal` es uno de los siguientes:

- Nombre de recurso de Amazon (ARN) para un usuario o AWS Identity and Access Management rol (IAM)
- ARN para un usuario o grupo que se autentica a través de un proveedor SAML, como Microsoft Active Directory Federation Service (AD FS)
- ARN de un QuickSight usuario o grupo de Amazon
- Para los permisos entre cuentas, un identificador de AWS cuenta, un identificador de organización o un identificador de unidad organizativa

A continuación encontrará la sintaxis y ejemplos para todos los tipos de `--principal`.

La entidad principal es un usuario de IAM

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
```

Ejemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1
```

La entidad principal es un rol de IAM

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
```

Ejemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/workflowrole
```

La entidad principal es un usuario que se autentica a través de un proveedor SAML

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:user/<user-name>
```

Ejemplos:

```
--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
```

La entidad principal es un grupo que se autentica a través de un proveedor SAML

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:group/<group-name>
```

Ejemplos:

```
--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 es usuario de Amazon QuickSight Enterprise Edition

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-id>:user/<namespace>/<user-name>
```

> **Note**
>
> En *<namespace>*, debe especificar `default`.

Ejemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-east-1:111122223333:user/default/bi_user1
```

Principal es un grupo de Amazon QuickSight Enterprise Edition

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-id>:group/<namespace>/<group-name>
```

> **Note**
>
> En *<namespace>*, debe especificar `default`.

Ejemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-east-1:111122223333:group/default/data_scientists
```

Principal es una AWS cuenta

Sintaxis:

```
--principal DataLakePrincipalIdentifier=<account-id>
```

Ejemplo:

```
--principal DataLakePrincipalIdentifier=111122223333
```

La entidad principal es una organización

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-id>:organization/<organization-id>
```

Ejemplo:

```
--principal
 DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/o-abcdefghijkl
```

La entidad principal es una unidad organizativa

Sintaxis:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-id>:ou/<organization-id>/<organizational-unit-id>
```

Ejemplo:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:ou/o-abcdefghijkl/ou-ab00-cdefghij
```

Principal es un usuario o grupo con identidad del Centro de Identidad de IAM

Ejemplo: usuario

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserID>
```

Ejemplo: Grupo:

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::group/<GroupID>
```

El director es un grupo de IAM - **`IAMAllowedPrincipals`**

Lake Formation otorga `Super` permisos en todas las bases de datos y tablas del catálogo de datos a un grupo llamado de forma `IAMAllowedPrincipals` predeterminada. Si este permiso de grupo existe en una base de datos o una tabla, todos los responsables de su cuenta tendrán acceso al recurso mediante las políticas principales de IAM. AWS Glue Proporciona compatibilidad con versiones anteriores al empezar a utilizar los permisos de Lake Formation para proteger los recursos del catálogo de datos para los que anteriormente estaban protegidos por AWS Glue las políticas de IAM.

Cuando usa Lake Formation para administrar los permisos de los recursos de su catálogo de datos, primero debe revocar el `IAMAllowedPrincipals` permiso de los recursos o optar por el modo de acceso híbrido entre los principales y los recursos para que los permisos de Lake Formation funcionen.

Ejemplo:

```
--principal DataLakePrincipalIdentifier=IAM_Allowed_Principals
```

Principal es un grupo de IAM - **`ALLIAMPrincipals`**

Al conceder permisos para `ALLIAMPrincipals` agrupar en un recurso del catálogo de datos, todos los directores de la cuenta tienen acceso al recurso del catálogo de datos mediante los permisos de Lake Formation y los permisos de IAM.

Ejemplo:

```
--principal DataLakePrincipalIdentifier=123456789012:IAMPrincipals
```

Permisos de Lake Formation

Esta sección contiene los permisos disponibles de Lake Formation que puede conceder a las entidades principales.

ALTER

Permiso	Concedido sobre este recurso	El beneficiario también necesita
ALTER	DATABASE	glue:UpdateDatabase
ALTER	TABLE	glue:UpdateTable
ALTER	LF-Tag	lakeformation:UpdateLFTag

Una entidad principal con este permiso puede alterar los metadatos de una base de datos o tabla del Catálogo de datos. Para las tablas, puede cambiar el esquema de columnas y añadir parámetros de columna. No puede alterar las columnas de los datos subyacentes a los que apunta una tabla de metadatos.

Si la propiedad que se está modificando es una ubicación registrada de Amazon Simple Storage Service (Amazon S3), la entidad principal debe tener permisos de ubicación de datos en la nueva ubicación.

Example

En el siguiente ejemplo, se concede el ALTER permiso al usuario datalake_user1 de la base de datos de la AWS cuenta retail 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "ALTER" --resource '{ "Database": {"Name":"retail"}}'
```

Example

El siguiente ejemplo concede ALTER al usuario datalake_user1 en la tabla inventory de la base de datos retail.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

CREATE_DATABASE

Permiso	Concedido sobre este recurso	El beneficiario también necesita
CREATE_DATABASE	Data Catalog	glue:CreateDatabase

Una entidad principal con este permiso puede crear una base de metadatos o un enlace de recursos en el Catálogo de datos. La entidad principal también puede crear tablas en la base de datos.

Example

En el siguiente ejemplo, se concede CREATE_DATABASE al usuario datalake_user1 de la cuenta 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

Cuando una entidad principal crea una base de datos en el Catálogo de datos, no se conceden permisos a los datos subyacentes. Se conceden los siguientes permisos adicionales para metadatos (junto con la posibilidad de conceder estos permisos a otros):

- CREATE_TABLE en la base de datos
- Base de datos de ALTER
- Base de datos de DROP

Al crear una base de datos, la entidad principal puede especificar de forma opcional una ubicación de Amazon S3. Dependiendo de si la entidad principal tiene permisos de localización de datos, el permiso CREATE_DATABASE podría no ser suficiente para crear bases de datos en todos los casos. Es importante tener en cuenta los siguientes puntos.

Caso práctico de creación de una base de datos	Permisos necesarios
La propiedad de ubicación no está especific ada.	CREATE_DATABASE es suficiente.

Caso práctico de creación de una base de datos	Permisos necesarios
Se especifica la propiedad de ubicación, y la ubicación no está administrada por Lake Formation (no está registrada).	`CREATE_DATABASE` es suficiente.
Se especifica la propiedad de ubicación, y la ubicación es administrada por Lake Formation (está registrada).	Se requiere `CREATE_DATABASE` , además de permisos de localización de datos en la ubicación especificada.

`CREATE_TABLE`

Permiso	Concedido sobre este recurso	El beneficiario también necesita
`CREATE_TABLE`	`DATABASE`	`glue:CreateTable`

Una entidad principal con este permiso puede crear una tabla de metadatos o un enlace de recursos en el Catálogo de datos dentro de la base de datos especificada.

Example

El siguiente ejemplo concede al usuario `datalake_user1` permiso para crear tablas en la `retail` base de datos de la cuenta 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Cuando una entidad principal crea una tabla en el Catálogo de datos, se le conceden todos los permisos de Lake Formation sobre la tabla, con la posibilidad de conceder estos permisos sobre otros.

Subsidios entre cuentas

Si una cuenta propietaria de una base de datos concede CREATE_TABLE a una cuenta de destinatario, y un usuario de la cuenta de destinatario crea con éxito una tabla en la base de datos de la cuenta propietaria, se aplican las siguientes reglas:

- El usuario y los administradores del lago de datos de la cuenta de destinatario tienen todos los permisos de Lake Formation sobre la tabla. Pueden conceder permisos sobre la tabla a otras entidades principales de su cuenta. No pueden conceder permisos a entidades principales en la cuenta del propietario ni en ninguna otra cuenta.
- Los administradores del lago de datos de la cuenta del propietario pueden conceder permisos sobre la tabla a otras entidades principales de su cuenta.

Permisos de ubicación de datos

Cuando intente crear una tabla que apunte a una ubicación de Amazon S3, dependiendo de si dispone de permisos de ubicación de datos, es posible que el permiso CREATE_TABLE no sea suficiente para crear una tabla. Es importante tener en cuenta los tres casos siguientes.

Caso práctico de creación de tablas	Permisos necesarios
Lake Formation no administra la ubicación especificada (no está registrada).	CREATE_TABLE es suficiente.
Lake Formation administra la ubicación especificada (está registrada), y la base de datos que la contiene no tiene propiedad de ubicación o tiene una propiedad de ubicación que no es un prefijo de Amazon S3 de la ubicación de la tabla.	Se requiere CREATE_TABLE , además de permisos de localización de datos en la ubicación especificada.
Lake Formation administra la ubicación especificada (está registrada), y la base de datos contenedora tiene una propiedad de ubicación que apunta a una ubicación que está registrada y es un prefijo de Amazon S3 de la ubicación de la tabla.	CREATE_TABLE es suficiente.

DATA_LOCATION_ACCESS

Permiso	Concedido sobre este recurso	El beneficiario también necesita
DATA_LOCATION_ACCESS	Ubicación de Amazon S3	(Permisos de Amazon S3 en la ubicación, que deben especificarse en el rol utilizado para registrar la ubicación).

Este es el único permiso de ubicación de datos. Una entidad principal con este permiso puede crear una base de datos o tabla de metadatos que apunte a la ubicación de Amazon S3 especificada. La ubicación debe estar registrada. Una entidad principal que tiene permisos de localización de datos en una localización también tiene permisos de localización en las localizaciones secundarias.

Example

El siguiente ejemplo concede permisos de ubicación de datos en `s3://products/retail` al usuario `datalake_user1` en la cuenta de AWS 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` no es necesario para consultar o actualizar los datos subyacentes. Este permiso se aplica únicamente a la creación de recursos del Catálogo de datos.

Para obtener más información sobre permisos de ubicación de datos, consulte Underlying data access control.

DELETE

Permiso	Concedido sobre este recurso	El beneficiario también necesita
DELETE	TABLE	(No se necesitan permisos de IAM adicionales si la ubicación está registrada).

Una entidad principal con este permiso puede eliminar datos subyacentes en la ubicación de Amazon S3 especificada por la tabla. La entidad principal también puede ver la tabla en la consola de Lake Formation y recuperar información sobre la tabla con la API AWS Glue.

Example

El siguiente ejemplo concede el DELETE permiso al usuario datalake_user1 de la tabla de la base de datos de inventory la AWS cuenta retail 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DELETE" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Este permiso se aplica solo a los datos de Amazon S3 y no a los de otros almacenes de datos como Amazon Relational Database Service (Amazon RDS).

DESCRIBE

Permiso	Concedido sobre este recurso	El beneficiario también necesita
DESCRIBE	Enlace de recurso a tabla	glue:GetTable
	Enlace de recurso a base de datos	glue:GetDatabase
DESCRIBE	DATABASE	glue:GetDatabase
DESCRIBE	TABLE	glue:GetTable

Permiso	Concedido sobre este recurso	El beneficiario también necesita

`DESCRIBE`	`LF-Tag`	`glue:GetTable`
		`glue:GetDatabase`
		`lakeformation:GetResourceLFTags`
		`lakeformation:ListLFTags`
		`lakeformation:GetLFTag`
		`lakeformation:SearchTablesByLFTags`
		`lakeformation:SearchDatabasesByLFTags`

Una entidad principal con este permiso puede ver la base de datos, tabla o enlace de recursos especificados. No se concede implícitamente ningún otro permiso del Catálogo de datos ni ningún permiso de acceso a los datos. Las bases de datos y las tablas aparecen en los editores de consultas de los servicios integrados, pero no se puede consultar en ellas salvo que se concedan otros permisos de Lake Formation (por ejemplo, `SELECT`).

Por ejemplo, un usuario que tiene `DESCRIBE` en una base de datos puede ver la base de datos y todos los metadatos de la base de datos (descripción, ubicación, etc.). Sin embargo, el usuario no puede averiguar qué tablas contiene la base de datos y no puede eliminar, modificar o crear tablas en la base de datos. Del mismo modo, un usuario que tiene `DESCRIBE` en una tabla puede ver la tabla y los metadatos de la tabla (descripción, esquema, ubicación, etc.), pero no puede soltar, alterar o ejecutar consultas contra la tabla.

A continuación se detallan algunas normas adicionales para `DESCRIBE`:

- Si un usuario tiene otros permisos de Lake Formation sobre una base de datos, tabla o enlace de recursos, `DESCRIBE` se le concede implícitamente.

- Si un usuario tiene `SELECT` solo en un subconjunto de columnas de una tabla (`SELECT` parcial), el usuario estará restringido a ver solo esas columnas.
- No puede conceder `DESCRIBE` a un usuario que tiene selección parcial en una tabla. Por el contrario, no puede especificar listas de inclusión o exclusión de columnas para las tablas sobre las que se concede `DESCRIBE`.

Example

En el siguiente ejemplo, se concede el `DESCRIBE` permiso al usuario `datalake_user1` en el enlace de recursos de la tabla de la base `retail` de datos de la cuenta `inventory-link` 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory-link"}}'
```

DROP

Permiso	Concedido sobre este recurso	El beneficiario también necesita
`DROP`	`DATABASE`	`glue:DeleteDatabase`
`DROP`	`TABLE`	`glue:DeleteTable`
`DROP`	`LF-Tag`	`lakeformation:DeleteLFTag`

`DROP`	Enlace de recurso a base de datos
	
	Enlace de recurso a tabla
		`glue:DeleteDatabase`
		
		`glue:DeleteTable`

Una entidad principal con este permiso puede borrar una base de datos, tabla o enlace de recurso en el Catálogo de datos. No puede conceder el permiso DROP en una base de datos a una cuenta u organización externa.

> **Warning**
>
> Al eliminar una base de datos se eliminan todas sus tablas.

Example

En el siguiente ejemplo, se concede el DROP permiso al usuario de la base de datos de la cuenta `datalake_user1` 1111-2222-3333`retail`. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Database": {"Name":"retail"}}'
```

Example

El siguiente ejemplo concede DROP al usuario `datalake_user1` sobre la tabla `inventory` en la base de datos `retail`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Example

El siguiente ejemplo concede DROP al usuario `datalake_user1` sobre el enlace de recursos de tabla `inventory-link` en la base de datos `retail`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail", "Name":"inventory-
link"}}'
```

INSERT

Permiso	Concedido sobre este recurso	El beneficiario también necesita
INSERT	TABLE	(No se necesitan permisos de IAM adicionales si la ubicación está registrada).

Una entidad principal con este permiso puede insertar, actualizar y leer datos subyacentes en la ubicación de Amazon S3 especificada por la tabla. La entidad principal también puede ver la tabla en la consola de Lake Formation y recuperar información sobre la tabla con la API AWS Glue.

Example

El siguiente ejemplo concede el INSERT permiso al usuario datalake_user1 de la tabla de la base retail de datos de la cuenta inventory 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "INSERT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Este permiso se aplica únicamente a los datos de Amazon S3 y no a los datos de otros almacenes de datos, como Amazon RDS.

SELECT

Permiso	Concedido sobre este recurso	El beneficiario también necesita
SELECT	• TABLE	(No se necesitan permisos de IAM adicionales si la ubicación está registrada).

Una entidad principal con este permiso puede ver una tabla en el Catálogo de datos y consultar los datos subyacentes en Amazon S3 en la ubicación especificada por la tabla. La entidad principal

puede ver la tabla en la consola de Lake Formation y recuperar información sobre la tabla con la API AWS Glue. Si se aplicó el filtrado por columnas cuando se concedió este permiso, la entidad principal puede ver los metadatos solo de las columnas incluidas y consultar los datos solo de las columnas incluidas.

> **Note**
>
> Es responsabilidad del servicio de análisis integrado aplicar el filtrado de columnas al procesar una consulta.

Example

En el siguiente ejemplo, se concede el `SELECT` permiso al usuario `datalake_user1` de la tabla de la base `retail` de datos de la cuenta `inventory` 1111-2222-3333. AWS

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "SELECT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Este permiso se aplica únicamente a los datos de Amazon S3 y no a los datos de otros almacenes de datos, como Amazon RDS.

Puede filtrar (restringir el acceso a) columnas específicas con una lista de inclusión opcional o una lista de exclusión. Una lista de inclusión especifica las columnas a las que se puede acceder. Una lista de exclusión especifica las columnas a las que no se puede acceder. En ausencia de una lista de inclusión o exclusión, todas las columnas de la tabla son accesibles.

Los resultados de `glue:GetTable` devuelven solo las columnas que la persona que llama tiene permiso para ver. Los servicios integrados como Amazon Athena y Amazon Redshift respetan las listas de inclusión y exclusión de columnas.

Example

El ejemplo siguiente concede `SELECT` al usuario `datalake_user1` sobre la tabla `inventory` utilizando una lista de inclusión.

```
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

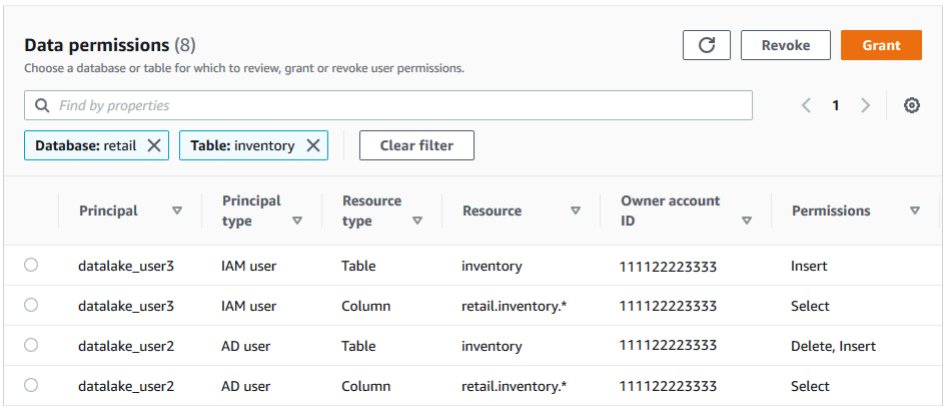
El ejemplo siguiente concede SELECT en la tabla `inventory` utilizando una lista de exclusión.

```
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"]}}}'
```

Se aplican las siguientes restricciones al permiso SELECT:

- Al conceder SELECT, no puede incluir la opción de concesión si se aplica el filtrado por columnas.
- No puede restringir el control de acceso en columnas que son claves de partición.
- A una entidad principal con el permiso SELECT sobre un subconjunto de columnas de una tabla no se le puede conceder el permiso ALTER, DROP, DELETE o INSERT sobre esa tabla. Del mismo modo, a una entidad principal con el permiso ALTER, DROP, DELETE, o INSERT en una tabla no se le puede conceder el permiso SELECT con el filtrado de columnas.

El permiso SELECT siempre aparece en la página Permisos de datos de la consola de Lake Formation como una fila separada. En la imagen siguiente se muestra que SELECT se concede a los usuarios `datalake_user2` y `datalake_user3` en todas las columnas de la tabla `inventory`.



Super

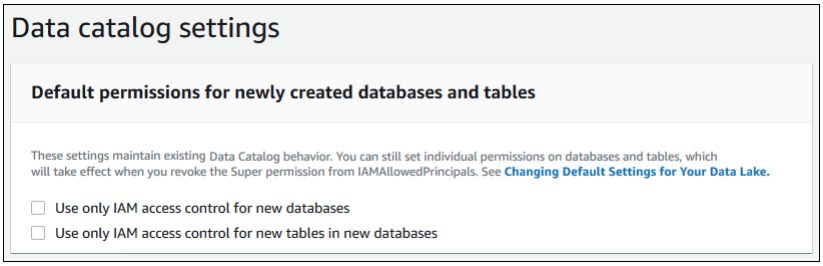
Permiso	Concedido sobre este recurso	El beneficiario también necesita
Super	DATABASE	glue:*Database*
Super	TABLE	glue:*Table*, glue:*Partition*

Este permiso permite a una entidad principal efectuar todas las operaciones compatibles con Lake Formation en la base de datos o en la tabla. No puede conceder Super en una base de datos a una cuenta externa.

Este permiso puede coexistir con los demás permisos de Lake Formation. Por ejemplo, puede conceder los permisos Super, SELECT y INSERT sobre una tabla de metadatos. La entidad principal puede entonces efectuar todas las operaciones admitidas en la tabla. Cuando revoca Super, los permisos SELECT y INSERT permanecen, y la entidad principal solo puede efectuar las operaciones de selección e inserción.

En lugar de conceder Super a una entidad principal individual, puede concederla al grupo IAMAllowedPrincipals. El grupo IAMAllowedPrincipals se crea automáticamente e incluye a todos los usuarios y roles de IAM a los que sus políticas de IAM permiten el acceso a los recursos de su Catálogo de datos. Cuando se concede Super a IAMAllowedPrincipals para un recurso del Catálogo de datos, el acceso al recurso queda efectivamente controlado únicamente por las políticas de IAM.

Puede Super obtener el permiso para que se le conceda automáticamente los nuevos recursos del catálogo si aprovecha las opciones de la página de configuración de la consola de Lake Formation. IAMAllowedPrincipals



- Para conceder `Super` a `IAMAllowedPrincipals` para todas las bases de datos nuevas, seleccione Usar solo control de acceso IAM para bases de datos nuevas.
- Para conceder `Super` a `IAMAllowedPrincipals` para todas las tablas nuevas en bases de datos nuevas, seleccione Usar solo control de acceso IAM para tablas nuevas en bases de datos nuevas.

> **Note**
>
> Esta opción hace que se marque de manera predeterminada la casilla Utilizar sólo el control de acceso IAM para las nuevas tablas de esta base de datos en el cuadro de diálogo Crear base de datos. No hace nada más que eso. Es la casilla de verificación del cuadro de diálogo Crear base de datos que permite la concesión de `Super` a `IAMAllowedPrincipals`.

Estas opciones de la página Configuración están habilitadas de forma predeterminada. Para más información, consulte los siguientes temas:

- the section called “Cambiar la configuración predeterminada de su lago de datos”
- the section called “Actualización de los permisos de datos AWS Glue al modelo de Lake Formation”

ASSOCIATE

Permiso	Concedido sobre este recurso	El beneficiario también necesita

`ASSOCIATE`	`LF-Tag`	`glue:GetDatabase`
		`glue:GetTable`
		`lakeformation:AddLFTagsToResource"`
		`lakeformation:RemoveLFTagsFromResource"`

Permiso	Concedido sobre este recurso	El beneficiario también necesita

		`lakeformation:GetResourceLFTags`
		`lakeformation:ListLFTags`
		`lakeformation:GetLFTag`
		`lakeformation:SearchTablesByLFTags`
		`lakeformation:SearchDatabasesByLFTags`

Una entidad principal con este permiso en una etiqueta LF puede asignar esta a un recurso del Catálogo de datos. Al conceder `ASSOCIATE` está otorgando `DESCRIBE` de manera implícita.

Example

En este ejemplo se concede al usuario `datalake_user1` el permiso `ASSOCIATE` sobre las etiquetas LF con la clave `module`. Concede permisos para ver y asignar todos los valores de esa clave, como indica el asterisco (*).

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Integración de IAM Identity Center

Con él AWS IAM Identity Center, puede conectarse a los proveedores de identidad (IdPs) y administrar de forma centralizada el acceso de los usuarios y grupos a todos los servicios de AWS análisis. Puede integrar proveedores de identidad como Okta, Ping y Microsoft Entra ID (anteriormente Azure Active Directory) con IAM Identity Center para que los usuarios de su

organización accedan a los datos mediante una experiencia de inicio de sesión único. IAM Identity Center también permite conectar otros proveedores de identidad de terceros.

Para obtener más información, consulte la sección Proveedores de identidad compatibles en la Guía del AWS IAM Identity Center usuario.

Puede configurarlo AWS Lake Formation como una aplicación habilitada en el Centro de identidades de IAM, y los administradores del lago de datos pueden conceder permisos detallados a los usuarios y grupos autorizados sobre los recursos. AWS Glue Data Catalog

Los usuarios de la organización pueden iniciar sesión en cualquier aplicación habilitada para Identity Center mediante el proveedor de identidad de su organización y consultar conjuntos de datos aplicando los permisos de Lake Formation. Con esta integración, puede gestionar el acceso a los AWS servicios sin crear varias funciones de IAM.

> **Note**
>
> La propagación de identidades fiable permite a los usuarios y grupos actuales de los usuarios acceder a los datos de todos los servicios de AWS análisis. Con la propagación de identidades fiable, un usuario puede iniciar sesión en una aplicación y la aplicación puede transmitir la identidad del usuario en las solicitudes de acceso a los datos de los AWS servicios. No es necesario realizar ninguna configuración de proveedor de identidad específica del servicio ni de funciones de IAM. Para obtener más información, consulte Propagación confiable de identidades en todas las aplicaciones en la Guía del usuario. AWS IAM Identity Center

Para conocer las limitaciones, consulte Limitaciones de la integración de IAM Identity Center.

Temas

Requisitos previos

A continuación, se indican los requisitos previos para integrar IAM Identity Center con Lake Formation.

1. Habilitar IAM Identity Center: la habilitación de IAM Identity Center es un requisito previo para permitir la autenticación y la propagación de identidades.
2. Elegir su origen de identidades: después de habilitar IAM Identity Center, debe tener un proveedor de identidades para administrar los usuarios y grupos. Puede usar el directorio integrado del Identity Center como origen de identidad o usar un IdP externo, como Microsoft Entra ID u Okta.

   Para obtener más información, consulte Administrar la fuente de identidad y Conectarse a un proveedor de identidad externo en la Guía del AWS IAM Identity Center usuario.
3. Crear un rol de IAM: el rol que crea la conexión a IAM Identity Center requiere permisos para crear y modificar la configuración de la aplicación en Lake Formation e IAM Identity Center, como se indica en la siguiente política en línea.

   Debe añadir permisos según las prácticas recomendadas de IAM. Los permisos específicos se detallan en los siguientes procedimientos. Para obtener más información, consulte Primeros pasos con IAM Identity Center.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:CreateLakeFormationIdentityCenterConfiguration",
                "sso:CreateApplication",
                "sso:PutApplicationAssignmentConfiguration",
                "sso:PutApplicationAuthenticationMethod",
                "sso:PutApplicationGrant",
                "sso:PutApplicationAccessScope"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

Las siguientes políticas en línea contienen los permisos específicos necesarios para ver, actualizar y eliminar las propiedades de la integración de Lake Formation con IAM Identity Center.

- Utilice la siguiente política en línea para permitir que un rol de IAM vea una integración de Lake Formation con IAM Identity Center.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- Utilice la siguiente política en línea para permitir que un rol de IAM actualice una integración de Lake Formation con IAM Identity Center.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:UpdateLakeFormationIdentityCenterConfiguration",
                "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication",
                "sso:UpdateApplication"
            ],
            "Resource": [
                "*"
            ]
```

```
        }
    ]
}
```

- Utilice la siguiente política en línea para permitir que un rol de IAM elimine una integración de Lake Formation con IAM Identity Center.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:DeleteLakeFormationIdentityCenterConfiguration",
                "sso:DeleteApplication"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- Para conocer los permisos de IAM necesarios para conceder o revocar los permisos de lago de datos para los usuarios y grupos de IAM Identity Center, consulte Permisos de IAM necesarios para conceder o revocar permisos de Lake Formation.

Descripción de los permisos

- `lakeformation:CreateLakeFormationIdentityCenterConfiguration` – Crea la configuración IdC de Lake Formation.
- `lakeformation:DescribeLakeFormationIdentityCenterConfiguration` – Describe una configuración de IdC existente.
- `lakeformation:DeleteLakeFormationIdentityCenterConfiguration` – Ofrece la posibilidad de eliminar una configuración de IdC de Lake Formation existente.
- `lakeformation:UpdateLakeFormationIdentityCenterConfiguration` – Se usa para cambiar una configuración existente de Lake Formation.
- `sso:CreateApplication` – Se utiliza para crear una aplicación de IAM Identity Center.

- `sso:DeleteApplication`: se utiliza para eliminar una aplicación IAM Identity Center.
- `sso:UpdateApplication`: se utiliza para actualizar una aplicación IAM Identity Center.
- `sso:PutApplicationGrant` – Se utiliza para cambiar la información del emisor de tokens de confianza.
- `sso:PutApplicationAuthenticationMethod` – Otorga acceso de autenticación a Lake Formation.
- `sso:GetApplicationGrant` – Se utiliza para enumerar la información del emisor de tokens de confianza.
- `sso:DeleteApplicationGrant` – Elimina la información del emisor de tokens de confianza.
- `sso:PutApplicationAccessScope` – Añade o actualiza la lista de objetivos autorizados para el ámbito de acceso de una aplicación a IAM Identity Center.
- `sso:PutApplicationAssignmentConfiguration` – Se utiliza para configurar cómo acceden los usuarios a una aplicación.

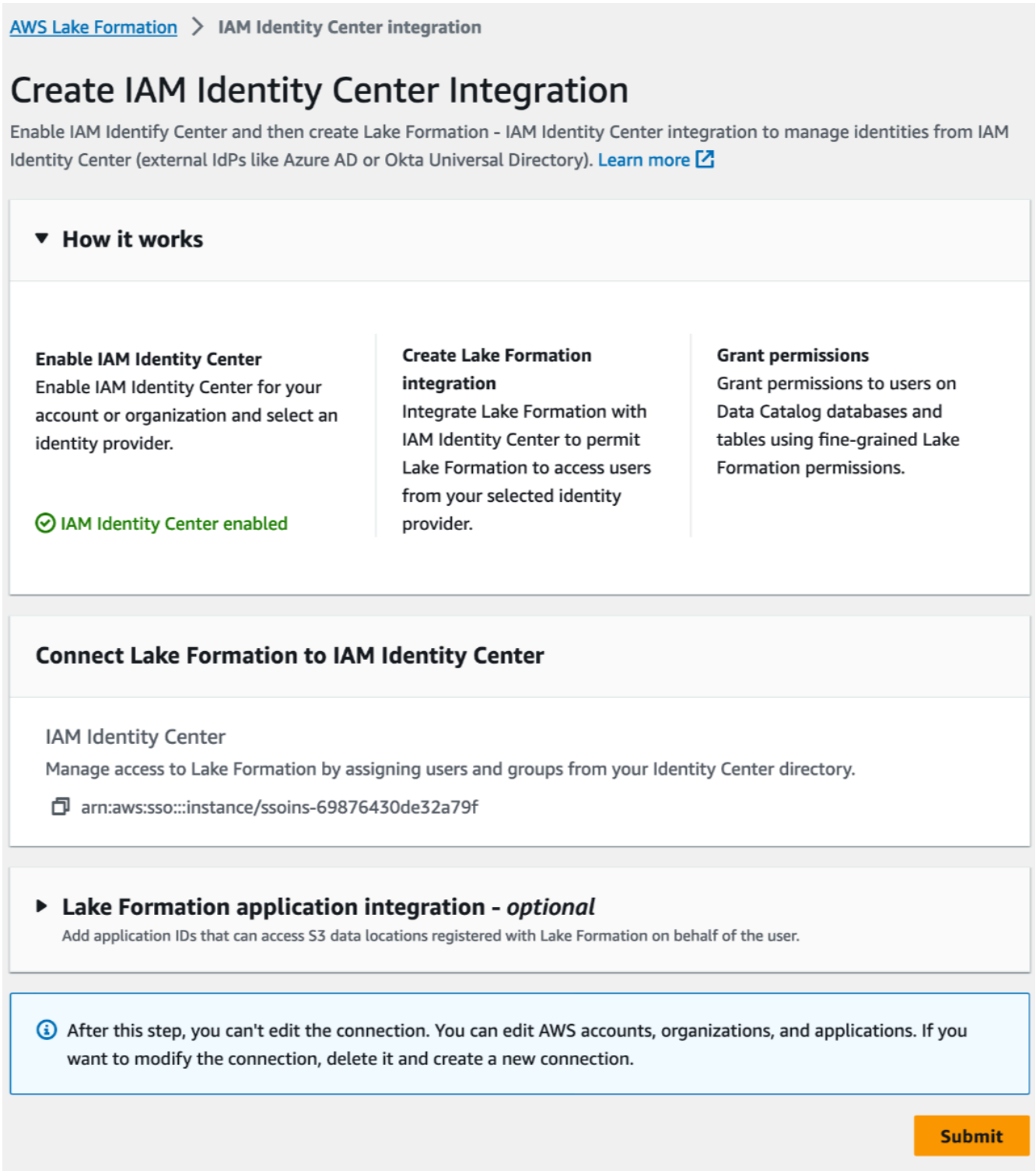
Conexión de Lake Formation con IAM Identity Center

Antes de poder usar IAM Identity Center para administrar identidades y conceder acceso a los recursos del catálogo de datos mediante Lake Formation, debe realizar los siguientes pasos. Puede crear la integración de IAM Identity Center mediante la consola o AWS CLI de Lake Formation.

AWS Management Console

Para conectar Lake Formation con IAM Identity Center

1. Inicie sesión en la AWS Management Console consola de Lake Formation y ábrala en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación izquierdo, seleccione Integración de IAM Identity Center.



3. (Opcional) En la pantalla Crear integración de Lake Formation, especifique los ARN de las aplicaciones de terceros que pueden acceder a los datos de las ubicaciones de Amazon S3 registradas en Lake Formation. Lake Formation vende credenciales temporales limitadas en forma de AWS STS tokens a las ubicaciones registradas de Amazon S3 en función de los

permisos vigentes, de modo que las aplicaciones autorizadas puedan acceder a los datos en nombre de los usuarios.

4. Seleccione Enviar.

   Una vez que el administrador de Lake Formation finalice los pasos y cree la integración, las propiedades de IAM Identity Center aparecen en la consola de Lake Formation. Al realizar estas tareas, Lake Formation se convierte en una aplicación habilitada para IAM Identity Center. Las propiedades de la consola incluyen el estado de la integración. Cuando se completa, el estado de la integración muestra `Success`. Este estado indica si la configuración de IAM Identity Center se ha completado.

AWS CLI

- En el siguiente ejemplo se muestra cómo crear la integración de Lake Formation con IAM Identity Center. También puede especificar el `Status` (`ENABLED`, `DISABLED`) de las aplicaciones.

```
aws lakeformation create-lake-formation-identity-center-configuration \
    --catalog-id <123456789012> \
    --instance-arn <arn:aws:sso:::instance/ssoins-112111f12ca1122p> \
    --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"],
 "Status": "ENABLED"}'
```

- En el siguiente ejemplo se muestra cómo ver una integración de Lake Formation con IAM Identity Center.

```
aws lakeformation describe-lake-formation-identity-center-configuration
 --catalog-id <123456789012>
```

Actualización de una integración de IAM Identity Center

Tras crear la conexión, puede añadir aplicaciones de terceros para que la integración de IAM Identity Center se integre con Lake Formation y obtener acceso a los datos de Amazon S3 en nombre de los usuarios. También puede eliminar las aplicaciones existentes de la integración de IAM Identity Center. Puede agregar o eliminar aplicaciones mediante la consola Lake Formation y mediante la UpdateLakeFormationIdentityCenterConfigurationoperación. AWS CLI

> **Note**
>
> Tras crear la integración de IAM Identity Center, no podrá actualizar el ARN de la instancia.

AWS Management Console

Para actualizar una conexión existente de IAM Identity Center con Lake Formation

1. Inicie sesión en la AWS Management Console consola de Lake Formation y ábrala en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación izquierdo, seleccione Integración de IAM Identity Center.
3. Seleccione Añadir en la página Integración de IAM Identity Center.
4. En la pantalla Añadir aplicaciones, introduzca los ID de las aplicaciones de terceros que desee integrar con Lake Formation.
5. Seleccione Añadir.

AWS CLI

Puede añadir o eliminar aplicaciones de terceros para la integración del IAM Identity Center ejecutando el siguiente AWS CLI comando. Cuando se establece el estado de filtrado externo en ENABLED, IAM Identity Center puede proporcionar administración de identidades para que las aplicaciones de terceros accedan a los datos administrados por Lake Formation. También puede activar o desactivar la integración de IAM Identity Center configurando el estado de la aplicación.

```
aws lakeformation update-lake-formation-identity-center-configuration \
 --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"], "Status":
 "ENABLED"}'\
 --application-status ENABLED
```

Eliminación de una conexión de Lake Formation con IAM Identity Center

Si desea eliminar una integración existente de IAM Identity Center, puede hacerlo mediante la consola o la DeleteLakeFormationIdentityCenterConfigurationoperación de Lake Formation. AWS CLI

AWS Management Console

Para eliminar una conexión existente de IAM Identity Center con Lake Formation

1. Inicie sesión en la AWS Management Console consola de Lake Formation y ábrala en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación izquierdo, seleccione Integración de IAM Identity Center.
3. Seleccione Eliminar en la página Integración de IAM Identity Center.
4. En la pantalla Confirmar integración, confirme la acción y seleccione Eliminar.

AWS CLI

Puede eliminar la integración de IAM Identity Center ejecutando el siguiente AWS CLI comando.

```
aws lakeformation delete-lake-formation-identity-center-configuration \
    --catalog-id <123456789012>
```

Concesión de permisos a usuarios y grupos

El administrador del lago de datos puede conceder permisos a los usuarios y grupos de IAM Identity Center sobre los recursos del catálogo de datos (bases de datos, tablas y vistas) para permitir un fácil acceso a los datos. Para conceder o revocar los permisos del lago de datos, el otorgante necesita permisos para las siguientes acciones de IAM Identity Center.

- DescribeUser
- DescribeGroup
- DescribeInstance

Puede conceder permisos utilizando la consola de Lake Formation, la API o la AWS CLI.

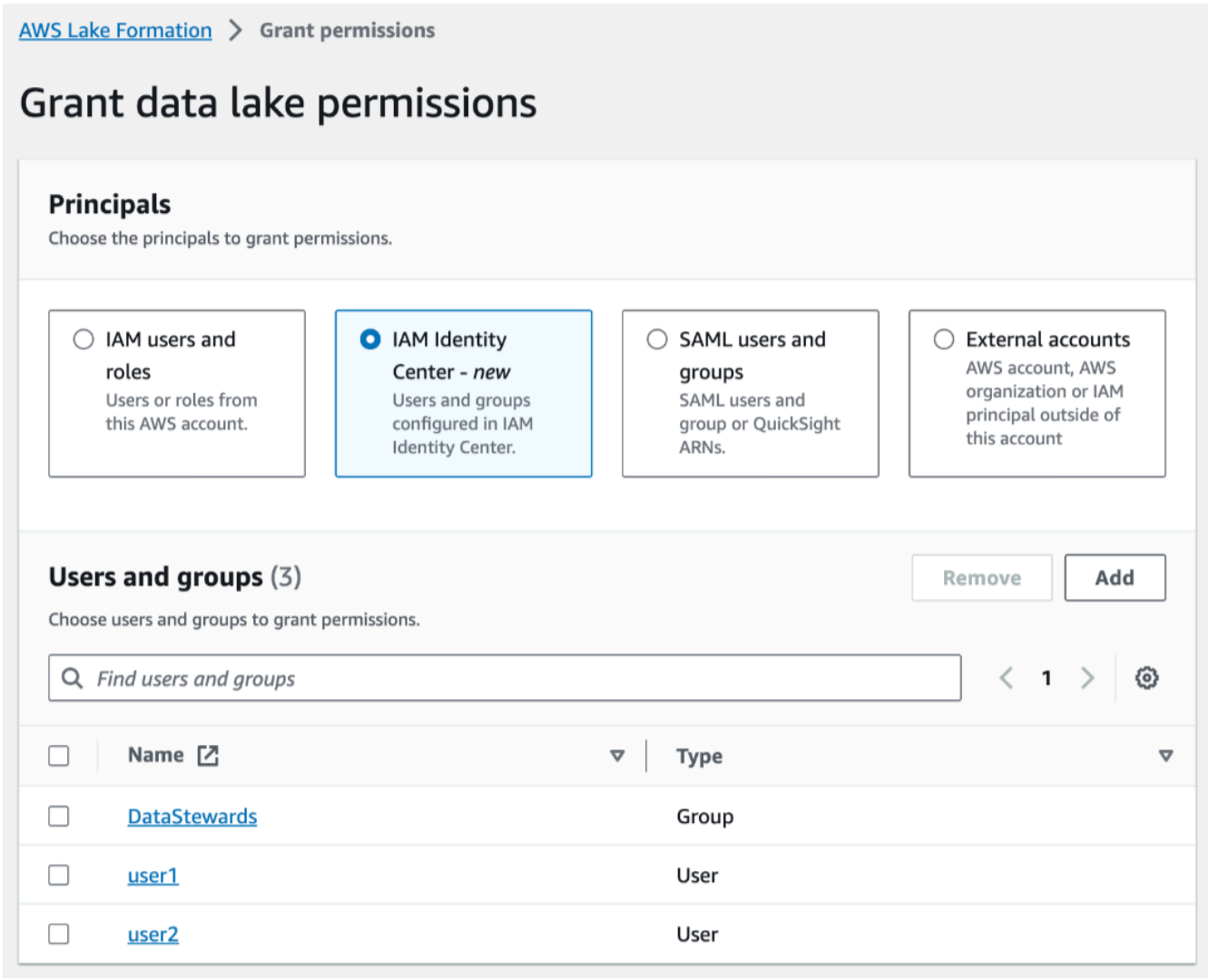
Para obtener más información sobre cómo conceder permisos, consulte the section called "Concesión y revocación de permisos del catálogo de datos".

> **Note**
>
> Solo puede conceder permisos a los recursos de su cuenta. Para transferir los permisos en cascada a los usuarios y grupos sobre los recursos que comparte con usted, debe utilizar AWS RAM los recursos compartidos.

AWS Management Console

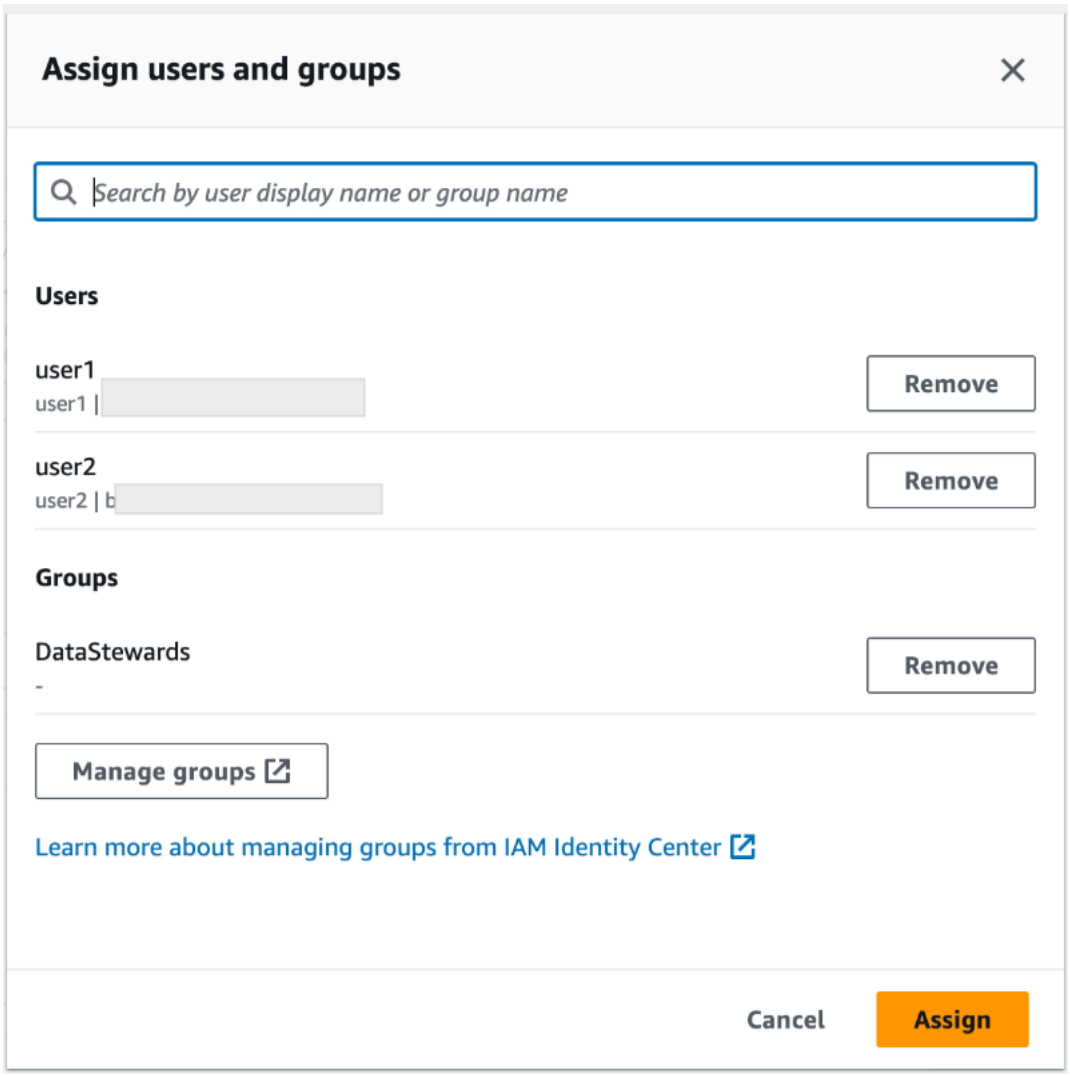
Para conceder permisos a usuarios y grupos

1. Inicie sesión en la AWS Management Console consola de Lake Formation y ábrala en https://console.aws.amazon.com/lakeformation/.
2. Seleccione Permisos de lago de datos bajo Permisos in en la consola de Lake Formation.
3. Seleccione Conceder.
4. En la página Conceder permisos de lago de datos, seleccione usuarios y grupos de SSM.
5. Seleccione Añadir para elegir los usuarios y grupos a los que va a conceder permisos.



6. En la pantalla Asignar usuarios y grupos, elija los usuarios o grupos a los que va a conceder los permisos.

   Seleccione Asignar.



7. A continuación, elija el método para conceder los permisos.

   Para obtener instrucciones sobre cómo conceder permisos mediante el método de recursos con nombre, consulte Concesión de permisos de lago de datos mediante el método de recurso con nombre.

   Para obtener instrucciones sobre cómo conceder permisos mediante etiquetas LF, consulte Conceder permisos de lago de datos mediante el método LF-TBAC.

8. Elija los recursos del catálogo de datos de los que desea conceder permisos.
9. Elija los permisos del catálogo de datos que desee conceder.
10. Seleccione Conceder.

AWS CLI

En el siguiente ejemplo, se muestra cómo conceder el permiso `SELECT` de usuario de IAM Identity Center en una tabla.

```
aws lakeformation grant-permissions \
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserId> \
--permissions "SELECT" \
--resource '{ "Table": { "DatabaseName": "retail", "TableWildcard": {} } }'
```

Para recuperarlo `UserId` desde el Centro de Identidad de IAM, consulte el GetUserIdfuncionamiento en la Referencia de la API del Centro de Identidad de IAM.

Añadir una ubicación de Amazon S3 a su lago de datos

Para añadir una ubicación de Amazon Simple Storage Service (Amazon S3) como almacenamiento en su lago de datos, debe registrar la ubicación con. AWS Lake Formation Luego, puede usar los permisos de Lake Formation para un control de acceso detallado a AWS Glue Data Catalog los objetos que apuntan a esta ubicación y a los datos subyacentes de la ubicación.

Lake Formation también permite registrar una ubicación de datos en modo de acceso híbrido y le proporciona la flexibilidad de habilitar selectivamente los permisos de Lake Formation para bases de datos y tablas en su Catálogo de datos. Con el modo de acceso híbrido, ahora tiene una ruta incremental que le permite establecer los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes.

Para más información sobre la configuración del modo de acceso híbrido, consulte Modo de acceso híbrido

Al registrar una ubicación, se registran esa ruta de Amazon S3 y todas las carpetas incluidas en esa ruta.

Por ejemplo, supongamos que tiene una organización de rutas de Amazon S3 como la siguiente:

`/mybucket/accounting/sales/`

Si registra `S3://mybucket/accounting`, la carpeta `sales` también estará registrada y bajo la administración de Lake Formation.

Para obtener más información sobre cómo registrar ubicaciones, consulte Underlying data access control.

> **Note**
>
> Se recomiendan los permisos de Lake Formation para los datos estructurados (organizados en tablas con filas y columnas). Si sus datos contienen datos no estructurados basados en objetos, plantéese la posibilidad de utilizar el permiso de IAM para que Amazon S3 administre el acceso a los datos.

Temas

Requisitos de los roles utilizados para registrar ubicaciones

Debe especificar un rol AWS Identity and Access Management (IAM) al registrar una ubicación de Amazon Simple Storage Service (Amazon S3). AWS Lake Formation asume esa función al acceder a los datos en esa ubicación.

Para registrar una ubicación, puede utilizar uno de los siguientes tipos de rol:

- El rol vinculado al servicio de Lake Formation. Este rol concede los permisos necesarios sobre la ubicación. Utilizar este rol es la forma más sencilla de registrar la ubicación. Para obtener más información, consulte Uso de roles vinculados a servicios para Lake Formation.
- Un rol definido por el usuario. Utilice un rol definido por el usuario cuando necesite conceder más permisos de los que proporciona el rol vinculado al servicio.

  Debe utilizar un rol definido por el usuario en las circunstancias siguientes:

  - Al registrar una ubicación en otra cuenta.

Para obtener más información, consulte the section called “Registrar una ubicación de Amazon S3 en otra cuenta AWS” y the section called “Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS”.

- Si utilizó una CMK (`aws/s3`) AWS administrada para cifrar la ubicación de Amazon S3.

  Para obtener más información, consulte Registro de una ubicación cifrada de Amazon S3.

- Si tiene previsto acceder a la ubicación mediante Amazon EMR.

  Si ya ha registrado una ubicación con el rol vinculado al servicio y desea comenzar a acceder a la ubicación con Amazon EMR, deberá anular el registro de la ubicación y volver a registrarla con un rol definido por el usuario. Para obtener más información, consulte the section called “Dar de baja el registro de una ubicación de Amazon S3”.

Uso de roles vinculados a servicios para Lake Formation

AWS Lake Formation utiliza un rol vinculado al servicio de AWS Identity and Access Management (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que está vinculado directamente a Lake Formation. El rol vinculado al servicio está predefinido por Lake Formation e incluye todos los permisos que el servicio requiere para llamar a otros servicios de AWS en su nombre.

Un rol vinculado a un servicio simplifica la configuración de Lake Formation porque ya no tendrá que agregar manualmente los permisos necesarios. Lake Formation define los permisos de su rol vinculado al servicio y, a menos que se defina lo contrario, solo Lake Formation puede asumir sus roles. Los permisos definidos incluyen las políticas de confianza y de permisos, y que la política de permisos no se puede adjuntar a ninguna otra entidad de IAM.

Este rol vinculado a servicios confía en los siguientes servicios para asumir el rol:

- `lakeformation.amazonaws.com`

Permisos de rol vinculados al servicio para Lake Formation

Lake Formation utiliza el rol vinculado al servicio denominado `AWSServiceRoleForLakeFormationDataAccess`. Este rol proporciona un conjunto de permisos de Amazon Simple Storage Service (Amazon S3) que permiten al servicio integrado de Lake Formation (como Amazon Athena) acceder a las ubicaciones registradas. Al registrar una ubicación de lago de datos, debe proporcionar un rol que tenga los permisos de lectura/escritura de Amazon

S3 necesarios en esa ubicación. En lugar de crear un rol con los permisos de Amazon S3 que se requieren, puede utilizar este rol vinculado con un servicio.

La primera vez que nombre el rol vinculado al servicio como rol con el que registrar una ruta, el rol vinculado al servicio y una nueva política de IAM se crearán en su nombre. Lake Formation añade la ruta a la política insertada y la adjunta al rol vinculado al servicio. Cuando registre rutas posteriores con el rol vinculado al servicio, Lake Formation añade la ruta a la política existente.

Inicie sesión como administrador del lago de datos y registre una ubicación del lago de datos. A continuación, en la consola de IAM, busque el rol `AWSServiceRoleForLakeFormationDataAccess` y vea sus políticas adjuntas.

Por ejemplo, después de registrar la ubicación `s3://my-kinesis-test/logs`, Lake Formation crea la siguiente política en línea y la adjunta a `AWSServiceRoleForLakeFormationDataAccess`.

```
{
    "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"
```

```
            ]
        }
    ]
}
```

Se requieren los permisos siguientes para poder registrar ubicaciones con este rol vinculado al servicio:

- `iam:CreateServiceLinkedRole`
- `iam:PutRolePolicy`

El administrador del lago de datos suele tener estos permisos.

Los siguientes son los requisitos para un rol definido por el usuario:

- Al crear el nuevo rol, en la página Crear rol de la consola de IAM, elija el Servicio de AWS y, a continuación, en Elija un caso de uso, seleccione Lake Formation.

  Si crea el rol utilizando una ruta diferente, asegúrese de que el rol tenga una relación de confianza con `lakeformation.amazonaws.com`. Para más información, consulte Modificación de una política de confianza de rol (consola).
- El rol debe tener relaciones de confianza con las siguientes entidades:
  - `glue.amazonaws.com`
  - `lakeformation.amazonaws.com`

  Para más información, consulte Modificación de una política de confianza de rol (consola).
- El rol debe tener una política en línea que conceda permisos de lectura/escritura de Amazon S3 en la ubicación. La siguiente es una política característica.

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

- El administrador del lago de datos que registra la ubicación debe tener el permiso `iam:PassRole` para el rol.

  La siguiente es una política insertada que concede este permiso. <account-id>Sustitúyalo por un número de AWS cuenta válido y <role-name>sustitúyalo por el nombre del rol.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "PassRolePermissions",
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::<account-id>:role/<role-name>"
            ]
        }
    ]
}
```

- Para permitir que Lake Formation añada CloudWatch registros en Logs y publique métricas, añada la siguiente política en línea.

> **Note**
>
> Escribir en CloudWatch Logs conlleva un cargo.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Sid1",
            "Effect": "Allow",
            "Action": [
                "logs:CreateLogStream",
                "logs:CreateLogGroup",
                "logs:PutLogEvents"
            ],
            "Resource": [
                 "arn:aws:logs:<region>:<account-id>:log-group:/aws-lakeformation-
acceleration/*",
                 "arn:aws:logs:<region>:<account-id>:log-group:/aws-lakeformation-
acceleration/*:log-stream:*"
            ]
        }
    ]
}
```

Registro de una ubicación de Amazon S3

Debe especificar un rol AWS Identity and Access Management (IAM) al registrar una ubicación de Amazon Simple Storage Service (Amazon S3). Lake Formation asume esa función cuando otorga credenciales temporales a los AWS servicios integrados que acceden a los datos en esa ubicación.

> **Important**
>
> Evite registrar un bucket de Amazon S3 que tenga activada la opción El solicitante paga. Para los buckets registrados en Lake Formation, el rol utilizado para registrar el bucket se considera siempre como el solicitante. Si otra AWS cuenta accede al depósito, se le cobrará

al propietario del depósito por el acceso a los datos si el rol pertenece a la misma cuenta que el propietario del depósito.

Puede usar la AWS Lake Formation consola, la API de Lake Formation o AWS Command Line Interface (AWS CLI) para registrar una ubicación de Amazon S3.

Antes de empezar

Revise los requisitos del rol utilizado para registrar la ubicación.

Para registrar una ubicación (consola)

Important

En los siguientes procedimientos se supone que la ubicación de Amazon S3 se encuentra en la misma AWS cuenta que el catálogo de datos y que los datos de la ubicación no están cifrados. Otras secciones de este capítulo tratan sobre el registro entre cuentas y el registro de ubicaciones cifradas.

1. Abra la AWS Lake Formation consola en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con el permiso de `lakeformation:RegisterResource` IAM.
2. En el panel de navegación, vaya a Registrar e ingerir y seleccione Ubicaciones de los lagos de datos.
3. Elija Registrar ubicación y, a continuación, seleccione Examinar para seleccionar una ruta de Amazon Simple Storage Service (Amazon S3).
4. (Opcional, pero muy recomendable) Seleccione Revisar permisos de ubicación para ver una lista de todos los recursos existentes en la ubicación de Amazon S3 seleccionada y sus permisos.

   El registro de la ubicación seleccionada podría dar lugar a que sus usuarios de Lake Formation accedan a los datos que ya se encuentran en esa ubicación. Revisar esta lista ayuda a garantizar que los datos existentes permanecen seguros.
5. Para el rol de IAM, elija el rol vinculado al servicio `AWSServiceRoleForLakeFormationDataAccess` (valor predeterminado) o un rol de IAM personalizado que cumpla los requisitos de the section called “Requisitos de los roles utilizados para registrar ubicaciones”.

Puede actualizar una ubicación registrada u otros detalles solo si la registra con un rol de IAM personalizado. Para editar una ubicación registrada con un rol vinculado a un servicio, debe anular el registro de la ubicación y volver a registrarla.

6. Elija la opción Habilitar la federación de catálogos de datos para permitir que Lake Formation asuma un rol y venda credenciales temporales a AWS servicios integrados para acceder a las tablas de bases de datos federadas. Si una ubicación está registrada en Lake Formation y desea utilizar la misma ubicación para una tabla en una base de datos federada, deberá registrar la misma ubicación con la opción Habilitar federación del Catálogo de datos.
7. Seleccione el modo de acceso híbrido para no habilitar los permisos de Lake Formation de forma predeterminada. Cuando registre la ubicación de Amazon S3 en modo de acceso híbrido, puede habilitar los permisos de Lake Formation optando por entidades principales para las bases de datos y las tablas bajo esa ubicación.

   Para más información sobre la configuración del modo de acceso híbrido, consulte Modo de acceso híbrido.
8. Seleccione Registrar ubicación.

Para registrar una ubicación (AWS CLI)

1. Registro de una nueva ubicación en Lake Formation

   Este ejemplo usa el rol vinculado a un servicio para registrar la ubicación. En su lugar, puede utilizar el argumento `--role-arn` para proporcionar su propio rol.

   <s3-path>Sustitúyalo por una ruta de Amazon S3 válida, el número de AWS cuenta por una cuenta válida y por <s3-access-role>un rol de IAM que tenga permisos para registrar una ubicación de datos.

   > **Note**
   >
   > No puede editar las propiedades de una ubicación registrada si está registrada con un rol vinculado a un servicio.

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
```

```
 --use-service-linked-role
```

En el siguiente ejemplo, se utiliza un rol personalizado para registrar la ubicación.

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>
```

2. Para actualizar una ubicación registrada con Lake Formation

   Puede editar una ubicación registrada solo si está registrada con un rol de IAM personalizado. en el caso de una ubicación registrada con un rol vinculado a un servicio, debe anular el registro de la ubicación y volver a registrarla. Para obtener más información, consulte the section called “Dar de baja el registro de una ubicación de 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. Registrar una ubicación de datos en modo de acceso híbrido con federación

```
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
```

Para obtener más información, consulte Funcionamiento de RegisterResourcela API.

> **Note**
>
> Una vez que registre una ubicación de Amazon S3, cualquier AWS Glue tabla que apunte a la ubicación (o a cualquiera de sus ubicaciones secundarias) devolverá el valor del `IsRegisteredWithLakeFormation` parámetro tal y como `true` aparece en la `GetTable` llamada. Existe una limitación conocida por la que las operaciones de la API del Catálogo de datos como `GetTables` y `SearchTables` no actualizan el valor del parámetro `IsRegisteredWithLakeFormation` y devuelven el valor predeterminado, que es falso. Se recomienda utilizar la API`GetTable` para ver el valor correcto del parámetro `IsRegisteredWithLakeFormation`.

Registro de una ubicación cifrada de Amazon S3

Lake Formation se integra con AWS Key Management Service (AWS KMS) para permitirle configurar más fácilmente otros servicios integrados para cifrar y descifrar datos en ubicaciones de Amazon Simple Storage Service (Amazon S3).

Ambos son gestionados por el cliente AWS KMS keys y Claves administradas por AWS cuentan con soporte. Actualmente, el cifrado/descifrado del lado del cliente solo es compatible con Athena.

Debe especificar un rol AWS Identity and Access Management (de IAM) al registrar una ubicación de Amazon S3. En el caso de las ubicaciones cifradas de Amazon S3, el rol debe tener permiso para cifrar y descifrar datos con el AWS KMS key, o bien la política de claves de KMS debe conceder permisos sobre la clave del rol.

> **Important**
>
> Evite registrar un bucket de Amazon S3 que tenga activada la opción El solicitante paga. Para los buckets registrados en Lake Formation, el rol utilizado para registrar el bucket se considera siempre como el solicitante. Si otra AWS cuenta accede al bucket, se le cobrará al propietario del bucket por el acceso a los datos si el rol pertenece a la misma cuenta que el propietario del bucket.

La forma más sencilla de registrar la localización es utilizar el rol vinculado al servicio Lake Formation. Este rol concede los permisos de lectura y escritura necesarios sobre la ubicación. También puede usar un rol personalizado para registrar la ubicación, siempre que cumpla con los requisitos de the section called “Requisitos de los roles utilizados para registrar ubicaciones”.

> **Important**
>
> Si ha utilizado un Clave administrada de AWS (`aws/s3`) para cifrar la ubicación de Amazon S3, no puede utilizar la función vinculada al servicio Lake Formation. Debe usar un rol personalizado y añadir permisos de IAM a la clave del rol. Los detalles se proporcionan más adelante en esta sección.

Los siguientes procedimientos explican cómo registrar una ubicación de Amazon S3 cifrada con una clave administrada por el cliente o una Clave administrada de AWS.

- Registrar una ubicación cifrada con una clave administrada por el cliente
- Registrar una ubicación cifrada con un Clave administrada de AWS

Antes de empezar

Revise los requisitos del rol utilizado para registrar la ubicación.

Para registrar una ubicación de Amazon S3 cifrada con una clave administrada por el cliente

> **Note**
>
> Si la clave de KMS o la ubicación de Amazon S3 no están en la misma AWS cuenta que el catálogo de datos, siga las instrucciones que se indican en the section called “Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS” su lugar.

1. Abra la AWS KMS consola en https://console.aws.amazon.com/kms e inicie sesión como usuario administrativo AWS Identity and Access Management (IAM) o como usuario que puede modificar la política de claves de la clave de KMS utilizada para cifrar la ubicación.
2. En el panel de navegación, elija Claves administradas por el cliente y, a continuación, el nombre de la clave de KMS deseada.

3. En la página de detalles de la clave KMS, elija la pestaña Política de claves y, a continuación, siga una de las instrucciones siguientes para añadir su rol personalizado o el rol vinculado al servicio de Lake Formation como usuario de la clave KMS:

   - Si se muestra la vista predeterminada (con las secciones Administradores clave, Eliminación de claves, Usuarios clave y Otras AWS cuentas), en la sección Usuarios clave, agregue su rol personalizado o el rol vinculado al servicio Lake Formation. `AWSServiceRoleForLakeFormationDataAccess`
   - Si se muestra la política de claves (JSON). Edite la política para añadir su rol personalizado o el rol `AWSServiceRoleForLakeFormationDataAccess` vinculado al servicio de Lake Formation al objeto "Permitir el uso de la clave", como se muestra en el siguiente ejemplo.

   > **Note**
   >
   > Si falta ese objeto, agréguelo con los permisos que se muestran en el ejemplo. El ejemplo utiliza el rol vinculado al servicio.

```
        ...
        {
            "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. Abra la AWS Lake Formation consola en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con el permiso de `lakeformation:RegisterResource` IAM.
5. En el panel de navegación, vaya a Registrar e ingerir y seleccione Ubicaciones del lago de datos.
6. Elija Registrar ubicación y, a continuación, seleccione Examinar para seleccionar una ruta de Amazon Simple Storage Service (Amazon S3).
7. (Opcional, pero muy recomendable) Seleccione Revisar permisos de ubicación para ver una lista de todos los recursos existentes en la ubicación de Amazon S3 seleccionada y sus permisos.

   El registro de la ubicación seleccionada podría dar lugar a que sus usuarios de Lake Formation accedan a los datos que ya se encuentran en esa ubicación. Revisar esta lista ayuda a garantizar que los datos existentes permanecen seguros.
8. Para el rol de IAM, elija el rol vinculado al servicio `AWSServiceRoleForLakeFormationDataAccess` (el predeterminado) o su rol personalizado que cumpla con el the section called “Requisitos de los roles utilizados para registrar ubicaciones”.
9. Seleccione Registrar ubicación.

Para obtener más información sobre el rol vinculado a servicios, consulte Permisos de rol vinculados al servicio para Lake Formation.

Para registrar una ubicación de Amazon S3 cifrada con un Clave administrada de AWS

> **Important**
>
> Si la ubicación de Amazon S3 no está en la misma AWS cuenta que el catálogo de datos, siga las instrucciones que se indican en the section called “Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS” su lugar.

1. Cree un rol de IAM que se utilizará para registrar la ubicación. Asegúrese de que cumple los requisitos que figuran en the section called “Requisitos de los roles utilizados para registrar ubicaciones”.
2. Añada la siguiente política insertada al rol. Concede permisos sobre la clave al rol. La especificación `Resource` debe designar el nombre de recurso de Amazon (ARN) del Clave

administrada de AWS. Puede obtener el ARN desde la AWS KMS consola. Para obtener el ARN correcto, asegúrese de iniciar sesión en la AWS KMS consola con la misma AWS cuenta y región Clave administrada de AWS que utilizó para cifrar la ubicación.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "<Clave administrada de AWS ARN>"
    }
  ]
}
```

3. Abra la AWS Lake Formation consola en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con el permiso de `lakeformation:RegisterResource` IAM.
4. En el panel de navegación, vaya a Registrar e ingerir y seleccione Ubicaciones del lago de datos.
5. Elija Registrar ubicación y, a continuación, seleccione Examinar para seleccionar una ruta de Amazon S3.
6. (Opcional, pero muy recomendable) Seleccione Revisar permisos de ubicación para ver una lista de todos los recursos existentes en la ubicación de Amazon S3 seleccionada y sus permisos.

   El registro de la ubicación seleccionada podría dar lugar a que sus usuarios de Lake Formation accedan a los datos que ya se encuentran en esa ubicación. Revisar esta lista ayuda a garantizar que los datos existentes permanecen seguros.
7. Para Rol de IAM, elija el rol que ha creado en el Paso 1.
8. Seleccione Registrar ubicación.

Registrar una ubicación de Amazon S3 en otra cuenta AWS

AWS Lake Formation le permite registrar las ubicaciones AWS de Amazon Simple Storage Service (Amazon S3) en todas las cuentas. Por ejemplo, si AWS Glue Data Catalog está en la cuenta A, un usuario de la cuenta A puede registrar un bucket de Amazon S3 en la cuenta B.

Para registrar un bucket de Amazon S3 en la AWS cuenta B con un rol AWS Identity and Access Management (IAM) en la AWS cuenta A se requieren los siguientes permisos:

- El rol de la cuenta A debe conceder permisos sobre el bucket de la cuenta B.
- La política de bucket de la cuenta B debe conceder permisos de acceso al rol de la cuenta A.

> **Important**
>
> Evite registrar un bucket de Amazon S3 que tenga activada la opción El solicitante paga. Para los buckets registrados en Lake Formation, el rol utilizado para registrar el bucket se considera siempre como el solicitante. Si otra AWS cuenta accede al bucket, se le cobrará al propietario del bucket por el acceso a los datos si el rol pertenece a la misma cuenta que el propietario del bucket.
>
> No puede utilizar el rol vinculado al servicio Lake Formation para registrar una ubicación en otra cuenta. En su lugar, debe utilizar un rol definido por el usuario. El rol debe cumplir los requisitos de the section called “Requisitos de los roles utilizados para registrar ubicaciones”. Para obtener más información sobre el rol vinculado a servicios, consulte Permisos de rol vinculados al servicio para Lake Formation.

Antes de empezar

Revise los requisitos del rol utilizado para registrar la ubicación.

Para registrar una ubicación en otra AWS cuenta

> **Note**
>
> Si la ubicación está cifrada, siga en su lugar las instrucciones de the section called “Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS”.

El siguiente procedimiento supone que una entidad principal de la cuenta 1111-2222-3333, que contiene el Catálogo de datos, desea registrar el bucket `awsexamplebucket1` de Amazon S3, que se encuentra en la cuenta 1234-5678-9012.

1. En la cuenta 1111-2222-3333, inicie sesión en AWS Management Console y abra la consola de IAM en. https://console.aws.amazon.com/iam/
2. Cree un nuevo rol o consulte uno existente que cumpla con los requisitos de the section called "Requisitos de los roles utilizados para registrar ubicaciones". Asegúrese de que el rol concede permisos de Amazon S3 `aawsexamplebucket1`.
3. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3. Inicie sesión con la cuenta 1234-5678-9012.
4. En la lista Nombre del bucket, seleccione el nombre del bucket, `awsexamplebucket1`.
5. Elija Permisos.
6. En la página Permisos, elija Política de bucket.
7. En el editor de políticas de bucket, pegue la política siguiente. Sustituya *`<nombre del rol>`* por el nombre de su rol.

```
{
    "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. Seleccione Guardar.
9. Abra la consola en https://console.aws.amazon.com/lakeformation/. AWS Lake Formation Inicie sesión en la cuenta 1111-2222-3333 como administrador del lago de datos o como usuario con permisos suficientes para registrar ubicaciones.
10. En el panel de navegación, bajo Administración, seleccione Ubicaciones de los lagos de datos.
11. En la página de Ubicaciones de los lagos de datos, seleccione Registrar ubicación.
12. En la página Registrar ubicación, para la ruta de Amazon S3, introduzca el nombre del bucket `s3://awsexamplebucket1`.

> **Note**
>
> Debe escribir el nombre del bucket porque los buckets entre cuentas no aparecen en la lista cuando selecciona Examinar.

13. En Rol de IAM, seleccione su rol.
14. Seleccione Registrar ubicación.

Registro de una ubicación cifrada de Amazon S3 entre cuentas AWS

AWS Lake Formation se integra con AWS Key Management Service(AWS KMS) para permitirle configurar más fácilmente otros servicios integrados para cifrar y descifrar datos en las ubicaciones de Amazon Simple Storage Service (Amazon S3).

Ambas son claves administradas por el cliente y Claves administradas por AWS son compatibles. No es compatible el cifrado/descifrado del cliente.

> **Important**
>
> Evite registrar un bucket de Amazon S3 que tenga activada la opción El solicitante paga. Para los buckets registrados en Lake Formation, el rol utilizado para registrar el bucket se considera siempre como el solicitante. Si otra AWS cuenta accede al depósito, se le cobrará al propietario del depósito por el acceso a los datos si el rol pertenece a la misma cuenta que el propietario del depósito.

En esta sección se explica cómo registrar una ubicación de Amazon S3 en las siguientes circunstancias:

- Los datos de la ubicación de Amazon S3 se cifran con una clave KMS creada en AWS KMS.
- La ubicación de Amazon S3 no está en la misma AWS cuenta que la AWS Glue Data Catalog.
- La clave de KMS está o no en la misma AWS cuenta que el catálogo de datos.

Para registrar un bucket de Amazon S3 AWS KMS cifrado en la AWS cuenta B con un rol AWS Identity and Access Management (IAM) en la AWS cuenta A se requieren los siguientes permisos:

- El rol de la cuenta A debe conceder permisos sobre el bucket de la cuenta B.
- La política de bucket de la cuenta B debe conceder permisos de acceso al rol de la cuenta A.
- Si la clave de KMS está en la cuenta B, la política de claves debe conceder el acceso al rol de la cuenta A y el rol de la cuenta A debe conceder permisos sobre la clave de KMS.

En el siguiente procedimiento, se crea un rol en la AWS cuenta que contiene el catálogo de datos (la cuenta A en el análisis anterior). A continuación, utilice este rol para registrar la ubicación. Lake Formation asume este rol al acceder a los datos subyacentes en Amazon S3. El rol asumido tiene los permisos necesarios en la clave de KMS. Como resultado, no tendrá que conceder permisos sobre la clave KMS a las entidades principales que accedan a los datos subyacentes con trabajos ETL o con servicios integrados como Amazon Athena.

> **Important**
>
> No puede utilizar el rol vinculado al servicio Lake Formation para registrar una ubicación en otra cuenta. En su lugar, debe utilizar un rol definido por el usuario. El rol debe cumplir los requisitos de the section called “Requisitos de los roles utilizados para registrar ubicaciones”. Para obtener más información sobre el rol vinculado a servicios, consulte Permisos de rol vinculados al servicio para Lake Formation.

Antes de empezar

Revise los requisitos del rol utilizado para registrar la ubicación.

Para registrar una ubicación de Amazon S3 cifrada en todas AWS las cuentas

1. En la misma AWS cuenta que el catálogo de datos, inicie sesión en la consola de IAM AWS Management Console y ábrala enhttps://console.aws.amazon.com/iam/.
2. Cree un nuevo rol o consulte uno existente que cumpla con los requisitos de the section called "Requisitos de los roles utilizados para registrar ubicaciones". Asegúrese de que el rol incluye una política que concede permisos de Amazon S3 en la ubicación.
3. Si la clave de KMS no está en la misma cuenta que el Catálogo de datos, añada al rol una política integrada que conceda los permisos necesarios para la clave de KMS. A continuación, se muestra una política de ejemplo. Sustituya <cmk-region>*< cmk-account-id >* por la región y el número de cuenta de la clave KMS. Sustituya *<id-clave>* por el ID de la llave.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
        "Effect": "Allow",
        "Action": [
            "kms:Encrypt",
            "kms:Decrypt",
            "kms:ReEncrypt*",
            "kms:GenerateDataKey*",
            "kms:DescribeKey"
         ],
        "Resource": "arn:aws:kms:<cmk-region>:<cmk-account-id>:key/<key-id>"
        }
    ]
}
```

4. En la consola de Amazon S3, agregue una política de bucket que conceda los permisos de Amazon S3 necesarios para el rol. A continuación se muestra un ejemplo de política de bucket. Sustituya *< catalog-account-id >* por el número de AWS cuenta del catálogo de datos, por <role-name>el nombre de su función y por <bucket-name>el nombre del depósito.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect":"Allow",
            "Principal": {
                "AWS":"arn:aws:iam::<catalog-account-id>:role/<role-name>"
```

```
            },
            "Action":"s3:ListBucket",
            "Resource":"arn:aws:s3:::<bucket-name>"
        },
        {
            "Effect":"Allow",
            "Principal": {
                "AWS":"arn:aws:iam::<catalog-account-id>:role/<role-name>"
            },
            "Action": [
                "s3:DeleteObject",
                "s3:GetObject",
                "s3:PutObject"
            ],
            "Resource":"arn:aws:s3:::<bucket-name>/*"
        }
    ]
}
```

5. En AWS KMS, agregue el rol como usuario de la clave KMS.

   a. Abra la AWS KMS consola en https://console.aws.amazon.com/kms. A continuación, inicie sesión como usuario administrador o como usuario que puede modificar la política de claves de la clave de KMS utilizada para cifrar la ubicación.

   b. En el panel de navegación, seleccione Claves administradas por el cliente y, a continuación, elija el nombre de la clave KMS.

   c. En la página de detalles de la clave KMS, en la pestaña Política de claves, si no se muestra la vista JSON de la política de claves, seleccione Cambiar a la vista de políticas.

   d. En la sección Política de claves, seleccione Editar y añada el Nombre de recurso de Amazon (ARN) del rol al objeto `Allow use of the key`, como se muestra en el siguiente ejemplo.

> **Note**
>
> Si falta ese objeto, agréguelo con los permisos que se muestran en el ejemplo.

```
        ...
        {
            "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": "*"
        },
        ...
```

Para más información, consulte Permitir a los usuarios de otras cuentas utilizar una clave KMS en la Guía para desarrolladores de AWS Key Management Service .

6. Abra la AWS Lake Formation consola en https://console.aws.amazon.com/lakeformation/. Inicie sesión en la cuenta AWS del Catálogo de datos como administrador del lago de datos.
7. En el panel de navegación, en Registrar e ingerir, seleccione las Ubicaciones del lago de datos.
8. Seleccione Registrar ubicación.
9. En la página Registrar ubicación, para la ruta de Amazon S3, introduzca el nombre del bucket como **s3://*<bucket>*/*<prefix>***. Sustituya *<bucket>* por el nombre del bucket y *<prefijo>* por el resto de la ruta para la ubicación.

> **Note**
>
> Debe escribir la ruta porque los buckets entre cuentas no aparecen en la lista cuando selecciona Examinar.

10. Para el rol de IAM, elija el rol del paso 2.
11. Seleccione Registrar ubicación.

Dar de baja el registro de una ubicación de Amazon S3

Puede anular el registro de una ubicación de Amazon Simple Storage Service (Amazon S3) si desea que Lake Formation deje de administrarla. Dar de baja una ubicación no afecta a los permisos de ubicación de datos de Lake Formation que se conceden sobre esa ubicación. Puede volver a registrar una ubicación que haya dado de baja y los permisos de ubicación de datos seguirán vigentes. Puede utilizar un rol diferente para volver a registrar la ubicación.

Para dar de baja una ubicación (consola)

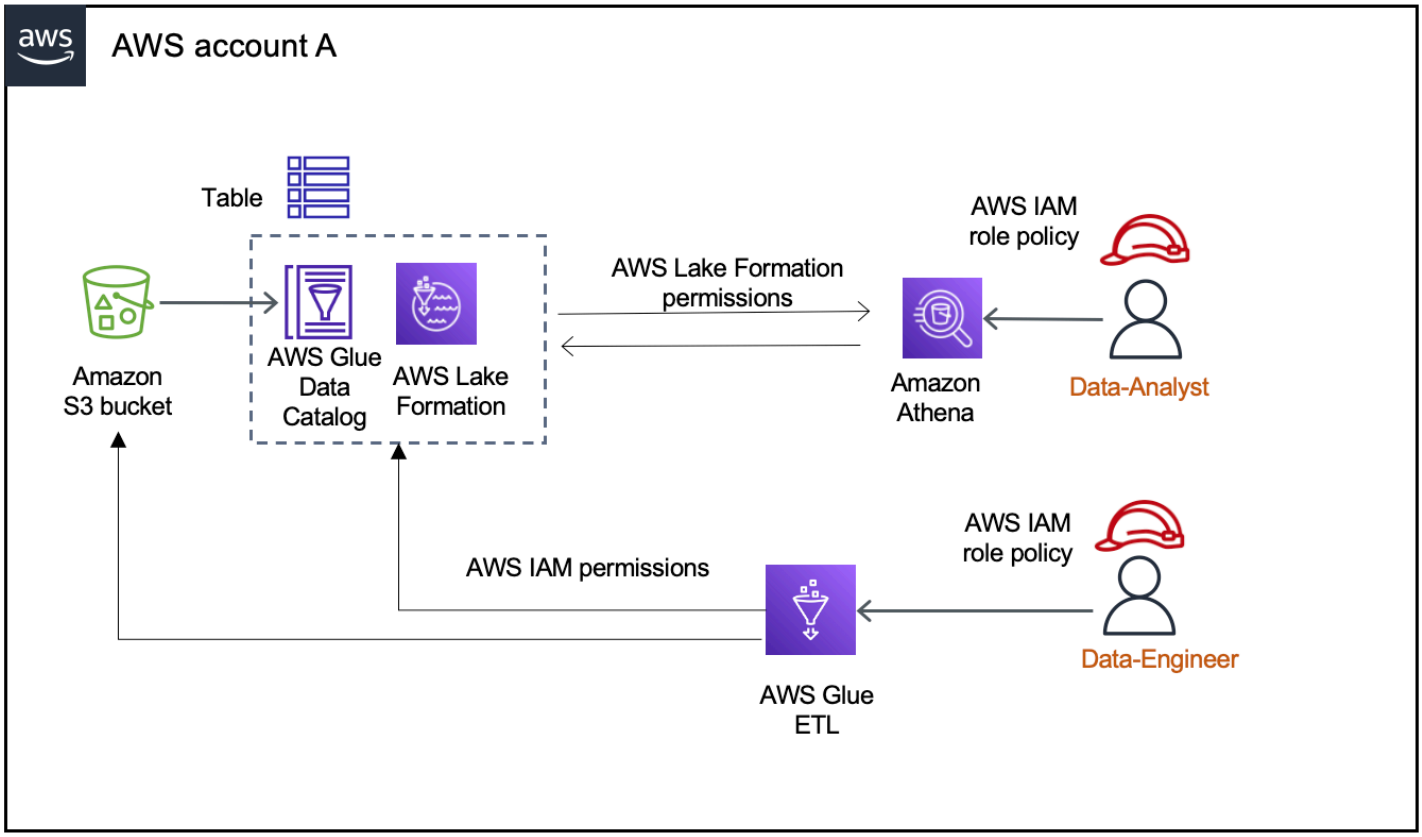
1. Abra la AWS Lake Formation consola en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con el permiso de `lakeformation:RegisterResource` IAM.
2. En el panel de navegación, vaya a Registrar e ingerir y seleccione Ubicaciones del lago de datos.
3. Seleccione una ubicación y, en el menú Acciones, seleccione Editar.
4. Cuando se le indique que confirme, elija Quitar.

Modo de acceso híbrido

AWS Lake Formation el modo de acceso híbrido admite dos rutas de permisos a las mismas AWS Glue Data Catalog bases de datos y tablas.
En la primera opción, Lake Formation te permite seleccionar directores específicos y concederles permisos de Lake Formation para acceder a bases de datos y tablas si optas por participar. La segunda vía permite a todos los demás responsables acceder a estos recursos a través de las políticas principales de IAM predeterminadas para Amazon S3 y AWS Glue sus acciones.

Al registrar una ubicación de Amazon S3 en Lake Formation, tiene la opción de aplicar los permisos de Lake Formation para todos los recursos de esta ubicación o utilizar el modo de acceso híbrido. El modo de acceso híbrido solo aplica los permisos `CREATE_TABLE`, `CREATE_PARTITION` y `UPDATE_TABLE` de forma predeterminada. Cuando una ubicación de Amazon S3 se encuentra en el modo híbrido, puede habilitar los permisos de Lake Formation optando por entidades principales para bases de datos y tablas de esa ubicación.

Así, el modo de acceso híbrido proporciona la flexibilidad necesaria para habilitar de forma selectiva Lake Formation para bases de datos y tablas de su Catálogo de datos para un conjunto específico de usuarios sin interrumpir el acceso para otros usuarios o cargas de trabajo existentes.



Para ver las consideraciones y limitaciones, consulte Consideraciones y limitaciones del modo de acceso híbrido .

Términos y definiciones

A continuación se detallan las definiciones de los recursos del Catálogo de datos según la configuración de los permisos de acceso:

Recurso de Lake Formation

Un recurso registrado con Lake Formation. Los usuarios necesitan permisos de Lake Formation para acceder al recurso.

AWS Glue recurso

Un recurso que no está registrado con Lake Formation. Los usuarios solo necesitan permisos de IAM para acceder al recurso porque tiene permisos de grupo `IAMAllowedPrincipals`. Los permisos de Lake Formation no se aplican.

Para obtener más información sobre los permisos de grupo de `IAMAllowedPrincipals`, consulte Permisos de metadatos.

Recursos híbridos

Recursos registrados en modo de acceso híbrido. En función de los usuarios que accedan al recurso, este cambia dinámicamente entre ser un recurso de Lake Formation o un recurso de AWS Glue .

Casos de uso comunes del modo de acceso híbrido

Puede utilizar el modo de acceso híbrido para proporcionar acceso en escenarios de uso compartido de datos de una sola cuenta y entre cuentas:

Escenarios de una sola cuenta

- Convertir un AWS Glue recurso en un recurso híbrido: en este escenario, actualmente no está utilizando Lake Formation, pero desea adoptar los permisos de Lake Formation para las bases de datos y tablas de Data Catalog. Al registrar la ubicación de Amazon S3 en modo de acceso híbrido, puede conceder permisos de Lake Formation a los usuarios que opten por acceder a bases de datos y tablas específicas que apunten a esa ubicación.
- Convierta un recurso de Lake Formation en un recurso híbrido: actualmente, utiliza los permisos de Lake Formation para controlar el acceso a una base de datos del catálogo de datos, pero desea proporcionar acceso a nuevos directores mediante los permisos de IAM para Amazon S3 y AWS Glue sin interrumpir los permisos de Lake Formation existentes.

  Al actualizar el registro de una ubicación de datos al modo de acceso híbrido, las nuevas entidades principales pueden acceder a la base de datos del Catálogo de datos que apunta a la ubicación de Amazon S3 mediante las políticas de permisos de IAM sin interrumpir los permisos de Lake Formation de los usuarios existentes.

  Antes de actualizar el registro de ubicación de datos para habilitar el modo de acceso híbrido, primero debe seleccionar las entidades principales que acceden ahora al recurso con permisos de Lake Formation.
  Esto tiene como objetivo evitar una posible interrupción del flujo de trabajo actual.
  También debe conceder permiso de `Super` sobre las tablas de la base de datos al grupo `IAMAllowedPrincipal`.

Escenarios de uso compartido de datos entre cuentas

- Comparta AWS Glue recursos mediante el modo de acceso híbrido: en este escenario, la cuenta del productor tiene tablas en una base de datos que actualmente se comparten con una cuenta de consumidor mediante políticas de permisos de IAM para Amazon S3 y AWS Glue acciones. La ubicación de los datos de la base de datos no está registrada en Lake Formation.

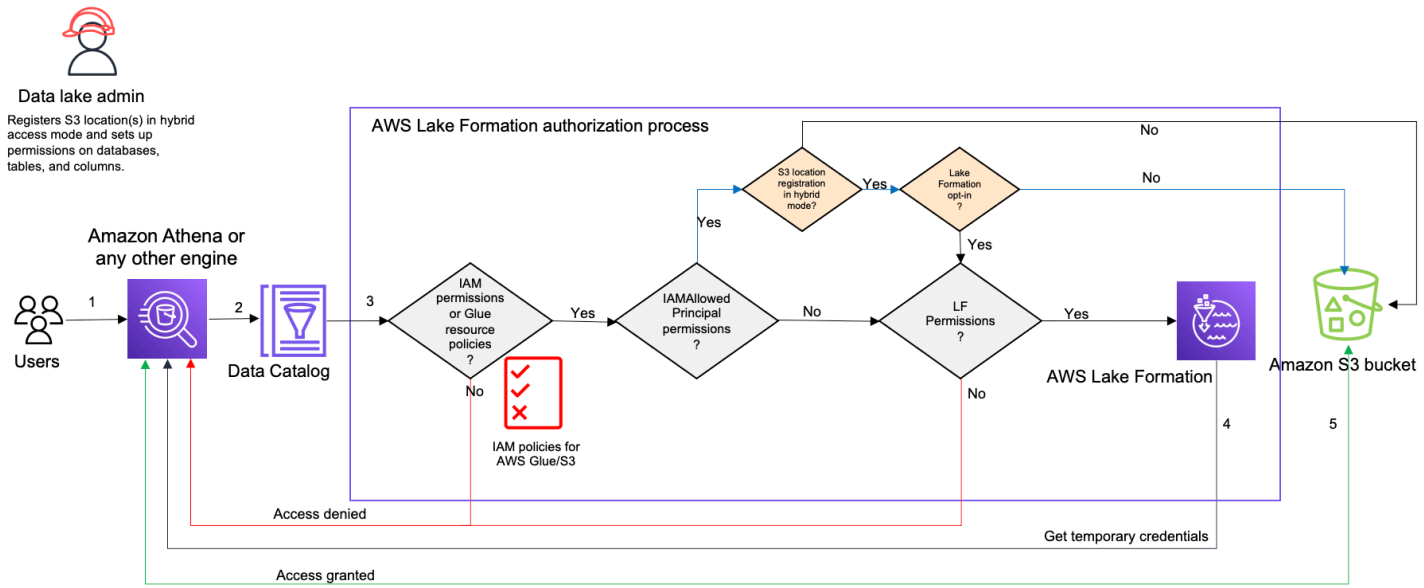
  Antes de registrar la ubicación de los datos en el modo de acceso híbrido, debe actualizar la configuración de la versión entre cuentas a la versión 4. La versión 4 proporciona las nuevas políticas de AWS RAM permisos necesarias para compartir entre cuentas cuando el `IAMAllowedPrincipal` grupo tiene `Super` permiso sobre el recurso. Para los recursos con permisos de grupo `IAMAllowedPrincipal`, puede conceder permisos de Lake Formation a cuentas externas y optar por que utilicen los permisos de Lake Formation. El administrador del lago de datos de la cuenta receptora puede conceder permisos de Lake Formation a entidades principales de la cuenta y optar por ellos para aplicar los permisos de Lake Formation.
- Compartir recursos de Lake Formation utilizando el modo de acceso híbrido: en la actualidad, la cuenta de productor tiene tablas en una base de datos que se comparten con una cuenta de consumidor que aplica los permisos de Lake Formation. La ubicación de los datos de la base de datos no está registrada en Lake Formation.

  En este caso, puede actualizar el registro de ubicación de Amazon S3 al modo de acceso híbrido y compartir los datos de Amazon S3 y los metadatos del Catálogo de datos mediante las políticas de bucket de Amazon S3 y las políticas de recursos del Catálogo de datos con las entidades principales de la cuenta del consumidor. Debe volver a conceder los permisos existentes de Lake Formation y optar por las entidades principales antes de actualizar el registro de la ubicación de Amazon S3. También debe conceder permiso de `Super` sobre las tablas de la base de datos al grupo `IAMAllowedPrincipals`.

Temas

Cómo funciona el modo de acceso híbrido

El diagrama siguiente muestra cómo funciona la autorización de Lake Formation en el modo de acceso híbrido al consultar los recursos del Catálogo de datos.



Antes de acceder a los datos del lago de datos, un administrador de este o un usuario con permisos administrativos configura políticas de usuario individuales de las tablas del Catálogo de datos para permitir o denegar el acceso a las tablas del Catálogo de datos. Luego, una entidad principal con permisos para la operación de `RegisterResource` registra en Lake Formation la ubicación de Amazon S3 de la tabla en modo de acceso híbrido. El administrador concede permisos de Lake Formation a usuarios específicos sobre las bases de datos y tablas del Catálogo de datos y les da la opción de utilizar los permisos de Lake Formation para esas bases de datos y tablas en modo de acceso híbrido.

1. Envía una consulta: un director envía una consulta o un script de ETL mediante un servicio integrado como Amazon Athena, Amazon EMR o AWS Glue Amazon Redshift Spectrum.
2. Solicita datos: el motor analítico integrado identifica la tabla solicitada y envía la solicitud de metadatos al Catálogo de datos (`GetTable`, `GetDatabase`)
3. Comprueba los permisos: el Catálogo de datos verifica los permisos de acceso de la entidad principal que hace la consulta con Lake Formation.
   a. Si la tabla no tiene permisos de grupo `IAMAllowedPrincipals` adjuntos, se aplican los permisos de Lake Formation.

b. Si la entidad principal ha optado por utilizar los permisos de Lake Formation en el modo de acceso híbrido y la tabla tiene adjuntos permisos de grupo `IAMAllowedPrincipals`, se aplicarán los permisos de Lake Formation. El motor de consulta aplica los filtros que recibió de Lake Formation y devuelve los datos al usuario.

c. Si la ubicación de la tabla no está registrada en Lake Formation y la entidad principal no ha optado por utilizar los permisos de Lake Formation en el modo de acceso híbrido, el Catálogo de datos comprueba si la tabla tiene permisos de grupo `IAMAllowedPrincipals` adjuntos. Si existe este permiso sobre la tabla, todas las entidades principales de la cuenta obtienen los permisos `Super` o `All` sobre la tabla.

4. Obtener credenciales: el Catálogo de datos comprueba y permite al motor saber si la ubicación de la tabla está registrada en Lake Formation o no. Si los datos subyacentes están registrados en Lake Formation, el motor analítico solicita a Lake Formation credenciales temporales para acceder a los datos del bucket de Amazon S3.
5. Obtener datos: si la entidad principal está autorizada para acceder a los datos de la tabla, Lake Formation proporciona acceso temporal al motor analítico integrado. Mediante el acceso temporal, el motor analítico obtiene los datos de Amazon S3 y aplica el filtrado necesario, como el de columnas, filas o celdas. Cuando el motor termina de ejecutar el trabajo, devuelve los resultados al usuario. Este proceso se denomina “expedición de credenciales”. Para obtener más información, consulte Integración con Lake Formation.
6. Si la ubicación de los datos de la tabla no está registrada en Lake Formation, la segunda llamada desde el motor analítico se hace directamente a Amazon S3. Para el acceso a los datos se evalúan la política de buckets de Amazon S3 y la política de usuarios de IAM correspondientes. Siempre que utilice políticas de IAM, compruebe que sigue las mejores prácticas IAM. Para obtener más información, consulte la sección Prácticas recomendadas de IAM en la Guía del usuario de IAM.

Configuración del modo de acceso híbrido: escenarios comunes

Al igual que con los permisos de Lake Formation, generalmente hay dos tipos de escenarios en los que puede usar el modo de acceso híbrido para administrar el acceso a los datos: proporcionar acceso a los directores dentro de uno Cuenta de AWS y proporcionar acceso a uno externo Cuenta de AWS o principal.

En esta sección se proporcionan instrucciones para configurar el modo de acceso híbrido en los escenarios siguientes:

Administre los permisos en el modo de acceso híbrido desde un solo lugar Cuenta de AWS

- Convertir un AWS Glue recurso en un recurso híbrido — Actualmente proporciona acceso a las tablas de una base de datos a todos los directores de su cuenta mediante permisos de IAM para Amazon S3, AWS Glue pero desea adoptar Lake Formation para administrar los permisos de forma incremental.

- Convertir un recurso de Lake Formation en un recurso híbrido — Actualmente utilizas Lake Formation para administrar el acceso a las tablas de una base de datos para todos los directores de tu cuenta, pero quieres usar Lake Formation solo para directores específicos. Desea proporcionar acceso a los nuevos directores mediante el uso de permisos de IAM para AWS Glue Amazon S3 en la misma base de datos y tablas.

Administre los permisos en modo de acceso híbrido en s Cuenta de AWS

- Compartir un AWS Glue recurso mediante el modo de acceso híbrido— Actualmente no utilizas Lake Formation para gestionar los permisos de una tabla, pero quieres aplicar los permisos de Lake Formation para permitir el acceso a los directores de otra cuenta.
- Compartir un recurso de Lake Formation mediante el modo de acceso híbrido— Utiliza Lake Formation para administrar el acceso a una tabla, pero desea proporcionar acceso a los directores de otra cuenta mediante permisos de IAM para Amazon S3 AWS Glue y en la misma base de datos y tablas.

Configuración del modo de acceso híbrido: pasos generales

1. Registre la ubicación de datos de Amazon S3 en Lake Formation seleccionando el modo de acceso híbrido.
2. Las entidades principales deben tener permisos de `DATA_LOCATION` en una ubicación de lago de datos para crear tablas o bases de datos del Catálogo de datos que apunten a esa ubicación.
3. Establezca la configuración de la versión entre cuentas en la Versión 4.
4. Conceda permisos específicos a roles o usuarios de IAM concretos en bases de datos y tablas. Al mismo tiempo, asegúrese de establecer permisos de `Super` o `All` para el grupo `IAMAllowedPrincipals` de la base de datos y para todas las tablas de la base de datos o para algunas de ellas.

5. Elija las entidades principales y los recursos. Los demás responsables de la cuenta pueden seguir accediendo a las bases de datos y tablas mediante las políticas de permisos de IAM AWS Glue y las acciones de Amazon S3.
6. También puede limpie las políticas de permisos de IAM para Amazon S3 para las entidades principales que hayan optado por usar los permisos de Lake Formation.

Requisitos previos para configurar el modo de acceso híbrido

Los siguientes son los requisitos previos para configurar el modo de acceso híbrido:

> **Note**
>
> Recomendamos que un administrador de Lake Formation registre la ubicación de Amazon S3 en modo de acceso híbrido y opte por entidades principales y recursos.

1. Otorgue el permiso de ubicación de datos (`DATA_LOCATION_ACCESS`) para crear recursos del Catálogo de datos que apunten a las ubicaciones de Amazon S3. Los permisos de ubicación de datos controlan la capacidad de crear bases de datos y tablas de Data Catalog que apuntan a ubicaciones concretas de Amazon S3.
2. Para compartir los recursos del Catálogo de datos con otra cuenta en modo de acceso híbrido (sin eliminar los permisos de `IAMAllowedPrincipals` de grupo del recurso), debe actualizar la configuración de la versión entre cuentas a la Versión 4. Para actualizar la versión mediante la consola de Lake Formation, elija la Versión 4 en la configuración de la versión entre cuentas en la página de configuración del Catálogo de datos.

   También puede usar el `put-data-lake-settings` AWS CLI comando para establecer el `CROSS_ACCOUNT_VERSION` parámetro en la versión 4:

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
 file://settings
{
"DataLakeAdmins": [
        {
"DataLakePrincipalIdentifier": "arn:aws:iam::<111122223333>:user/<user-name>"
        }
    ],
    "CreateDatabaseDefaultPermissions": [],
```

```
    "CreateTableDefaultPermissions": [],
    "Parameters": {
"CROSS_ACCOUNT_VERSION": "4"
    }
}
```

3. Para conceder permisos entre cuentas en el modo de acceso híbrido, el otorgante debe tener los permisos de IAM y los servicios necesarios. AWS Glue AWS RAM La política AWS gestionada `AWSLakeFormationCrossAccountManager` concede los permisos necesarios.
Para permitir el intercambio de datos entre cuentas en el modo de acceso híbrido, hemos actualizado la política administrada `AWSLakeFormationCrossAccountManager` añadiendo dos nuevos permisos de IAM:

   - RAM: ListResourceSharePermissions
   - RAM: AssociateResourceSharePermission

   > **Note**
   >
   > Si no utilizas la política AWS gestionada para el rol de otorgante, añade las políticas anteriores a tus políticas personalizadas.

Convertir un AWS Glue recurso en un recurso híbrido

Siga estos pasos para registrar una ubicación de Amazon S3 en modo de acceso híbrido e incorporar nuevos usuarios de Lake Formation sin interrumpir el acceso a los datos de los usuarios actuales del Catálogo de datos.

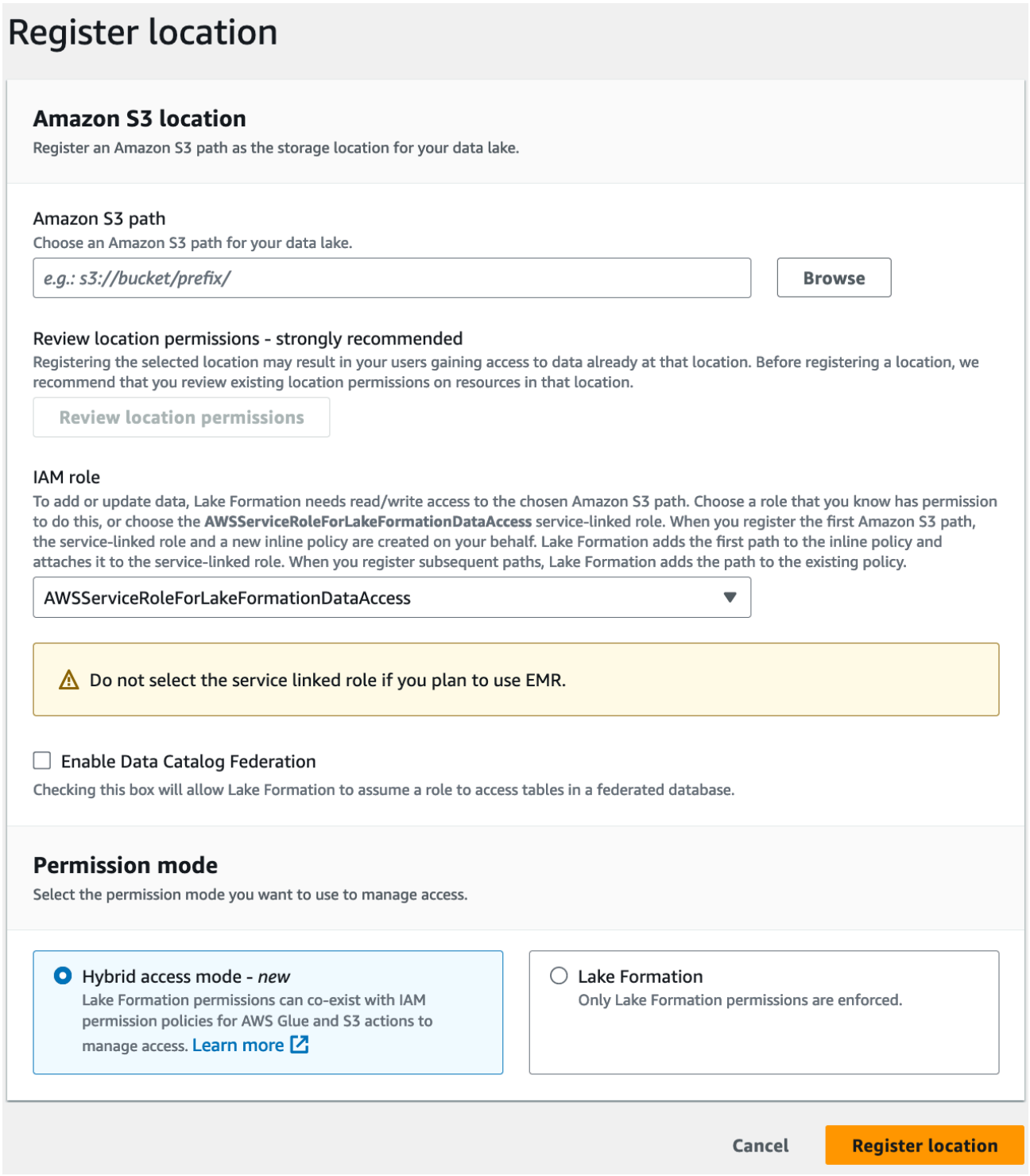
Descripción del escenario: la ubicación de los datos no está registrada en Lake Formation y el acceso de los usuarios a la base de datos y las tablas del Catálogo de datos se determina mediante las políticas de permisos de IAM para acciones de AWS Glue y Amazon S3.
De forma predeterminada, el grupo `IAMAllowedPrincipals` tiene permisos de `Super` en todas las tablas de la base de datos.

Para habilitar el modo de acceso híbrido para una ubicación de datos no registrada en Lake Formation

1. Registre una ubicación de Amazon S3 que habilite el modo de acceso híbrido.

   Console

   1. Inicie sesión en la consola de Lake Formation como administrador del lago de datos.
   2. En Administración del panel de navegación, seleccione las Ubicaciones de los lagos de datos.
   3. Seleccione Registrar ubicación.



4. En la ventana Registrar ubicación, elija la ruta de Amazon S3 que desee registrar en Lake Formation.
5. Para el rol de IAM, elija el rol vinculado al servicio `AWSServiceRoleForLakeFormationDataAccess` (valor predeterminado) o un rol de

IAM personalizado que cumpla los requisitos de Requisitos de los roles utilizados para registrar ubicaciones.

6. Elija el modo de acceso híbrido para aplicar políticas específicas de control de acceso de Lake Formation a las entidades principales y a las bases de datos y tablas del Catálogo de datos que apuntan a la ubicación registrada.

   Seleccione Lake Formation para permitir que Lake Formation autorice las solicitudes de acceso a la ubicación registrada.

7. Seleccione Registrar ubicación.

AWS CLI

El siguiente es un ejemplo para registrar una ubicación de datos con Lake Formation HybridAccessEnabled con:true/false. El valor predeterminado para el parámetro `HybridAccessEnabled` es false. Sustituya la ruta, el nombre del rol y el identificador de AWS cuenta de Amazon S3 por valores válidos.

```
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. Conceda permisos y seleccione entidades principales para utilizar los permisos de Lake Formation para los recursos en modo de acceso híbrido.

   Antes de optar por entidades principales y recursos en el modo de acceso híbrido, verifique que la concesión de permisos `Super` o `All` al grupo `IAMAllowedPrincipals` existe en las bases de datos y tablas que tienen ubicación registrada con Lake Formation en el modo de acceso híbrido.

> **Note**
>
> No puede conceder al grupo `IAMAllowedPrincipals` permiso sobre `All tables` dentro de una base de datos. Debe seleccionar cada tabla por separado en el menú desplegable y conceder los permisos. Además, al crear nuevas tablas en la base

de datos, puede optar por utilizarlas `Use only IAM access control for new tables in new databases` en la configuración del catálogo de datos. Esta opción concede el permiso de `Super` al grupo `IAMAllowedPrincipals` automáticamente al crear nuevas tablas en la base de datos.

Console

1. En la consola de Lake Formation, en Catálogo de datos, elija Bases de datos o Tablas.
2. Seleccione una base de datos o una tabla de la lista y elija Otorgar en el menú Acciones.
3. Elija entidades principales a las que conceder permisos sobre la base de datos, las tablas y las columnas utilizando el método de recursos con nombre o las etiquetas LF.

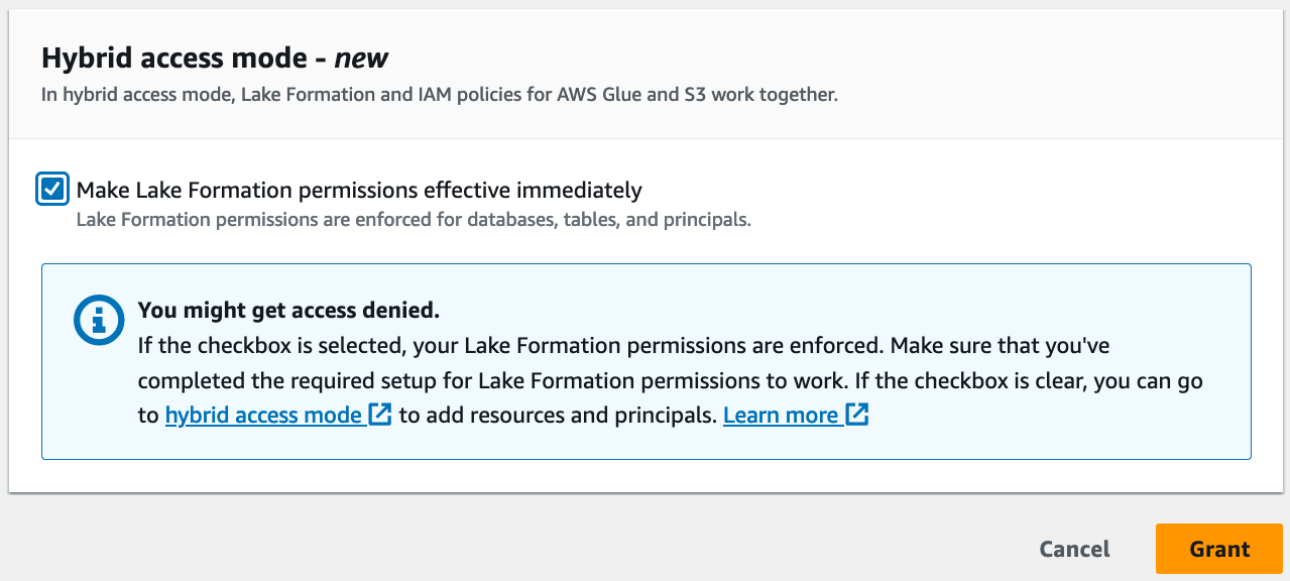
   Como alternativa, elija Permisos de lago de datos, seleccione las entidades principales a las que conceder permisos de la lista y elija Conceder.

   Para obtener más información sobre la concesión de permisos de datos, consulte Concesión y revocación de permisos sobre los recursos del catálogo de datos.

   > **Note**
   >
   > Si concede a una entidad principal el permiso para crear una tablas, también deberá concederle permisos de ubicación de datos (`DATA_LOCATION_ACCESS`). Este permiso no es necesario para actualizar las tablas.
   > Para obtener más información, consulte Conceder permisos de ubicación de datos.

4. Si utiliza el método de recurso con nombre para conceder permisos, la opción de incluir entidades principales y recursos está disponible en la sección inferior de la página de concesión de permisos de datos.

   Seleccione Hacer efectivos inmediatamente los permisos de Lake Formation para habilitar los permisos de Lake Formation para las entidades principales y los recursos.



5. Elija Conceder.

   Al optar por la entidad principal A en la tabla A que apunta a una ubicación de datos, permite que la entidad principal A tenga acceso a la ubicación de esta tabla utilizando los permisos de Lake Formation si la ubicación de datos está registrada en modo híbrido.

AWS CLI

A continuación se muestra un ejemplo para optar por una entidad principal y una tabla en modo de acceso híbrido. Sustituya el nombre del rol, el id de la cuenta de AWS , el nombre de la base de datos y el nombre de la tabla por valores aceptables.

```
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>"
            }
        }
```

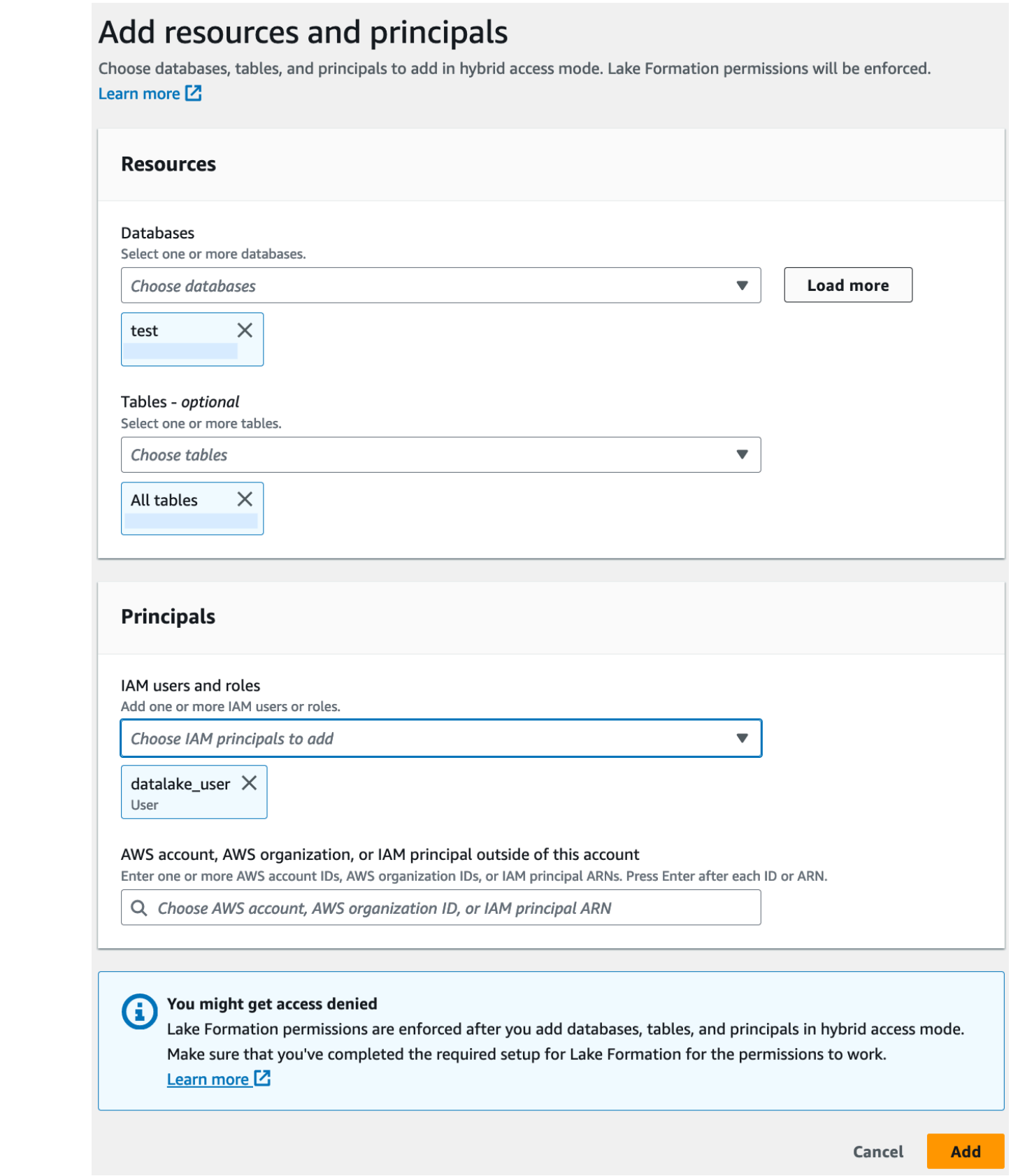
```
    }
```

a. (Optional) Si elige las etiquetas LF para la concesión de permisos, puede optar por que las entidades principales utilicen los permisos de Lake Formation en un paso aparte. Para ello, elija el modo de acceso Híbrido en Permisos de la barra de navegación izquierda.

b. En la sección inferior de la página Modo de acceso híbrido, seleccione Añadir para añadir recursos y entidades principales al modo de acceso híbrido.

c. En la página Añadir recursos y entidades principales, elija las bases de datos y las tablas registradas en el modo de acceso híbrido. Elija entidades principales para optar a utilizar los permisos de Lake Formation en el modo de acceso híbrido.

   Puede elegir `All tables` bajo una base de datos a la que conceder acceso.


Add resources and principals
Choose databases, tables, and principals to add in hybrid access mode. Lake Formation permissions will be enforced.
Learn more
Resources
Databases
Select one or more databases.
Choose databases
Load more
test
Tables - optional
Select one or more tables.
Choose tables
All tables
Principals
IAM users and roles
Add one or more IAM users or roles.
Choose IAM principals to add
datalake_user
User
AWS account, AWS organization, or IAM principal outside of this account
Enter one or more AWS account IDs, AWS organization IDs, or IAM principal ARNs. Press Enter after each ID or ARN.
Choose AWS account, AWS organization ID, or IAM principal ARN
You might get access denied
Lake Formation permissions are enforced after you add databases, tables, and principals in hybrid access mode.
Make sure that you've completed the required setup for Lake Formation for the permissions to work.
Learn more
Cancel
Add

Convertir un recurso de Lake Formation en un recurso híbrido

En los casos en los que esté utilizando permisos de Lake Formation para sus bases de datos y tablas del Catálogo de datos, puede editar las propiedades de registro de la ubicación para activar el modo de acceso híbrido. Esto le permite proporcionar a los nuevos directores acceso a los mismos recursos mediante las políticas de permisos y AWS Glue acciones de IAM para Amazon S3 sin interrumpir los permisos existentes de Lake Formation.

Descripción del escenario: en los siguientes pasos se asume que tiene una ubicación de datos registrada en Lake Formation y que ha configurado permisos para entidades principales en bases de datos, tablas o columnas que apuntan a esa ubicación. Si la ubicación se registró con un rol vinculado a un servicio, no podrá actualizar los parámetros de ubicación ni habilitar el modo de acceso híbrido. De forma predeterminada, el grupo `IAMAllowedPrincipals` tiene permisos Super sobre la base de datos y todas sus tablas.

> **Important**
>
> No actualice el registro de una ubicación al modo de acceso híbrido sin optar por los principales que acceden a los datos de esta ubicación.

Activación del modo de acceso híbrido para una ubicación de datos registrada en Lake Formation

1.
> **Warning**
>
> No recomendamos convertir una ubicación de datos gestionada por Lake Formation en un modo de acceso híbrido para evitar interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes.

   Opte por las entidades principales existentes que tengan permisos de Lake Formation.

   1. Recopile y revise los permisos que ha concedido a las entidades principales sobre las bases de datos y las tablas. Para obtener más información, consulte Consulta de los permisos de bases de datos y tablas en Lake Formation.
   2. Elija el modo de acceso Híbrido en Permisos de la barra de navegación izquierda y seleccione Añadir.

3. En la página Añadir entidades principales y recursos, elija las bases de datos y las tablas de la ubicación de datos de Amazon S3 que desee utilizar en el modo de acceso híbrido. Elija las entidades principales que ya tienen permisos de Lake Formation.
4. Elija Añadir para optar a que las entidades principales utilicen los permisos de Lake Formation en el modo de acceso híbrido.

2. Actualice el registro del bucket o prefijo de Amazon S3 seleccionando la opción de modo de acceso Híbrido.

   Console

   1. Inicie sesión en la consola de Lake Formation como administrador del lago de datos.
   2. En el panel de navegación, vaya a Registrar e ingerir y seleccione las ubicaciones de los lagos de datos.
   3. Seleccione una ubicación y, en el menú Acciones, seleccione Editar.
   4. Elija Modo de acceso híbrido.
   5. Seleccione Guardar.
   6. En el Catálogo de datos, seleccione la base de datos o tabla y conceda permisos `Super` o `All` al grupo virtual denominado `IAMAllowedPrincipals`.
   7. Compruebe que el acceso de sus usuarios actuales de Lake Formation no se interrumpa al actualizar las propiedades de registro de la ubicación. Inicie sesión en la consola de Athena como entidad principal de Lake Formation y ejecute una consulta de ejemplo en una tabla que apunte a la ubicación actualizada.

      Del mismo modo, compruebe el acceso de AWS Glue los usuarios que utilizan las políticas de permisos de IAM para acceder a la base de datos y a las tablas.

   AWS CLI

   El siguiente es un ejemplo para registrar una ubicación de datos con Lake Formation HybridAccessEnabled con:true/false. El valor predeterminado para el parámetro `HybridAccessEnabled` es false. Sustituya la ruta, el nombre del rol y el identificador de AWS cuenta de Amazon S3 por valores válidos.

   ```
   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
}
```

Compartir un AWS Glue recurso mediante el modo de acceso híbrido

Comparta datos con otra persona Cuenta de AWS o con un director en otra persona Cuenta de AWS haciendo cumplir los permisos de Lake Formation sin interrumpir el acceso basado en IAM de los usuarios existentes del Catálogo de Datos.

Descripción del escenario: la cuenta del productor tiene una base de datos del catálogo de datos cuyo acceso está controlado mediante las principales políticas y AWS Glue acciones de IAM para Amazon S3. La ubicación de los datos de la base de datos no está registrada en Lake Formation. El `IAMAllowedPrincipals` grupo, de forma predeterminada, tiene `Super` permisos en la base de datos y en todas sus tablas.

Conceder permisos entre cuentas de Lake Formation en modo de acceso híbrido

1. Configuración de una cuenta de productor

    1. Inicie sesión en la consola de Lake Formation con un rol que tenga permiso `lakeformation:PutDataLakeSettings` de IAM.
    2. Vaya a la configuración del Catálogo de datos y seleccione `Version 4` para la configuración de la versión entre cuentas.

        Si utiliza la versión 1 o 2, consulte las instrucciones de Actualización de los ajustes de la versión entre cuentas para compartir datos para actualizar a la versión 3.

        No es necesario hacer cambios en la política de permisos para actualizar de la versión 3 a la 4.
    3. Registre la ubicación en Amazon S3 de la base de datos o tabla que planea compartir en el modo de acceso híbrido.
    4. Compruebe que existe permiso `Super` para el grupo `IAMAllowedPrincipals` en las bases de datos y tablas en las que registró la ubicación de datos en modo de acceso híbrido en el paso anterior.

5. Otorgue permisos de Lake Formation a AWS organizaciones, unidades organizativas (OU) o directamente con un director de IAM en otra cuenta.
6. Si concede los permisos directamente a una entidad principal de IAM, opte por que la entidad principal de la cuenta de consumidor aplique los permisos de Lake Formation en el modo de acceso híbrido activando la opción Hacer efectivos inmediatamente los permisos de Lake Formation.

   Si vas a conceder permisos multicuenta a otra AWS cuenta, al activar la cuenta, los permisos de Lake Formation solo se aplican a los administradores de esa cuenta. El administrador del lago de datos de la cuenta de destinatario tiene que aplicar en cascada los permisos y optar por las entidades principales de la cuenta para hacer cumplir los permisos de Lake Formation para los recursos compartidos que están en modo de acceso híbrido.

   Si elige la opción Recursos emparejados por etiquetas LF para conceder permisos entre cuentas, deberá completar primero el paso de concesión de permisos. Puede optar por el modo de acceso híbrido para entidades principales y recursos como un paso separado, seleccionando el modo de acceso híbrido en Permisos en la barra de navegación izquierda de la consola de Lake Formation. Luego, seleccione Agregar para agregar los recursos y las entidades principales a los que desea aplicar los permisos de Lake Formation.

2. Configuración de una cuenta de consumidor

   1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ como administrador del lago de datos.
   2. Ve a https://console.aws.amazon.com/ram y acepta la invitación a compartir recursos. La pestaña Compartido conmigo de la AWS RAM consola muestra la base de datos y las tablas que se comparten con su cuenta.
   3. Cree un enlace de recursos a la base de datos o tabla compartidas en Lake Formation.
   4. Conceda permiso `Describe` en el enlace de recursos y permiso `Grant on target` (para el recurso compartido original) a las entidades principales de IAM de su cuenta (de consumidor).
   5. Conceda a las entidades principales de su cuenta permisos de Lake Formation en la base de datos o tabla compartida. Indique a las entidades principales y a los recursos que apliquen los permisos de Lake Formation en el modo de acceso híbrido activando la opción Hacer efectivos inmediatamente los permisos de Lake Formation.
   6. Pruebe los permisos de Lake Formation de la entidad principal ejecutando consultas Athena de ejemplo. Pruebe el acceso actual de sus AWS Glue usuarios con las políticas principales de IAM para Amazon S3 y AWS Glue sus acciones.

(Opcional) Elimine la política de bucket de Amazon S3 para el acceso a los datos y las políticas principales de IAM para AWS Glue y el acceso a los datos de Amazon S3 para las entidades principales que configuró para usar los permisos de Lake Formation.

Compartir un recurso de Lake Formation mediante el modo de acceso híbrido

Autorice a los nuevos usuarios del Catálogo de datos de una cuenta externa a acceder a las bases de datos y tablas del Catálogo de datos mediante políticas basadas en IAM sin interrumpir los permisos existentes de uso compartido entre cuentas de Lake Formation.

Descripción del escenario. La cuenta del productor tiene una base de datos y tablas administradas por Lake Formation que se comparten con una cuenta externa (de consumidor) a nivel de cuenta o a nivel de entidad principal de IAM. La ubicación de los datos de la base de datos está registrada en Lake Formation. El grupo `IAMAllowedPrincipals` no tiene permisos `Super` en la base de datos ni en sus tablas.

Conceder acceso entre cuentas a nuevos usuarios del Catálogo de datos mediante políticas basadas en IAM sin interrumpir el permiso existente de Lake Formation.

1. Configuración de una cuenta de productor

    1. Inicie sesión en la consola de Lake Formation con un rol que `lakeformation:PutDataLakeSettings`.
    2. Vaya a la configuración del Catálogo de datos y seleccione `Version 4` para la configuración de la versión entre cuentas.

        Si utiliza la versión 1 o 2, consulte las instrucciones de Actualización de los ajustes de la versión entre cuentas para compartir datos para actualizar a la versión 3.

        No es necesario hacer cambios en la política de permisos para actualizar de la versión 3 a la 4.
    3. Recopile y revise los permisos que ha concedido a las entidades principales sobre las bases de datos y las tablas. Para obtener más información, consulte Consulta de los permisos de bases de datos y tablas en Lake Formation.
    4. Vuelva a conceder los permisos entre cuentas existentes de Lake Formation optando por entidades principales y recursos.

> **Note**
>
> Antes de actualizar el registro de una ubicación de datos al modo de acceso híbrido para conceder permisos entre cuentas, debe volver a conceder al menos un recurso compartido de datos entre cuentas por cuenta. Este paso es necesario para actualizar los permisos AWS RAM administrados adjuntos al AWS RAM recurso compartido.
> En julio de 2023, Lake Formation actualizó los permisos AWS RAM gestionados que se utilizan para compartir bases de datos y tablas:
>
> - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueAllTablesReadWriteForDatabase` (política de uso compartido a nivel de base de datos)
> - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueTableReadWrite` (política de uso compartido a nivel de tabla)
>
> Las concesiones de permisos multicuenta realizadas antes de julio de 2023 no tienen estos AWS RAM permisos actualizados.
> Si ha concedido permisos para varias cuentas directamente a las entidades principales, tendrá que volver a concederlos individualmente. Si omite este paso, las entidades principales que accedan al recurso compartido podrían obtener un error de combinación ilegal.

5. Ve a https://console.aws.amazon.com/ram.
6. La pestaña Compartidos por mí de la AWS RAM consola muestra los nombres de bases de datos y tablas que ha compartido con una cuenta o entidad principal externa.

   Asegúrese de que los permisos adjuntos al recurso compartido tengan el ARN correcto.
7. Comprueba que los recursos del AWS RAM recurso compartido estén en `Associated` estado. Si el estado es `Associating`, espere a que pasen al estado `Associated`. Si el estado pasa a ser `Failed`, deténgase y póngase en contacto con el equipo de servicio de Lake Formation.
8. Elija el modo de acceso Híbrido en Permisos de la barra de navegación izquierda y seleccione Añadir.
9. La página Agregar entidades principales y recursos muestra las bases de datos o tablas y las entidades principales a las que se puede acceder. Para efectuar las actualizaciones necesarias, añada o quite entidades principales y recursos.

10. Elija las entidades principales con permisos de Lake Formation para la base de datos y las tablas que desea cambiar al modo de acceso híbrido. Elija las bases de datos y tablas.
11. Elija Añadir para que las entidades principales puedan utilizar los permisos de Lake Formation en el modo de acceso híbrido.
12. Conceda permiso `Super` al grupo virtual `IAMAllowedPrincipals` de su base de datos y de las tablas seleccionadas.
13. Edite el registro de Lake Formation de la ubicación de Amazon S3 en modo de acceso híbrido.
14. Otorgue permisos a los AWS Glue usuarios de la cuenta externa (de consumidor) mediante las políticas de permisos de IAM para las AWS Glue acciones de Amazon S3.

2. Configuración de una cuenta de consumidor

    1. Inicie sesión en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ como administrador del lago de datos.
    2. Vaya a https://console.aws.amazon.com/ram y acepte la invitación para compartir recursos. La pestaña Recursos compartidos conmigo de la AWS RAM página muestra los nombres de las bases de datos y las tablas que se comparten con su cuenta.

       Para AWS RAM compartir, asegúrate de que el permiso adjunto tenga el ARN correcto de la invitación compartida AWS RAM . Comprueba si los recursos del recurso AWS RAM compartido están en `Associated` estado. Si el estado es `Associating`, espere a que pasen al estado `Associated`. Si el estado pasa a ser `Failed`, deténgase y póngase en contacto con el equipo de servicio de Lake Formation.
    3. Cree un enlace de recursos a la base de datos o tabla compartidas en Lake Formation.
    4. Conceda permiso `Describe` en el enlace de recursos y permiso `Grant on target` (para el recurso compartido original) a las entidades principales de IAM de su cuenta (de consumidor).
    5. A continuación, configure los permisos de Lake Formation para las entidades principales de su cuenta en la base de datos o tabla compartida.

       En la barra de navegación de la izquierda, en Permisos, elija el modo de acceso Híbrido.
    6. Seleccione Añadir en la sección inferior de la página del modo de acceso híbrido para optar por las entidades principales y la base de datos o tabla compartida de la cuenta de productor.
    7. Conceda permisos a los AWS Glue usuarios de su cuenta mediante las políticas de permisos de IAM para las AWS Glue acciones de Amazon S3.

8. Pruebe los permisos y AWS Glue permisos de Lake Formation de los usuarios ejecutando consultas de ejemplo independientes en la tabla con Athena

   (Opcional) Limpie las políticas de permisos de IAM para Amazon S3 para las entidades principales que se encuentran en el modo de acceso híbrido.

Eliminación de entidades principales y recursos del modo de acceso híbrido

Siga estos pasos para eliminar bases de datos, tablas y entidades principales del modo de acceso híbrido.

Console

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En Permisos, seleccione el modo de acceso híbrido.
3. En la página Modo de acceso híbrido, marque la casilla situada junto al nombre de la base de datos o tabla y elija `Remove`.
4. Un mensaje de advertencia le solicitará que confirme la acción. Elija Eliminar.

   Lake Formation ya no exige permisos para esos recursos, y el acceso a este recurso se controlará mediante IAM y AWS Glue permisos. Esto puede provocar que el usuario deje de tener acceso a este recurso si no tiene los permisos de IAM adecuados.

AWS CLI

En el siguiente ejemplo, se muestra cómo eliminar recursos del modo de acceso híbrido.

```
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>"
```

```
            }
        }
}
```

Eliminación de entidades principales y recursos del modo de acceso híbrido

Siga estos pasos para eliminar bases de datos, tablas y entidades principales del modo de acceso híbrido.

Console

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En Permisos, seleccione el modo de acceso híbrido.
3. La página Modo de acceso híbrido muestra los recursos y entidades principales presentes en el modo de acceso híbrido.

AWS CLI

En el ejemplo siguiente, se muestra cómo enumerar todas las entidades principales y recursos que están en el modo de acceso híbrido.

```
aws lakeformation list-lake-formation-opt-ins
```

El siguiente ejemplo muestra cómo enumerar un par específico de entidad principal-recurso.

```
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>"
          }
    }
}
```

Recursos adicionales de

En la siguiente entrada del blog, le guiaremos a través de las instrucciones para incorporar los permisos de Lake Formation en modo de acceso híbrido para los usuarios seleccionados, mientras que la base de datos ya es accesible para otros usuarios a través de los permisos de IAM y Amazon S3. Revisaremos las instrucciones para configurar el modo de acceso híbrido dentro de una AWS cuenta y entre dos cuentas.

- Presentamos el modo de acceso híbrido AWS Glue Data Catalog para proteger el acceso mediante Lake Formation y las políticas de IAM y Amazon S3.

Creación de tablas y bases de datos del Catálogo de datos

AWS Lake Formation utiliza el Catálogo de datos AWS Glue para almacenar metadatos sobre lagos de datos, orígenes de datos, transformaciones y objetivos. Los metadatos sobre orígenes de datos y objetivos se presentan en forma de bases de datos y tablas. Las tablas almacenan información sobre los datos subyacentes, incluida la información sobre esquemas, particiones y ubicaciones de datos. Las bases de datos son colecciones de tablas. El Catálogo de datos también contiene enlaces a recursos, que son enlaces a bases de datos y tablas compartidas en cuentas externas, y se utilizan para el acceso entre cuentas a los datos del lago de datos.

Cada cuenta AWS tiene un Catálogo de datos por región AWS.

Temas

- Creación de una base de datos
- Creación de tablas
- Uso de vistas

Creación de una base de datos

Las tablas de metadatos del Catálogo de datos se almacenan en una base de datos. Puede crear tantas bases de datos como necesite, y conceder diferentes permisos de Lake Formation en cada base de datos.

Las bases de datos pueden tener una propiedad de ubicación opcional. Esta ubicación suele estar dentro de una ubicación de Amazon Simple Storage Service (Amazon S3) registrada en Lake Formation. Cuando se especifica una ubicación, las entidades principales no necesitan permisos de ubicación de datos para crear tablas del Catálogo de datos que apunten a ubicaciones dentro de la ubicación de la base de datos. Para obtener más información, consulte Underlying data access control.

Para crear una base de datos mediante la consola de Lake Formation, debe iniciar sesión como administrador del lago de datos o creador de la base de datos. El creador de una base de datos es una entidad principal a la que se le ha otorgado el permiso de `CREATE_DATABASE` de Lake Formation. Puede ver una lista de los creadores de bases de datos en la página Roles y tareas administrativas de la consola de Lake Formation. Para ver esta lista, debe tener el permiso de `lakeformation:ListPermissions` IAM e iniciar sesión como administrador de un lago de datos o como creador de bases de datos con la opción de conceder el permiso `CREATE_DATABASE`.

Para crear una base de datos

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/, e Inicie sesión como administrador del lago de datos o creador de la base de datos.
2. En el panel de navegación, bajo Catálogo de datos, elija Bases de datos.
3. Elija Crear base de datos.
4. En el cuadro de diálogo Crear base de datos, introduzca un nombre de base de datos, una ubicación opcional y una descripción opcional.
5. Seleccione de forma opcional Utilizar solo el control de acceso IAM para las nuevas tablas de esta base de datos.

   Para obtener más información acerca de esta opción, consulte the section called “Cambiar la configuración predeterminada de su lago de datos”.
6. Elija Crear base de datos.

Creación de tablas

Las tablas de metadatos de AWS Lake Formation contienen información sobre los datos del lago de datos, incluida información sobre el esquema, la partición y la ubicación de datos. Estas tablas se almacenan en el Catálogo de datos de AWS Glue. Se utilizan para acceder a los datos subyacentes del lago de datos y administrarlos con los permisos de Lake Formation. Las tablas se almacenan dentro de bases de datos en el Catálogo de datos.

Hay varias formas de crear tablas del Catálogo de datos:

- Ejecutar un rastreador en AWS Glue. Consulte Definición de rastreadores en la Guía para desarrolladores de AWS Glue.
- Crear y ejecutar un flujo de trabajo. Consulte the section called “Importación de datos mediante flujos de trabajo”.
- Cree una tabla manualmente utilizando la consola de Lake Formation, la API de AWS Glue o la AWS Command Line Interface (AWS CLI).
- Cree una tabla mediante Amazon Athena.
- Crea un enlace de recurso a una tabla en una cuenta externa. Consulte the section called “Creación de enlaces de recursos”.

Creación de tablas de Apache Iceberg

AWS Lake Formation es compatible con la creación de tablas Apache Iceberg que utilizan el formato de datos Apache Parquet en el AWS Glue Data Catalog con datos que residen en Amazon S3. Una tabla en el Catálogo de datos es la definición de metadatos que representa los datos en un almacén de datos. De forma predeterminada, Lake Formation crea tablas Iceberg v2. Para ver la diferencia entre las tablas v1 y v2, consulte Format version changes (Cambios de versión de formato) en la documentación de Apache Iceberg.

Apache Iceberg es un formato de tabla abierto para conjuntos de datos analíticos muy grandes. Iceberg permite modificar fácilmente su esquema, o evolución del esquema, de manera que los usuarios pueden añadir, renombrar o eliminar columnas de una tabla de datos sin alterar los datos subyacentes. Iceberg también ofrece compatibilidad con el control de versiones de datos, que permite a los usuarios hacer un seguimiento de los cambios en los datos a lo largo del tiempo. Esto habilita la característica de viaje en el tiempo, con la que los usuarios pueden acceder a versiones históricas de los datos y consultarlas, así como analizar los cambios en los datos entre actualizaciones y eliminaciones.

Puede utilizar la consola de Lake Formation o la operación `CreateTable` en la API de AWS Glue para crear una tabla Iceberg en el Catálogo de datos. Para obtener más información, consulte CreateTable action (Python: create_table).

Cuando cree una tabla Iceberg en el Catálogo de datos, deberá especificar el formato de la tabla y la ruta del archivo de metadatos en Amazon S3 para poder hacer lecturas y escrituras.

Puede utilizar Lake Formation para asegurar su tabla Iceberg utilizando permisos de control de acceso específicos cuando registre la ubicación de datos de Amazon S3 con AWS Lake Formation. Para los datos de origen en Amazon S3 y los metadatos que no están registrados en Lake Formation, el acceso se determina mediante las políticas de permisos de IAM para Amazon S3 y acciones de AWS Glue. Para obtener más información, consulte Administrar los permisos de Lake Formation.

> **Note**
>
> El Catálogo de datos no admite la creación de particiones ni la adición de propiedades de tablas de iceberg.

Temas

Requisitos previos

Para crear tablas Iceberg en el Catálogo de datos y configurar los permisos de acceso a los datos de Lake Formation, debe cumplir los siguientes requisitos:

1. Se requieren permisos para crear tablas de Iceberg sin datos registrados en Lake Formation.

   Además de los permisos necesarios para crear una tabla en el Catálogo de datos, el creador de la tabla requiere los siguientes permisos:

   - `s3:PutObject` en el recurso arn:aws:s3:::{bucketName}
   - `s3:GetObject` en el recurso arn:aws:s3:::{bucketName}
   - `s3:DeleteObject` en el recurso arn:aws:s3:::{bucketName}

2. Se requieren permisos para crear tablas de Iceberg con datos registrados en Lake Formation:

   Para utilizar Lake Formation para administrar y asegurar los datos de su lago de datos, registre su ubicación de Amazon S3 que tiene los datos de las tablas con Lake Formation. De este modo, Lake Formation puede suministrar credenciales a servicios analíticos de AWS como Athena, Redshift Spectrum y Amazon EMR para acceder a los datos. Para obtener más información sobre el registro de una ubicación de Amazon S3, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

   Una entidad principal que lee y escribe los datos subyacentes que están registrados en Lake Formation requiere los siguientes permisos:

   - `lakeformation:GetDataAccess`
   - `DATA_LOCATION_ACCESS`

     Una entidad principal que tiene permisos de localización de datos en una localización también tiene permisos de localización en todas las ubicaciones secundarias.

     Para obtener más información sobre permisos de ubicación de datos, consulte Control de acceso a los datos subyacentes.

Para habilitar la compactación, el servicio debe asumir un rol de IAM que tenga permisos para actualizar las tablas del Catálogo de datos. Para obtener más información, consulte Requisitos previos para la optimización de tablas

Creación de tablas de Iceberg

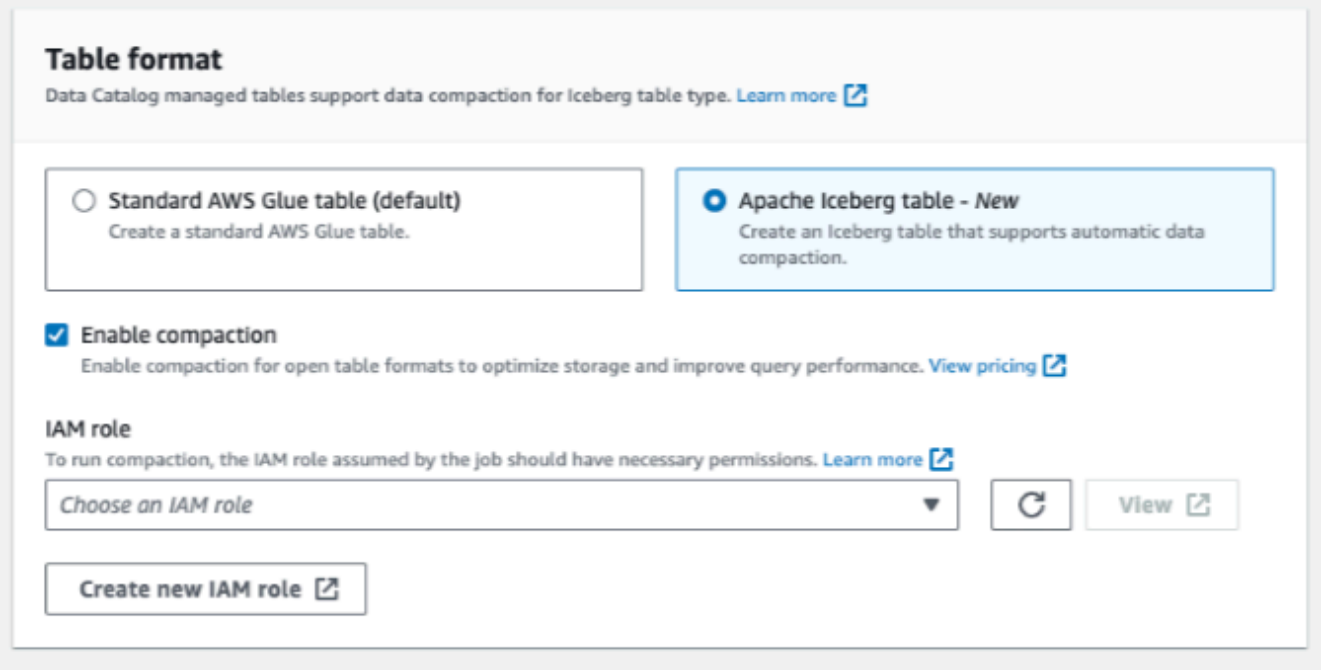
Puede crear tablas de Iceberg v1 y v2 con la consola Lake Formation o la AWS Command Line Interface tal como se documenta en esta página. También puede crear tablas de Iceberg utilizando la consola de AWS Glue o Rastreador de AWS Glue. Para más información, consulte Catálogo de datos y rastreadores en la Guía para desarrolladores de AWS Glue.

Para crear una tabla de Iceberg

Console

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

2. En Catálogo de datos, seleccione Tablas y utilice el botón Crear tabla para especificar los siguientes atributos:
   - Nombre de tabla. Escriba un nombre para la tabla. Si utiliza Athena para acceder a las tablas, utilice los consejos para nombres recogidos en la Guía del usuario de Amazon Athena.
   - Base de datos. Elija una base de datos existente o cree una nueva.
   - Descripción. Descripción de la tabla. Puede escribir una descripción para ayudarle a entender el contenido de la tabla.
   - Formato de tabla. Para el formato de la tabla, elija Apache Iceberg.



   - Habilitar la compactación. Elija Habilitar la compactación para compactar los objetos de Amazon S3 pequeños en objetos más grandes.
   - Rol de IAM. Para ejecutar la compactación, el servicio asume un rol de IAM en su nombre. Puede elegir un rol de IAM mediante el menú desplegable. Asegúrese de que el rol tenga los permisos necesarios para habilitar la compactación.

     Para obtener más información sobre los permisos necesarios, consulte Requisitos previos para la optimización de tablas .
   - Ubicación. Especifique la ruta a la carpeta de Amazon S3 que almacena la tabla de metadatos. Iceberg necesita un archivo de metadatos y una ubicación en el Catálogo de datos para poder hacer lecturas y escrituras.

- Esquema. Seleccione Agregar columnas para añadir columnas y tipos de datos de las columnas. Tiene la opción de crear una tabla vacía y actualizar el esquema más adelante. El Catálogo de datos admite los tipos de datos de Hive. Para obtener más información, consulte Tipos de datos de Hive.

  Con Iceberg podrá desarrollar el esquema y la partición después de crear la tabla. Puede utilizar consultas de Athena para actualizar el esquema de la tabla y consultas de Spark para actualizar las particiones.

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/"
           }
       }'
```

Optimización de las tablas de Iceberg

Los lagos de datos de Amazon S3 que utilizan formatos de tablas abiertas, como Apache Iceberg, almacenan los datos como objetos de Amazon S3. Tener miles de objetos pequeños de Amazon S3 en una tabla de lago de datos aumenta la sobrecarga de metadatos en las tablas Iceberg y afecta al rendimiento de lectura. Para mejorar el rendimiento de lectura de los servicios de análisis de

AWS, como Amazon Athena y Amazon EMR, y los trabajos de AWS Glue ETL, AWS Glue Data Catalog proporciona una compactación administrada (un proceso que compacta objetos pequeños de Amazon S3 para convertirlos en objetos más grandes) para las tablas de Iceberg del Catálogo de datos. Puede usar la consola de AWS Glue, la consola de Lake Formation, AWS CLI, o la API de AWS para habilitar o deshabilitar la compactación de las tablas de Iceberg individuales que están en el Catálogo de datos.

El optimizador de tablas supervisa constantemente las particiones de las tablas e inicia el proceso de compactación cuando se supera el umbral de cantidad y tamaño de los archivos. En el catálogo de datos, el valor límite predeterminado para iniciar la compactación se establece en 384 MB, mientras que en la biblioteca Iceberg el umbral de compactación es aproximadamente el 75 % del tamaño del archivo objetivo. El Catálogo de datos efectúa la compactación sin interferir con las consultas simultáneas. El catálogo de datos admite la compactación de datos solo para tablas en formato Parquet.

Para conocer los tipos de datos, los formatos de compresión y las limitaciones compatibles, consulte Formatos compatibles y limitaciones de la compactación de datos administrada .

Temas

Requisitos previos para la optimización de tablas

El optimizador de tablas asume los permisos del rol de AWS Identity and Access Management (IAM) que especifica al habilitar la compactación de una tabla. El rol de IAM debe tener los permisos para leer los datos y actualizar los metadatos en el catálogo de datos. Cree un rol de IAM y adjunte las siguientes políticas integradas:

- Agregue la siguiente política en línea que conceda a Amazon S3 permisos de lectura y escritura en la ubicación para los datos que no estén registrados en Lake Formation. Esta política también incluye permisos para actualizar la tabla en el catálogo de datos y permitir que AWS Glue agregue registros en los registros Amazon CloudWatch y publicar métricas. Para los datos de origen en

Amazon S3 que no estén registrados en Lake Formation, el acceso se determina mediante las políticas de permisos de IAM para Amazon S3 y acciones de AWS Glue.

En las siguientes políticas en línea, sustituya `bucket-name` por el nombre de su bucket de Amazon S3, `aws-account-id` y `region` por un número de cuenta y región del catálogo de datos de AWS válidos, `database_name` por el nombre de su base de datos y `table_name` por el nombre de la tabla.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::<bucket-name>/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::<bucket-name>"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "glue:UpdateTable",
                "glue:GetTable"
            ],
            "Resource": [
                "arn:aws:glue:<region>:<aws-account-id>:table/<database-name>/<table-
name>",
                "arn:aws:glue:<region>:<aws-account-id>:database/<database-name>",
                "arn:aws:glue:<region>:<aws-account-id>:catalog"
            ]
```

```
        },
        {
            "Effect": "Allow",
            "Action": [
                "logs:CreateLogGroup",
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Resource": "arn:aws:logs:<region>:<aws-account-id>:log-group:/aws-glue/
iceberg-compaction/logs:*"
        }
    ]
}
```

- Utilice la siguiente política para habilitar la compactación de los datos registrados en Lake Formation.

  Si el rol de compactación no tiene permisos de grupo `IAM_ALLOWED_PRINCIPALS` otorgados en la tabla, el rol requiere los permisos de Lake Formation `ALTER`, `DESCRIBE`, `INSERT` y `DELETE` de la tabla.

  Para obtener más información sobre cómo registrar un bucket de Amazon S3 en Lake Formation, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

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

        "Resource": [
```

```
            "arn:aws:glue:<region>:<aws-account-
id>:table/<databaseName>/<tableName>",
            "arn:aws:glue:<region>:<aws-account-id>:database/<database-name>",
            "arn:aws:glue:<region>:<aws-account-id>:catalog"
        ]
    },
    {
        "Effect": "Allow",
        "Action": [
            "logs:CreateLogGroup",
            "logs:CreateLogStream",
            "logs:PutLogEvents"
        ],
        "Resource": "arn:aws:logs:<region>:<aws-account-id>:log-group:/aws-glue/
iceberg-compaction/logs:*"
        }
    ]
 }
```

- (Opcional) Para compactar tablas Iceberg con datos de buckets de Amazon S3 cifrados mediante cifrado del lado del servidor, el rol de compactación requiere permisos para descifrar los objetos de Amazon S3 y generar una nueva clave de datos para escribir los objetos en los buckets cifrados. Agregue la siguiente política a la clave de AWS KMS deseada. Solo admitimos el cifrado a nivel de bucket.

```
{
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::<aws-account-id>:role/<compaction-role-name>"
    },
    "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
    ],
    "Resource": "*"
}
```

- (Opcional) Para la ubicación de datos registrada en Lake Formation, el rol utilizado para registrar la ubicación requiere permisos para descifrar los objetos de Amazon S3 y generar una nueva clave de datos para escribir los objetos en los buckets cifrados. Para obtener más información, consulte Registro de una ubicación cifrada de Amazon S3.

- (Opcional) Si la clave AWS KMS está almacenada en una cuenta AWS diferente, debe incluir los siguientes permisos para el rol de compactación.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:GenerateDataKey"
            ],
            "Resource": ["arn:aws:kms:<REGION>:<KEY_OWNER_ACCOUNT_ID>:key/<KEY_ID>" ]
        }
    ]
}
```

- El rol que utilice para ejecutar la compactación debe tener el permiso `iam:PassRole` correspondiente al rol.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::<account-id>:role/<compaction-role-name>"
            ]
        }
    ]
}
```

- Agregue la siguiente política de confianza al rol para que el servicio AWS Glue asuma el rol de IAM para ejecutar el proceso de compactación.

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

```
    "Statement": [
        {
            "Sid": "",
            "Effect": "Allow",
            "Principal": {
                "Service": "glue.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

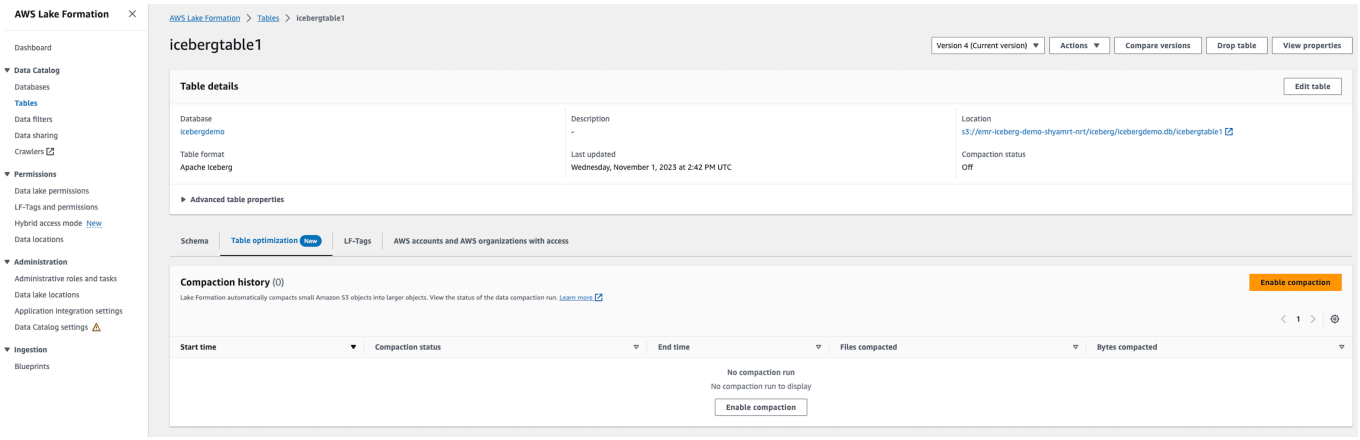
Habilitar la compactación

Puede usar la consola de Lake Formation, la consola de AWS Glue, AWS CLI, o la API de AWS para habilitar la compactación de las tablas de Apache Iceberg en el Catálogo de datos. Para las tablas nuevas, puede elegir Apache Iceberg como formato de tabla y habilitar la compactación al crear la tabla. La compactación está deshabilitada de forma predeterminada para las tablas nuevas.

Console

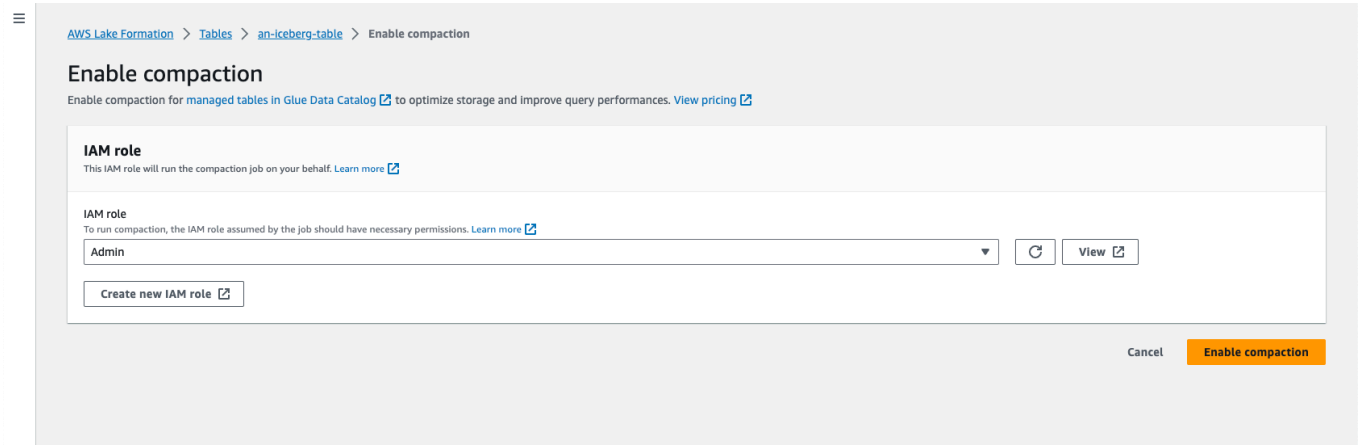
Para habilitar la compactación

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/ e inicie sesión como administrador del lago de datos, creador de la tabla o usuario al que se le hayan concedido los permisos `glue:UpdateTable` y `lakeformation:GetDataAccess` de la tabla.
2. En el panel de navegación, en Catálogo de datos, elija Tablas.
3. En la página Tablas, elija una tabla en formato de tabla abierta para la que desee activar la compactación y, a continuación, en el menú Acciones, seleccione Habilitar la compactación.
4. También puede activar la compactación seleccionando la tabla y abriendo la página de Detalles de la tabla. Seleccione la pestaña Optimización de tablas en la sección inferior de la página y elija Habilitar la compactación.



5. A continuación, seleccione un rol de IAM existente en el menú desplegable con los permisos que se muestran en la sección Requisitos previos para la optimización de tablas .

   Al elegir la opción Crear un nuevo rol de IAM, el servicio crea un rol personalizado con los permisos necesarios para ejecutar la compactación.



   Siga los pasos que se indican a continuación para actualizar un rol de IAM existente:

   a. Para actualizar la política de permisos del rol de IAM, en la consola de IAM, vaya al rol de IAM que se está utilizando para ejecutar la compactación.

   b. En la sección Agregar permisos, elija Crear política. En la ventana del navegador que se acaba de abrir, cree una nueva política para utilizarla con su rol.

   c. En la página Crear política, elija la pestaña JSON. Copie el código JSON que se muestra en la sección Requisitos previos en el campo del editor de políticas.

AWS CLI

En el ejemplo siguiente se muestra cómo habilitar la compactación. Sustituya el ID de cuenta por un ID de cuenta de AWS válido. Sustituya el nombre de la base de datos y el nombre de la tabla por el nombre real de la tabla Iceberg y el nombre de la base de datos. Sustituya el `roleArn` por el nombre de recurso de AWS (ARN) del rol de IAM y el nombre del rol de IAM que tiene los permisos necesarios para ejecutar la compactación.

```
aws glue create-table-optimizer \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --table-optimizer-configuration
 '{"roleArn":"arn:aws:iam::123456789012:role/compaction_role", "enabled":'true'}' \
  --type compaction
```

AWS API

Llame a la operación `CreateTableOptimizer` para habilitar la compactación de una tabla.

Después de activar la compactación, la pestaña de Optimización de la tabla muestra los siguientes detalles de compactación (después de aproximadamente 15 a 20 minutos):

Hora de inicio

Hora a la que se inició el proceso de compactación en Lake Formation. El valor es una marca en la hora UTC.

Hora de finalización

Hora a la que terminó el proceso de compactación en el Catálogo de datos. El valor es una marca en la hora UTC.

Status

Estado del ciclo de compactación. Los valores indican éxito o fracaso.

Archivos compactados

Número de archivos compactados.

Bytes compactados

Número de bytes compactados.

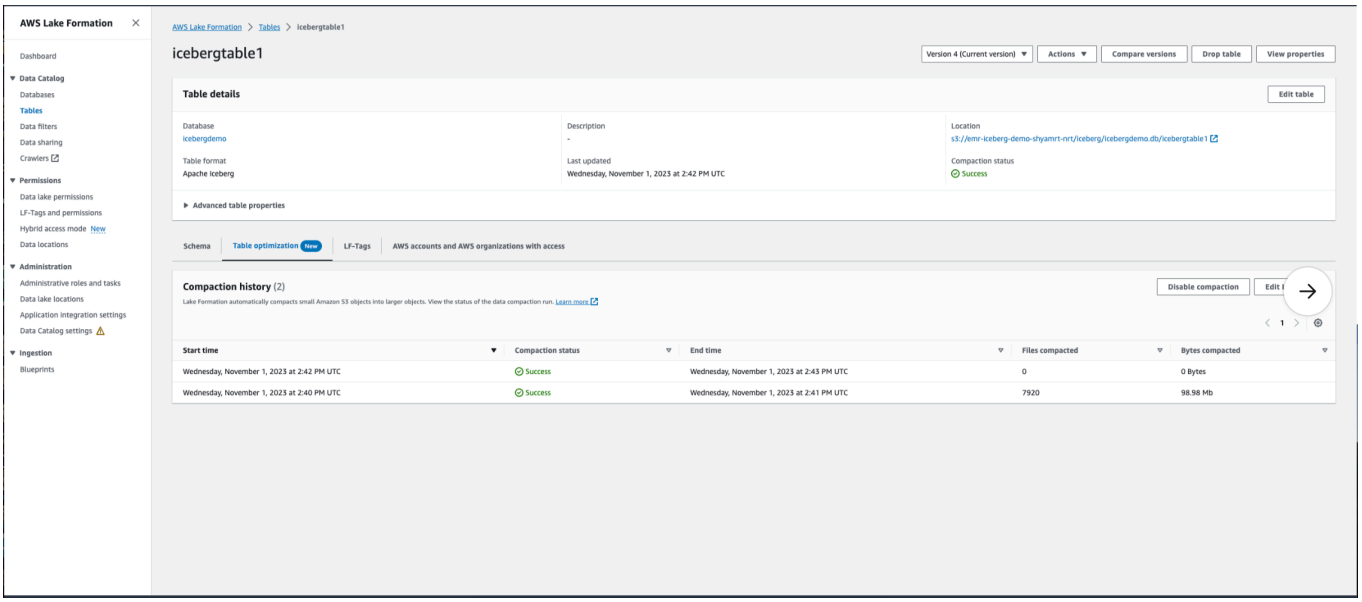
Deshabilitación de la compactación

Puede deshabilitar la compactación automática de una tabla Apache Iceberg concreta mediante la consola AWS Glue o AWS CLI.

Console

1. Elija la Base de datos y las Tablas. En la lista de tablas, elija la tabla en formato de tabla abierta en la que desee deshabilitar la compactación.
2. Puede elegir una tabla Iceberg y elegir Desactivar la compactación en Acciones.

   También puede deshabilitar la compactación de la tabla seleccionando Desactivar la compactación en la sección inferior de la página de Detalles de las tablas.



3. Selecciona Desactivar la compactación en el mensaje de confirmación. Puede volver a habilitar la compactación más adelante.

   Tras la confirmación, la compactación se desactiva y el estado de compactación de la tabla vuelve a ser el siguiente `Off`.

AWS CLI

En el siguiente ejemplo, reemplace el ID de cuenta con un ID de AWS válido. Sustituya el nombre de la base de datos y el nombre de la tabla por el nombre real de la tabla Iceberg y el nombre de la base de datos. Sustituya el `roleArn` por el nombre de recurso de AWS (ARN) del rol de IAM y el nombre real del rol de IAM que tiene los permisos necesarios para ejecutar la compactación.

```
aws glue update-table-optimizer \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --table-optimizer-configuration
 '{"roleArn":"arn:aws:iam::123456789012:role/compaction_role", "enabled":'false'}'\
  --type compaction
```

AWS API

UpdateTableOptimizer Operación de llamada para deshabilitar la compactación de una tabla específica.

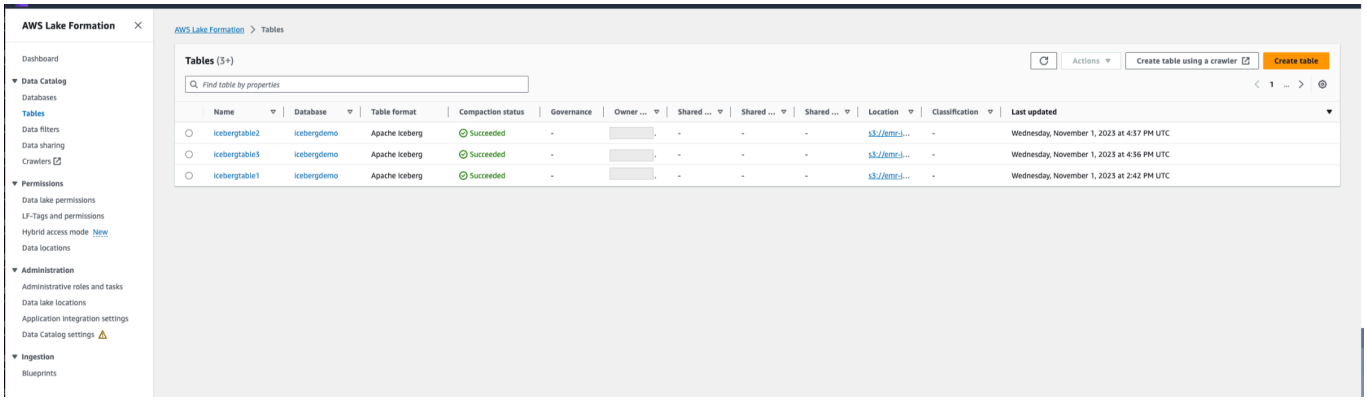
Consultar los detalles de compactación

Puede ver el estado de compactación de Apache Iceberg mediante la consola de AWS CLI o mediante las operaciones de la API de AWS.

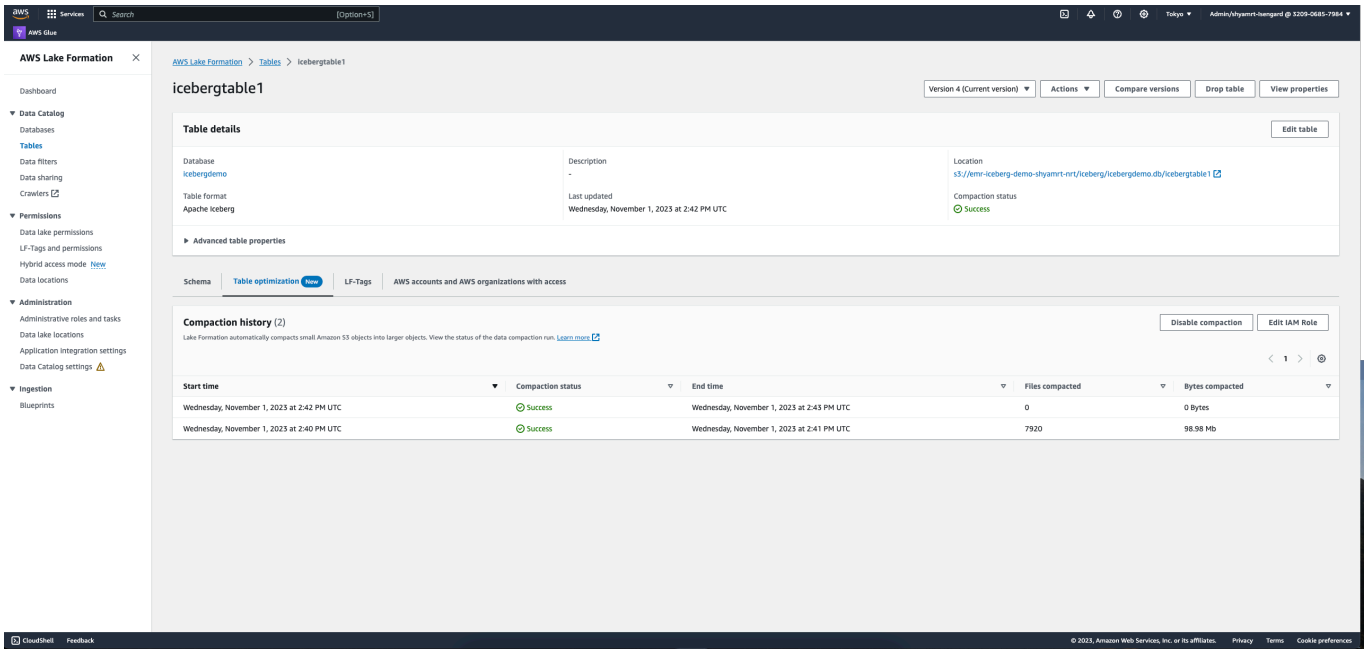
Console

Para ver el estado de compactación de las tablas Iceberg (consola)

- Puede ver el estado de compactación de las tablas de Iceberg en la consola de Lake Formation seleccionando Tablas en Catálogo de datos. El campo Estado de compactación muestra el estado de la operación de compactación. Puede mostrar el formato de la tabla y el estado de compactación mediante las preferencias de la tabla.



- Para ver el historial de compactación de una tabla específica, seleccione Tablas en AWS Glue Data Catalog y elija una tabla para ver los detalles de la tabla. La pestaña Optimización de la tabla muestra el historial de compactación de la tabla.



AWS CLI

Puede ver los detalles de compactación utilizando AWS CLI.

En los siguientes ejemplos, sustituya el identificador de cuenta por un identificador de cuenta de AWS válido, el nombre de la base de datos y el nombre de la tabla por el nombre real de la tabla de Iceberg.

- Para obtener los detalles de la última tanda de compactación de una tabla

```
aws get-table-optimizer \
```

```
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --type compaction
```

- Utilice el siguiente ejemplo para recuperar el historial de un optimizador de una tabla específica.

```
aws list-table-optimizer-runs \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --type compaction
```

- En el siguiente ejemplo se muestra cómo recuperar la ejecución de compactación y los detalles de configuración de varios optimizadores. Puede especificar un máximo de 20 optimizadores.

```
aws glue batch-get-table-optimizer \
--entries '[{"catalogId":"123456789012", "databaseName":"iceberg_db",
 "tableName":"iceberg_table", "type":"compaction"}]'
```

AWS API

- Utilice la operación `GetTableOptimizer` para recuperar los detalles de la última ejecución de un optimizador.
- Utilice la operación `ListTableOptimizerRuns` para recuperar el historial de un optimizador determinado en una tabla específica. Puede especificar 20 optimizadores en una sola llamada a la API.
- Utilice la operación `BatchGetTableOptimizer` para recuperar los detalles de configuración de varios optimizadores de su cuenta. Esta operación no permite hacer llamadas entre cuentas.

Visualización de métricas de Amazon CloudWatch

Tras ejecutar la compactación correctamente, el servicio crea métricas Amazon CloudWatch sobre el rendimiento del trabajo de compactación. Puedes ir a las CloudWatch métricas y elegir Métricas, Todas las métricas. Puede filtrar las métricas por el espacio de nombres específico (por ejemplo AWS Glue), el nombre de la tabla o el nombre de la base de datos.

Para obtener más información, consulte Ver métricas disponibles en la Guía del usuario de Amazon CloudWatch.

- Número de bytes compactados
- Número de archivos compactados
- Número de DPU asignadas a los trabajos
- Duración del trabajo (horas)

Eliminar un optimizador

Puede eliminar un optimizador y los metadatos asociados a la tabla mediante AWS CLI o una operación de API de AWS.

Ejecute el siguiente comando AWS CLI para eliminar el historial de compactación de una tabla.

```
aws glue delete-table-optimizer \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --type compaction
```

Utilice la operación `DeleteTableOptimizer` para eliminar un optimizador de una tabla.

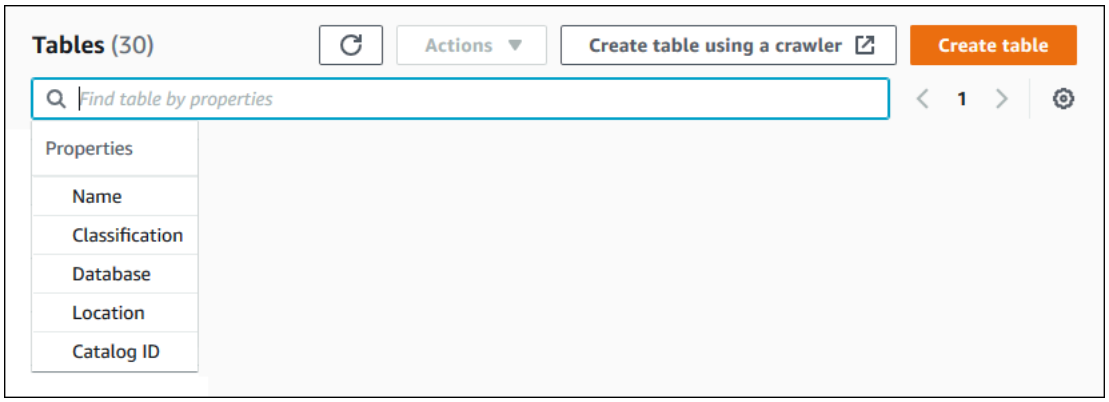
Búsqueda de tablas

Use la consola de AWS Lake Formation para buscar tablas del Catálogo de datos por nombre, ubicación, base de datos contenedora y más. Los resultados de la búsqueda muestran solo las tablas en las que tiene permisos de Lake Formation.

Para buscar tablas (consola)

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, elija Tablas.
3. Coloque el cursor en el campo de búsqueda en la parte superior de la página. El campo tiene el texto marcador de posición Buscar tabla por propiedades.

Aparece el menú Propiedades, que muestra las distintas propiedades de la tabla por las que buscar.



4. Realice una de las acciones siguientes:

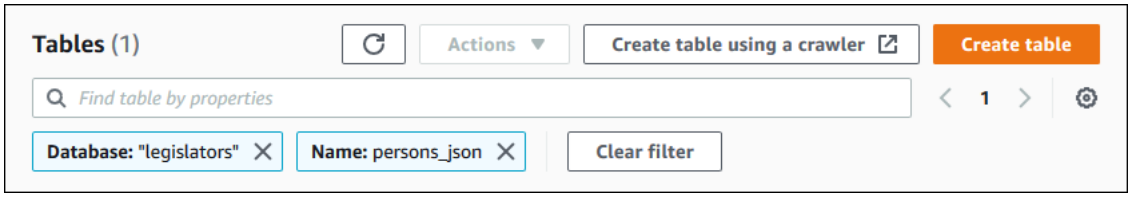
- Buscar por base de datos contenedora.

  1. Seleccione Base de datos en el menú Propiedades y, a continuación, elija una base de datos en el menú Bases de datos que aparece o escriba un nombre de base de datos y pulse Intro.

     Se muestran las tablas sobre las que tiene permisos en la base de datos.

  2. (Opcional) Para reducir la lista a una sola tabla de la base de datos, vuelva a colocar el cursor en el campo de búsqueda, elija Nombre en el menú Propiedades y elija un nombre de tabla en el menú Tablas que aparece o escriba un nombre de tabla y pulse Intro.

     Aparece la tabla individual y tanto el nombre de la base de datos como el nombre de la tabla aparecen como mosaicos debajo del campo de búsqueda.



     Para ajustar el filtro, cierre cualquiera de los mosaicos o seleccione Borrar filtro.

- Busque por otras propiedades.

  1. Seleccione una propiedad de búsqueda en el menú Propiedades.

     Para buscar por ID de cuenta AWS, seleccione ID de catálogo en el menú Propiedades, introduzca un ID de cuenta AWS válido (por ejemplo, 111122223333) y pulse Intro.

Para buscar por ubicación, elija Ubicación en el menú Propiedades y seleccione una en el menú Ubicaciones que aparece. Se devuelven todas las tablas de la ubicación raíz de la ubicación seleccionada (por ejemplo, Amazon S3).

Uso compartido de tablas y bases de datos del Catálogo de datos entre cuentas AWS

Puede compartir los recursos del Catálogo de datos (bases de datos y tablas) con AWS cuentas externas concediendo permisos de Lake Formation sobre los recursos a las cuentas externas. A continuación, los usuarios pueden ejecutar consultas y trabajos para unir y consultar tablas en varias cuentas. Con algunas restricciones, cuando comparte un recurso del Catálogo de datos con otra cuenta, las entidades principales de esa cuenta pueden operar con ese recurso como si estuviera en su Catálogo de datos.

Los recursos no se comparten con entidades principales específicas en cuentas AWS externas, sino que se comparten los recursos con una cuenta AWS u organización. Cuando comparte un recurso con una organización AWS, está compartiendo el recurso con todas las cuentas de todos los niveles de esa organización. A continuación, el administrador del lago de datos de cada cuenta externa debe conceder permisos sobre los recursos compartidos a las entidades principales de su cuenta.

Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation y Concesión y revocación de permisos sobre los recursos del catálogo de datos.

Consulte también:

- Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos
- Requisitos previos

Uso de vistas

Esta característica está en versión preliminar y está sujeta a cambios. Para obtener más informaci ón, consulte la sección Betas y versiones preliminares del documento Términos de servicio de AWS.

En AWS Glue Data Catalog, una vista es una tabla virtual en la que el contenido se define mediante una consulta que hace referencia a una o más tablas. Puede crear una vista que haga referencia a un máximo de 10 tablas mediante editores de SQL para Amazon Athena, Amazon Redshift o Amazon EMR. Las tablas de referencia subyacentes de una vista pueden pertenecer a la misma base de datos o a bases de datos distintas dentro de la misma Cuenta de AWS.

SQL es un lenguaje de programación que se utiliza para consultar tablas y cada motor analítico de AWS utiliza su propia variante de SQL, o dialecto de SQL. El catálogo de datos permite crear vistas con distintos dialectos de SQL siempre que cada dialecto haga referencia al mismo conjunto de tablas, columnas y tipos de datos. Al definir un esquema de vista y un objeto de metadatos comunes que puede consultar desde varios motores, las vistas del catálogo de datos le permiten utilizar vistas uniformes de todo el lago de datos.

Cuando administra vistas en el catálogo de datos, puede utilizar AWS Lake Formation para conceder permisos detallados mediante el método de recurso con nombre o mediante etiquetas LF y compartirlos entreCuentas de AWS, organizaciones AWS y unidades organizativas. También puede compartir vistas del catálogo de datos entre Regiones de AWS. Esto permite a los usuarios proporcionar acceso a los datos entre Regiones de AWS sin duplicar el origen de datos.

Para obtener más información el uso compartido de datos entre cuentas y el acceso de datos entre regiones, consulte:

- Compartir datos entre cuentas en Lake Formation
- Acceso a las tablas entre regiones

Puede utilizar las vistas del catálogo de datos para:

- Crear y administrar permisos en un esquema de vista única. Esto le ayuda a evitar el riesgo de que se existan permisos incoherentes en vistas duplicadas creadas en varios motores.
- Conceda permisos a los usuarios en una vista que haga referencia a varias tablas sin conceder permisos directamente en las tablas de referencia subyacentes.

Para conocer las limitaciones, consulte Vistas, consideraciones y limitaciones del catálogo de datos

Temas

- Requisitos previos para crear vistas
- Creación de vistas

- Concesión de permisos del vistas del catálogo de datos

Requisitos previos para crear vistas

- Para crear vistas en el catálogo de datos, debe registrar las ubicaciones de datos subyacentes de Amazon S3 de las tablas de referencia en Lake Formation.

  Para obtener más información sobre el registro de datos con Lake Formation, consulteAñadir una ubicación de Amazon S3 a su lago de datos.
- El definidor de vistas debe ser un rol de IAM. Otras identidades de IAM no pueden crear vistas del catálogo de datos.
- El rol de IAM que defina la vista debe tener los siguientes permisos:
  - Permiso `SELECT` de Lake Formation completo con opción `Grantable` en todas las tablas de referencia.
  - Una política de confianza para que Lake Formation y los servicios de AWS Glue asuman el rol.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "DataCatalogViewDefinerAssumeRole1",
            "Effect": "Allow",
            "Principal": {
               "Service": [
                    "glue.amazonaws.com",
                    "lakeformation.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

  - El permiso iam:PassRole para AWS Glue y Lake Formation.

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

```
            "Sid": "DataCatalogViewDefinerPassRole1",
            "Action": [
                "iam:PassRole"
            ],
            "Effect": "Allow",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": [
                        "glue.amazonaws.com",
                        "lakeformation.amazonaws.com"
                      ]
                }
            }
        }
    ]
}
```

- Permisos de AWS Glue y Lake Formation

```
{
    "Version": "2012-10-17",
                 "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "Glue:GetDatabase",
                "Glue:GetDatabases",
                "Glue:CreateTable",
                "Glue:GetTable",
                "Glue:UpdateTable",
                "Glue:DeleteTable",
                "Glue:GetTables",
                "Glue:SearchTables",
                "Glue:BatchGetPartition",
                "Glue:GetPartitions",
                "Glue:GetPartition",
                "Glue:GetTableVersion",
                "Glue:GetTableVersions",
                "lakeFormation:GetDataAccess",
                "lakeFormation:GetTemporaryTableCredentials",
                "lakeFormation:GetTemporaryGlueTableCredentials",
                "lakeFormation:GetTemporaryUserCredentialsWithSAML"
            ],
```

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

- No puede crear vistas si la base de datos en la que se va a crear la vista tiene el permiso `Super` o `ALL` otorgado al mismo grupo de `IAMAllowedPrincipals`. Para revocar el permiso `Super` de un grupo de `IAMAllowedPrincipals` en una base de datos, consulte Paso 4: Cambiar sus almacenes de datos al modelo de permisos de Lake Formation.

  Si la configuración existente del lago de datos no le permite ajustar `CreateTableDefaultPermissions` en vacío para el grupo `IAMAllowedPrincipals`, puede crear una nueva base de datos y codificar el ajuste del lago de datos con la siguiente estructura.

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ],
        CreateTableDefaultPermissions": [
            {
                "Principal": {
                    "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
                },
                "Permissions": []
            }
        ]
 }
```

Creación de vistas

Puede usar editores de SQL para Athena, Amazon Redshift o Amazon EMR para crear vistas en AWS Glue Data Catalog.

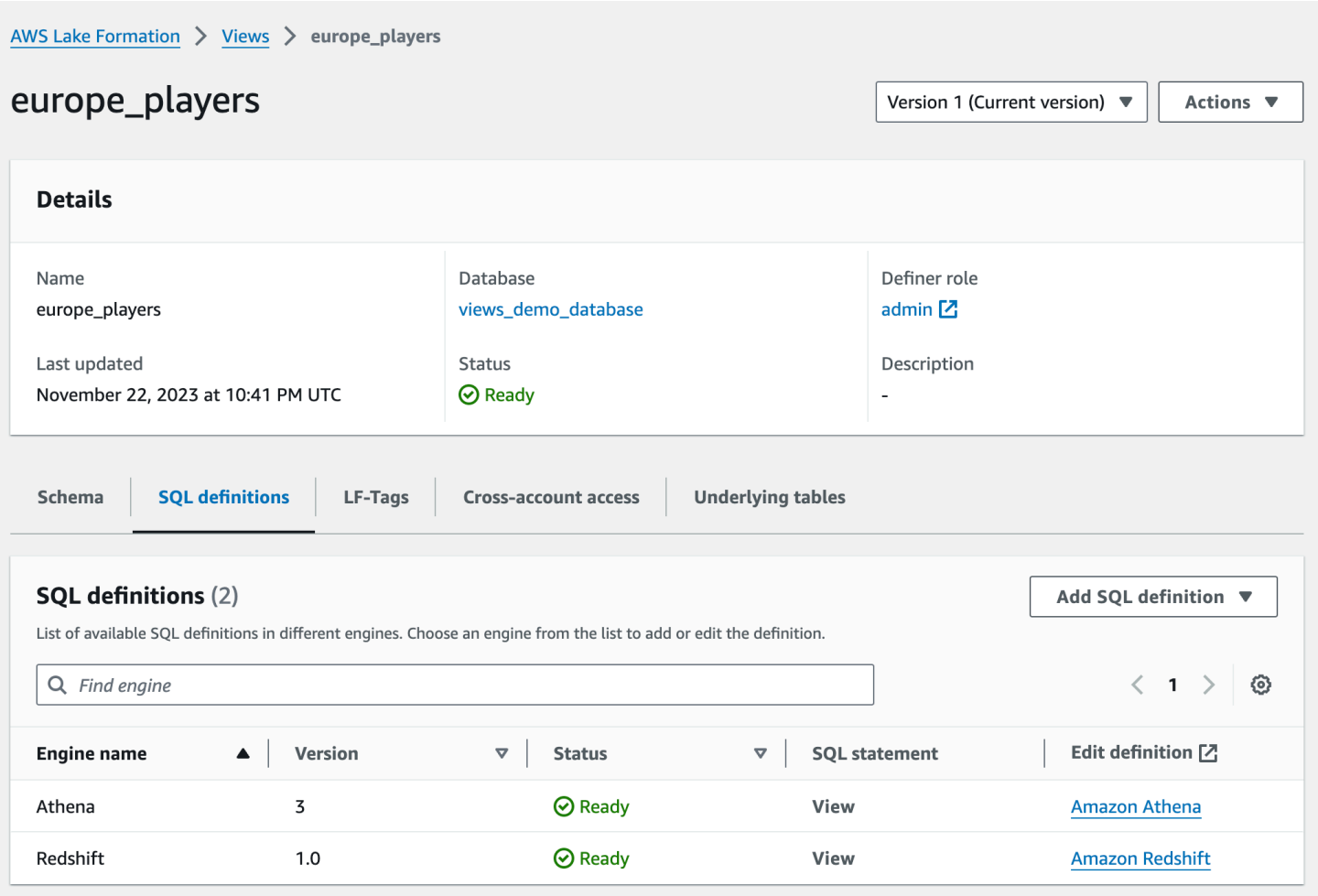
Para obtener más información sobre la sintaxis para crear y administrar vistas del catálogo de datos, consulte:

- Uso de las vistas de AWS Glue Data Catalog de la Guía del usuario de Amazon Athena.

- Creación de vistas en AWS Glue Data Catalog en la Guía para desarrolladores de bases de datos de Amazon Redshift.
- Trabajo con vistas de AWS Glue Data Catalog en la Guía de administración de Amazon EMR.

Tras crear una vista del catálogo de datos, los detalles de la vista en la consola de Lake Formation.

1. Seleccione Vistas en el catálogo de datos en la consola de Lake Formation.
2. Aparece una lista de las vistas disponibles en la página de vistas.
3. Seleccione una vista de la lista y la página de detalles mostrará los atributos de la vista.



Esquema

Elija una fila `Column` y seleccione Editar etiquetas LF para actualizar los valores de las etiquetas o asignar nuevas etiquetas LF.

Definiciones de SQL

Puede ver una lista completa de las definiciones de SQL disponibles. Seleccione Añadir definición de SQL y elija un motor de consultas para añadir una definición de SQL. Elija un motor de consultas (Athena o Amazon Redshift) en la columna `Edit definition` para actualizar una definición de SQL.

Etiquetas LF

Seleccione Editar etiquetas LF para editar los valores de una etiqueta o asignar etiquetas nuevas. Puede utilizar etiquetas LF para conceder permisos sobre las vistas.

Acceso entre cuentas

Puede ver una lista de Cuentas de AWS, organizaciones y unidades organizativas (OU) con las que ha compartido la vista del catálogo de datos.

Tablas subyacentes

Las tablas subyacentes a las que se hace referencia en la definición de SQL utilizadas para crear la vista se muestran en esta pestaña.

Concesión de permisos del vistas del catálogo de datos

Tras crear vistas, puede conceder permisos de lago de datos en las vistas a las entidades principales entre Cuentas de AWS, organizaciones y unidades organizativas. Para obtener más información sobre cómo conceder permisos, consulte Concesión de permisos para vistas mediante el método de recursos con nombre .

Importación de datos mediante flujos de trabajo en Lake Formation

Con AWS Lake Formation, puede importar sus datos mediante flujos de trabajo. Un flujo de trabajo define el origen de datos y la programación para importar los datos a su lago de datos. Es un contenedor para rastreadores AWS Glue, trabajos y desencadenadores que se utilizan para orquestar los procesos de carga y actualización del lago de datos.

Temas

- Esquemas y flujos de trabajo en Lake Formation
- Creación de un flujo de trabajo
- Ejecución de un flujo de trabajo

Esquemas y flujos de trabajo en Lake Formation

Un flujo de trabajo encapsula una actividad compleja de extracción, transformación y carga (ETL) de múltiples tareas. Los flujos de trabajo generan rastreadores de AWS Glue, trabajos y desencadenadores para orquestar la carga y actualización de datos. Lake Formation ejecuta y rastrea un flujo de trabajo como una única entidad. Puede configurar un flujo de trabajo para que se ejecute bajo demanda o de forma programada.

Los flujos de trabajo que cree en Lake Formation son visibles en la consola de AWS Glue como un gráfico acíclico dirigido (DAG). Cada nodo del DAG es una tarea, un rastreador o un disparador. Para supervisar el progreso y solucionar problemas, puede hacer un seguimiento del estado de cada nodo del flujo de trabajo.

Cuando se completa un flujo de trabajo de Lake Formation, el usuario que lo ejecutó recibe el permiso `SELECT` de Lake Formation en las tablas del catálogo de datos que crea el flujo de trabajo.

También puede crear flujos de trabajo en AWS Glue. Sin embargo, como Lake Formation le permite crear un flujo de trabajo a partir de un esquema, la creación de flujos de trabajo es mucho más sencilla y automatizada en Lake Formation. Lake Formation proporciona los siguientes tipos de esquemas:

- Instantánea de la base de datos. Carga o recarga los datos de todas las tablas en el lago de datos desde una fuente JDBC. Puede excluir algunos datos de la fuente en función de un patrón de exclusión.
- Base de datos incremental. Carga solo los datos nuevos en el lago de datos desde una fuente JDBC, en función de los marcadores establecidos anteriormente. El usuario especifica las tablas individuales de la base de datos de origen de JDBC que desee incluir. Para cada tabla, elige las columnas de marcadores y el orden de clasificación de los marcadores para hacer un seguimiento de los datos que se han cargado previamente. La primera vez que ejecuta un esquema incremental de base de datos sobre un conjunto de tablas, el flujo de trabajo carga todos los datos de las tablas y establece los marcadores para la siguiente ejecución del esquema incremental de base de datos. Por lo tanto, puede utilizar un esquema de base de datos incremental en lugar del esquema de instantánea de base de datos para cargar todos los datos, siempre que especifique cada tabla de los orígenes de datos como parámetro.
- Archivo de registro. Carga de forma masiva datos de fuentes de archivos de registro, incluidos AWS CloudTrail, registros de Equilibrador de carga elástico y registros de Equilibrador de carga de aplicación

Utilice la siguiente tabla como ayuda para decidir si debe utilizar una instantánea de base de datos o un esquema incremental de base de datos.

Utilice la instantánea de la base de datos cuando...	Utilice la base de datos incremental cuando...
• La evolución del esquema es flexible. (Se cambia el nombre de las columnas, se eliminan las columnas anteriores y se añaden nuevas columnas en su lugar). • Se necesita una coherencia total entre el origen y el destino.	• La evolución del esquema es incremental. (Solo hay adición sucesiva de columnas). • Solo se añaden nuevas filas; las filas anteriores no se actualizan.

> **Note**
>
> Los usuarios no pueden editar los esquemas y flujos de trabajo creados por Lake Formation.

Creación de un flujo de trabajo

Antes de empezar, asegúrese de que ha concedido los permisos de datos y de ubicación de datos necesarios al rol `LakeFormationWorkflowRole`. Esto es para que el flujo de trabajo pueda crear tablas de metadatos en el catálogo de datos y escribir datos en ubicaciones de destino en Amazon S3. Para obtener más información, consulte (Opcional) Cree un rol de IAM para los flujos de trabajo y Descripción general de los permisos de Lake Formation .

Para crear un flujo de trabajo a partir de un esquema

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con permisos de ingeniero de datos. Para obtener más información, consulte Personas de Lake Formation y referencia de permisos IAM.
2. En el panel de navegación, seleccione Esquemas y, a continuación, seleccione Utilizar esquema.
3. En la página Usar un esquema, elija un mosaico para seleccionar el tipo de esquema.
4. En Origen de importación, especifique el origen de datos.

Si está importando desde un origen JDBC, especifique lo siguiente:

- Conexión a la base de datos. Elija una conexión de la lista. Cree conexiones adicionales utilizando la consola de AWS Glue. El nombre de usuario JDBC y la contraseña de la conexión determinan los objetos de la base de datos a los que tiene acceso el flujo de trabajo.
- Ruta de origen de los datos. Introduzca *<base de datos>*/*<esquema>*/*<tabla>* o *<base de datos>*/*<tabla>*, en función del producto de base de datos. Oracle Database y MySQL no permiten utilizar un esquema en la ruta. Puede sustituir *<esquema>* o *<tabla>* por el carácter de porcentaje (%). Por ejemplo, para una base de datos Oracle con un identificador del sistema (SID) de `orcl`, introduzca `orcl/%` para importar todas las tablas a las que tenga acceso el usuario nombrado en la conexión.

  > **Important**
  >
  > Este campo distingue entre mayúsculas y minúsculas. El flujo de trabajo fallará si no coincide entre mayúsculas y minúsculas en alguno de los componentes.

  Si especifica una base de datos MySQL, AWS Glue ETL utiliza el controlador JDBC Mysql5 de forma predeterminada, por lo que MySQL8 no es compatible de forma nativa. Puede editar el script del trabajo ETL para utilizar un parámetro `customJdbcDriverS3Path` como el descrito en Valores de tipo de conexión JDBC en la Guía para desarrolladores de AWS Glue para utilizar un controlador JDBC diferente que sea compatible con MySQL8.

Si está importando desde un archivo de registro, asegúrese de que el rol que especifique para el flujo de trabajo (el "rol del flujo de trabajo") tenga los permisos IAM necesarios para acceder a los orígenes de datos. Por ejemplo, para importar registros de AWS CloudTrail, el usuario debe tener los permisos `cloudtrail:DescribeTrails` y `cloudtrail:LookupEvents` para ver la lista de registros de CloudTrail al crear el flujo de trabajo, y el rol del flujo de trabajo debe tener permisos sobre la ubicación de CloudTrail en Amazon S3.

5. Haga una de las siguientes acciones:

   - Para el tipo de esquema Instantánea de base de datos, identifique de forma opcional un subconjunto de datos a importar especificando uno o varios patrones de exclusión. Estos patrones de exclusión son patrones `glob` de estilo Unix. Se almacenan como una propiedad de las tablas creadas por el flujo de trabajo.

Para obtener más información sobre los patrones de exclusión disponibles, consulte Incluir y excluir patrones en la Guía para desarrolladores de AWS Glue.

- Para el tipo de esquema de base de datos incremental, especifique los siguientes campos. Agregue una fila para cada tabla que desee importar.

  Nombre de la tabla

  Tabla que se va a importar. Debe estar en minúscula.

  Claves de marcadores

  Lista de nombres de columnas delimitados por comas que definen las claves de los marcadores. Si está en blanco, la clave principal se utiliza para determinar los datos nuevos. Las mayúsculas y minúsculas de cada columna deben coincidir con las definidas en los orígenes de datos.

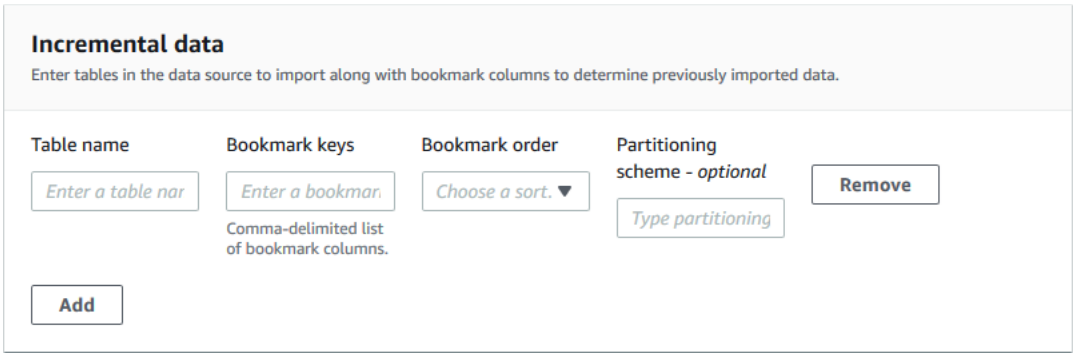
  > **Note**
  >
  > La clave principal se califica como clave marcadora predeterminada solo si es secuencialmente creciente o decreciente (sin huecos). Si desea utilizar la clave principal como clave de marcador y tiene huecos, debe nombrar la columna de clave principal como clave de marcador.

  Orden de los marcadores

  Si elige Ascendente, las filas con valores superiores a los marcados se identifican como nuevas filas. Si elige Descendente, las filas con valores inferiores a los marcados se identifican como nuevas filas.

  Esquema de partición

  (Opcional) Lista de columnas de claves de partición, delimitadas por barras (/). Ejemplo: `year/month/day`.



Para más información, consulte Seguimiento de los datos procesados mediante marcadores de trabajo en la Guía para desarrolladores de AWS Glue.

6. En Importar destino, especifique la base de datos de destino, la ubicación de Amazon S3 de destino y el formato de datos.

   Asegúrese de que el rol de flujo de trabajo tenga los permisos de Lake Formation necesarios en la base de datos y en la ubicación de destino de Amazon S3.

   > **Note**
   >
   > Actualmente, los esquemas no admiten el cifrado de datos en el destino.

7. Elija una frecuencia de importación.

   Puede especificar una expresión `cron` con la opción Personalizada.

8. En Opciones de importación:

   a. Introduzca un nombre de flujo de trabajo.

   b. Para el rol, elija el rol `LakeFormationWorkflowRole`, que creó en (Opcional) Cree un rol de IAM para los flujos de trabajo.

   c. Si lo desea, especifique un prefijo de tabla. El prefijo se antepone a los nombres de las tablas del catálogo de datos que crea el flujo de trabajo.

9. Seleccione Crear y espere a que la consola informe de que el flujo de trabajo se ha creado correctamente.

   > **Tip**
   >
   > ¿Ha recibido este mensaje de error?

`User: arn:aws:iam::<account-id>:user/<username> is not authorized to perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/<rolename>...`
En caso afirmativo, compruebe que ha sustituido *<id-cuenta>* por un número de cuenta AWS válido en todas las pólizas.

Véase también:

- Esquemas y flujos de trabajo en Lake Formation

Ejecución de un flujo de trabajo

Puede ejecutar un flujo de trabajo desde la consola de Lake Formation, la consola AWS Glue, la interfaz de línea de comandos de AWS Glue (AWS CLI) o la API.

Para ejecutar un flujo de trabajo (consola de Lake Formation)

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como usuario con permisos de ingeniero de datos. Para obtener más información, consulte Personas de Lake Formation y referencia de permisos IAM.
2. En el panel de navegación, elija Blueprints (Esquemas).
3. En la página Esquemas, seleccione el flujo de trabajo. Después, en el menú Acciones, seleccione Comenzar.
4. Conforme se ejecuta el flujo de trabajo, puede ver su progreso en la columna Estado de la última ejecución. Pulse el botón de actualización de vez en cuando.

   El estado pasa de EN EJECUCIÓN a Detectando, Importando y FINALIZADO.

   Cuando finalice el flujo de trabajo:

   - El catálogo de datos tendrá nuevas tablas de metadatos.
   - Sus datos se incorporan al lago de datos.

   Si el flujo de trabajo falla, haga lo siguiente:

a. Seleccione el flujo de trabajo. Elija Acciones y, a continuación, elija Ver gráfico.

   El flujo de trabajo se abre en la consola de AWS Glue.

b. Asegúrese de que se seleccione el flujo de trabajo y elija la pestaña History (Historial).

c. En Historial, seleccione la ejecución más reciente y seleccione Ver detalles de la ejecución.

d. Seleccione un trabajo o un rastreador fallidos en el gráfico dinámico (tiempo de ejecución) y revise el mensaje de error. Los nodos con errores aparecen en rojo o amarillo.

Véase también:

- Esquemas y flujos de trabajo en Lake Formation

Administrar los permisos de Lake Formation

Lake Formation proporciona controles de acceso centralizados para los datos de su lago de datos. Puede definir reglas basadas en políticas de seguridad para sus usuarios y aplicaciones por roles en Lake Formation, y la integración con AWS Identity and Access Management autentica a esos usuarios y roles. Una vez definidas las reglas, Lake Formation aplica sus controles de acceso a nivel de tabla y columna para los usuarios de Amazon Redshift Spectrum y Amazon Athena.

Temas

Conceder permisos de ubicación de datos

Los permisos de ubicación de datos en AWS Lake Formation permiten a las entidades principales crear y alterar recursos del catálogo de datos que apunten a ubicaciones registradas designadas de Amazon S3. Los permisos de ubicación de datos funcionan con los permisos de datos de Lake Formation para asegurar la información en su lago de datos.

Lake Formation no utiliza el servicio AWS Resource Access Manager (AWS RAM) para las concesiones de permisos de localización de datos, por lo que no necesita aceptar invitaciones para compartir recursos para los permisos de localización de datos.

Puede conceder permisos de localización de datos utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

> **Note**
>
> Para que una concesión tenga éxito, primero debe registrar la ubicación de los datos en Lake Formation.

> **Consulte también:**
>
> - Underlying data access control

Temas

Concesión de permisos de localización de datos (misma cuenta)

Siga estos pasos para conceder permisos de ubicación de datos a las entidades principales de su cuenta AWS. Puede conceder permisos utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

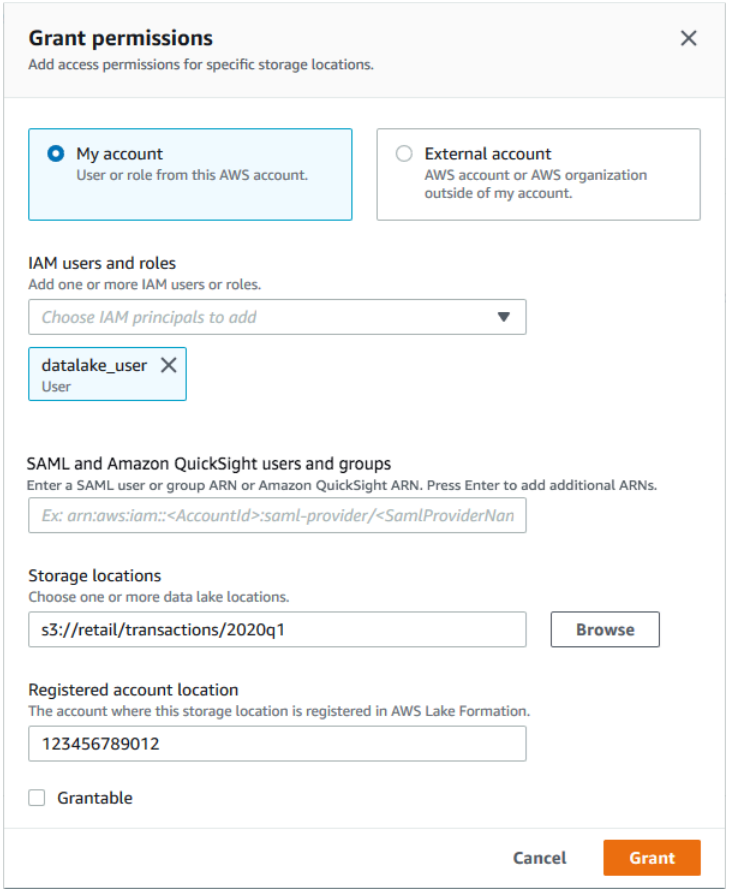
Para conceder permisos de ubicación de datos (misma cuenta, consola)

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos o como una entidad principal que dispone de permisos de concesión en la ubicación de datos deseada.
2. En el panel de navegación, seleccione Ubicaciones de datos.
3. Elija Conceder.
4. En el cuadro de diálogo Conceder permisos, asegúrese de que el icono Mi cuenta esté seleccionado. A continuación, facilite la siguiente información:

   - Para Usuarios y roles de IAM, elija una o varias entidades principales.
   - Para los usuarios y grupos de SAML y Amazon QuickSight, introduzca uno o varios nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de SAML o ARN para los usuarios o grupos de Amazon QuickSight.

Introduzca un ARN cada vez y pulse Intro después de cada uno. Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

- Para las ubicaciones de almacenamiento, elija Examinar y busque una ubicación de Amazon Simple Storage Service (Amazon S3). La ubicación debe estar registrada en Lake Formation. Vuelva a seleccionar Examinar para añadir otra ubicación. También puede escribir la ubicación, pero asegúrese de anteponerle `s3://`.
- En Ubicación de la cuenta registrada, introduzca el ID de la cuenta de AWS donde está registrada la ubicación. De forma predeterminada, es su ID de cuenta. En un escenario entre cuentas, los administradores del lago de datos de una cuenta destinataria pueden especificar aquí la cuenta propietaria al conceder el permiso de ubicación de datos a otras entidades principales de la cuenta destinataria.
- (Opcional) Para permitir que las entidades principales seleccionadas concedan permisos de ubicación de datos en la ubicación seleccionada, seleccione Concedible.



5. Elija Conceder.

Para conceder permisos de ubicación de datos (misma cuenta, AWS CLI)

- Ejecute un comando `grant-permissions` y conceda `DATA_LOCATION_ACCESS` a la entidad principal, especificando la ruta de Amazon S3 como recurso.

  Example

  En el siguiente ejemplo se conceden permisos de ubicación de datos en `s3://retail` al usuario `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

  En el siguiente ejemplo se conceden permisos de ubicación de datos en `s3://retail` al grupo `ALLIAMPrincipals`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "111122223333"}}'
```

> Consulte también:
>
> - Referencia de permisos de Lake Formation

Concesión de permisos de ubicación de datos (cuenta externa)

Siga estos pasos para conceder permisos de ubicación de datos a una cuenta AWS u organización externa.

Puede conceder permisos utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Antes de empezar

Asegúrese de que se cumplan todos los requisitos previos de acceso entre cuentas. Para obtener más información, consulte Requisitos previos.

Para conceder permisos de ubicación de datos (cuenta externa, consola)

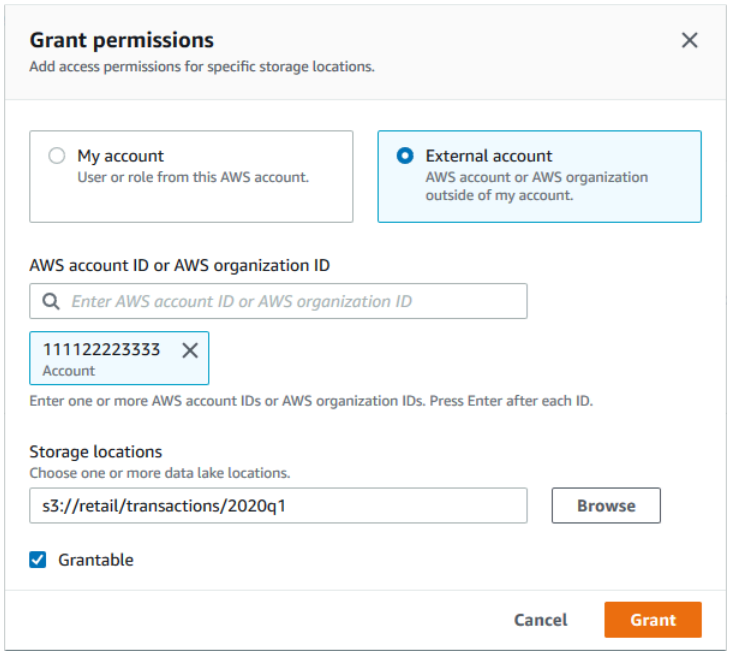
1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, seleccione Permisos y, a continuación, Conceder.
3. En el cuadro de diálogo Conceder permisos, elija el mosaico Cuenta externa.
4. Proporcione la información siguiente:

   - Para el ID de cuenta de AWS o el ID de organización de AWS, introduzca números de cuenta AWS, ID de organización o ID de unidad organizativa válidos.

     Pulse Intro después de cada ID.

     El ID de una organización está formado por en "o-" seguida de 10 a 32 letras minúsculas o dígitos.

     Un ID de unidad organizativa consta de "ou-" seguido de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo "-" (guión) y de 8 a 32 letras minúsculas o dígitos adicionales.
   - En las ubicaciones de almacenamiento, elija Examinar y busque una ubicación de almacenamiento de Amazon Simple Storage Service (Amazon S3). La ubicación debe estar registrada en Lake Formation.



5. Seleccione Concedible.
6. Elija Conceder.

Para conceder permisos de ubicación de datos (cuenta externa, AWS CLI)

- Para conceder permisos a una cuenta AWS externa, introduzca un comando similar al siguiente.

```
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"}}'
```

Este comando concede `DATA_LOCATION_ACCESS` con la opción de concesión a la cuenta 1111-2222-3333 de la ubicación de Amazon S3 `s3://retail/transactions/2020q1`, propiedad de la cuenta 1234-5678-9012.

Para conceder permisos a una organización, introduzca un comando similar al siguiente.

```
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"}}'
```

Este comando concede `DATA_LOCATION_ACCESS` con la opción de concesión a organización `s3://retail/transactions/2020q1` de la ubicación de Amazon S3 `o-abcdefghijkl`, propiedad de la cuenta 1234-5678-9012.

Para conceder permisos a una entidad principal en una cuenta AWS externa, introduzca un comando similar al siguiente.

```
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"}}'
```

Este comando concede `DATA_LOCATION_ACCESS` a una entidad principal en la cuenta 1111-2222-3333 en la ubicación de Amazon S3 `s3://retail/transactions/2020q1`, propiedad de la cuenta 1234-5678-9012.

Example

En el siguiente ejemplo se conceden permisos de ubicación de datos en `s3://retail` al grupo `ALLIAMPrincipals` en una cuenta externa.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "123456789012"}}'
```

Consulte también:

- Referencia de permisos de Lake Formation

Concesión de permisos sobre una ubicación de datos compartida con su cuenta

Después de compartir un recurso del catálogo de datos con su cuenta AWS, como administrador del lago de datos, puede conceder permisos sobre el recurso a otras entidades principales de su cuenta. Si el permiso `ALTER` se concede en una tabla compartida y la tabla apunta a una ubicación registrada de Amazon S3, también deberá conceder permisos de ubicación de datos en la ubicación. Del mismo modo, si se concede el permiso `CREATE_TABLE` o `ALTER` en una base de datos compartida y la base de datos tiene una propiedad de ubicación que apunta a una ubicación registrada, también deberá conceder permisos de ubicación de datos en la ubicación.

Para conceder permisos de ubicación de datos en una ubicación compartida a una entidad principal de su cuenta, ésta debe haber obtenido el permiso `DATA_LOCATION_ACCESS` en la ubicación con la opción de concesión. Cuando después conceda `DATA_LOCATION_ACCESS` a otra entidad principal de su cuenta, deberá incluir el ID del catálogo de datos (ID de cuenta AWS) de la cuenta propietaria. La cuenta del propietario es la cuenta en la que se registró la ubicación.

Puede usar la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI) para conceder los permisos de ubicación de datos.

Para conceder permisos en una ubicación de datos compartida con su cuenta (consola)

- Siga los pasos de Concesión de permisos de localización de datos (misma cuenta).

  Para las ubicaciones de almacenamiento, debe escribirlas. Para Ubicación de la cuenta registrada, introduzca el ID AWS de la cuenta propietaria.

Para conceder permisos sobre una ubicación de datos compartida con su cuenta (AWS CLI)

- Introduzca uno de los siguientes comandos para conceder permisos a un usuario o a un rol.

```
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>"}}'
```

Concesión y revocación de permisos sobre los recursos del catálogo de datos

Puede conceder permisos de lago de datos a las entidades principales en AWS Lake Formation para que estas puedan crear y administrar los recursos del catálogo de datos y acceder a los datos subyacentes. Puede conceder permisos de lago de datos en bases de datos, tablas y vistas. Al conceder permisos en tablas, puede limitar el acceso a columnas o filas específicas de la tabla para un control de acceso aún más específico.

Puede conceder permisos en tablas y vistas individuales, o bien con una única operación de concesión, en todas las tablas y vistas de una base de datos. Si concede permisos sobre todas las tablas de una base de datos, estará concediendo implícitamente el permiso `DESCRIBE` sobre la base de datos. A continuación, la base de datos aparece en la página Bases de datos de la consola y es devuelta por la operación `GetDatabases` de la API.

Puede conceder permisos utilizando el método de recursos con nombre o el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC).

Puede conceder permisos a entidades principales de la misma cuenta de Cuenta de AWS o a cuentas u organizaciones externas. Cuando los concede a cuentas u organizaciones externas, está compartiendo recursos de su propiedad con esas cuentas u organizaciones. A continuación, las entidades principales de esas cuentas u organizaciones podrán acceder a los recursos del catálogo de datos de su propiedad y a los datos subyacentes.

> **Note**
>
> Actualmente, el método LF-TBAC es compatible con la concesión de permisos entre cuentas a entidades principales de IAM, Cuentas de AWS, organizaciones y unidades organizativas (UO).

Al conceder permisos a cuentas u organizaciones externas, debe incluir la opción de concesión. Solo el administrador del lago de datos de la cuenta externa puede acceder a los recursos compartidos hasta que el administrador conceda permisos sobre los recursos compartidos a otras entidades principales de la cuenta externa.

Puede conceder permisos para el catálogo de datos utilizando la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

> **Note**
>
> Al eliminar un recurso del catálogo de datos, todos los permisos asociados al recurso dejan de ser válidos. Si se vuelve a crear el mismo recurso con el mismo nombre, no se recuperarán los permisos de Lake Formation. Los usuarios deberán volver a configurar los permisos nuevos.

> **Véase también:**
>
> - Uso compartido de tablas y bases de datos del Catálogo de datos entre cuentas AWS
> - Control de acceso a los metadatos
> - Referencia de permisos de Lake Formation

Permisos de IAM necesarios para conceder o revocar permisos de Lake Formation

Todas las entidades principales, incluido el administrador del lago de datos, necesitan los siguientes permisos de AWS Identity and Access Management (IAM) para conceder o revocar permisos de AWS Lake Formation en el catálogo de datos o permisos de ubicación de datos con la API de Lake Formation o la AWS CLI:

- `lakeformation:GrantPermissions`
- `lakeformation:BatchGrantPermissions`
- `lakeformation:RevokePermissions`
- `lakeformation:BatchRevokePermissions`
- `glue:GetTable` o `glue:GetDatabase` para una tabla o base de datos a la que esté concediendo permisos con el método de recursos con nombre.

> **Note**
>
> Los administradores del lago de datos tienen permisos implícitos de Lake Formation para conceder y revocar permisos de Lake Formation. Pero aún necesitan los permisos IAM en la concesión de Lake Formation y revocar las operaciones de API.

> Los roles de IAM con política administrada `AWSLakeFormationDataAdmin` de AWS no pueden añadir nuevos administradores de lago de datos porque esta política contiene una denegación explícita para la operación API Lake Formation, `PutDataLakeSetting`.

La siguiente política de IAM se recomienda para las entidades principales que no son administradores del lago de datos y que desean conceder o revocar permisos utilizando la consola de Lake Formation.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:ListPermissions",
                "lakeformation:GrantPermissions",
                "lakeformation:BatchGrantPermissions",
                "lakeformation:RevokePermissions",
                "lakeformation:BatchRevokePermissions",
                "glue:GetDatabases",
                "glue:SearchTables",
                "glue:GetTables",
                "glue:GetDatabase",
                "glue:GetTable",
                "iam:ListUsers",
                "iam:ListRoles",
                "sso-directory:DescribeUser",
                "sso-directory:DescribeGroup",
                "sso:DescribeInstance"
            ],
            "Resource": "*"
        }
    ]
}
```

Todos los permisos `glue:` e `iam:` de esta política están disponibles en la política administrada `AWSGlueConsoleFullAccess` de AWS.

Para conceder permisos utilizando el control de acceso basado en etiquetas (LF-TBAC), las entidades principales necesitan permisos IAM adicionales. Para más información, consulte Prácticas

recomendadas y consideraciones sobre el control de acceso basado en etiquetas de Lake Formation y Personas de Lake Formation y referencia de permisos IAM.

Permisos entre cuentas

Los usuarios que deseen conceder permisos entre cuentas de Lake Formation utilizando el método de recursos con nombre deben tener también los permisos en la política administrada `AWSLakeFormationCrossAccountManager` de AWS.

Los administradores del lago de datos necesitan esos mismos permisos para conceder permisos entre cuentas, además del permiso AWS Resource Access Manager (AWS RAM) para la concesión de permisos a organizaciones. Para obtener más información, consulte Permisos de administrador del lago de datos.

El usuario administrativo

Una entidad principal con permisos administrativos —por ejemplo, con la política administrada `AdministratorAccess` de AWS— tiene permisos para conceder permisos de Lake Formation y crear administradores de lago de datos. Para denegar a un usuario o rol el acceso a las operaciones del administrador de Lake Formation, adjunte o añada a su política una instrucción `Deny` para las operaciones de la API del administrador.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "lakeformation:GetDataLakeSettings",
                "lakeformation:PutDataLakeSettings"
            ],
            "Effect": "Deny",
            "Resource": [
                "*"
            ]
        }
    ]
}
```

> **Important**
>
> Para evitar que los usuarios se añadan como administradores con un script de extracción, transformación y carga (ETL), asegúrese de que todos los usuarios y roles que no sean administradores tengan denegado el acceso a estas operaciones de la API. La política administrada `AWSLakeFormationDataAdmin` de AWS contiene una denegación explícita para la operación de la API Lake Formation, `PutDataLakeSetting` que impide a los usuarios añadir nuevos administradores del lago de datos.

Concesión de permisos de lago de datos mediante el método de recurso con nombre

Puede utilizar el método de recurso con nombre para conceder permisos de Lake Formation sobre bases de datos, tablas y vistas específicas del catálogo de datos. Puede conceder permisos utilizando la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Temas

Concesión de permisos de base de datos mediante el método de recurso con nombre

Los siguientes pasos explican cómo conceder permisos de base de datos utilizando el método de recursos con nombre.

Console

Utilice la página de Conceder permisos de lago de datos en la consola de Lake Formation. La página está dividida en las secciones siguientes:

- Entidades principales – Los usuarios de IAM, roles, usuarios y grupos de IAM Identity Center, usuarios y grupos de SAML, cuentas AWS, organizaciones u unidades organizativas para conceder permisos.
- Etiquetas LF o recursos del catálogo – Las bases de datos, tablas, vistas o enlaces a recursos para los que se conceden permisos.

- Permisos. Los permisos de Lake Formation que se conceden.

> **Note**
>
> Para conceder permisos en un enlace a un recurso de base de datos, consulte Conceder permisos de enlace de recursos.

1. Abra la página Conceder permisos de lagos de datos.

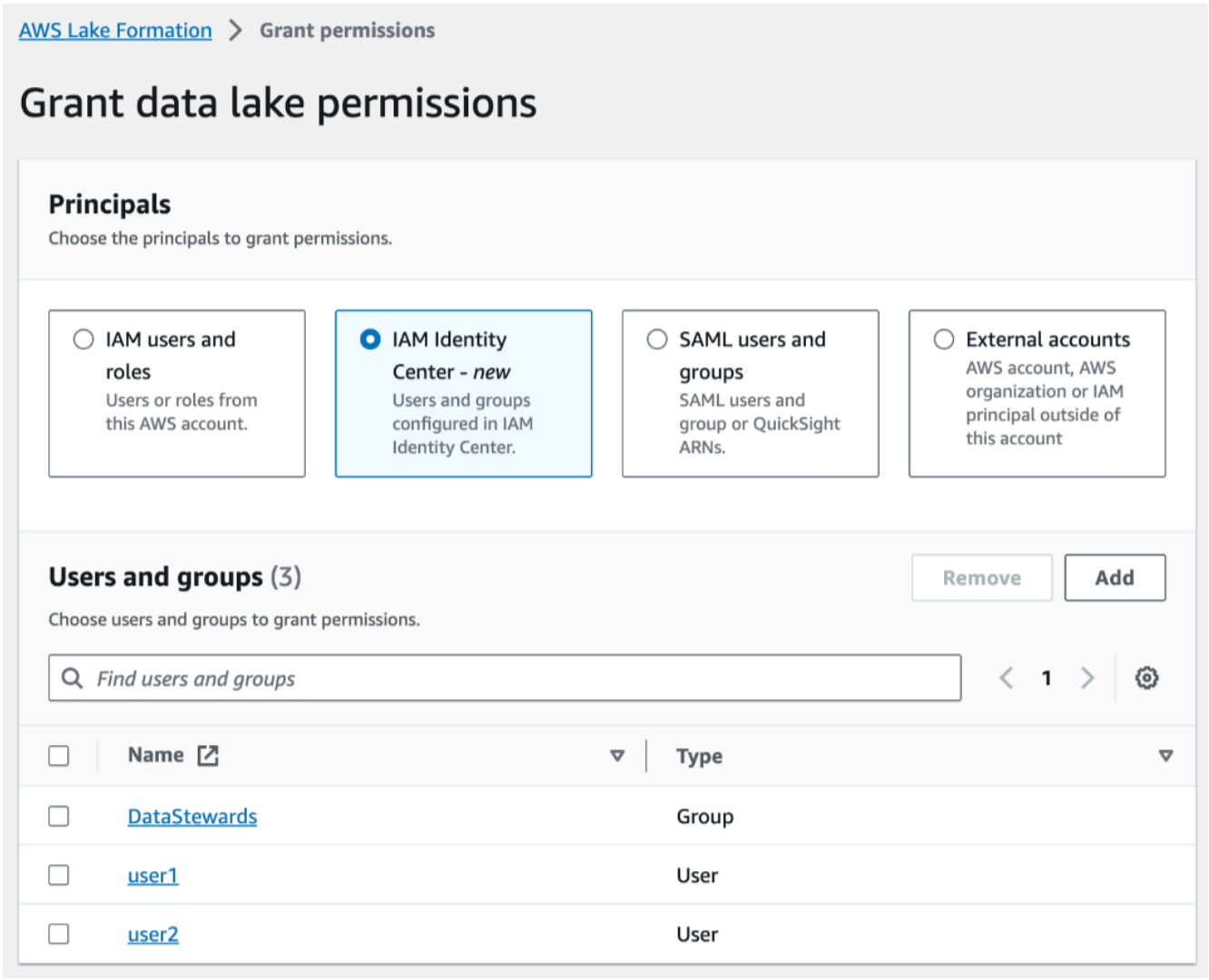
   Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/, e inicie sesión como administrador del lago de datos, creador de la base de datos o usuario de IAM que tenga permisos concedibles en la base de datos.

   Haga una de las acciones siguientes:

   - En el panel de navegación, en Permisos, seleccione Permisos de lago de datos. A continuación, seleccione Conceder.
   - En el panel de navegación, elija Bases de datos en Catálogo de datos. A continuación, en la página Bases de datos, elija una y, en el menú Acciones, en Permisos, seleccione Conceder.

   > **Note**
   >
   > Puede conceder permisos en una base de datos desde su enlace de recursos. Para ello, en la página Bases de datos, elija un enlace de recursos y, en el menú Acciones, Conceder en el destino. Para obtener más información, consulte Cómo funcionan los enlaces de recursos en Lake Formation.

2. A continuación, en la sección Entidades principales, elija un tipo de entidad principal y, a continuación, especifique las que van a recibir los permisos concedidos.



Usuarios y roles de IAM

Elija uno o varios usuarios o roles en la lista de usuarios y roles de IAM.

IAM Identity Center

Elija uno o varios usuarios o grupos en la lista de Usuarios y grupos. Seleccione Añadir para añadir más usuarios o grupos.

Usuarios y grupos de SAML

Para los usuarios y grupos de SAML y Amazon QuickSight, introduzca uno o varios nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de SAML, o ARN para los usuarios o grupos de Amazon QuickSight. Pulse Intro después de cada ARN.

Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

> **Note**
>
> La integración de Lake Formation con Amazon QuickSight solo es compatible con Amazon QuickSight Enterprise Edition.

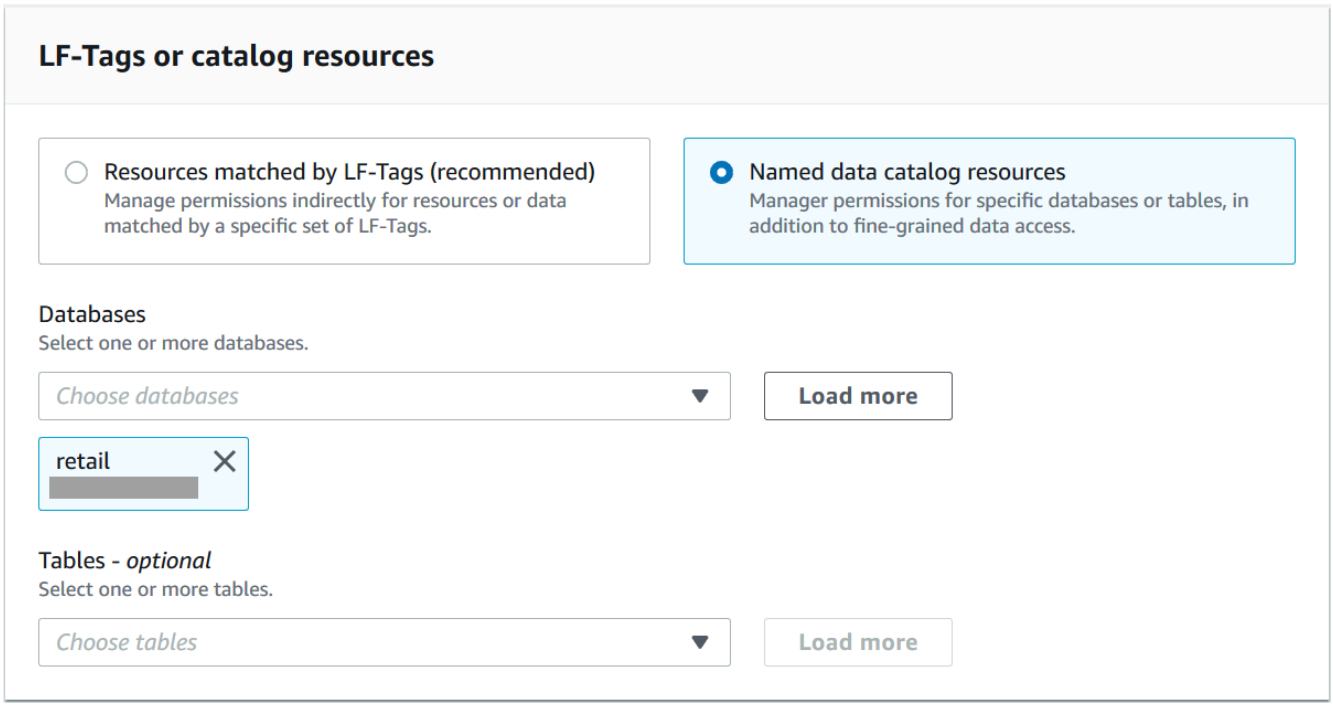
Cuentas externas

Para Cuenta de AWS, organización de AWS o la entidad principal de IAM, introduzca uno o varios identificadores de cuenta de AWS, ID de organización, ID de unidad organizativa o ARN válidos para el rol o usuario de IAM. Pulse Intro después de cada ID.

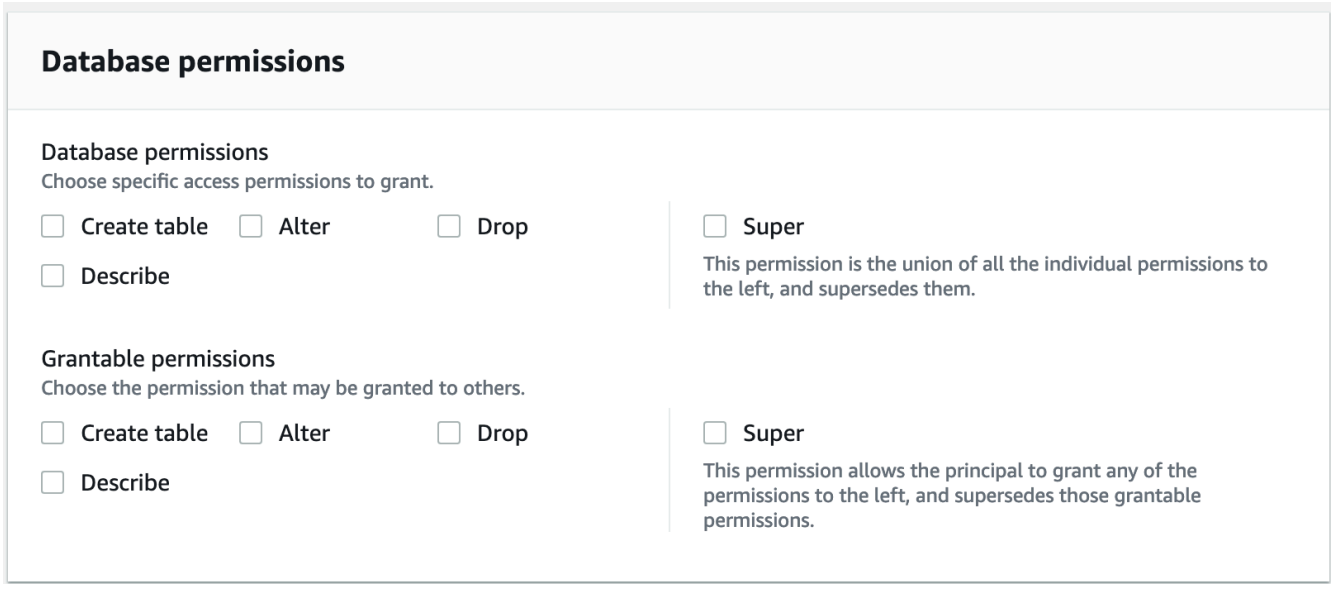
El ID de una organización consta de una «o-» seguida de 10 a 32 letras minúsculas o dígitos.

Un ID de una unidad organizativa comienza por «ou-» seguidos de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo guion «-» y de 8 a 32 letras minúsculas o dígitos adicionales.

3. En la sección de etiquetas LF o recursos del catálogo, seleccione Recursos del catálogo de datos con nombre.



4. Elija una o más bases de datos de la lista Bases de datos. También puede elegir una o más tablas o filtros de datos.
5. En la sección Permisos, seleccione los permisos y los permisos concedibles. En Permisos de base de datos, seleccione uno o más permisos para conceder.



> **Note**
>
> Después de conceder `Create Table` o `Alter` en una base de datos que tenga una propiedad de ubicación que apunte a una ubicación registrada, asegúrese de

conceder también permisos de ubicación de datos en la ubicación a las entidades principales. Para obtener más información, consulte Conceder permisos de ubicación de datos.

6. (Opcional) En Permisos concedibles, seleccione los permisos que el destinatario de la concesión puede conceder a otras entidades principales de su cuenta AWS. Esta opción no es compatible cuando se conceden permisos a una entidad principal de IAM desde una cuenta externa.
7. Elija Conceder.

AWS CLI

Para conceder permisos a la base de datos, utilice el método de recursos con nombre y la AWS Command Line Interface (AWS CLI).

Para conceder permisos a la base de datos mediante la AWS CLI

- Ejecute un comando `grant-permissions` y especifique como recurso una base de datos o el catálogo de datos, en función del permiso que se conceda.

  En los siguientes ejemplos, sustituya *<id-cuenta>* por un ID de cuenta de AWS válido.

  Example — Concesión para crear una base de datos

  En este ejemplo se concede `CREATE_DATABASE` al usuario `datalake_user1`. Dado que el recurso sobre el que se concede este permiso es el catálogo de datos, la orden especifica una estructura `CatalogResource` vacía como parámetro `resource`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

  Example — Concesión para crear tablas en una base de datos designada

  En el siguiente ejemplo, se concede `CREATE_TABLE` sobre la base de datos `retail` al usuario `datalake_user1`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Example — Concesión a una cuenta AWS externa con la opción Conceder

El siguiente ejemplo concede `CREATE_TABLE` con la opción de concesión de la base de datos `retail` a la cuenta externa 1111-2222-3333.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333 --permissions "CREATE_TABLE"
 --permissions-with-grant-option "CREATE_TABLE" --resource '{ "Database":
 {"Name":"retail"}}'
```

Example — Concesión a una organización

En el siguiente ejemplo se concede `ALTER` con la opción de concesión en la base de datos `issues` a la organización `o-abcdefghijkl`.

```
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 — Concesión de **`ALLIAMPrincipals`** en la misma cuenta

En el siguiente ejemplo se concede permiso de `CREATE_TABLE` sobre la base de datos `retail` a todas las entidades principales de la misma cuenta. Esta opción permite a todas las entidades principales de la cuenta crear una tabla en la base de datos y crear un enlace de recursos de tabla que permita a los motores de consulta integrados acceder a las bases de datos y tablas compartidas. Esta opción es especialmente útil cuando una entidad principal recibe una concesión entre cuentas y no tiene permiso para crear enlaces de recursos. En este escenario, el administrador del lago de datos puede crear una base de datos de marcadores de posición y conceder permiso `CREATE_TABLE` al grupo `ALLIAMPrincipal`, lo que permitirá a todas las entidades principales de IAM de la cuenta crear enlaces de recursos en la base de datos de marcadores de posición.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"temp","CatalogId":"111122223333"}}'
```

Example — Concesión a **`ALLIAMPrincipals`** en una cuenta externa

En el siguiente ejemplo se concede `CREATE_TABLE` sobre la base de datos `retail` a todas las entidades principales de la misma cuenta. Esta opción permite a cada entidad principal de la cuenta crear una tabla en la base de datos.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"retail","CatalogId":"123456789012"}}'
```

> **Note**
>
> Después de conceder `CREATE_TABLE` o `ALTER` en una base de datos que tenga una propiedad de ubicación que apunte a una ubicación registrada, asegúrese de conceder también permisos de ubicación de datos en la ubicación a las entidades principales. Para obtener más información, consulte Conceder permisos de ubicación de datos.

> **Véase también**
>
> - Referencia de permisos de Lake Formation
> - Conceder permisos en una base de datos o tabla compartida con su cuenta
> - Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos

Concesión de permisos de tabla mediante el método de recursos con nombre

Puede utilizar la consola de Lake Formation o la AWS CLI para conceder permisos de Lake Formation en las tablas del catálogo de datos. Puede conceder permisos sobre tablas individuales o, con una única operación de concesión, sobre todas las tablas de una base de datos.

Si concede permisos sobre todas las tablas de una base de datos, estará concediendo implícitamente el permiso `DESCRIBE` sobre la base de datos. A continuación, la base de datos aparece en la página Bases de datos de la consola y es devuelta por la operación `GetDatabases` de la API.

Cuando elija `SELECT` como permiso para conceder, tendrá la opción de aplicar un filtro de columnas, de filas o de celdas.

Console

Los siguientes pasos explican cómo conceder permisos de tabla utilizando el método de recursos con nombre y la página Conceder permisos de lago de datos de la consola de Lake Formation. La página está dividida en las estas secciones:

- Entidades principales. Usuarios, roles, cuentas de AWS, organizaciones o unidades organizativas a los que se conceden los permisos.
- Etiquetas LF o recursos del catálogo. Bases de datos, tablas o enlaces a recursos sobre los que se conceden los permisos.
- Permisos. Los permisos de Lake Formation que se conceden.

> **Note**
>
> Para conceder permisos sobre un enlace de recursos de tabla, consulte Conceder permisos de enlace de recursos.

1. Abra la página Conceder permisos de lagos de datos.

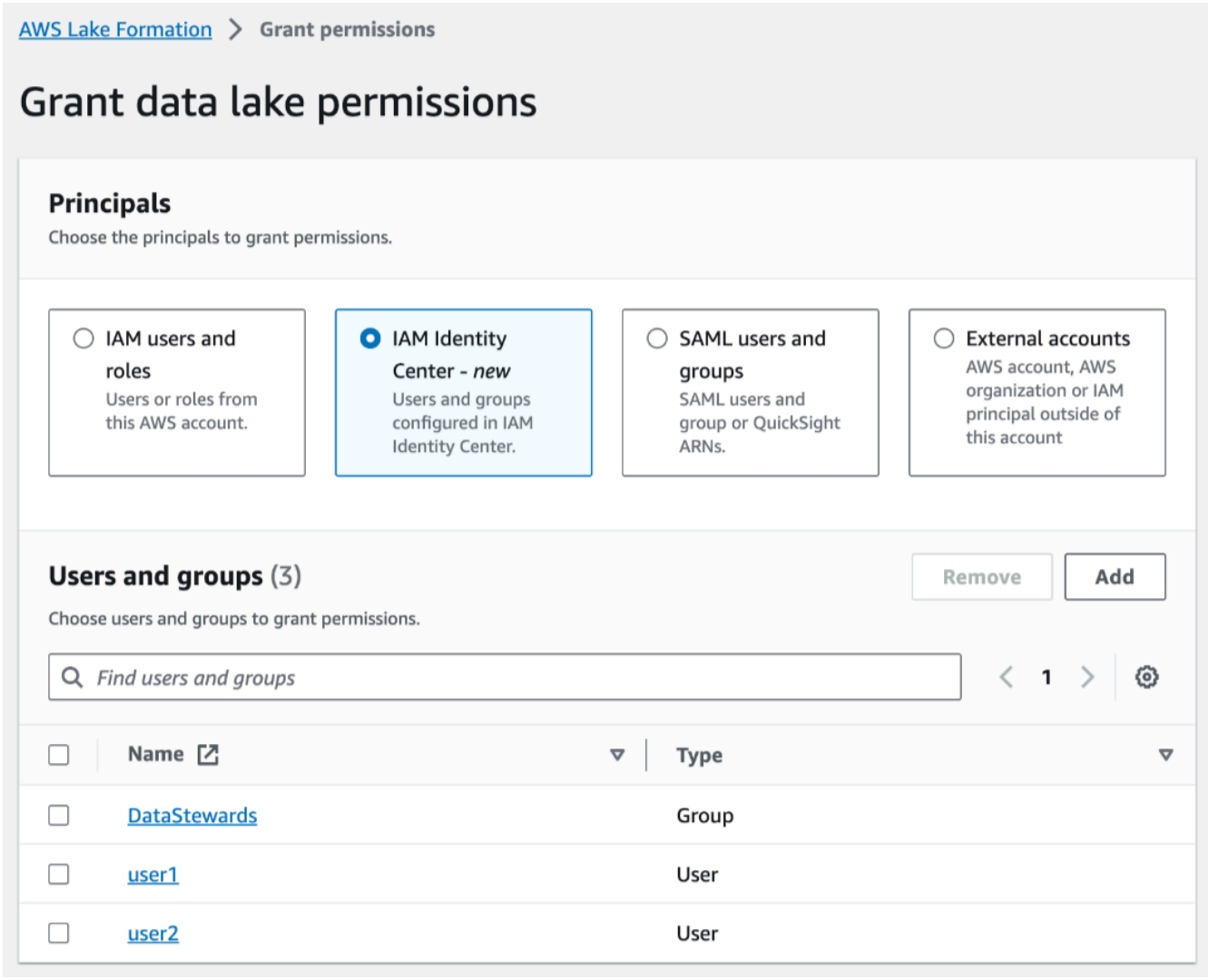
   Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/, e inicie sesión como administrador del lago de datos, como creador de la tabla o como usuario al que se le han concedido permisos sobre la tabla con la opción de concesión.

   Haga una de las acciones siguientes:

   - En el panel de navegación, elija Permisos de lago de datos en Permisos. A continuación, seleccione Conceder.
   - En el panel de navegación, elija Tablas. A continuación, en la página Tablas, elija una y, en el menú Acciones, Permisos, seleccione Conceder.

> **Note**
>
> Puede conceder permisos sobre una tabla a través de su enlace de recursos. Para ello, en la página Tablas, elija un enlace de recursos y, en el menú Acciones, Conceder en el destino. Para obtener más información, consulte Cómo funcionan los enlaces de recursos en Lake Formation.

2. A continuación, en la sección Entidades principales, elija un tipo de entidad principal y especifique las que van a recibir los permisos concedidos.



Usuarios y roles de IAM

Elija uno o varios usuarios o roles en la lista de usuarios y roles de IAM.

IAM Identity Center

Elija uno o varios usuarios o grupos en la lista de Usuarios y grupos.

Usuarios y grupos de SAML

Para los usuarios y grupos de SAML y Amazon QuickSight, introduzca uno o varios nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de SAML, o ARN para los usuarios o grupos de Amazon QuickSight. Pulse Intro después de cada ARN.

Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

> **Note**
>
> La integración de Lake Formation con Amazon QuickSight solo es compatible con Amazon QuickSight Enterprise Edition.

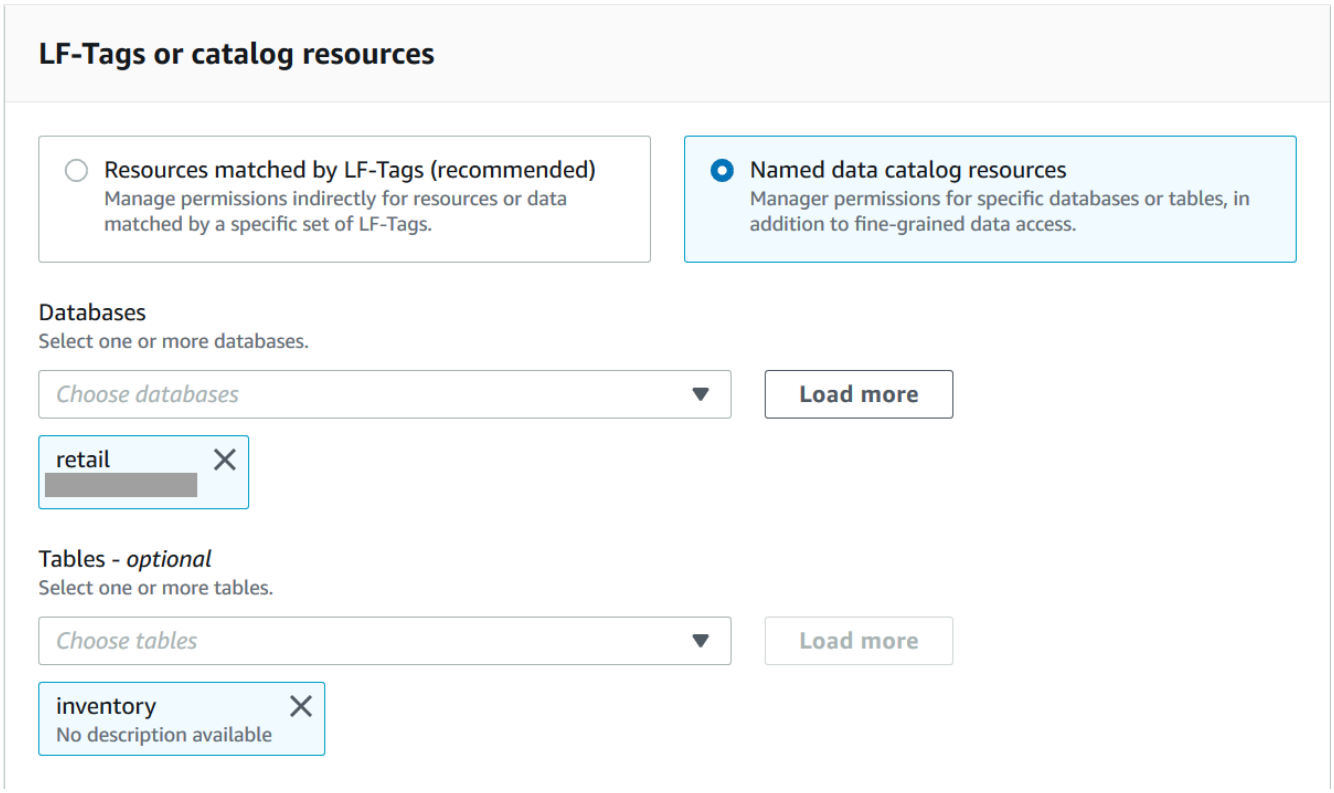
Cuentas externas

Para Cuenta de AWS, organización de AWS o entidad principal de IAM, introduzca uno o varios ID de Cuenta de AWS, ID de organización, ID de unidad organizativa o ARN válidos para el usuario o rol de IAM. Pulse Intro después de cada ID.

El ID de una organización consta de una «o-» seguida de 10 a 32 letras minúsculas o dígitos.

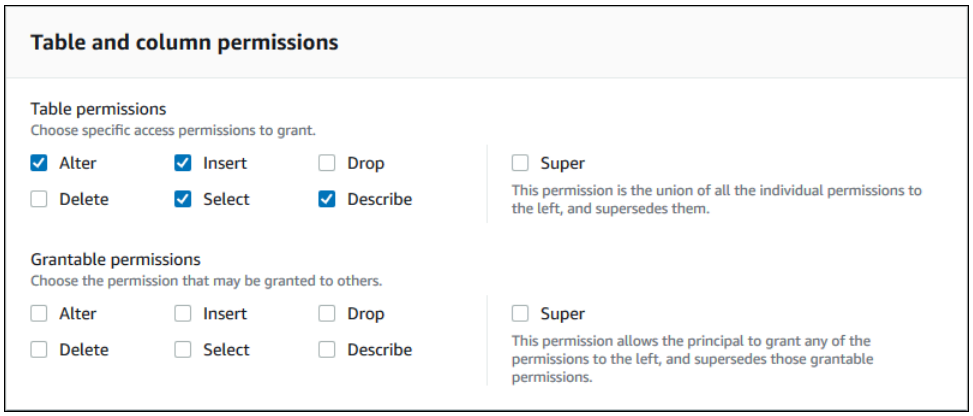
Un ID de una unidad organizativa comienza por «ou-» seguidos de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo carácter «-» y de 8 a 32 letras minúsculas o dígitos adicionales.

3. En la sección de Etiquetas LF o recursos del catálogo, seleccione una base de datos. A continuación, seleccione una o más tablas o Todas las tablas.

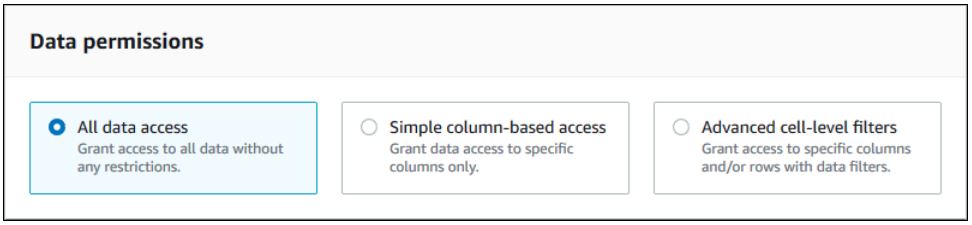


4. Especifique los permisos sin filtrado de datos

   En la sección Permisos, seleccione los permisos de tabla que desee conceder y, opcionalmente, seleccione los permisos que se pueden conceder.



   Si concede Seleccionar, la sección Permisos de datos aparece debajo de la sección Permisos de tabla y columna, con la opción Todos los accesos a datos seleccionada por defecto. Acepte los valores predeterminados.



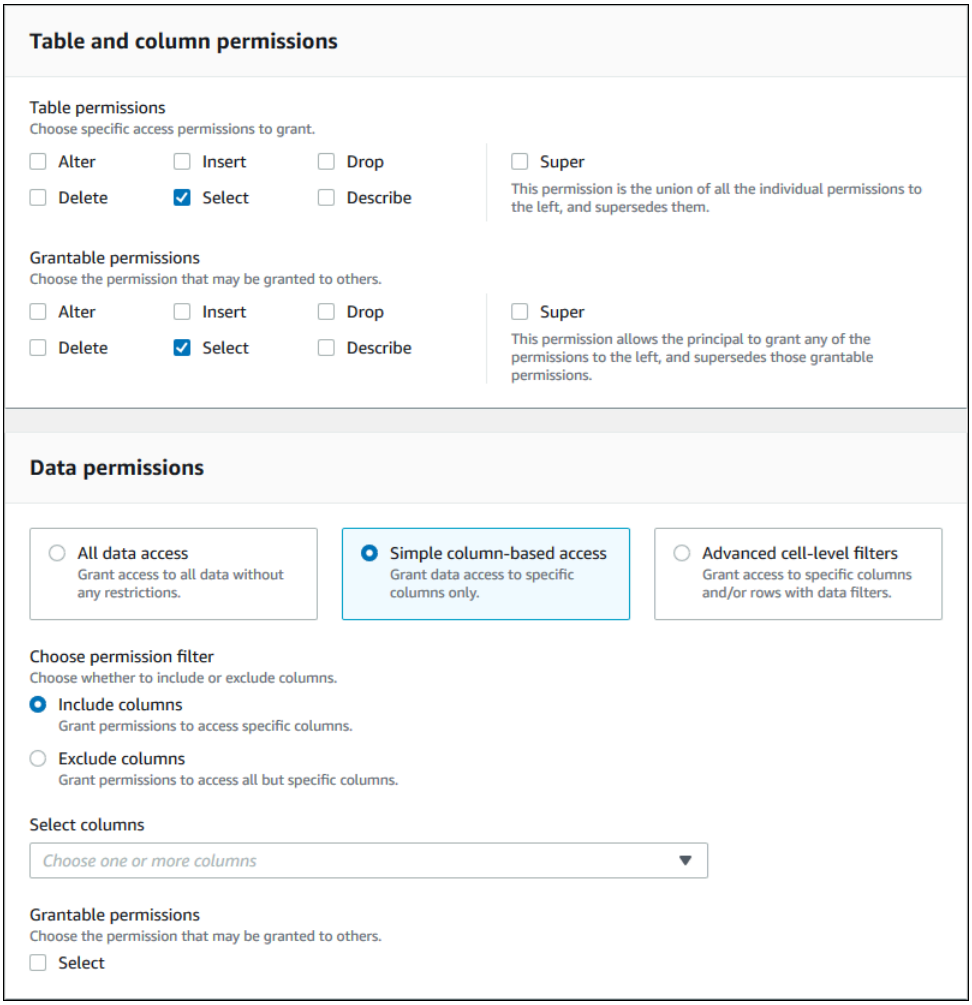
5. Elija Conceder.
6. Especifique el permiso Seleccionar con filtrado de datos

   Elija el permiso Seleccionar. No seleccione ningún otro permiso.

   La sección de permisos de datos aparece debajo de la sección de permisos de tabla y columna.
7. Haga una de las acciones siguientes:

   - Aplique solo un filtrado de columnas simple.

     1. Elija Acceso simple basado en columnas.



2. Elija si desea incluir o excluir columnas y, a continuación, elija las que desea incluir o excluir.

   Solo se admiten listas de inclusión cuando se conceden permisos a una cuenta AWS u organización externa.

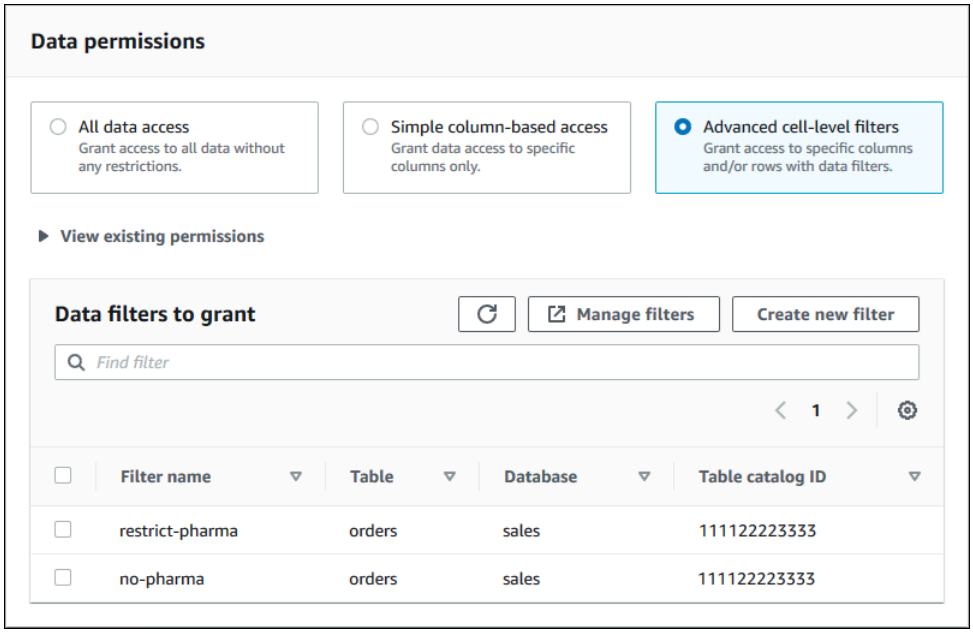
3. (Opcional) En Permisos concedibles, active la opción de concesión para el permiso Seleccionar.

   Si incluye la opción de concesión, el destinatario de esta podrá conceder permisos solo sobre las columnas que le conceda.

> **Note**
>
> También puede aplicar el filtrado por columnas solo creando un filtro de datos que especifique un filtro de columnas y especifique todas las filas como filtro de filas. Sin embargo, esto requiere más pasos.

- Aplicar filtros de columna, fila o celda.

  1. Seleccione Filtros avanzados a nivel de celda.



  2. (Opcional) Amplíe Ver los permisos existentes.
  3. (Opcional) Seleccione Crear filtro nuevo.
  4. (Opcional) Para ver los detalles de los filtros de la lista o para crear filtros nuevos o eliminar los existentes, seleccione Administrar filtros.

     La página Filtros de datos se abre en una nueva ventana del navegador.

     Cuando haya terminado de acceder a la página Filtros de datos, vuelva a la página Conceder permisos y, si es necesario, actualícela para ver los filtros de datos nuevos que haya creado.
  5. Seleccione uno o más filtros de datos para aplicarlos a la concesión.

> **Note**
>
> Si no hay filtros de datos en la lista, significa que no se creó ningún filtro de datos para la tabla seleccionada.

8. Elija Conceder.

AWS CLI

Para conceder permisos de tabla, utilice el método de recursos con nombre y la AWS Command Line Interface (AWS CLI).

Para conceder permisos de tabla mediante la AWS CLI

- Introduzca un comando `grant-permissions` y especifique una tabla como recurso.

Example . Concesión en una sola tabla, sin filtrado

En el ejemplo siguiente concede el permiso `SELECT` y `ALTER` al usuario `datalake_user1` en la cuenta AWS 1111-2222-3333 de la tabla `inventory` en la base de datos `retail`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

> **Note**
>
> Si concede el permiso `ALTER` en una tabla cuyos datos subyacentes están en una ubicación registrada, asegúrese de conceder también permisos de ubicación de datos en la ubicación a las entidades principales. Para obtener más información, consulte Conceder permisos de ubicación de datos.

Example – Concesión en todas las tablas con la opción Concesión, sin filtrado

En el siguiente ejemplo, se concede `SELECT` con la opción de concesión en todas las tablas de la base de datos `retail`.

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" --permissions-with-grant-option "SELECT" --resource '{ "Table":
 { "DatabaseName": "retail", "TableWildcard": {} } }'
```

Example – Concesión con filtrado simple de columnas

El siguiente ejemplo concede SELECT en un subconjunto de columnas de la tabla `persons`. Utiliza un filtrado de columnas simple.

```
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 – Concesión con filtro de datos

En este ejemplo se concede SELECT en la tabla `orders` y se aplica el filtro de datos `restrict-pharma`.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

A continuación se muestra el contenido del archivo `grant-params.json`.

```
{
    "Principal": {"DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"},
    "Resource": {
        "DataCellsFilter": {
            "TableCatalogId": "111122223333",
            "DatabaseName": "sales",
            "TableName": "orders",
            "Name": "restrict-pharma"
        }
    },
    "Permissions": ["SELECT"],
    "PermissionsWithGrantOption": ["SELECT"]
}
```

Véase también

- Descripción general de los permisos de Lake Formation
- Filtrado de datos y seguridad de celda en Lake Formation
- Personas de Lake Formation y referencia de permisos IAM
- Conceder permisos de enlace de recursos
- Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos

Concesión de permisos para vistas mediante el método de recursos con nombre

Los siguientes pasos explican cómo conceder permisos para vistas utilizando el método de recursos con nombre y la página Conceder permisos de lago de datos. La página está dividida en las secciones siguientes:

- Entidades principales – Los usuarios de IAM, roles, usuarios y grupos de IAM Identity Center, usuarios y grupos de SAML, cuentas Cuentas de AWS, organizaciones u unidades organizativas para conceder permisos.
- Etiquetas LF o recursos del catálogo – Las bases de datos, tablas, vistas o enlaces a recursos para los que se conceden permisos.
- Permisos – Los permisos de lagos de datos que se conceden.

Abra la página Conceder permisos de lagos de datos

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/, e inicie sesión como administrador del lago de datos, creador de la base de datos o usuario de IAM que tenga permisos concedibles en la base de datos.
2. Haga una de las acciones siguientes:
   - En el panel de navegación, en Permisos, seleccione Permisos de lago de datos. A continuación, seleccione Conceder.
   - En el panel de navegación, en Catálogo de datos, elija Vistas. A continuación, en la página Vistas, elija una y, en el menú Acciones, Permisos, seleccione Conceder.

> **Note**
>
> Puede conceder permisos sobre una vista a través de su enlace de recursos. Para ello, en la página Vistas, elija un enlace de recursos y, en el menú Acciones, elija Conceder en el destino. Para obtener más información, consulte Cómo funcionan los enlaces de recursos en Lake Formation.

Especificar las entidades principales

En la sección Entidades principales, elija uno de los tipos y, a continuación, especifique las que van a recibir los permisos concedidos.

Usuarios y roles de IAM

Elija uno o varios usuarios o roles en la lista de usuarios y roles de IAM.

IAM Identity Center

Elija uno o varios usuarios o grupos en la lista de Usuarios y grupos.

Usuarios y grupos de SAML

Para los usuarios y grupos de SAML y Amazon QuickSight, introduzca uno o varios nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de SAML, o ARN para los usuarios o grupos de Amazon QuickSight. Pulse Intro después de cada ARN.

Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

> **Note**
>
> La integración de Lake Formation con Amazon QuickSight solo es compatible con Amazon QuickSight Enterprise Edition.

Cuentas externas

Para Cuenta de AWS, organización de AWS o la entidad principal de IAM, introduzca uno o varios identificadores de cuenta de AWS, ID de organización, ID de unidad organizativa o ARN válidos para el rol o usuario de IAM. Pulse Intro después de cada ID.

El ID de una organización consta de una «o-» seguida de 10 a 32 letras minúsculas o dígitos.

Un ID de una unidad organizativa comienza por «ou-» seguidos de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo guion «-» y de 8 a 32 letras minúsculas o dígitos adicionales.

Véase también

- Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos

Especifique las vistas

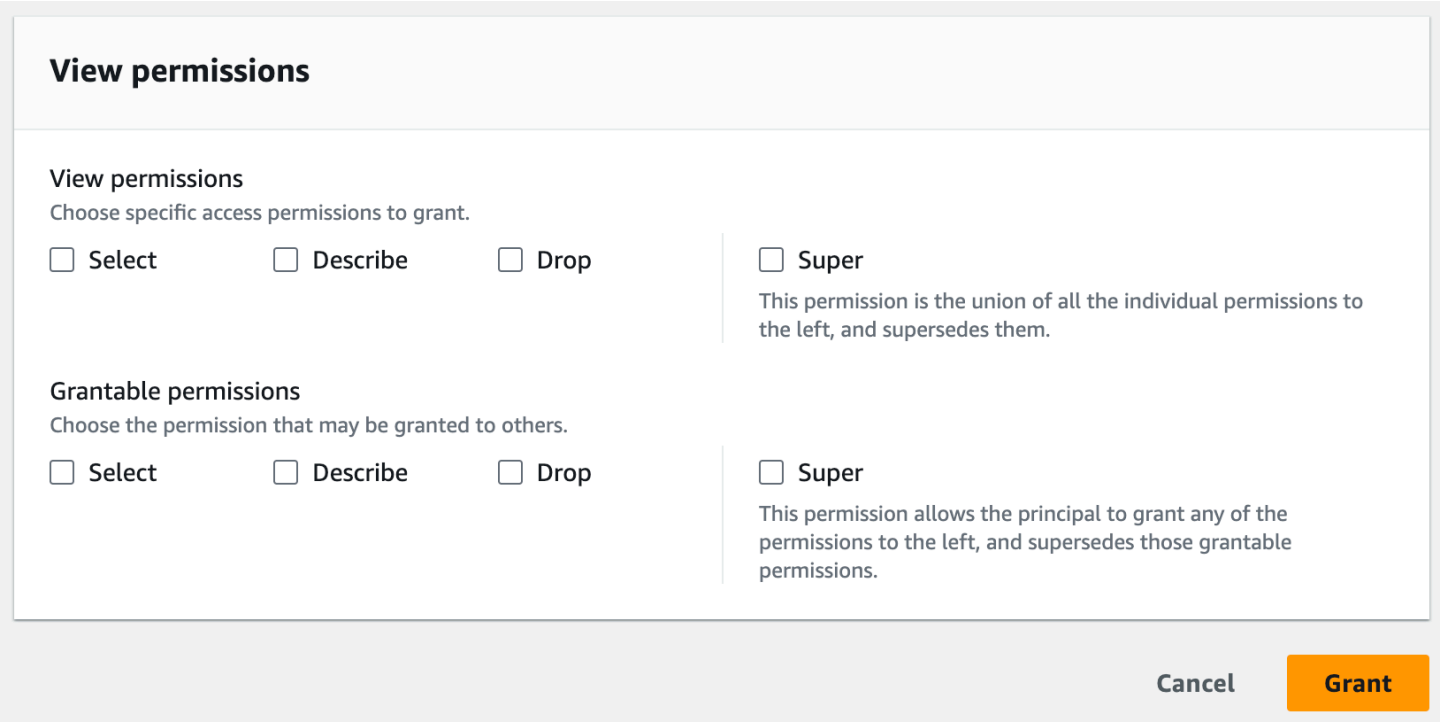
En la sección de etiquetas LF o recursos del catálogo, elija una o más vistas para las que conceder permisos.

1. Seleccione un recurso de catálogo de datos con nombre.
2. Elija una o más vistas de la lista Vistas. También puede elegir una o más bases de datos, tablas o filtros de datos.

   Al conceder permisos de lago de datos a `All views` dentro de una base de datos, el beneficiario tendrá permisos en todas las tablas y vistas de la base de datos.

Especificar los permisos

En la sección Permisos, seleccione los permisos y los permisos concedibles.



1. En Ver de permisos, seleccione uno o más permisos para conceder.
2. (Opcional) En Permisos concedibles, seleccione los permisos que el destinatario de la concesión puede conceder a otras entidades principales de su cuenta Cuenta de AWS. Esta opción no es compatible cuando se conceden permisos a una entidad principal de IAM desde una cuenta externa.
3. Elija Conceder.

Véase también

- Referencia de permisos de Lake Formation
- Conceder permisos en una base de datos o tabla compartida con su cuenta

Control de acceso basado en etiquetas de Lake Formation

El control de acceso basado en etiquetas de Lake Formation (LF-TBAC) es una estrategia de autorización que define permisos basados en atributos. En Lake Formation, estos atributos se denominan etiquetas LF. Puede adjuntar etiquetas LF a los recursos del catálogo de datos y conceder permisos a los directores de Lake Formation sobre esos recursos mediante estas etiquetas

LF. Lake Formation permite realizar operaciones en esos recursos cuando el valor de la etiqueta del principal coincide con el valor de la etiqueta del recurso. LF-TBAC es útil en entornos que crecen rápidamente y ayuda en situaciones en las que la administración de políticas se vuelve engorrosa.

LF-TBAC es el método recomendado para conceder permisos de Lake Formation cuando hay un gran número de recursos del Catálogo de datos. LF-TBAC es más escalable que el método de recursos con nombre y requiere menos sobrecarga en la administración de permisos.

> **Note**
>
> Las etiquetas IAM no son lo mismo que las etiquetas LF. Estas etiquetas no son intercambiables. Las etiquetas LF se utilizan para conceder permisos de Lake Formation y las etiquetas de IAM se utilizan para definir las políticas de IAM.

Cómo funciona el control de acceso basado en etiquetas de Lake Formation

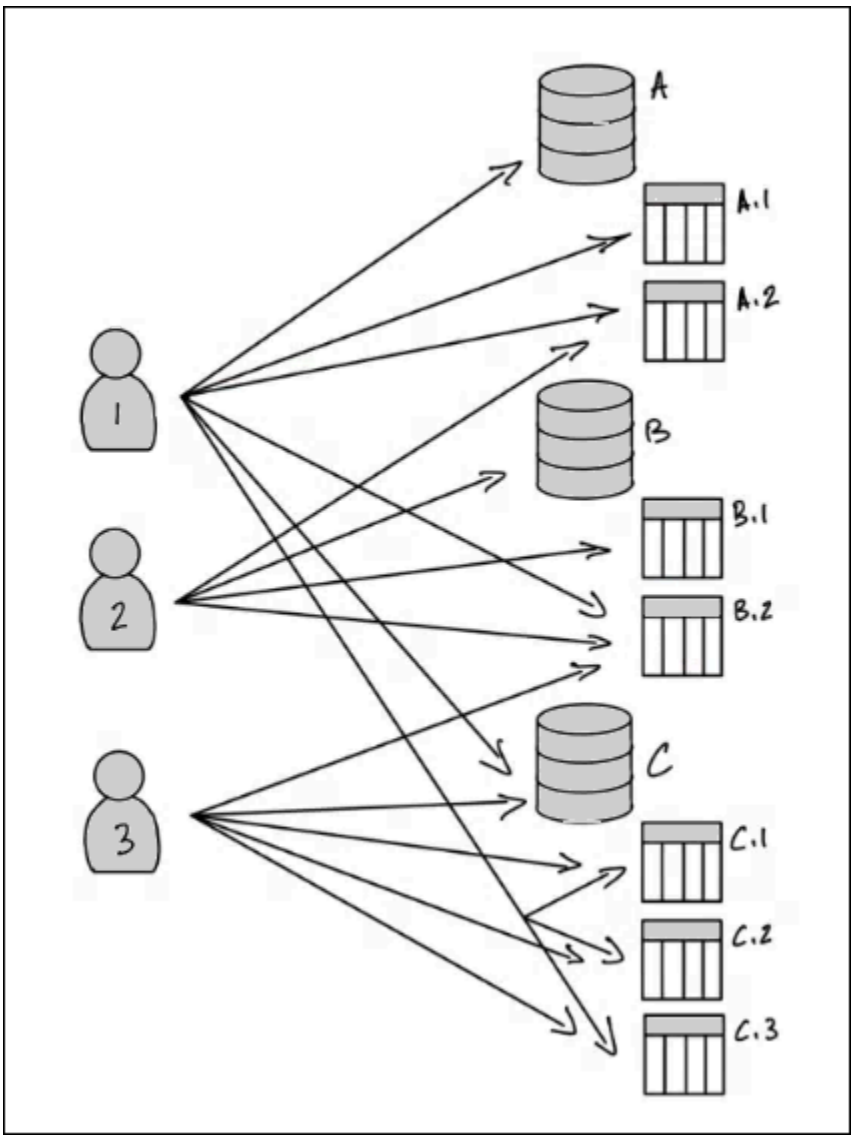
Cada etiqueta LF es un par clave-valor, como o `department=sales` o `classification=restricted`. Una clave puede tener varios valores definidos, como `department=sales,marketing,engineering,finance`.

Para utilizar el método LF-TBAC, los administradores del lago de datos y los ingenieros de datos llevan a cabo las tareas siguientes.

Tarea	Detalles de la tarea
1. Definir las propiedades y las relaciones de las etiquetas LF.	-
2. Generar los creadores de etiquetas LF en Lake Formation.	Añadir creadores de etiquetas LF
3. Generar la etiqueta LF en Lake Formation.	Crear etiquetas LF
4. Asignar etiquetas LF a recursos del Catálogo de datos.	Asignación de varias etiquetas LF a recursos del catálogo de datos
5. Conceder permisos a otras entidades principales para asignar etiquetas LF a los	Conceder, revocar y enumerar permisos de valores de etiqueta LF

Tarea	Detalles de la tarea
recursos, de forma opcional con la opción de concesión.	
6. Conceder expresiones de etiquetas LF a entidades principales, de forma opcional con la opción de concesión.	Conceder permisos de lago de datos mediante el método LF-TBAC
7. (Recomendado) Después de verificar que las entidades principales tienen acceso a los recursos correctos mediante el método LF-TBAC, revoque los permisos que se concedieron utilizando el método de recursos con nombre.	-

Imagine un caso en el que debe conceder permisos a tres entidades principales sobre tres bases de datos y siete tablas.



Para conseguir los permisos indicados en el diagrama anterior utilizando el método de recursos con nombre, tendría que efectuar 17 concesiones, como se indica a continuación (en pseudocódigo).

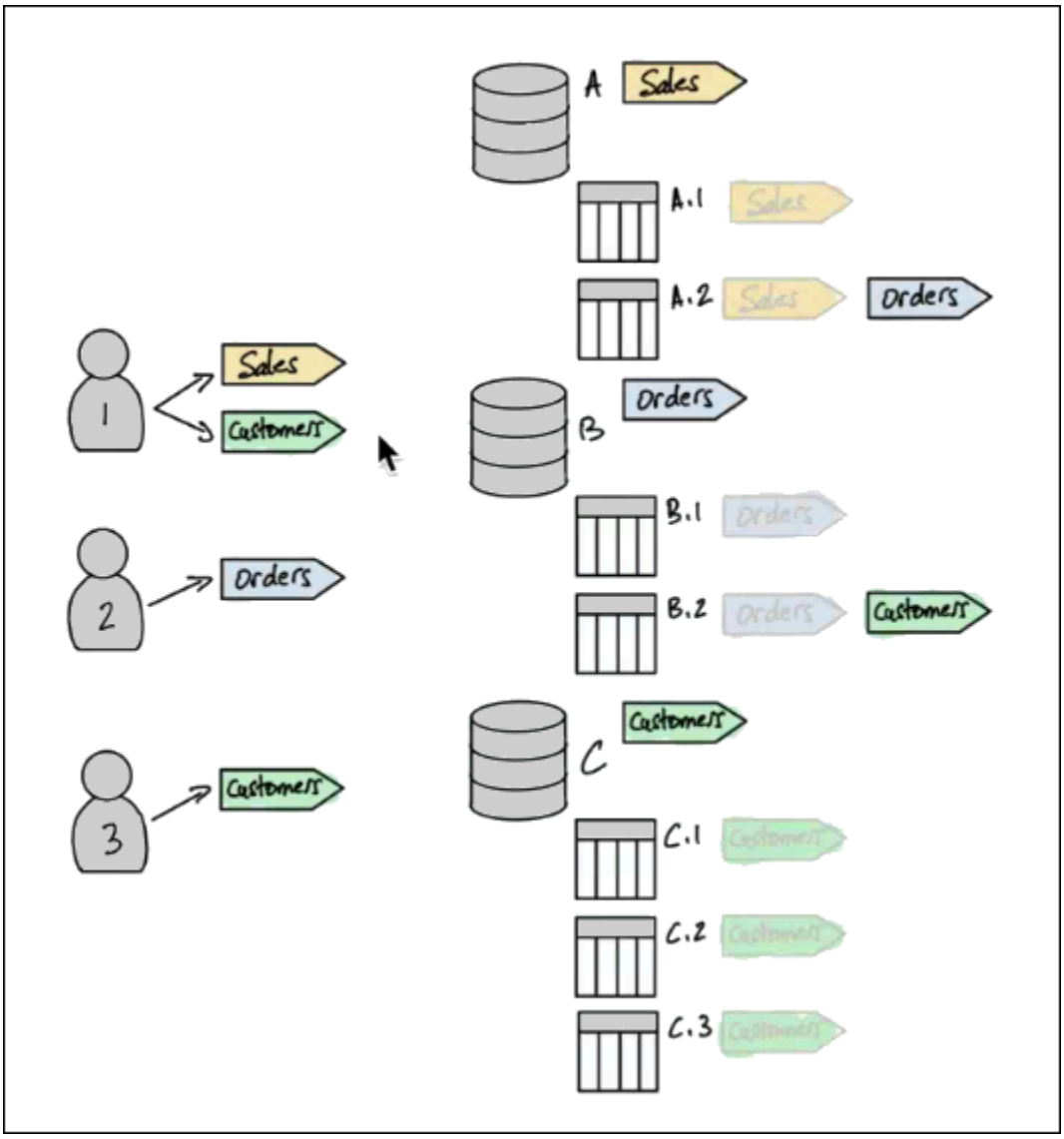
```
GRANT CREATE_TABLE ON Database A TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.1 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table B.2 TO PRINCIPAL 1
...
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 2
GRANT CREATE_TABLE ON Database B TO PRINCIPAL 2
...
GRANT SELECT, INSERT ON Table C.3 TO PRINCIPAL 3
```

Ahora piense cómo concedería los permisos mediante el método de LF-TBAC. El siguiente diagrama indica que ha asignado etiquetas LF a bases de datos y tablas, y ha concedido permisos sobre las etiquetas LF a las entidades principales.

En este ejemplo, las etiquetas LF representan áreas del lago de datos que contienen análisis para diferentes módulos de un conjunto de aplicaciones de planificación de recursos empresariales (ERP). Puede controlar el acceso a los datos analíticos de los distintos módulos. Todas las etiquetas LF tienen la clave `module` y los posibles valores `Sales`, `Orders`, y `Customers`. Un ejemplo de URL sería:

```
module=Sales
```

El diagrama muestra solo los valores de las etiquetas LF.



Asignación de etiquetas a los recursos del Catálogo de datos y herencia

Las tablas heredan las etiquetas LF de las bases de datos y las columnas heredan las etiquetas LF de las tablas. Los valores heredados pueden anularse. En el diagrama anterior, las etiquetas LF atenuadas se heredan.

Debido a la herencia, el administrador del lago de datos solo necesita hacer las cinco asignaciones de etiquetas LF siguientes a los recursos (en pseudocódigo).

```
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
```

Etiquetar las concesiones a las entidades principales

Tras asignar las etiquetas LF a las bases de datos y tablas, el administrador del lago de datos debe efectuar solo cuatro concesiones de etiquetas LF a las entidades principales, como se indica a continuación (en pseudocódigo).

```
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
```

Ahora, una entidad principal con la etiqueta LF `module=Sales` puede acceder a los recursos del Catálogo de datos con la etiqueta LF `module=Sales` (por ejemplo, la base de datos A), una entidad principal con la etiqueta LF `module=Customers` puede acceder a los recursos con la etiqueta LF `module=Customers`, y así sucesivamente.

Los comandos de concesión anteriores están incompletos. Esto se debe a que, aunque indican mediante etiquetas LF los recursos del Catálogo de datos sobre los que las entidades principales tienen permisos, no indican exactamente qué permisos de Lake Formation (como `SELECT`, `ALTER`) tienen las entidades principales sobre esos recursos. Por lo tanto, los siguientes comandos de pseudocódigo son una representación más exacta de cómo se conceden los permisos de Lake Formation sobre los recursos del Catálogo de datos a través de las etiquetas LF.

```
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Sales TO Principal 1
GRANT (SELECT, INSERT ON TABLES)  ON TAGS module=Sales TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 1
```

```
GRANT (SELECT, INSERT ON TABLES)  ON TAGS module=Customers TO Principal 1
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Orders TO Principal 2
GRANT (SELECT, INSERT ON TABLES)  ON TAGS module=Orders TO Principal 2
GRANT (CREATE_TABLE ON DATABASES) ON TAGS module=Customers TO Principal 3
GRANT (SELECT, INSERT ON TABLES)  ON TAGS module=Customers TO Principal 3
```

Recopilación. Permisos resultantes sobre los recursos

Dadas las etiquetas LF asignadas a las bases de datos y las tablas en el diagrama anterior, y las etiquetas LF concedidas a las entidades principales en el diagrama, la siguiente tabla enumera los permisos de Lake Formation que las entidades principales tienen sobre las bases de datos y las tablas.

Entidad principal	Permisos otorgados mediante etiquetas LF
Entidad principal 1	• `CREATE_TABLE` sobre la base de datos A • `SELECT`, `INSERT` sobre la tabla A.1 • `SELECT`, `INSERT` sobre la tabla B.2 • `CREATE_TABLE` sobre la base de datos C • `SELECT`, `INSERT` sobre la tabla C.1 • `SELECT`, `INSERT` sobre la tabla C.2 • `SELECT`, `INSERT` sobre la tabla C.3
Entidad principal 2	• `SELECT`, `INSERT` sobre la tabla A.2 • `CREATE_TABLE` sobre la base de datos B • `SELECT`, `INSERT` sobre la tabla B.1 • `SELECT`, `INSERT` sobre la tabla B.2
Entidad principal 3	• `SELECT`, `INSERT` sobre la tabla B.2 • `CREATE_TABLE` sobre la base de datos C • `SELECT`, `INSERT` sobre la tabla C.1 • `SELECT`, `INSERT` sobre la tabla C.2 • `SELECT`, `INSERT` sobre la tabla C.3

Conclusión

En este sencillo ejemplo, el administrador del lago de datos ha podido especificar 17 permisos mediante cinco operaciones de asignación y ocho de concesión. Cuando hay decenas de bases de datos y cientos de tablas, la ventaja del método LF-TBAC sobre el método de recursos con nombre se hace evidente. En el caso hipotético de la necesidad de conceder a cada entidad principal acceso a cada recurso, y donde `n(P)` es el número de entidades principales y `n(R)` es el número de recursos:

- Con el método de recurso con nombre, la cantidad de concesiones requeridas es $n(P) \times n(R)$.
- Con el método LF-TBAC, que utiliza una sola etiqueta LF, el número total de concesiones a entidades principales y asignaciones a recursos es $n(P) + n(R)$.

> Véase también
>
> - Gestión de etiquetas LF para el control de acceso a los metadatos
> - Conceder permisos de lago de datos mediante el método LF-TBAC

Temas

- Gestión de etiquetas LF para el control de acceso a los metadatos
- Conceder, revocar y enumerar permisos de valores de etiqueta LF

Gestión de etiquetas LF para el control de acceso a los metadatos

Para utilizar el método de control de acceso basado en etiquetas (LF-TBAC) de Lake Formation para proteger los recursos del catálogo de datos (bases de datos, tablas y columnas), debe crear etiquetas LF, asignarlas a los recursos y conceder permisos de etiquetas LF a las entidades principales.

Antes de poder asignar etiquetas LF a los recursos del catálogo de datos o conceder permisos a las entidades principales, debe definir las etiquetas LF. Solo un administrador del lago de datos o una entidad principal con permisos de creación de etiquetas LF pueden crear etiquetas LF.

Creadores de etiquetas LF

El creador de etiquetas LF es una entidad principal no administradora que tiene permisos para crear y administrar etiquetas LF. Los administradores de lagos de datos pueden añadir creadores de

etiquetas LF mediante la consola Lake Formation o la CLI. Los creadores de etiquetas LF tienen permisos implícitos de Lake Formation para actualizar y eliminar etiquetas LF, asignar etiquetas LF a recursos y conceder permisos de etiquetas LF y permisos de valores de etiquetas LF a otras entidades principales.

Con los roles de creador de etiquetas LF, los administradores de lagos de datos pueden delegar tareas de administración de etiquetas, como la creación y actualización de claves y valores de etiquetas, a entidades principales que no sean administradoras. Los administradores de los lagos de datos también pueden conceder permisos `Create LF-Tag` a los creadores de etiquetas LF. Luego, el creador de las etiquetas LF puede conceder el permiso para crear etiquetas LF a otras entidades principales.

Puede conceder dos tipos de permisos a las etiquetas LF:

- Permisos de etiqueta LF: `Create LF-Tag`, `Alter` y `Drop`. Estos permisos son necesarios para crear, actualizar y eliminar LF-tags.

  Los administradores de los lagos de datos y los creadores de etiquetas LF tienen implícitamente estos permisos en las etiquetas LF que crean y pueden concederlos de forma explícita a las entidades principales para que gestionen las etiquetas del lago de datos.
- Permisos de par clave-valor de etiqueta LF: `Assign`, `Describe` y `Grant with LF-Tag expressions`. Estos permisos son necesarios para asignar etiquetas LF a las bases de datos, tablas y columnas del catálogo de datos, y para conceder permisos sobre los recursos a las entidades principales mediante el control de acceso basado en etiquetas de Lake Formation. Los creadores de etiquetas LF reciben implícitamente estos permisos al crear etiquetas LF.

Tras recibir el permiso `Create LF-Tag` y crear correctamente las etiquetas LF, el creador de las etiquetas LF puede asignarlas a recursos y conceder permisos (`Create LF-Tag`, `Alter` y `Drop`) a otras entidades principales no administrativas para que gestionen las etiquetas en el lago de datos. Puede conceder permisos utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

> **Note**
>
> Los administradores de lagos de datos tienen permisos implícitos de Lake Formation para crear, actualizar y eliminar etiquetas LF, asignar etiquetas LF a los recursos y conceder permisos de etiquetas LF a las entidades principales.

Para conocer las prácticas recomendadas y consideraciones, consulte Prácticas recomendadas y consideraciones sobre el control de acceso basado en etiquetas de Lake Formation

Temas

Véase también

- Conceder, revocar y enumerar permisos de valores de etiqueta LF
- Conceder permisos de lago de datos mediante el método LF-TBAC
- Control de acceso basado en etiquetas de Lake Formation

Añadir creadores de etiquetas LF

De forma predeterminada, los administradores de lagos de datos pueden crear, actualizar y eliminar etiquetas LF, asignar etiquetas a los recursos del catálogo de datos y conceder permisos de etiquetas a las entidades principales. Si desea delegar las operaciones de creación y administración de etiquetas a entidades principales no administrativa, el administrador del lago de datos puede crear roles de creador de etiquetas LF y conceder permisos `Create LF-Tag` a Lake Formation para los roles. Con un permiso `Create LF-Tag` concedible, los creadores de etiquetas LF pueden delegar las tareas de creación y mantenimiento de etiquetas a otras entidades principales no administrativas.

> **Note**
>
> Las concesiones de permisos entre cuentas solo pueden incluir permisos `Describe` y `Associate`. No puede conceder permisos `Create LF-Tag`, `Drop`, `Alter` ni `Grant with LFTag expressions` a las entidades principales de una cuenta diferente.

Temas

- Permisos de IAM requeridos para crear etiquetas LF
- Añadir creadores de etiquetas LF

> **Véase también**
>
> - Conceder, revocar y enumerar permisos de valores de etiqueta LF
> - Conceder permisos de lago de datos mediante el método LF-TBAC
> - Control de acceso basado en etiquetas de Lake Formation

Permisos de IAM requeridos para crear etiquetas LF

Debe configurar permisos para que una entidad principal de Lake Formation pueda crear LF-tags. Añada la siguiente instrucción a la política de permisos de la entidad principal que necesita ser creadora de etiquetas LF.

> **Note**
>
> Si bien los administradores de lagos de datos tienen permisos implícitos de Lake Formation para crear, actualizar y eliminar etiquetas LF, asignar etiquetas LF a los recursos y conceder etiquetas LF a las entidades principales, los administradores de lagos de datos también necesitan los permisos de IAM siguientes.

Para obtener más información, consulte Personas de Lake Formation y referencia de permisos 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"
     ]
 }
```

Las entidades principales que asignan etiquetas LF a recursos y conceden etiquetas LF a entidades principales deben tener los mismos permisos, salvo `CreateLFTag`, `UpdateLFTag` y `DeleteLFTag`.

Añadir creadores de etiquetas LF

Un creador de etiquetas LF puede crear una etiqueta LF, actualizar la clave y los valores de la etiqueta, eliminar etiquetas, asociar etiquetas a los recursos del catálogo de datos y conceder permisos sobre los recursos del catálogo de datos a entidades principales mediante el método LF-TBAC. El creador de la etiqueta LF también puede conceder estos permisos a entidades principales.

Puede crear roles de creador de etiquetas LF mediante la consola AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

console

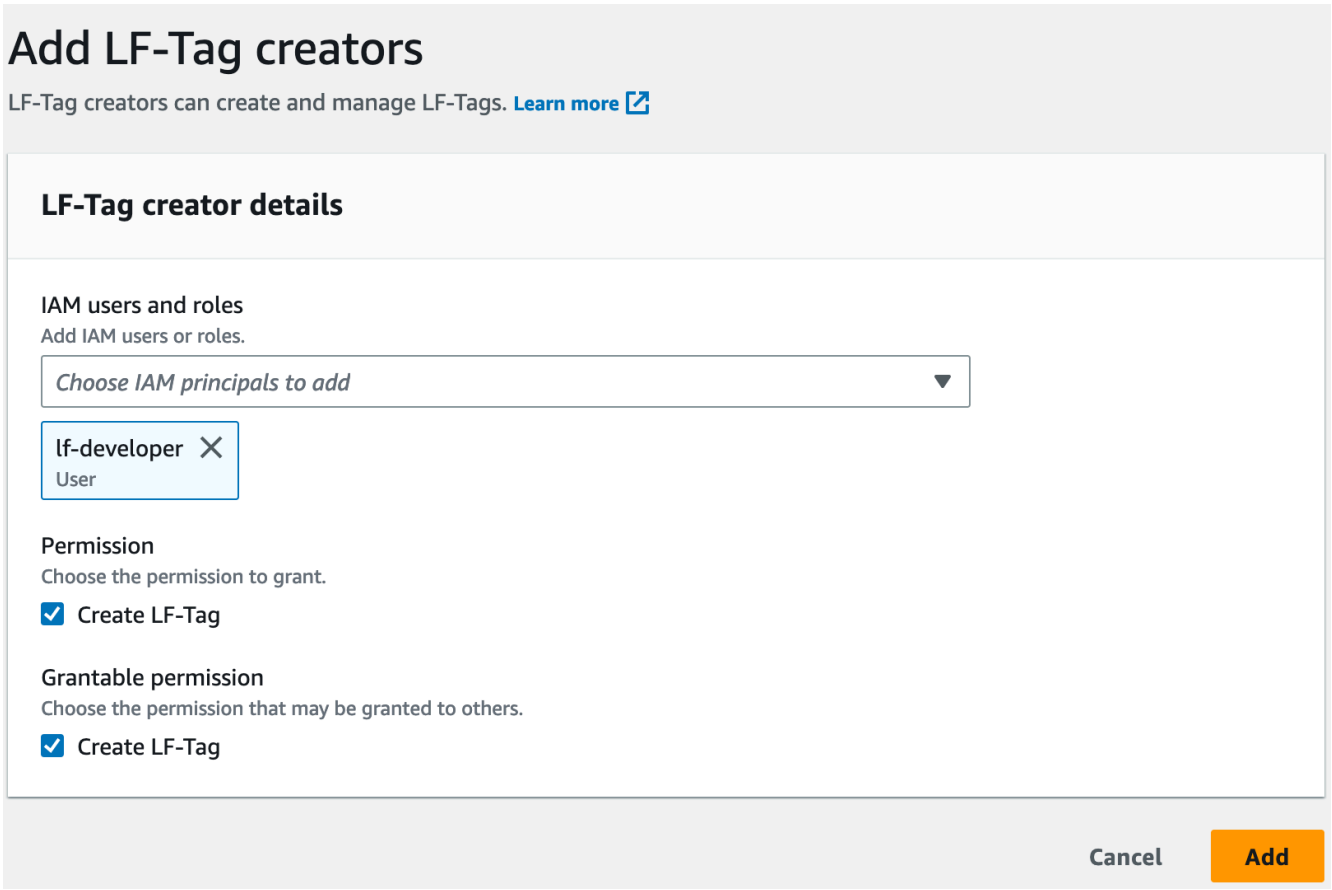
Para añadir un creador de etiquetas LF

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos.

2. En el panel de navegación, en Permisos, elija Etiquetas LF y permisos.

   En la página Permisos y etiquetas LF, seleccione la sección Creadores de etiquetas LF y seleccione Añadir creadores de etiquetas LF.



3. En la página Añadir creadores de etiquetas LF, seleccione un usuario o un rol de IAM que tenga los permisos necesarios para crear etiquetas LF.
4. Marque la casilla de verificación del permiso `Create LF-Tag`.
5. (Opcional) Para que las entidades principales seleccionadas puedan conceder permisos `Create LF-Tag` a las entidades principales, seleccione el permiso concedible `Create LF-Tag`.
6. Seleccione Agregar.

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"
    ]
}
```

Los siguientes son los permisos disponibles para un rol de creador de etiquetas LF:

Permiso	Descripción
`Drop`	Una entidad principal con este permiso en una etiqueta LF puede eliminar una etiqueta LF del lago de datos. La entidad principal obtiene un permiso implícito `Describe` sobre todos los valores de etiqueta de un recurso de etiqueta LF.
`Alter`	Una entidad principal con este permiso en una etiqueta LF puede añadir o eliminar el valor de una etiqueta LF. La entidad principal obtiene un permiso implícito `Alter` sobre todos los valores de una etiqueta LF.
`Describe`	Una entidad principal con este permiso en una etiqueta LF puede ver la etiqueta LF y sus valores al asignar etiquetas LF a los recursos o conceder permisos a las etiquetas LF. Puede conceder `Describe` en todos los valores clave o en valores específicos.
`Associate`	Una entidad principal con este permiso en una etiqueta LF puede asignar esta a un recurso del catálogo de datos. Al conceder `Associate` está otorgando `Describe` de manera implícita.
`Grant with LF-Tag expression`	Una entidad principal con este permiso en una etiqueta LF puede conceder permisos sobre los recursos de un catálogo de datos utilizando la clave y los valores de la etiqueta LF. Conceder `Grant with LF-Tag expression` otorga `Describe` de manera implícita.

Estos permisos se pueden conceder. Una entidad principal que haya recibido estos permisos con la opción de otorgarlos puede otorgarlos a otras entidades principales.

Crear etiquetas LF

Todas las etiquetas LF deben estar definidas en Lake Formation antes de poder utilizarlas. Una etiqueta LF está formada por una clave y uno o más valores posibles para dicha clave.

En cuanto el administrador del lago de datos haya configurado los permisos IAM necesarios y los permisos de Lake Formation para el rol de creador de la etiqueta LF, la entidad principal podrá crear una etiqueta LF. El creador de la etiqueta LF obtiene permiso implícito para actualizar o eliminar cualquier valor de la etiqueta LF y borrar la etiqueta LF.

Puede crear etiquetas LF utilizando la consola AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

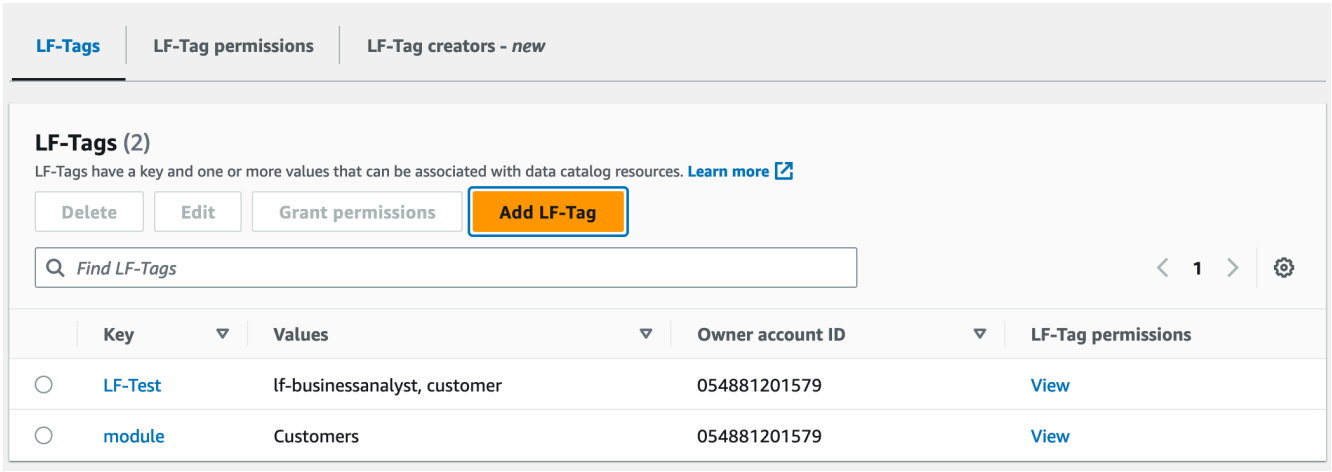
Console

Para crear una etiqueta LF

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como entidad principal con permisos de creador de etiquetas LF o como administrador del lago de datos.
2. En el panel de navegación, en permisos y etiquetas LF, seleccione Etiquetas LF.

   Aparece la página Etiquetas LF.



3. Seleccione Añadir etiqueta LF.
4. En el cuadro de diálogo Añadir etiqueta LF, introduzca una clave y uno o más valores.

Cada clave debe tener un valor como mínimo. Para introducir varios valores, introduzca una lista delimitada por comas y, a continuación, pulse Intro, o introduzca un valor cada vez y pulse Añadir después de cada uno. El número máximo de valores permitidos es 1000.

5. Seleccione Agregar etiqueta.

AWS CLI

Para crear una etiqueta LF

- Introduzca un comando `create-lf-tag`.

  En el siguiente ejemplo, se crea una etiqueta LF con clave `module` y valores `Customers` y `Orders`.

```
aws lakeformation create-lf-tag --tag-key module --tag-values Customers Orders
```

Como creadora de la etiqueta, la entidad principal obtiene el permiso `Alter` para utilizar esta etiqueta LF y puede actualizar o eliminar cualquier valor de esta etiqueta LF. La entidad principal creadora de la etiqueta LF también puede conceder permiso `Alter` a otra entidad principal para que actualice y elimine los valores de las etiquetas de esta etiqueta LF.

Actualización de las etiquetas LF

Para actualizar una etiqueta LF sobre la que tenga permiso `Alter`, añada o elimine los valores clave permitidos. No se puede cambiar la clave de la etiqueta LF. Para cambiar la clave, elimine la etiqueta LF y añada una con la clave necesaria. Además del permiso `Alter`, también necesita el permiso IAM `lakeformation:UpdateLFTag` para actualizar los valores.

Al eliminar un valor de etiqueta LF, no se comprueba la presencia de ese valor de etiqueta LF en ningún recurso del catálogo de datos. Si el valor de la etiqueta LF eliminada está asociado a un recurso, deja de estar visible para el recurso y las entidades principales a las que se les concedieron permisos sobre ese par clave-valor dejan de tener los permisos.

Antes de eliminar un valor de etiqueta LF, si lo desea, puede utilizar el comando `remove-lf-tags-from-resource` para eliminar la etiqueta LF de los recursos del catálogo de datos que tengan el valor que desee eliminar y, a continuación, volver a etiquetar el recurso con los valores que desee conservar.

Solo los administradores del lago de datos, el creador de la etiqueta LF y las entidades principales que tengan permisos `Alter` sobre la etiqueta LF pueden actualizar una etiqueta LF.

Para actualizar una etiqueta LF, utilice la consola AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Console

Para actualizar una etiqueta LF (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador de un lago de datos, creador de etiquetas LF o entidad principal con permiso `Alter` sobre la etiqueta LF.
2. En el panel de navegación, en permisos y etiquetas LF, seleccione Etiquetas LF.
3. En la página de etiquetas LF, seleccione una etiqueta LF y, a continuación, seleccione Editar.
4. En el cuadro de diálogo Editar etiqueta LF, añada o elimine valores de etiqueta LF.

   Para añadir varios valores, en el campo Valores introduzca una lista delimitada por comas y, a continuación, pulse Intro, o un valor cada vez y pulse Agregar después de cada uno.
5. Elija Guardar.

AWS CLI

Para actualizar una etiqueta LF (AWS CLI)

- Introduzca un comando `update-lf-tag`. Proporcione uno o los dos argumentos siguientes:
  - `--tag-values-to-add`
  - `--tag-values-to-delete`

Example

En el siguiente ejemplo, se reemplaza el valor `vp` por el valor `vice-president` de la clave LF-tag `level`.

```
aws lakeformation update-lf-tag --tag-key level --tag-values-to-add vice-president
--tag-values-to-delete vp
```

Eliminar etiquetas LF

Puede borrar las etiquetas LF que ya no estén en uso. No se comprueba la presencia de la etiqueta LF en un recurso del catálogo de datos. Si la etiqueta LF eliminada está asociada a un recurso, deja de estar visible para el recurso y las entidades principales a las que se concedieron permisos sobre esa etiqueta LF dejan de tener los permisos.

Antes de eliminar una etiqueta LF, si lo desea, puede utilizar el comando `remove-lf-tags-from-resource` para eliminar la etiqueta LF de todos los recursos.

Solo los administradores del lago de datos, el creador de la etiqueta LF y las entidades principales que tengan permisos `Drop` sobre la etiqueta LF pueden actualizar una etiqueta LF. Además del permiso `Drop`, la entidad principal también necesita permiso de IAM `lakeformation:DeleteLFTag` para eliminar una etiqueta LF.

Para eliminar una etiqueta LF, utilice la consola AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Console

Para eliminar una etiqueta LF (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos.
2. En el panel de navegación, en permisos y etiquetas LF, seleccione Etiquetas LF.
3. En la página de etiquetas LF, seleccione una etiqueta LF y, a continuación, seleccione Eliminar.
4. En el cuadro de diálogo para borrar el entorno de la etiqueta, para confirmar la eliminación, introduzca el valor de la clave LF-Tag en el campo designado y, a continuación, seleccione Eliminar.

AWS CLI

Para eliminar una etiqueta (AWS CLI)

- Introduzca un comando `delete-lf-tag`. Proporcione la clave de la etiqueta LF que desee eliminar.

Example

En el siguiente ejemplo se elimina la etiqueta LF con la clave `region`.

```
aws lakeformation delete-lf-tag --tag-key region
```

Listado de etiquetas

Puede consultar las etiquetas LF sobre las que tiene los permisos `Describe` o `Associate`. Los valores que aparecen con cada clave de etiqueta LF son los valores sobre los que tiene permisos.

El creador de etiquetas LF tiene permisos implícitos para ver las etiquetas LF que ha creado.

Los administradores del lago de datos pueden ver todas las etiquetas LF definidas en la cuenta local AWS y todas las etiquetas LF para las que se han otorgado permisos `Describe` y `Associate` a la cuenta local desde cuentas externas. El administrador del lago de datos puede ver todos los valores de todas las etiquetas LF.

Puede crear etiquetas LF utilizando la consola AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

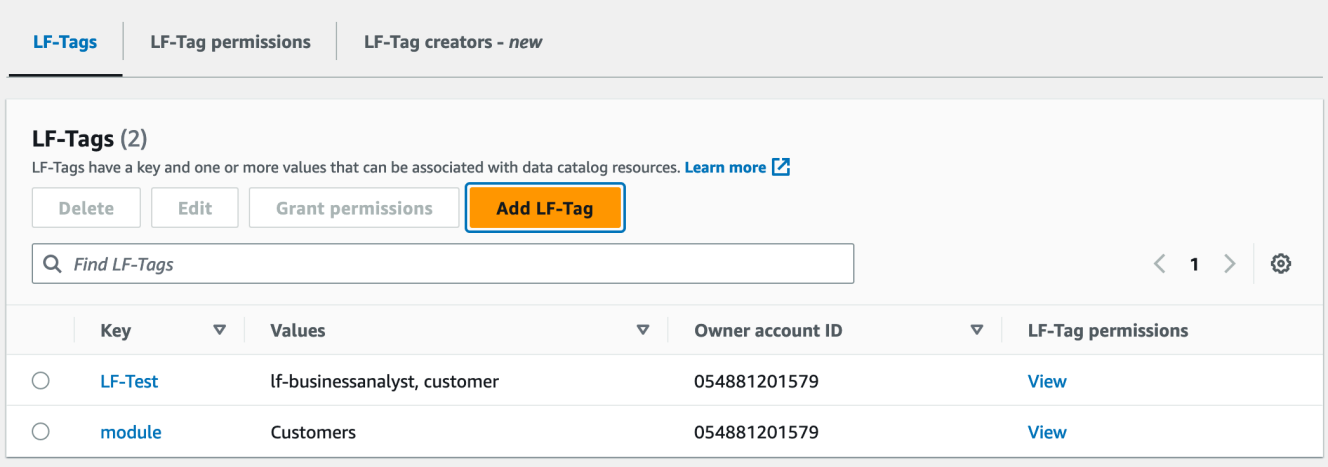
Console

Para generar una lista de etiquetas LF (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como creador de etiquetas LF, como administrador de un lago de datos o como entidad principal a la que se le hayan concedido permisos sobre las etiquetas LF y que tenga el permiso IAM `lakeformation:ListLFTags`.
2. En el panel de navegación, en permisos y etiquetas LF, seleccione Etiquetas LF.

   Aparece la página Etiquetas LF.



Consulte la columna del identificador de la cuenta del propietario para determinar las etiquetas LF que se compartieron con su cuenta desde una cuenta externa.

AWS CLI

Para mostrar las etiquetas LF (AWS CLI)

- Ejecute el comando siguiente como administrador de un lago de datos o como entidad principal a la que se le hayan concedido permisos sobre las etiquetas LF y que tenga el permiso de IAM `lakeformation:ListLFTags`.

```
aws lakeformation list-lf-tags
```

El resultado es similar al siguiente.

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
```

```
            "TagKey": "module",
            "TagValues": [
                "Orders",
                "Sales",
                "Customers"
            ]
        }
    ]
}
```

Para ver también las etiquetas LF concedidas desde cuentas externas, incluya la opción de comando `--resource-share-type ALL`.

```
aws lakeformation list-lf-tags --resource-share-type ALL
```

El resultado es similar al siguiente. Observe la clave `NextToken`, que indica que la lista continúa.

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "Orders",
                "Sales",
                "Customers"
            ]
        }
    ],
    "NextToken": "eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ=="
}
```

Repita el comando y añada el argumento `--next-token` para ver las etiquetas LF locales restantes y las que se hayan concedido desde cuentas externas. Las etiquetas LF de cuentas externas siempre se encuentran en una página separada.

```
aws lakeformation list-lf-tags --resource-share-type ALL
--next-token eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ==
```

```
{
    "LFTags": [
        {
            "CatalogId": "123456789012",
            "TagKey": "region",
            "TagValues": [
                "central",
                "south"
            ]
        }
    ]
}
```

API

Puede usar los SDK disponibles para Lake Formation para enumerar las etiquetas que el solicitante tiene permiso para ver.

```
import boto3

client = boto3.client('lakeformation')
...

response = client.list_lf_tags(
    CatalogId='string',
    ResourceShareType='ALL',
    MaxResults=50'
)
```

Este comando devuelve un objeto `dict` con la estructura siguiente:

```
{
```

```
    'LFTags': [
        {
            'CatalogId': 'string',
            'TagKey': 'string',
            'TagValues': [
                'string',
            ]
        },
    ],
    'NextToken': 'string'
}
```

Para obtener más información sobre los permisos necesarios, consulte Personas de Lake Formation y referencia de permisos IAM.

Asignación de varias etiquetas LF a recursos del catálogo de datos

Puede asignar etiquetas LF a recursos del catálogo de datos (bases de datos, tablas y columnas) para controlar acceso a esos recursos. Solo pueden acceder a los recursos las entidades principales a las que se concedan etiquetas LF coincidentes (y las entidades principales a las que se concede acceso con el método de recurso con nombre).

Si una tabla hereda una etiqueta LF de una base de datos o una columna hereda una etiqueta LF de una tabla, puede anular el valor heredado asignando un valor nuevo a la clave de etiqueta LF.

La cantidad máxima de etiquetas LF que puede asignar a un recurso es 50.

Temas

- Requisitos para administrar las etiquetas asignadas a los recursos
- Asignar etiquetas LF a una columna de una tabla
- Asignación de etiquetas LF a un recurso del catálogo de datos
- Actualización de etiquetas LF para un recurso
- Eliminación de una etiqueta LF de un recurso

Requisitos para administrar las etiquetas asignadas a los recursos

Para asignar una etiqueta LF a un recurso del catálogo de datos, debe:

- Tener el permiso de Lake Formation `ASSOCIATE` en la etiqueta LF.

- Tener el permiso de IAM `lakeformation:AddLFTagsToResource`.
- Tener el permiso Glue:GetDatabase en una base de datos de Glue.
- Sea el propietario (creador) del recurso, tener el permiso de Lake Formation `Super` sobre el recurso con la opción `GRANT` o tenga los siguientes permisos con la opción `GRANT`:
  - Para bases de datos de la misma cuenta AWS: `DESCRIBE`, `CREATE_TABLE`, `ALTER` y `DROP`
  - Para bases de datos de una cuenta externa: `DESCRIBE`, `CREATE_TABLE` y `ALTER`
  - Para tablas (y columnas): `DESCRIBE`, `ALTER`, `DROP`, `INSERT`, `SELECT` y `DELETE`

Además, la etiqueta LF y el recurso al que se está asignando deben estar en la misma cuenta AWS.

Para eliminar una etiqueta LF de un recurso del catálogo de datos, debe cumplir estos requisitos y, además, tener el permiso de IAM `lakeformation:RemoveLFTagsFromResource`.

Asignar etiquetas LF a una columna de una tabla

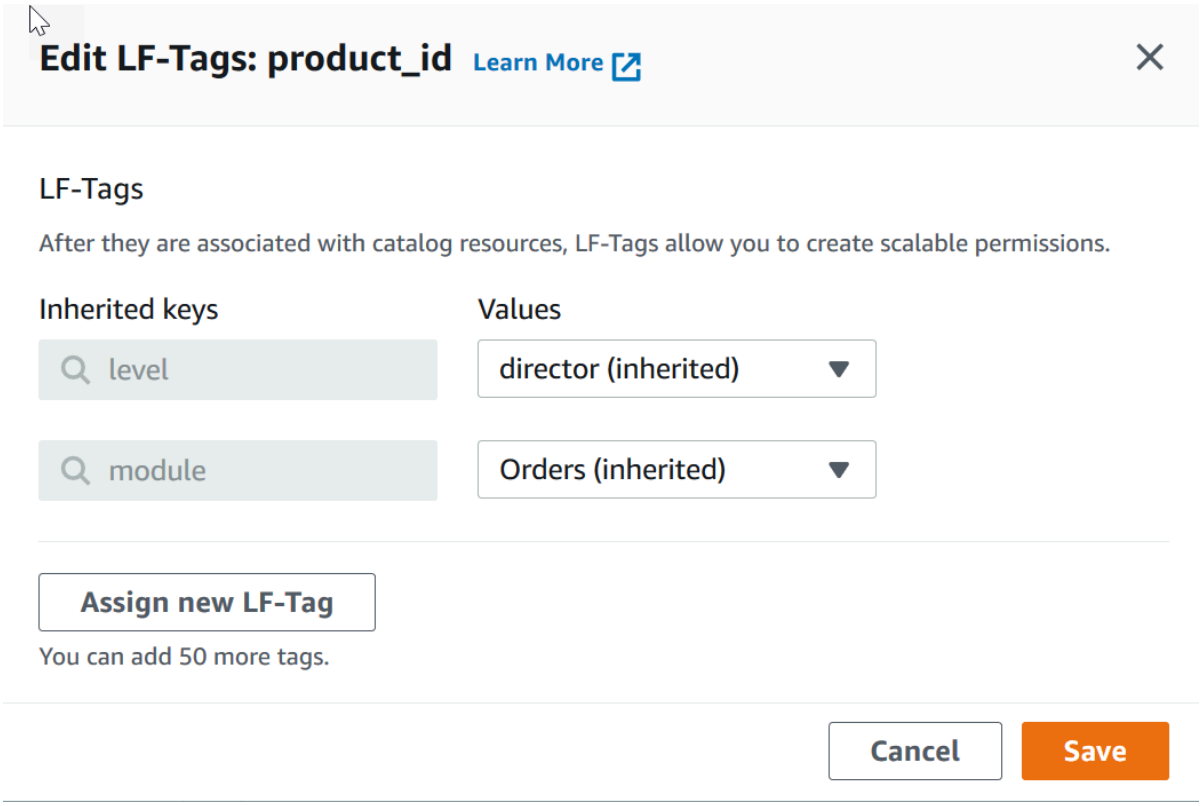
Para asignar etiquetas LF a una columna de una tabla (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

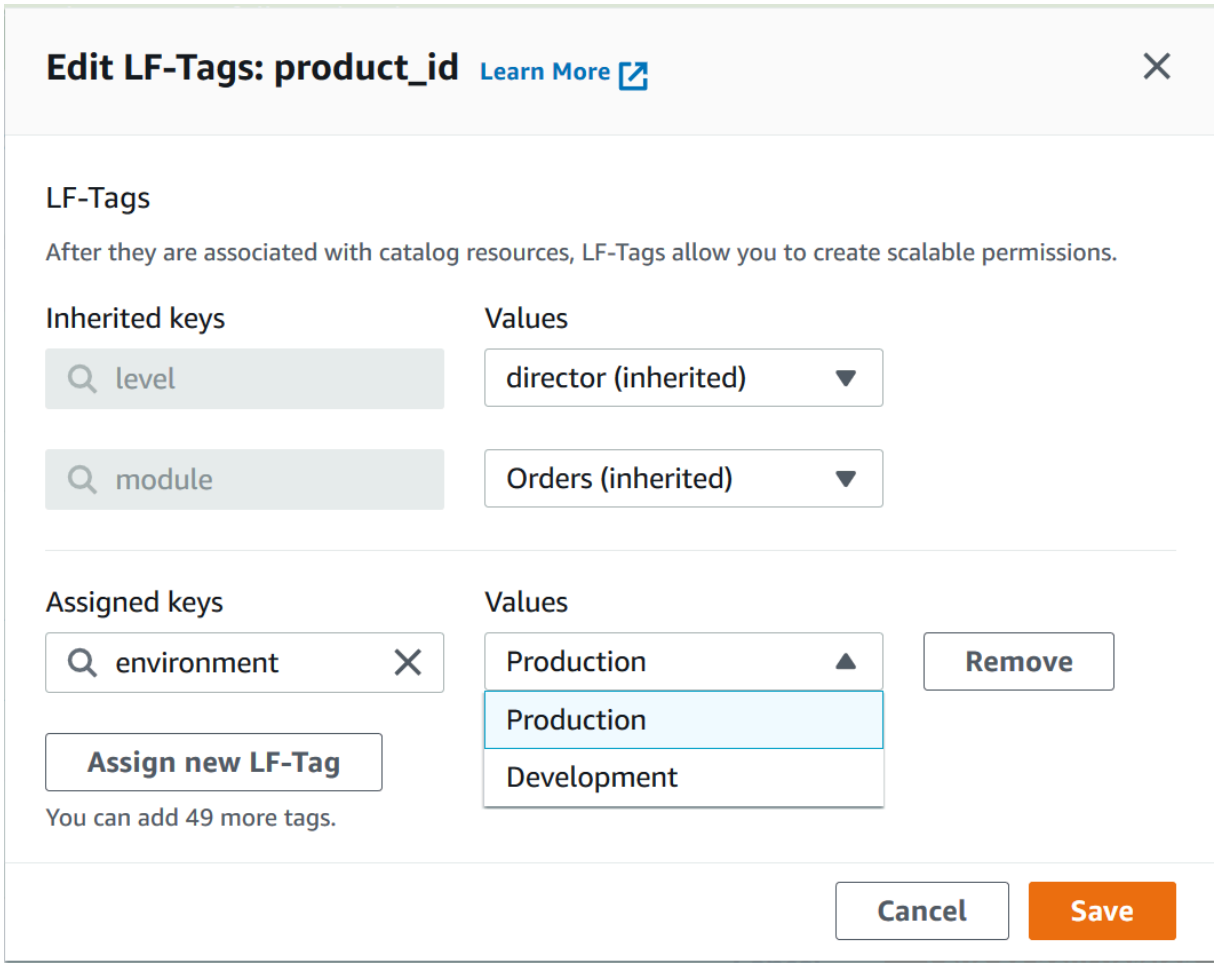
   Inicie sesión como usuario que cumpla con los requisitos enumerados anteriormente.
2. En el panel de navegación, elija Tablas.
3. Elija un nombre de tabla (no el botón de opción situado junto al nombre de tabla).
4. En la página de detalles de la tabla, en la sección Esquema, seleccione Editar esquema.
5. En la página Editar esquema, seleccione una o varias columnas y, a continuación, Editar etiquetas.

   > **Note**
   >
   > Si quiere añadir o eliminar columnas y guardar una nueva versión, hágalo en primer lugar. A continuación, edite las etiquetas LF.

   Aparece el cuadro de diálogo Editar etiquetas LF, en el que se muestran las etiquetas LF heredadas de la tabla.



6. (Opcional) En la lista de valores que hay junto a un campo de claves heredadas, seleccione un valor que sustituya al heredado.
7. (Opcional) Seleccione Asignar una nueva etiqueta LF. Después, en Teclas asignadas, seleccione una clave y, en Valores, elija un valor para ella.



8. (Opcional) Para añadir otra etiqueta LF, vuelva a elegir Asignar una nueva etiqueta LF.
9. Elija Guardar.

Asignación de etiquetas LF a un recurso del catálogo de datos

Console

Para asignar etiquetas LF a una base de datos o tabla del catálogo de datos

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

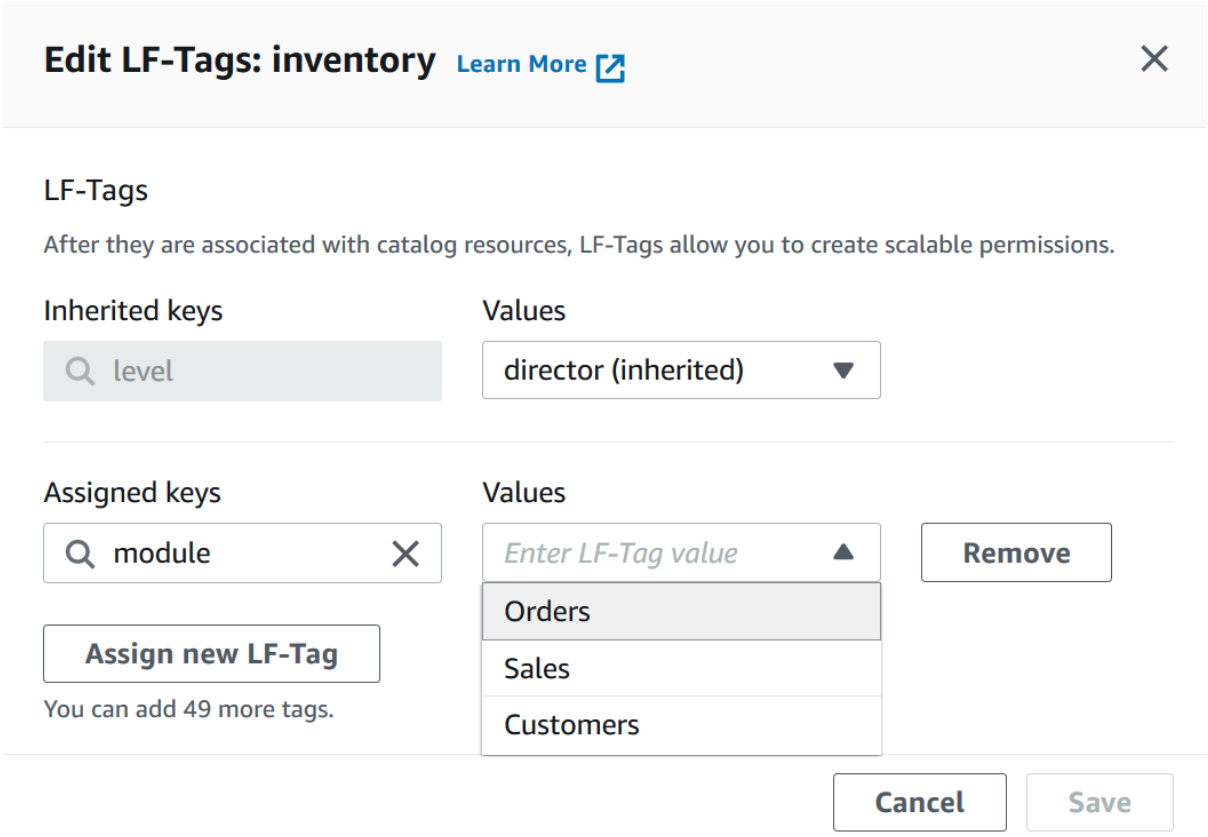
   Inicie sesión como usuario que cumple los requisitos enumerados anteriormente.
2. En el Catálogo de datos del panel de navegación, puede elegir entre:

   - Asignar etiquetas LF a las bases de datos, seleccione Bases de datos.
   - Asignar etiquetas LF a las tablas, seleccione Tablas.

3. Seleccione una base de datos o una tabla y, en el menú Acciones, seleccione Editar etiquetas.

   Aparece el cuadro de diálogo Editar etiquetas LF: ***nombre-del-recurso***.

   Si una tabla hereda las etiquetas LF de la base de datos que la contiene, la ventana muestra las etiquetas LF heredadas. De lo contrario, muestra el texto «No hay etiquetas LF heredadas asociadas al recurso».



4. (Opcional) Si una tabla ha heredado etiquetas LF, para la lista Valores junto a un campo de claves heredadas, puede elegir un valor para sobrescribir el valor heredado.
5. Para asignar nuevas etiquetas LF, siga estos pasos:

   a. Seleccione Asignar una nueva etiqueta LF.

   b. En el campo Claves asignadas, elija una clave de etiqueta LF y, en el campo Valores, elija un valor.

   c. (Opcional) Para añadir otra etiqueta LF, vuelva a elegir Asignar una nueva etiqueta LF.
6. Elija Guardar.

AWS CLI

Para asignar etiquetas LF a un recurso del catálogo de datos

- Ejecute el comando `add-lf-tags-to-resource`.

  En el siguiente ejemplo, se asigna la etiqueta LF `module=orders` a la tabla `orders` de la base de datos `erp`. Utiliza la sintaxis de atajos para el argumento `--lf-tags`. La propiedad `CatalogID` para `--lf-tags` es opcional. Si no se proporciona, se asume el ID de catálogo del recurso (en este caso, la tabla).

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":
 {"DatabaseName":"erp", "Name":"orders"}}' --lf-tags
 CatalogId=111122223333,TagKey=module,TagValues=orders
```

  Si el comando se ejecuta correctamente, verá el siguiente resultado:

```
{
    "Failures": []
}
```

  En el ejemplo siguiente, se asignan dos etiquetas LF a la tabla `sales` y la sintaxis JSON para el argumento `--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"]}]'
```

  En el ejemplo siguiente, se asigna la etiqueta LF `level=director` a la columna `total` de la tabla `sales`.

```
aws lakeformation add-lf-tags-to-resource --resource '{ "TableWithColumns":
 {"DatabaseName":"erp", "Name":"sales", "ColumnNames":["total"]}}' --lf-tags
 TagKey=level,TagValues=director
```

Actualización de etiquetas LF para un recurso

Para actualizar la etiqueta LF de un recurso del catálogo de datos (AWS CLI)

- Utilice el comando `add-lf-tags-to-resource`, según se describe en el procedimiento anterior.

  Al añadir una etiqueta LF con la misma clave que una etiqueta LF existente, pero con un valor diferente, se actualiza el valor existente.

Eliminación de una etiqueta LF de un recurso

Para eliminar la etiqueta LF de un recurso del catálogo de datos (AWS CLI)

- Ejecute el comando `remove-lf-tags-from-resource`.

  Si una tabla tiene un valor de etiqueta LF que anula el valor heredado de la base de datos principal, al eliminar esa etiqueta LF de la tabla se restaura el valor heredado. Este comportamiento también se aplica a una columna que anula los valores de clave heredados de la tabla.

  En el ejemplo siguiente, se elimina la etiqueta LF `level=director` de la columna `total` de la tabla `sales`. La propiedad `CatalogID` para `--lf-tags` es opcional. Si no se proporciona, se asume el ID de catálogo del recurso (en este caso, la tabla).

```
aws lakeformation remove-lf-tags-from-resource
--resource ' { "TableWithColumns":
{ "DatabaseName": "erp",  "Name": "sales",  "ColumnNames":[ "total"]}}'
--lf-tags  CatalogId=111122223333,TagKey=level,TagValues=director
```

Consulta de las etiquetas LF asignadas a un recurso.

Puede consultar las etiquetas LF asignadas a un recurso del catálogo de datos. Para ver la etiqueta LF, debe tener el permiso `DESCRIBE` o `ASSOCIATE` sobre ella.

Console

Para ver las etiquetas LF asignadas a un recurso (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

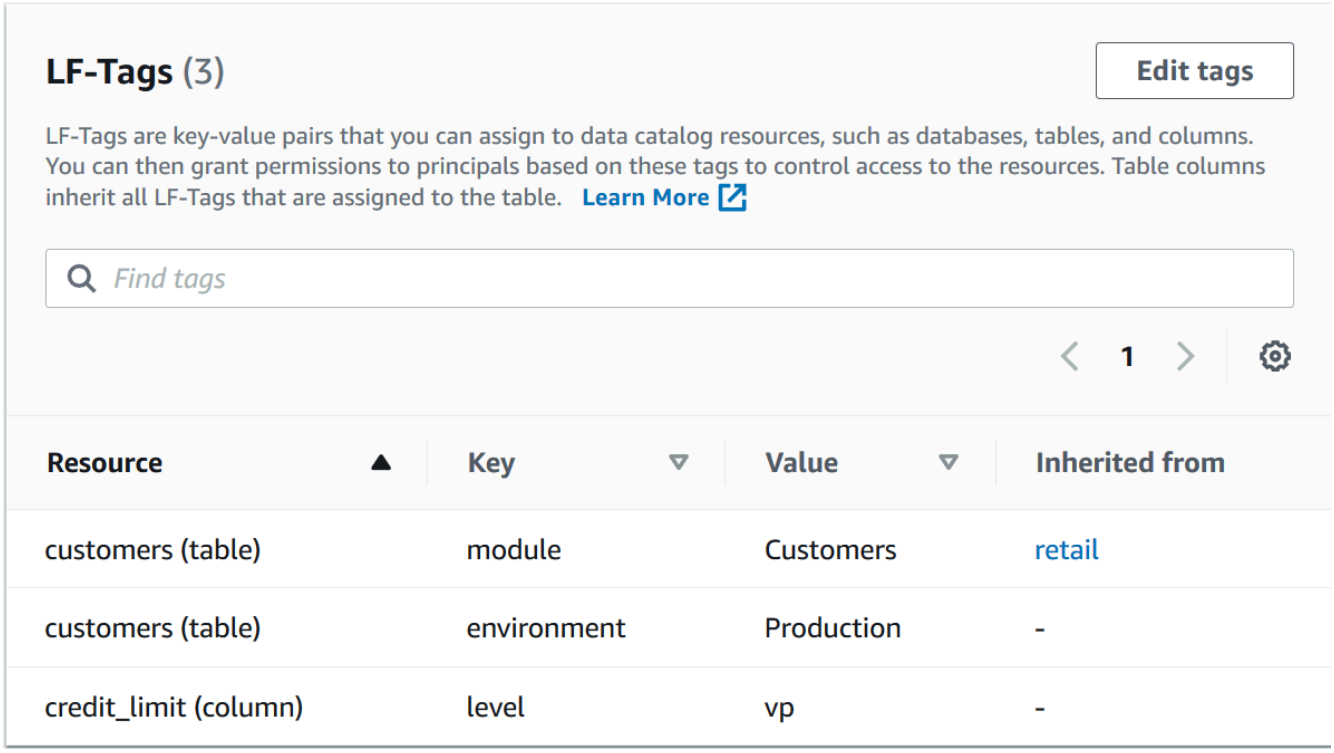
Inicie sesión como administrador del lago de datos, propietario del recurso o usuario al que se le hayan concedido permisos de Lake Formation sobre el recurso.

2. En el Catálogo de datos del panel de navegación, puede elegir entre:

   - Ver las etiquetas LF asignadas a una base de datos; para ello, seleccione Bases de datos.
   - Ver las etiquetas LF asignadas a una tabla; para ello, seleccione Tablas.

3. En la página Tablas o Bases de datos, elija el nombre de la base de datos o la tabla. En la página de detalles, desplácese hasta la sección Etiquetas LF.

   La siguiente captura de pantalla muestra las etiquetas LF asignadas a una tabla `customers`, incluida en la base de datos `retail`. La etiqueta LF `module` se hereda de la base de datos. La columna `credit_limit` tiene asignada la etiqueta LF `level=vp`.



AWS CLI

Para ver las etiquetas LF asignadas a un recurso (AWS CLI)

- Introduzca un comando similar al siguiente.

```
aws lakeformation get-resource-lf-tags --show-assigned-lf-tags --
resource '{ "Table": {"CatalogId":"111122223333", "DatabaseName":"erp",
 "Name":"sales"}}'
```

El comando devuelve el resultado siguiente.

```
{
    "TableTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "sales"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "environment",
            "TagValues": [
                "development"
            ]
        }
    ],
    "ColumnTags": [
        {
            "Name": "total",
            "Tags": [
                {
                    "CatalogId": "111122223333",
                    "TagKey": "level",
                    "TagValues": [
                        "director"
                    ]
                }
            ]
        }
    ]
}
```

Este resultado muestra solo las etiquetas LF asignadas de forma explícita, no las heredadas. Si desea ver todas las etiquetas LF de todas las columnas, incluidas las etiquetas LF heredadas, omita la opción `--show-assigned-lf-tags`.

Consulta de los recursos a los que está asignada una etiqueta LF

Puede ver todos los recursos del catálogo de datos a los que está asignada una clave de etiqueta LF concreta. Para ello, necesitará los siguientes permisos de Lake Formation:

- `Describe` o `Associate` sobre la etiqueta LF.
- `Describe` o cualquier otro sobre el recurso.

Además, necesitará los siguientes permisos AWS Identity and Access Management (IAM):

- `lakeformation:SearchDatabasesByLFTags`
- `lakeformation:SearchTablesByLFTags`

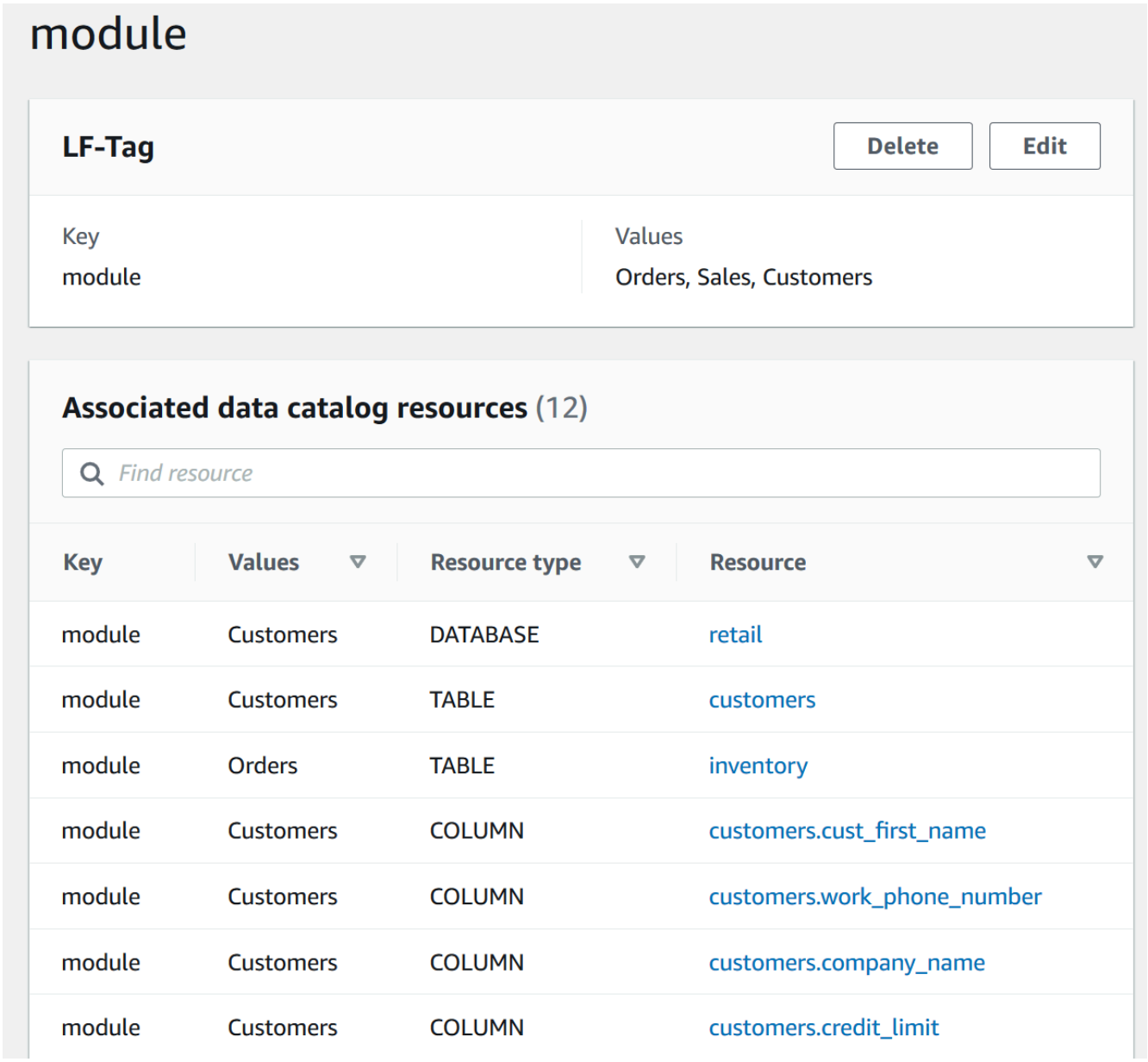
Console

Para ver los recursos a los que está asignada una etiqueta LF (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador de un lago de datos o como usuario que cumpla los requisitos enumerados anteriormente.
2. En el panel de navegación, bajo los encabezados Permisos y Roles y tareas administrativas, seleccione Etiquetas LF.
3. Seleccione una clave de etiqueta LF (no el botón de opción que hay junto al nombre de la clave).

   La página de detalles de la etiqueta LF muestra una lista de los recursos a los que se ha asignado la etiqueta LF.



AWS CLI

Para ver los recursos a los que está asignada una etiqueta LF

- Ejecute un comando `search-tables-by-lf-tags` o `search-databases-by-lf-tags`.

Example

En el siguiente ejemplo, se muestran las tablas y la columnas que tienen asignada la etiqueta LF `level=vp`. Para cada tabla y columna de la lista, se muestran todas las etiquetas LF asignadas a la tabla o columna, no solo la expresión de búsqueda.

```
aws lakeformation search-tables-by-lf-tags --expression
 TagKey=level,TagValues=vp
```

Para obtener más información sobre los permisos necesarios, consulte Personas de Lake Formation y referencia de permisos IAM.

Ciclo de vida de una etiqueta LF

1. Michael, el creador de la etiqueta LF, crea una etiqueta LF `module=Customers`.
2. A continuación, concede la etiqueta LF `Associate` al ingeniero de datos Eduardo. Al conceder `Associate` está otorgando `Describe` de manera implícita.
3. Michael concede `Super` sobre la tabla `Custs` a Eduardo con la opción de concesión, para que Eduardo pueda asignar etiquetas LF a la tabla. Para obtener más información, consulte Asignación de varias etiquetas LF a recursos del catálogo de datos.
4. Eduardo asigna la etiqueta LF `module=customers` a la tabla `Custs`.
5. Michael efectúa la siguiente concesión a la ingeniera de datos Sandra (en pseudocódigo).

   ```
   GRANT (SELECT, INSERT ON TABLES) ON TAGS module=customers TO Sandra WITH GRANT OPTION
   ```

6. Sandra hace la concesión siguiente a la analista de datos Maria.

   ```
   GRANT (SELECT ON TABLES) ON TAGS module=customers TO Maria
   ```

   Ahora María puede ejecutar consultas en la tabla `Custs`.

Véase también

- Control de acceso a los metadatos

Comparación del control de acceso basado en etiquetas de Lake Formation con el control de acceso basado en atributos de IAM

El control de acceso basado en atributos (ABAC) es una estrategia de autorización que define permisos en función de atributos. En AWS, estos atributos se denominan etiquetas. Puede asociar etiquetas a recursos de IAM, incluidas entidades de IAM (usuarios o roles) y recursos de AWS. Puede crear una única política ABAC o un conjunto pequeño de políticas para sus entidades principales de IAM. Estas políticas ABAC se pueden diseñar para permitir operaciones cuando la etiqueta de la entidad principal coincida con la etiqueta del recurso. ABAC es útil en entornos que crecen con rapidez y ayuda en situaciones en las que la administración de las políticas resulta engorrosa.

Los equipos de seguridad y gobernanza de la nube utilizan IAM para definir las políticas de acceso y los permisos de seguridad para todos los recursos, incluidos los buckets de Amazon S3, las instancias de Amazon EC2 y cualquier recurso al que pueda hacer referencia con un ARN. Las políticas de IAM definen permisos amplios (específicos) para los recursos de su lago de datos, por ejemplo, para permitir o denegar el acceso a nivel de bucket, prefijo o base de datos de Amazon S3. Para obtener más información sobre ABAC, consulte ¿Qué es ABAC para AWS¿ en la Guía del usuario de IAM.

Por ejemplo, puede crear tres roles con la clave de etiqueta `project-access`. Establezca el valor de etiqueta del primer rol en `Dev`, el segundo en `Support` y el tercero en `Marketing`. Asigne etiquetas con el valor adecuado a los recursos. A continuación, puede utilizar una única política que permita el acceso cuando el rol y el recurso estén etiquetados con el mismo valor para `project-access`.

Los equipos de gobierno de datos utilizan Lake Formation para definir permisos específicos para recursos de lagos de datos concretos. Las etiquetas LF se asignan a los recursos del catálogo de datos (bases de datos, tablas y columnas) y se otorgan a las entidades principales. Una entidad principal con etiquetas LF que coincidan con las etiquetas LF de un recurso puede acceder a ese recurso. Los permisos de Lake Formation son secundarios a los permisos de IAM. Por ejemplo, si los permisos de IAM no permiten a un usuario acceder a un lago de datos, Lake Formation no concede acceso a ningún recurso de ese lago de datos a ese usuario, aunque coincidan las etiquetas LF de la entidad principal y el recurso.

El control de acceso basado en etiquetas de Lake Formation (LF-TBAC) funciona con IAM ABAC para proporcionar niveles adicionales de permisos para sus datos y recursos de Lake Formation.

- Los permisos del TBAC de Lake Formation se escalan con innovación. Ya no es necesario que un administrador actualice las políticas existentes para permitir el acceso a nuevos recursos. Por ejemplo, supongamos que utiliza una estrategia ABAC de IAM con la etiqueta `project-access` para proporcionar acceso a bases de datos específicas de Lake Formation. Con LF-TBAC, la etiqueta LF `Project=SuperApp` se asigna a tablas o columnas específicas, y la misma etiqueta LF se otorga a un desarrollador para ese proyecto. A través de la IAM, el desarrollador puede acceder a la base de datos, y los permisos LF-TBAC conceden al desarrollador un mayor acceso a tablas o columnas específicas dentro de las tablas. Si se agrega una tabla nueva al proyecto, el administrador de Lake Formation solo necesita asignar la etiqueta a la nueva tabla para que el desarrollador tenga acceso a la tabla.
- TBAC de Lake Formation requiere menos políticas de IAM. Al utilizar las políticas de IAM para conceder un acceso de alto nivel a los recursos de Lake Formation y TBAC de Lake Formation para gestionar un acceso más preciso a los datos, crea menos políticas de IAM.
- Con el TBAC de Lake Formation, los equipos pueden cambiar y crecer rápidamente. Esto se debe a que los permisos para nuevos recursos se conceden automáticamente de acuerdo con los atributos. Por ejemplo, si un nuevo desarrollador se une al proyecto, es fácil concederle acceso asociando el rol de IAM al usuario y, a continuación, asignándole las etiquetas LF necesarias. No es necesario cambiar la política de IAM para respaldar un nuevo proyecto o para crear nuevas etiquetas LF.
- Se pueden obtener permisos más específicos con TBAC de Lake Formation. Las políticas de IAM otorgan acceso a los recursos de nivel superior, como las bases de datos o las tablas del catálogo de datos. Con TBAC de Lake Formation, puede conceder acceso a tablas o columnas específicas que contienen valores de datos específicos.

> **Note**
>
> Las etiquetas IAM no son lo mismo que las etiquetas LF. Estas etiquetas no son intercambiables. Las etiquetas LF se utilizan para conceder permisos de Lake Formation y las etiquetas de IAM se utilizan para definir las políticas de IAM.

Conceder, revocar y enumerar permisos de valores de etiqueta LF

Puede conceder los permisos `Drop`, `Alter` sobre las etiquetas LF a las entidades principales para administrar las expresiones de valor de las etiquetas LF. También puede conceder permisos `Describe`, `Associate` y `Grant with LF-Tag expressions` sobre las etiquetas LF a las

entidades principales para ver las etiquetas LF y asignarlas a los recursos del Catálogo de datos (bases de datos, tablas y columnas). Cuando se asignan etiquetas LF a los recursos del Catálogo de datos, puede utilizar el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC) para proteger dichos recursos. Para obtener más información, consulte Control de acceso basado en etiquetas de Lake Formation.

Puede conceder estos permisos con la opción de concesión para que otras entidades principales puedan otorgarlos. Los permisos `Grant with LF-Tag expressions`, `Describe` y `Associate` se explican en Añadir creadores de etiquetas LF.

Puede conceder los `Associate` permisos `Describe` y de una etiqueta LF a una cuenta externa AWS . Un administrador del lago de datos de esa cuenta puede entonces conceder esos permisos a otras entidades principales de la cuenta. Las entidades principales a las que el administrador del lago de datos de la cuenta externa conceda el permiso `Associate` podrán entonces asignar etiquetas LF a los recursos del Catálogo de datos que haya compartido con su cuenta.

Al hacer concesiones a una cuenta externa, debe incluir la opción de concesión.

Puede conceder permisos sobre etiquetas LF utilizando la consola de Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Temas

- Mostrar los permisos de las etiquetas LF mediante la consola
- Concesión de permisos de etiqueta LF desde la consola
- Otorgar, revocar y enumerar los permisos de la etiqueta LF mediante el AWS CLI

Para obtener más información, consulte Gestión de etiquetas LF para el control de acceso a los metadatos y Control de acceso basado en etiquetas de Lake Formation.

Mostrar los permisos de las etiquetas LF mediante la consola

Puede utilizar la consola de Lake Formation para ver los permisos concedidos sobre las etiquetas LF. Debe ser un creador de etiquetas LF, un administrador de lagos de datos o tener el permiso `Describe` o `Associate` sobre una etiqueta LF para verla.

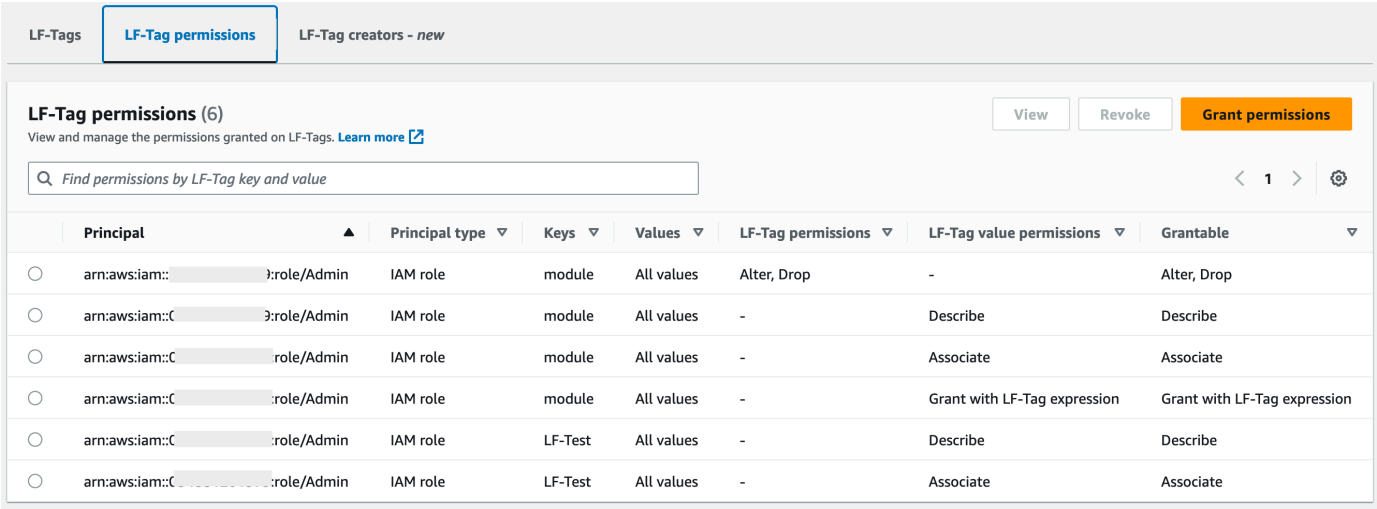
Para enumerar los permisos de las etiquetas LF (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

Inicie sesión como creador de etiquetas LF, administrador de un lago de datos o usuario al que se han concedido los permisos `Drop`, `Alter`, `Associate` o `Describe` sobre las etiquetas LF.

2. En el panel de navegación, en Permisos, seleccione Etiquetas LF y permisos, y elija la sección Permisos de etiquetas LF.

   La sección Permisos de etiquetas LF muestra una tabla que contiene la entidad principal, las claves de las etiquetas, los valores y los permisos.



Concesión de permisos de etiqueta LF desde la consola

Los pasos siguientes explican cómo conceder permisos a las etiquetas LF utilizando la página Conceder permisos a las etiquetas LF de la consola de Lake Formation. La página está dividida en las estas secciones:

- Tipos de permiso. El tipo de permiso que se va a conceder.
- Responsables: los usuarios, roles o AWS cuentas a los que se van a conceder permisos.
- Etiquetas LF. Las etiquetas LF sobre las que conceder permisos.
- Permisos. Los permisos que se van a conceder.

Abrir la página Conceder permisos de etiquetas LF

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

Inicie sesión como creador de etiquetas LF, como administrador de un lago de datos o como usuario al que se le han concedido permisos de etiquetas LF o permisos de par clave-valor de etiquetas LF en etiquetas LF con la opción `Grant`.

2. En el panel de navegación, seleccione Etiquetas LF y permisos, elija la sección Permisos de etiquetas LF.
3. Elija Grant permissions (Conceder permisos).

Especificar el tipo de permisos

En la sección Tipo de permisos, seleccione uno.

Permisos de etiqueta LF

Elija los permisos de las etiquetas LF para permitir a las entidades principales actualizar los valores de las etiquetas LF o eliminarlas.

Permisos de par clave-valor de etiquetas LF

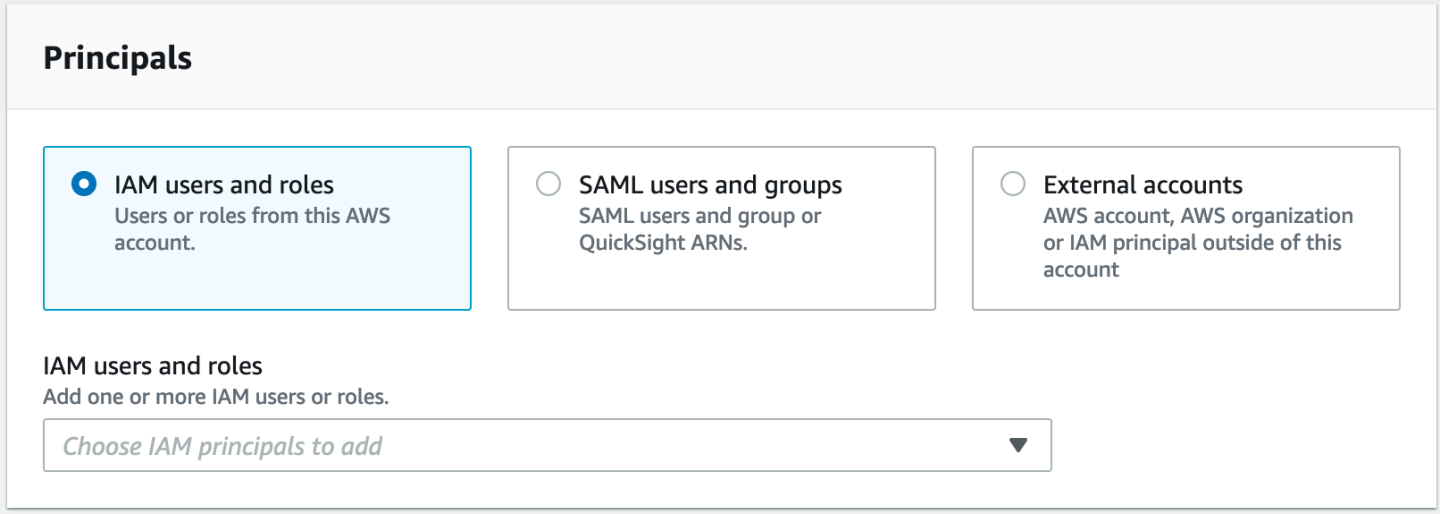
Elija los permisos de par clave-valor de etiquetas LF para permitir a las entidades principales asignar etiquetas LF a los recursos del Catálogo de datos, ver las etiquetas LF y sus valores, y conceder a las entidades principales permisos basados en etiquetas LF sobre los recursos del Catálogo de datos.

Las opciones disponibles en las siguientes secciones dependen del tipo de permiso.

Especificar las entidades principales

> **Note**
>
> No puede conceder permisos de etiqueta LF (`Alter` y `Drop`) a cuentas externas o a entidades principales de otra cuenta.

En la sección Entidades principales, elija uno de los tipos y, a continuación, especifique las que van a recibir los permisos concedidos.



Usuarios y roles de IAM

Elija uno o varios usuarios o roles en la lista de usuarios y roles de IAM.

Usuarios y grupos de SAML

Para QuickSight los usuarios y grupos de SAML y Amazon, introduzca uno o más nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de SAML, o ARN para los usuarios o grupos de Amazon. QuickSight Pulse Intro después de cada ARN.

Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

> **Note**
>
> La integración de Lake Formation con Amazon solo QuickSight es compatible con Amazon QuickSight Enterprise Edition.

Cuentas externas

Para la AWS cuenta, introduce uno o más identificadores de AWS cuenta válidos. Pulse Intro después de cada ID.

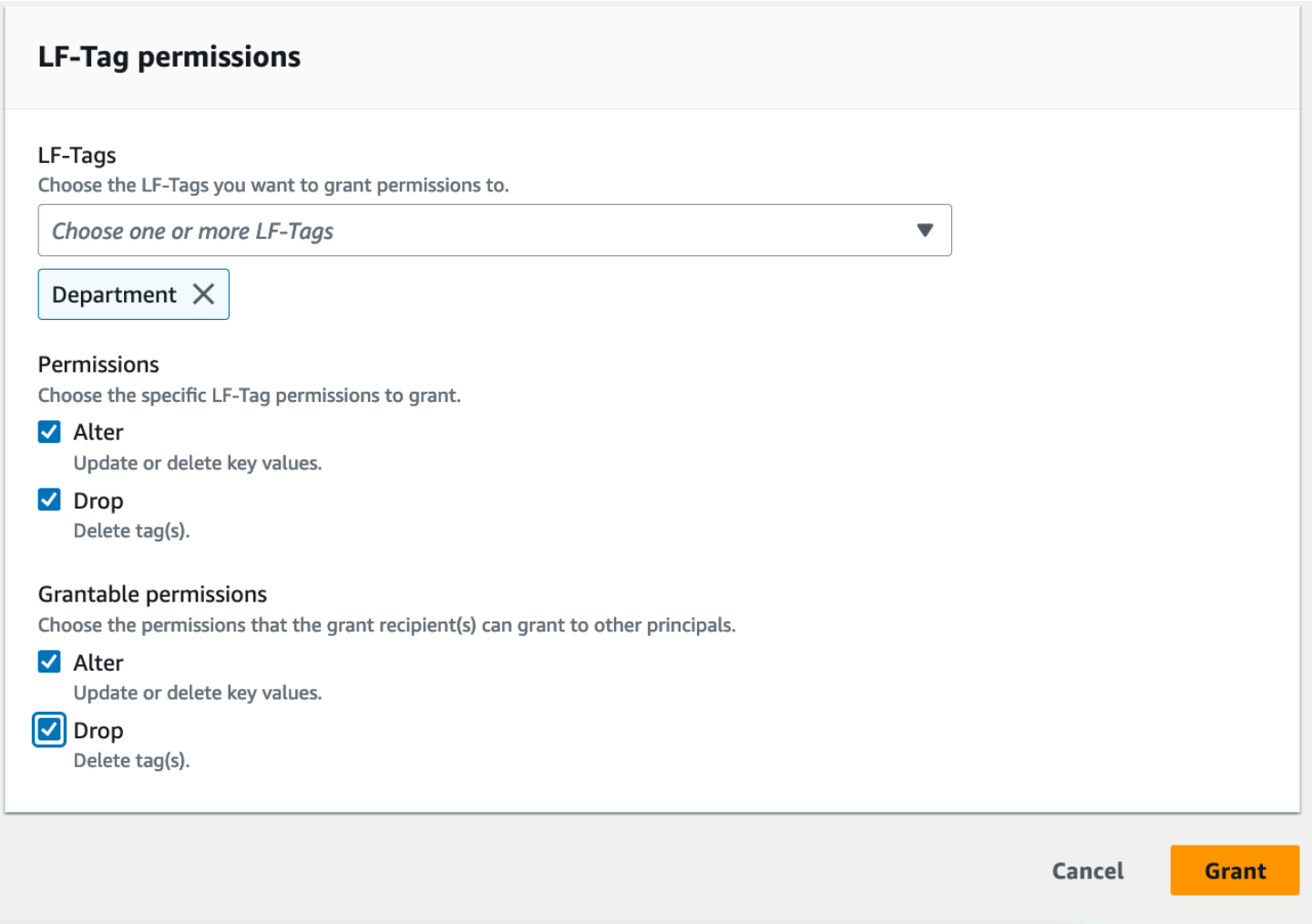
El ID de una organización está formado por una "o-" seguida de 10 a 32 letras minúsculas o dígitos.

Un ID de una unidad organizativa comienza por «ou-» seguidos de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo guion «-» y de 8 a 32 letras minúsculas o dígitos adicionales.

Para la entidad principal de IAM, escriba el ARN del usuario o el rol de IAM.

Especificar las etiquetas LF

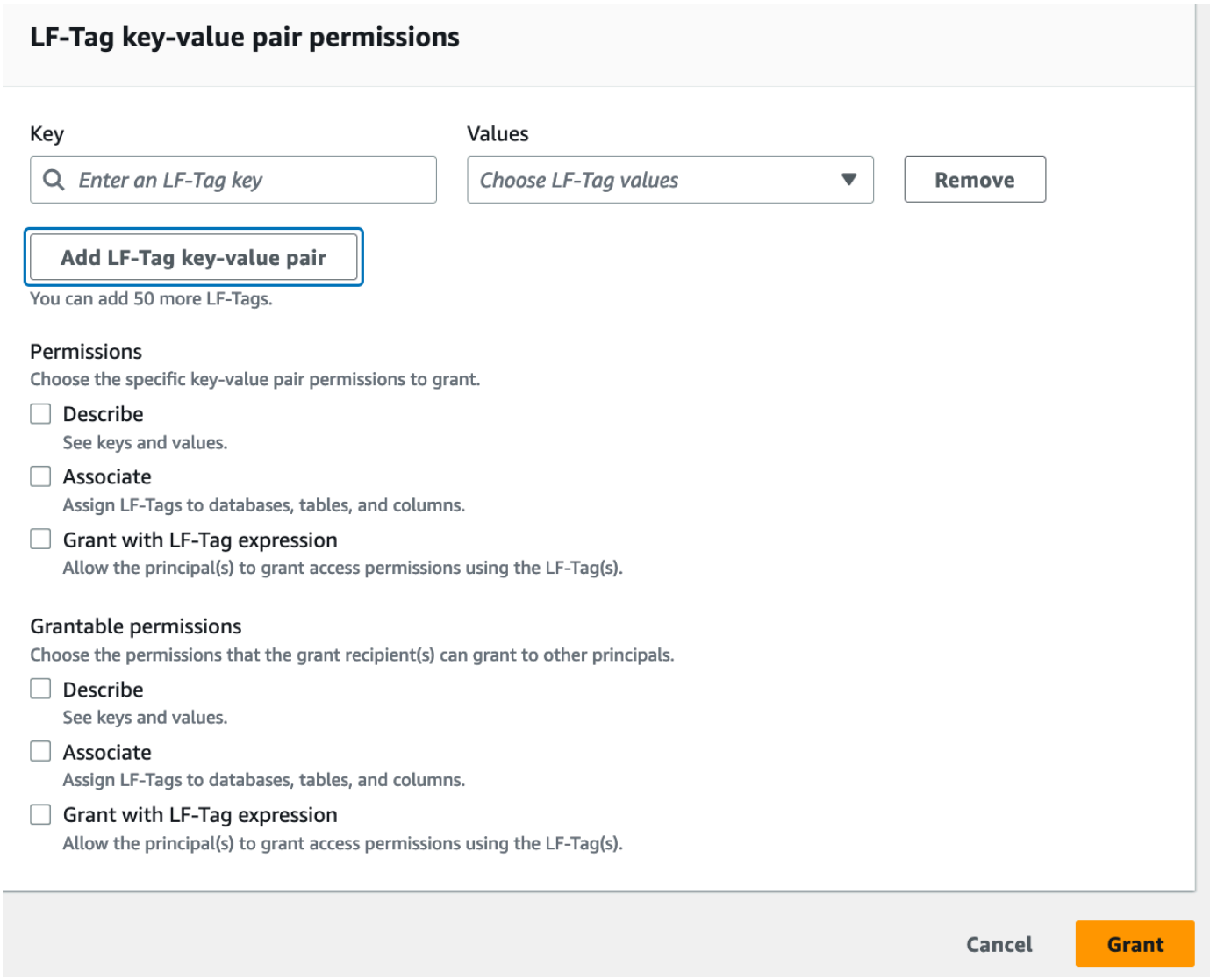
Para conceder permisos a las etiquetas LF, en la sección de permisos de las etiquetas LF, especifique las etiquetas LF sobre las que se van a conceder permisos.



- Elija una o más etiquetas LF en el menú desplegable.

Especificar los pares clave-valor de la etiqueta LF

1. Para conceder permisos sobre pares clave-valor de etiquetas LF, (primero debe elegir permisos de pares clave-valor de etiquetas LF como tipo de permiso) seleccione Añadir par clave-valor de etiquetas LF para revelar la primera fila de campos que especifica la clave y los valores de las etiquetas LF.



2. Sitúe el cursor en el campo Tecla, empiece a escribir —opcional— para delimitar la lista de selección, y seleccione una clave de etiqueta LF.
3. En la lista Valores, seleccione uno o varios y, a continuación, pulse la tecla Etiqueta, haga clic o pulse fuera del campo para guardar los valores seleccionados.

> **Note**
>
> Si una de las filas de la lista Valores tiene el foco, al pulsar Intro se selecciona o anula la selección de la casilla.

Los valores seleccionados aparecen como mosaicos debajo de la lista Valores. Elija la ✖ para eliminar un valor. Seleccione Eliminar para eliminar toda la etiqueta LF.

4. Para añadir otra etiqueta LF, vuelva a seleccionar Añadir etiqueta LF y repita los dos pasos anteriores.

Especificar los permisos

Esta sección muestra los permisos de etiquetas LF o los permisos de valores de etiquetas LF en función del tipo de permiso que haya elegido en el paso anterior.

En función del tipo de permiso que haya decidido conceder, seleccione los permisos LF-Tag o los permisos de par clave-valor etiqueta LF, y los permisos concedibles.

1. En Permisos para etiquetas LF, seleccione que va a conceder.

   Al conceder Eliminar y Modificar se concede implícitamente Describe.

   Debe conceder los permisos Alterar y Eliminar para todos los valores de las etiquetas.
2. En Permisos para valores clave de etiquetas LT, seleccione que desea conceder.

   Al conceder Asociar se concede implícitamente Describir. Seleccione Conceder con expresión de etiquetas LF para permitir que el destinatario conceda o revoque permisos de acceso a los recursos del Catálogo de datos utilizando el método LF-TBAC.
3. (Opcional) En Permisos concedibles, selecciona los permisos que el destinatario de la subvención puede conceder a otros directores de su cuenta. AWS
4. Elija Conceder.

Otorgar, revocar y enumerar los permisos de la etiqueta LF mediante el AWS CLI

Para conceder, revocar y publicar permisos sobre las etiquetas LF, utilice la tecla (). AWS Command Line Interface AWS CLI

Para enumerar los permisos de las etiquetas LF (AWS CLI)

- Introduzca un comando `list-permissions`. Debe ser el creador de etiquetas LF, un administrador de lagos de datos o tener el permiso `Drop`, `Alter`, `Describe`, `Associate` o `Grant with LF-Tag permissions` sobre una etiqueta LF para verla.

  El siguiente comando solicita todas las etiquetas LF sobre las que tiene permisos.

```
aws lakeformation list-permissions --resource-type LF_TAG
```

  A continuación se muestra un ejemplo de salida para un administrador de un lago de datos, que ve todas las etiquetas LF concedidas a todas las entidades principales. Los usuarios no administrativos solo ven las etiquetas LF que les han sido concedidas. Los permisos de etiquetas LF concedidos desde una cuenta externa aparecen en una página de resultados independiente. Para verlos, repita el comando y suministre el `--next-token` argumento con el token devuelto por la ejecución del comando anterior.

```
{
    "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"
}
```

Puede enumerar todas las concesiones para una clave de etiqueta LF específica. El siguiente comando devuelve todos los permisos concedidos en la etiqueta LF `module`.

```
aws lakeformation list-permissions --resource-type LF_TAG --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

También puede listar los valores de etiquetas LF concedidos a una entidad principal específica para una etiqueta LF concreta. Al proporcionar el argumento `--principal`, debe indicar el argumento `--resource`. Por lo tanto, el comando solo puede solicitar de forma efectiva los valores concedidos a una entidad principal específica para una clave de etiquetas LF concreta. El siguiente comando muestra cómo hacerlo para la entidad principal `datalake_user1` y la clave de etiquetas LF `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":["*"]}}'
```

A continuación, se muestra un ejemplo del resultado.

```
{
    "PrincipalResourcePermissions": [
        {
            "Principal": {
                "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
            },
            "Resource": {
                "LFTag": {
                    "CatalogId": "111122223333",
                    "TagKey": "module",
                    "TagValues": [
                        "Orders",
                        "Sales"
                    ]
                }
            },
            "Permissions": [
                "ASSOCIATE"
            ],
            "PermissionsWithGrantOption": []
        }
    ]
}
```

Para conceder permisos a las etiquetas LF ()AWS CLI

1. Introduzca un comando similar al siguiente. En este ejemplo se concede al usuario `datalake_user1` el permiso `Associate` sobre las etiquetas LF con la clave `module`. Concede permisos para ver y asignar todos los valores de esa clave, como indica el asterisco (*).

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
```

```
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Al conceder el permiso `Associate`, se otorga el permiso `Describe` de forma implícita.

En el siguiente ejemplo, se conceden `Associate` a la AWS cuenta externa 1234-5678-9012 de la etiqueta LF con la clave, con la opción de concesión. `module` Otorga permisos para ver y asignar solo los valores `sales` y `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. Al conceder el permiso `GrantWithLFTagExpression`, se otorga el permiso `Describe` de forma implícita.

   El siguiente ejemplo concede `GrantWithLFTagExpression` a un usuario en las etiquetas LF con la clave `module`, con la opción de concesión. Concede permisos para ver y conceder permisos sobre los recursos del Catálogo de datos utilizando solo los valores `sales` y `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. El siguiente ejemplo concede permisos `Drop` a un usuario en las etiquetas LF con la clave `module`, con la opción de concesión. Concede permisos para eliminar las etiquetas LF. Para borrar una etiqueta LF, necesita permisos sobre todos los valores de esa clave.

   ```
   aws lakeformation grant-permissions --principal
    DataLakePrincipalIdentifier=111122223333 --permissions "DROP"
    --permissions-with-grant-option "DROP" --resource '{ "LFTag":
    {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
   ```

4. El siguiente ejemplo concede permisos `Alter` al usuario sobre las etiquetas LF con la clave `module`, con la opción de concesión. Concede permisos para eliminar las etiquetas LF. Para actualizar una etiqueta LF, necesita permisos sobre todos los valores de esa clave.

   ```
   aws lakeformation grant-permissions --principal
    DataLakePrincipalIdentifier=111122223333 --permissions "ALTER"
   ```

```
 --permissions-with-grant-option "ALTER" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Para revocar permisos sobre las etiquetas LF (AWS CLI)

- Introduzca un comando similar al siguiente. En este ejemplo se revoca el permiso `Associate` sobre las etiquetas LF con la clave `module` al usuario `datalake_user1`.

```
aws lakeformation revoke-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

Conceder permisos de lago de datos mediante el método LF-TBAC

Puede conceder los permisos `DESCRIBE` y `ASSOCIATE` de Lake Formation sobre las etiquetas LF a las entidades principales para que puedan ver las etiquetas LF y asignarlas a los recursos del catálogo de datos (bases de datos, tablas, vistas y columnas). Cuando se asignan etiquetas LF a los recursos del catálogo de datos, puede utilizar el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC) para proteger dichos recursos. Para obtener más información, consulte Control de acceso basado en etiquetas de Lake Formation.

Al principio, solo el administrador del lago de datos puede conceder estos permisos. Si el administrador del lago de datos concede estos permisos con la opción de concesión, otras entidades principales podrán otorgarlos. Los permisos `DESCRIBE` y `ASSOCIATE` se explican en Prácticas recomendadas y consideraciones sobre el control de acceso basado en etiquetas de Lake Formation.

Puede conceder los permisos `DESCRIBE` y `ASSOCIATE` de una etiqueta LF a una cuenta externa AWS. Un administrador del lago de datos de esa cuenta puede entonces conceder esos permisos a otras entidades principales de la cuenta. Las entidades principales a las que el administrador del lago de datos de la cuenta externa conceda el permiso `ASSOCIATE` podrán entonces asignar etiquetas LF a los recursos del catálogo de datos que haya compartido con su cuenta.

Al hacer concesiones a una cuenta externa, debe incluir la opción de concesión.

Puede conceder permisos sobre etiquetas LF utilizando la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Temas

Véase también

- Conceder, revocar y enumerar permisos de valores de etiqueta LF
- Gestión de etiquetas LF para el control de acceso a los metadatos
- Control de acceso basado en etiquetas de Lake Formation

Concesión de permisos del catálogo de datos

Utilice la consola de Lake Formation o la AWS CLI para conceder permisos de Lake Formation en bases de datos, tablas y columnas del catálogo de datos con el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC).

Console

Los siguientes pasos explican cómo conceder permisos utilizando el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC) y la página Conceder permisos de lago de datos en la consola de Lake Formation. La página está dividida en las secciones siguientes:

- Entidades principales. Los usuarios, roles y las Cuentas de AWS a las que conceder permisos.
- Etiquetas LF o recursos del catálogo. Bases de datos, tablas o enlaces a recursos sobre los que se conceden los permisos.
- Permisos. Los permisos de Lake Formation que se conceden.

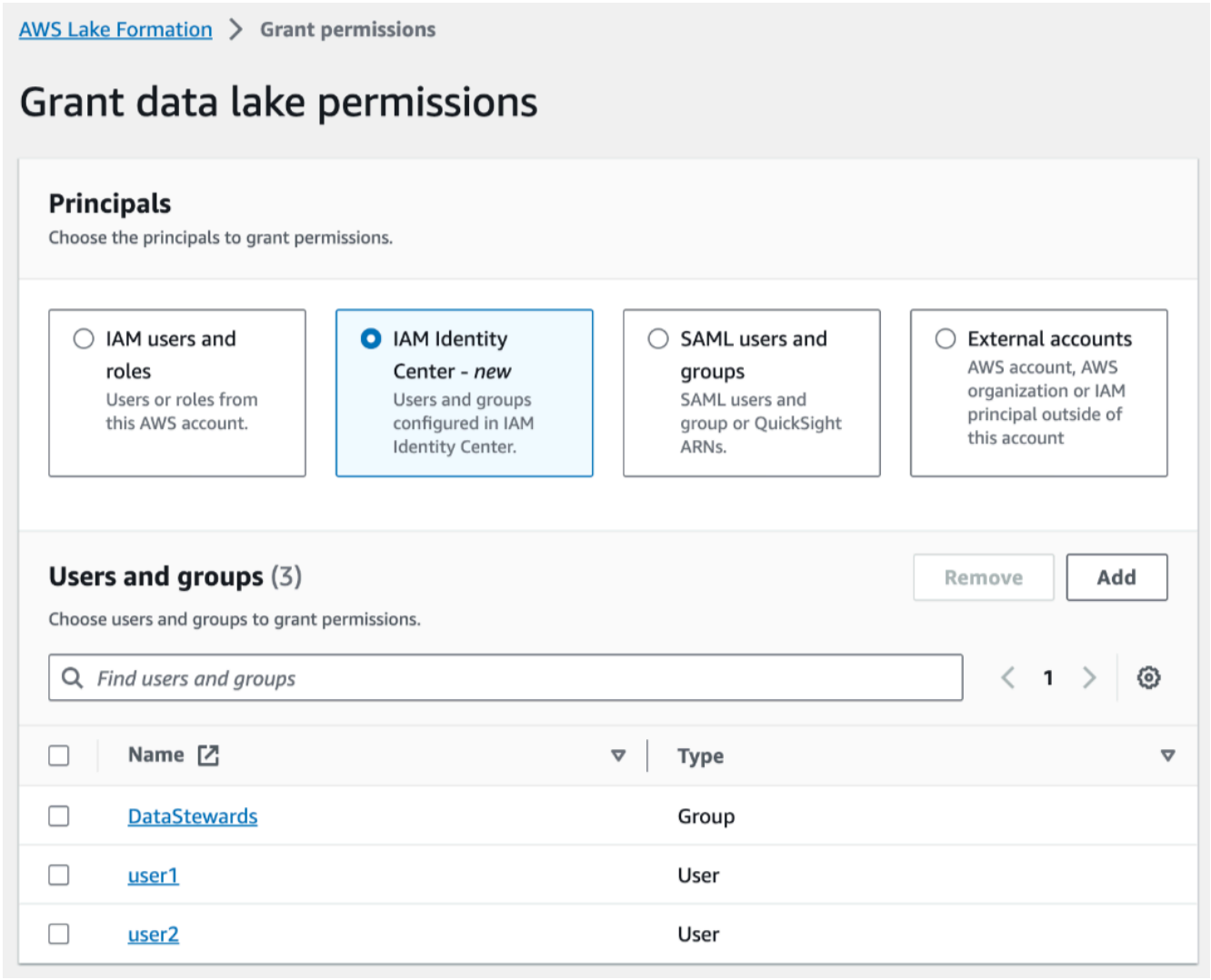
1. Abra la página Conceder permisos de lagos de datos.

   Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/, e inicie sesión como administrador del lago de datos o como usuario al que se le han concedido permisos de Lake Formation sobre los recursos del catálogo de datos mediante LF-TBAC con la opción de concesión.

   En el panel de navegación, en Permisos, seleccione Permisos de lago de datos. A continuación, seleccione Conceder.

2. Especifique las entidades principales.

   En la sección Entidades principales, elija uno de los tipos y, a continuación, especifique las que van a recibir los permisos concedidos.



   Usuarios y roles de IAM

   Elija uno o varios usuarios o roles en la lista de usuarios y roles de IAM.

   IAM Identity Center

   Elija uno o varios usuarios o en la lista de Usuarios y grupos.

   Usuarios y grupos de SAML

   Para los usuarios y grupos de SAML y Amazon QuickSight, introduzca uno o varios nombres de recursos de Amazon (ARN) para los usuarios o grupos federados a través de

SAML, o ARN para los usuarios o grupos de Amazon QuickSight. Pulse Intro después de cada ARN.

Para obtener información sobre cómo crear ARN, consulte Lake Formation otorga y revoca órdenes AWS CLI.

> **Note**
>
> La integración de Lake Formation con Amazon QuickSight solo es compatible con Amazon QuickSight Enterprise Edition.

Cuentas externas

Para Cuentas de AWS, organización de AWS o entidad principal de IAM, introduzca uno o varios ID de Cuenta de AWS, ID de organización, ID de unidad organizativa o ARN válidos para el usuario o rol de IAM. Pulse Intro después de cada ID.

El ID de una organización está formado por una "o-" seguida de 10 a 32 letras minúsculas o dígitos.

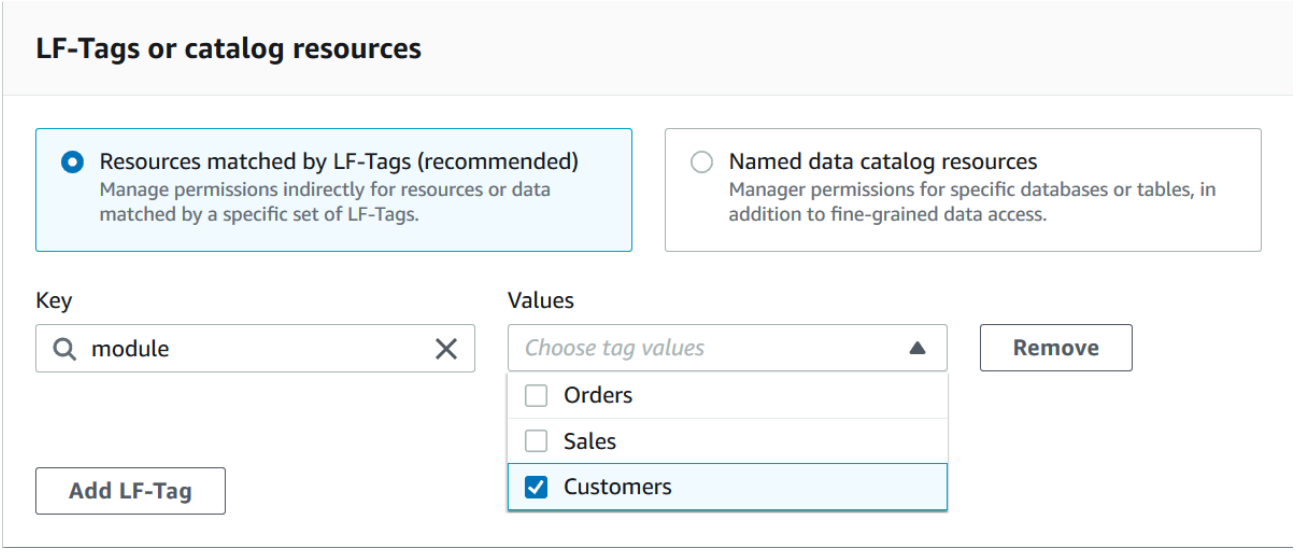
Un ID de una unidad organizativa comienza por «ou-» seguidos de 4 a 32 letras minúsculas o dígitos (el ID de la raíz que contiene la UO). Esta cadena va seguida de un segundo guion «-» y de 8 a 32 letras minúsculas o dígitos adicionales.

3. Especifique las etiquetas LF.

Asegúrese de elegir la opción Recursos que coinciden con las etiquetas LF. Seleccione Añadir etiqueta LF.

1. Elija una clave y valores de etiquetas LF.

Si elige más de un valor, está creando una expresión de etiqueta LF con un operador `OR`. Esto significa que si alguno de los valores de las etiquetas LF coincide con una etiqueta LF asignada a un recurso del catálogo de datos, se le conceden permisos sobre el recurso.



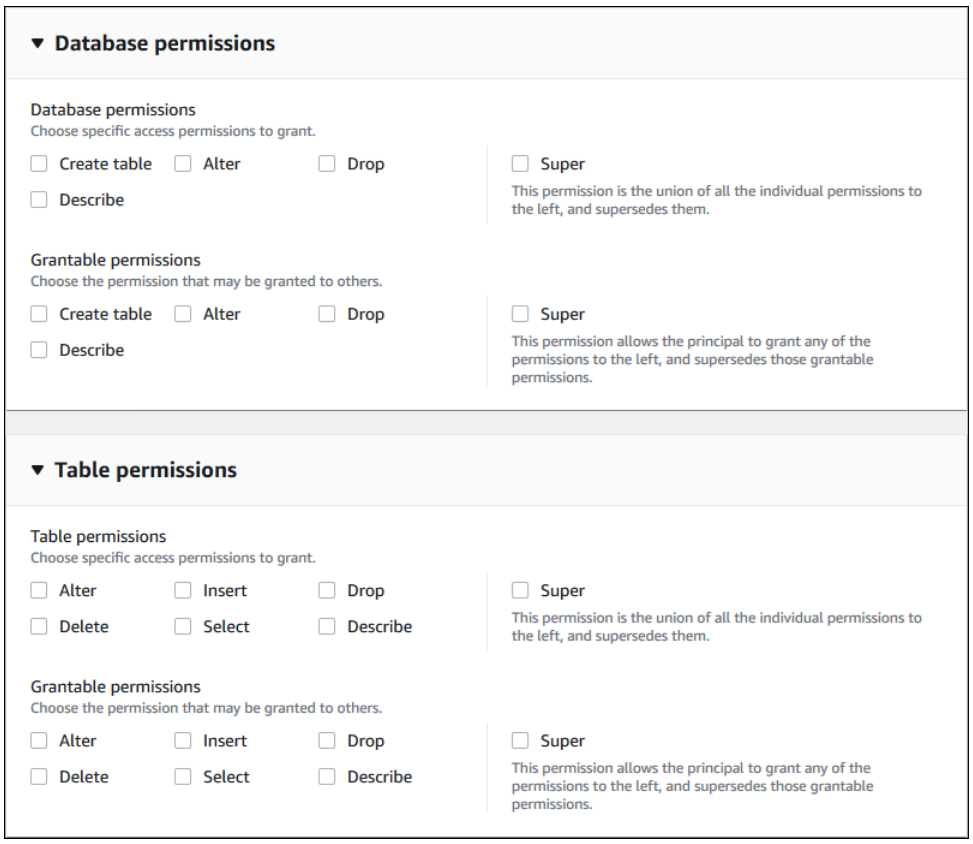
2. (Opcional) Seleccione Añadir etiqueta LF de nuevo para especificar otra etiqueta LF.

   Si elige más de un valor, está creando una expresión de etiqueta LF con un operador `AND`. A la entidad principal se le conceden permisos sobre un recurso del catálogo de datos solo si al recurso se le asignó una etiqueta LF coincidente para cada etiqueta LF en la expresión de la etiqueta LF.

4. Especifique los permisos.

   Especifique los permisos que se conceden a la entidad principal sobre los recursos coincidentes del catálogo de datos. Los recursos coincidentes son aquellos a los que se asignaron etiquetas LF que coinciden con una de las expresiones de etiquetas LF concedidas a la entidad principal.

   Puede especificar los permisos que se van a conceder sobre las bases de datos coincidentes, las tablas coincidentes y las vistas coincidentes.



En Permisos de base de datos, seleccione los permisos de base de datos que se concederán a la entidad principal en las bases de datos coincidentes.

En Permisos de tabla, seleccione los que se concederán a la entidad principal en las tablas y vistas coincidentes.

También puede elegir `Select`,`Describe` y permisos de `Drop` en Permisos de tabla y aplicarlos a las vistas.

5. Elija Conceder.

AWS CLI

Puede utilizar la AWS Command Line Interface (AWS CLI) y el método de control de acceso basado en etiquetas de Lake Formation (LF-TBAC) para conceder permisos de Lake Formation en bases de datos, tablas y columnas del catálogo de datos.

Conceder permisos de lago de datos mediante AWS CLI y el método LF-TBAC

- Utilice el comando `grant-permissions`.

Example

El siguiente ejemplo concede al usuario `datalake_user1` la expresión de etiquetas LF "`module=*`" (todos los valores de la clave de etiquetas LF `module`). Ese usuario tendrá el permiso `CREATE_TABLE` sobre todas las bases de datos coincidentes —bases de datos a las que se hayan asignado las etiquetas LF con la clave `module`, con cualquier valor.

```
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

El siguiente ejemplo otorga la expresión "`(level=director) AND (region=west OR region=south)`" de etiqueta LF al usuario `datalake_user1`. Ese usuario tendrá los permisos `SELECT`, `ALTER`, y `DROP` con la opción de concesión en las tablas que coincidan (tablas a las que se les haya asignado tanto `level=director` como (`region=west` o `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

En el siguiente ejemplo se concede la expresión de etiquetas LF "`module=orders`" a la cuenta AWS 1234-5678-9012. El administrador del lago de datos de esa cuenta puede entonces conceder la expresión "`module=orders`" a las entidades principales de su cuenta. Esas entidades principales tendrán entonces permiso `CREATE_TABLE` para hacer coincidir las bases de datos propiedad de la cuenta 1111-2222-3333 y compartidas con la cuenta 1234-5678-9012 utilizando el método de recursos con nombre o el método 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"]}]}}'
```

Ejemplo de escenario de permisos

El siguiente escenario ayuda a demostrar cómo se pueden configurar los permisos para proteger el acceso a los datos AWS Lake Formation.

Shirley es administradora de datos. Quiere configurar un lago de datos para su empresa, AnyCompany. Ahora mismo, todos los datos se almacenan en Amazon S3. John es responsable de marketing y necesita acceso de escritura a la información de compra de los clientes (contenida en `s3://customerPurchases`). Un analista de marketing, Diego, se une a John este verano. John necesita poder conceder a Diego acceso para consultar los datos sin involucrar a Shirley.

Mateo, de finanzas, necesita acceso para consultar datos contables (por ejemplo, `s3://transactions`). Quiere consultar los datos de las transacciones en las tablas de una base de datos (`Finance_DB`) utilizadas por el equipo financiero. Su jefe, Arnav, puede darle acceso a la `Finance_DB`. Aunque no debería poder modificar los datos contables, necesita la capacidad de convertir los datos a un formato (esquema) adecuado para la previsión. Estos datos se almacenarán en un bucket separado (`s3://financeForecasts`) que él podrá modificar.

Para resumir:

- Shirley es la administradora del lago de datos.
- John necesita permiso `CREATE_DATABASE` y `CREATE_TABLE` para crear nuevas bases de datos y tablas en el catálogo de datos.
- John también necesita los permisos `SELECT`, `INSERT` y `DELETE` en las tablas que crea.
- Diego necesita permiso `SELECT` sobre la tabla para ejecutar consultas.

Los empleados de AnyCompany proceden de la siguiente manera para configurar los permisos. Las operaciones de la API mostradas en este escenario muestran una sintaxis simplificada para mayor claridad.

1. Shirley registra en Lake Formation la ruta de Amazon S3 que contiene la información de compra del cliente.

```
RegisterResource(ResourcePath("s3://customerPurchases"), false, Role_ARN )
```

2. Shirley concede a John acceso a la ruta de Amazon S3 que contiene la información de compra del cliente.

```
GrantPermissions(John, S3Location("s3://customerPurchases"),
 [DATA_LOCATION_ACCESS]) )
```

3. Shirley concede a John permiso para crear bases de datos.

```
GrantPermissions(John, catalog, [CREATE_DATABASE])
```

4. John crea la base de datos `John_DB`. John tiene automáticamente permiso `CREATE_TABLE` sobre esa base de datos porque él la creó.

```
CreateDatabase(John_DB)
```

5. John crea la tabla `John_Table` que apunta a `s3://customerPurchases`. Como él creó la tabla, tiene todos los permisos sobre ella y puede conceder permisos sobre ella.

```
CreateTable(John_DB, John_Table)
```

6. Juan permite a su analista, Diego, acceder a la tabla `John_Table`.

```
 GrantPermissions(Diego, John_Table, [SELECT])
```

7. Juan permite a su analista, Diego, el acceso a `s3://customerPurchases/London/`. Porque Shirley ya registró `s3://customerPurchases`, sus subcarpetas están registradas con Lake Formation.

```
 GrantDataLakePrivileges( 123456789012/datalake, Diego, [DATA_LOCATION_ACCESS], [],
 S3Location("s3://customerPurchases/London/") )
```

8. John permite a su analista, Diego, crear tablas en la base de datos `John_DB`.

```
 GrantDataLakePrivileges( 123456789012/datalake, Diego, John_DB, [CREATE_TABLE],
 [] )
```

9. Diego crea una tabla en `John_DB` en `s3://customerPurchases/London/` y obtiene automáticamente permisos `ALTER`, `DROP`, `SELECT`, `INSERT` y `DELETE`.

```
CreateTable( 123456789012/datalake, John_DB, Diego_Table )
```

Filtrado de datos y seguridad de celda en Lake Formation

Al conceder permisos de Lake Formation en una tabla del catálogo de datos, puede incluir especificaciones de filtrado de datos para restringir el acceso a determinados datos en los resultados de las consultas y en los motores integrados con Lake Formation. Lake Formation utiliza el filtrado de datos para ofrecer una seguridad a nivel de columna, fila y celda. Puede definir y aplicar filtros de datos en las columnas anidadas si los datos de origen contienen estructuras anidadas.

Temas

Información general del filtrado de datos

Con las capacidades de filtrado de datos de Lake Formation, puede implementar los siguientes niveles de seguridad de datos.

Seguridad a nivel de columna

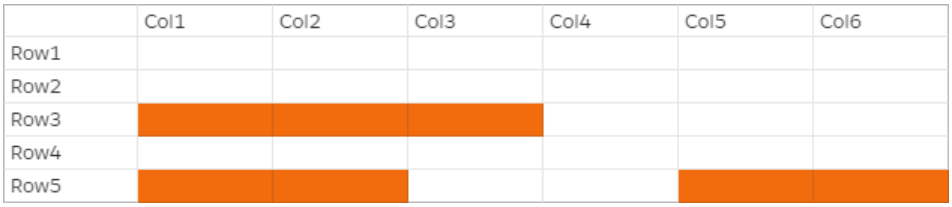
La concesión de permisos en una tabla del catálogo de datos con seguridad a nivel de columna (filtrado de columnas) permite a los usuarios ver solo las columnas específicas y columnas anidadas a las que tienen acceso en la tabla. Considere una tabla `persons` que se utiliza en varias aplicaciones para una gran empresa de comunicaciones multirregional. La concesión de permisos en las tablas del catálogo de datos con filtrado de columnas puede impedir que los usuarios que no trabajan en el departamento de Recursos Humanos vean información de identificación personal (PII), por ejemplo, un número de seguro social o fecha de nacimiento. También puede definir políticas de seguridad y conceder acceso solo a subestructuras parciales de las columnas anidadas.

Seguridad de nivel básico

La concesión de permisos en una tabla del catálogo de datos con seguridad a nivel de fila (filtrado de filas) permite a los usuarios ver solo las filas específicas de datos a las que tienen acceso en la tabla. El filtrado se basa en los valores de una o más columnas. Puede incluir estructuras de columnas anidadas al definir las expresiones de filtro de filas. Por ejemplo, si las diferentes oficinas regionales de la empresa de comunicaciones tienen sus propios departamentos de Recursos Humanos, puede limitar los registros de personas que los empleados de RH pueden ver a solo los registros de los empleados de su región.

Seguridad a nivel de celda

La seguridad a nivel de celda combina el filtrado de filas y el filtrado de columnas para lograr un modelo de permisos muy flexible. Si ve las filas y columnas de una tabla como una cuadrícula, mediante el uso de la seguridad a nivel de celda, puede restringir el acceso a los elementos individuales (celdas) de la cuadrícula en cualquier parte de las dos dimensiones. Es decir, puede restringir el acceso a diferentes columnas en función de la fila. Esto se ilustra en el siguiente diagrama, en el que las columnas restringidas aparecen sombreadas.



Siguiendo con el ejemplo de la tabla de personas, puede crear un filtro de datos a nivel de celda que restrinja el acceso a la columna de direcciones si la fila tiene la columna de país establecida en «Reino Unido», pero que permita el acceso a la columna de direcciones si la fila tiene la columna de país como «EE. UU.».

Los filtros se aplican solo a las operaciones de lectura. Por lo tanto, solo puede conceder el permiso de Lake Formation `SELECT` con filtros.

Seguridad a nivel de celda en las columnas anidadas

Lake Formation le permite definir y aplicar filtros de datos con seguridad a nivel de celda en columnas anidadas. Sin embargo, los motores analíticos integrados, como Amazon Athena, Amazon EMR y Amazon Redshift Spectrum, permiten ejecutar consultas en tablas anidadas administradas de Lake Formation con seguridad a nivel de filas y columnas.

Para conocer las limitaciones, consulte Limitaciones de filtrado de datos.

Filtros de datos en Lake Formation

Puede implementar seguridad a nivel de columna, de fila y de celda mediante el uso de filtros de datos. Selecciona un filtro de datos al conceder el permiso de Lake Formation de `SELECT` sobre las tablas. Si la tabla contiene estructuras de columnas anidadas, puede definir un filtro de datos incluyendo o excluyendo las columnas secundarias y definir expresiones de filtro a nivel de fila en los atributos anidados.

Cada filtro de celdas de datos pertenece a una tabla específica de un catálogo de datos. Un filtro de datos incluye la siguiente información:

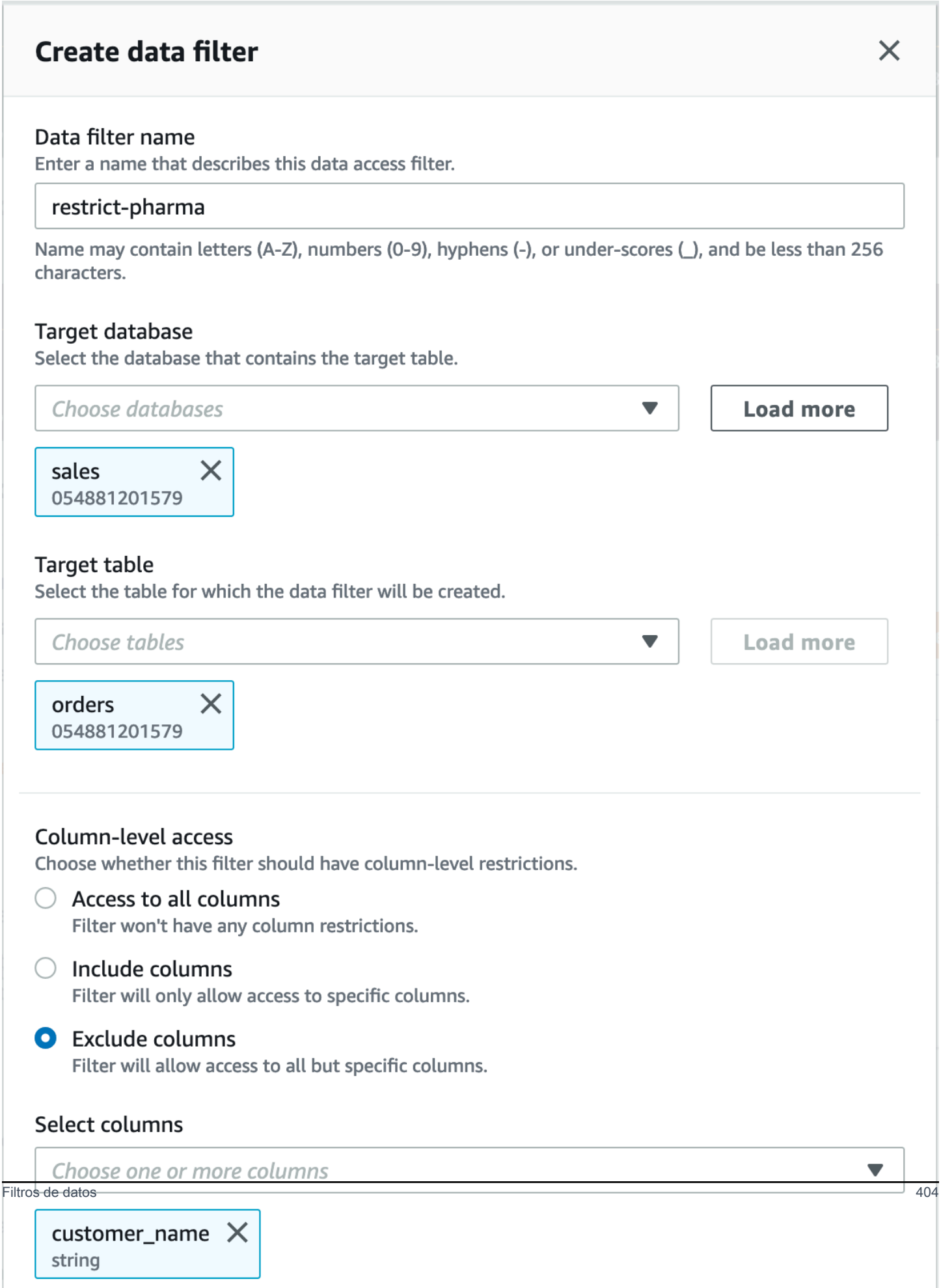
- Nombre del filtro
- Los ID del catálogo de la tabla asociada al filtro
- Nombre de la tabla
- Nombre de la base de datos que contiene las tablas que se enumerarán.
- Especificación de columna: una lista de columnas y columnas anidadas (con tipos de datos `struct`) para incluir o excluir en los resultados de la consulta.
- Expresión de filtro de filas: una expresión que especifica las filas que se van a incluir en los resultados de la consulta. Con algunas restricciones, la expresión tiene la sintaxis de una cláusula `WHERE` en el lenguaje PartiQL. Para especificar todas las filas, seleccione Acceso a todas las filas en Acceso a nivel de fila en la consola o utilice `AllRowsWildcard` en las llamadas a la API.

  Para obtener más información acerca de lo que se admite en las expresiones de filtro de fila, consulte Compatibilidad con PartiQL en expresiones de filtro de filas.

El nivel de filtrado que recibe depende de cómo rellene el filtro de datos.

- Al especificar el comodín “todas las columnas” y proporcionar una expresión de filtro de filas, solo se establece la seguridad de fila (filtrado de filas).
- Al incluir o excluir columnas específicas y columnas anidadas y especificar «todas las filas» con el comodín de todas las filas, solo establece la seguridad a nivel de columna (filtrado de columnas).
- Cuando se incluyen o excluyen columnas específicas y se proporciona también una expresión de filtro de fila, se establece la seguridad a nivel de celda (filtrado de celdas).

La siguiente captura de pantalla de la consola de Lake Formation muestra un filtro de datos que actúa a nivel de celda. En el caso de las consultas basadas en la tabla `orders`, restringe el acceso a la columna `customer_name` y los resultados de la consulta solo muestran las filas en las que la columna `product_type` contiene «farmacia».


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

Tenga en cuenta el uso de comillas simples para encerrar la cadena literal `'pharma'`.

Puede usar la consola de Lake Formation para crear este filtro de datos o proporcionar el siguiente objeto de solicitud a la operación de la API `CreateDataCellsFilter`.

```
{
     "Name": "restrict-pharma",
     "DatabaseName": "sales",
     "TableName": "orders",
     "TableCatalogId": "111122223333",
     "RowFilter": {"FilterExpression": "product_type='pharma'"},
     "ColumnWildcard": {
         "ExcludedColumnNames": ["customer_name"]
     }
}
```

Puede crear todos los filtros de datos que necesite para una tabla. Para ello, necesita un permiso `SELECT` con la opción de concesión que aparece en una tabla. De forma predeterminada, los administradores del lago de datos tienen permiso para crear filtros de datos en todas las tablas de esa cuenta. Normalmente solo se utiliza un subconjunto de los posibles filtros de datos cuando se conceden permisos sobre la tabla a una entidad principal. Por ejemplo, puede crear un segundo filtro de datos para la tabla `orders` que sea un filtro de datos solo con seguridad de filas. En referencia a la captura de pantalla anterior, puede elegir la opción Acceso a todas las columnas e incluir una expresión de filtro de filas de `product_type<>pharma`. El nombre de este filtro de datos podría ser `no-pharma`. Restringe el acceso a todas las filas cuya `product_type` columna esté configurada como «pharma».

El objeto de solicitud de la operación de la API `CreateDataCellsFilter` para este filtro de datos es el siguiente.

```
{
     "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"]
}
```

A continuación, puede conceder SELECT sobre la tabla `restrict-pharma` con el filtro de datos `orders` a un usuario administrativo y SELECT sobre la tabla `orders` con el filtro de datos `no-pharma` a los usuarios no administrativos. En el caso de los usuarios del sector sanitario, concedería SELECT en la tabla `orders` con acceso total a todas las filas y columnas (sin filtro de datos), o quizá con otro filtro de datos que restrinja el acceso a la información sobre precios.

Puede incluir o excluir columnas anidadas al especificar la seguridad a nivel de columna y de fila dentro de un filtro de datos. En el siguiente ejemplo, el acceso al campo `product.offer` se especifica mediante nombres de columna cualificados (entre comillas dobles). Esto es importante en el caso de los campos anidados para evitar que se produzcan errores cuando los nombres de las columnas contienen caracteres especiales y para mantener la compatibilidad con versiones anteriores de las definiciones de seguridad de nivel columnas superior.

```
{
     "Name": "example_dcf",
     "DatabaseName": "example_db",
     "TableName": "example_table",
     "TableCatalogId": "111122223333",
     "RowFilter": { "FilterExpression": "customer.customerName <> 'John'" },
     "ColumnNames": ["customer", "\"product\".\"offer\""]
}
```

Véase también

- Administrar filtros de datos

Compatibilidad con PartiQL en expresiones de filtro de filas

Puede crear expresiones de filtro de filas mediante un subconjunto de tipos de datos PartiQL, operadores y agregaciones. Lake Formation no permite ninguna función PartiQL estándar o definida por el usuario en la expresión del filtro. Puede usar operadores de comparación para comparar columnas con constantes (por ejemplo, `views >= 10000`), pero no puede comparar columnas con otras columnas.

Una expresión de filtro de filas puede ser una expresión simple o una expresión compuesta. La longitud total de la expresión debe ser inferior a 2048 caracteres.

Expresiones simples

Una expresión simple tendrá el siguiente formato: `<column name > <comparison operator ><value >`

- Nombre de la columna

  Puede ser una columna de datos de nivel superior, una columna de partición o una columna anidada presente en el esquema de la tabla y debe pertenecer a los Tipos de datos admitidos indicados a continuación.

- Operador de comparación

  Se admiten los siguientes operadores: `=, >, <, >=, <=, <>,!=, BETWEEN, IN, LIKE, NOT, IS [NOT] NULL`

- Todas las comparaciones de cadenas y coincidencias del patrón `LIKE` distinguen entre mayúsculas y minúsculas. No puede usar el operador IS [NOT] NULL en las columnas de partición.
- Valor de columna

  El valor de la columna debe coincidir con el tipo de datos del nombre de la columna.

Expresión compuesta

Una expresión simple tendrá el formato: `( <simple expression >) <AND/OR >(<simple expression >)`. Las expresiones compuestas se pueden combinar aún más mediante operadores lógicos `AND/OR`.

Tipos de datos admitidos

Los filtros de filas que hacen referencia a una tabla AWS Glue Data Catalog que contiene un tipo de datos no compatible generarán un error. Los siguientes son los tipos de datos admitidos para las columnas y constantes de la tabla, que se asignan a los tipos de datos Amazon Redshift:

- `STRING, CHAR, VARCHAR`
- `INT, LONG, BIGINT, FLOAT, DECIMAL, DOUBLE`
- `BOOLEAN`
- `STRUCT`

Para obtener más información acerca de los tipos de datos en Amazon Redshift, consulte Tipos de datos en la Guía para desarrolladores de bases de datos de Amazon Redshift.

Expresión de filtro de filas

Example

Los siguientes son ejemplos de expresiones de filtro de filas válidas para una tabla con columnas: `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

Los siguientes son ejemplos válidos de expresiones de filtro de filas para una tabla con columnas anidadas: `year > 2010 and customer.customerId <> 1`

No se debe hacer referencia a los campos anidados de las columnas de partición al definir expresiones anidadas a nivel de fila.

Las constantes de cadena se deben escribir entre comillas simples.

Palabras clave reservadas

Si la expresión de filtro de filas contiene palabras clave PartiQL, se producirá un error de análisis, ya que los nombres de las columnas pueden entrar en conflicto con las palabras clave. Cuando esto suceda, escape de los nombres de las columnas usando comillas dobles. Algunos ejemplos de palabras clave reservadas son «first», «last», «asc» y «missing». Consulte la especificación PartiQL para obtener una lista de palabras clave reservadas.

Referencia PartiQL

Para obtener más información acerca de PartiQL, consulte https://partiql.org/.

Notas y restricciones para el filtrado de nivel de columna

Hay tres formas de especificar el filtrado de columnas:

- Mediante filtros de datos como se ha descrito antes.
- Mediante el uso de un filtrado de columnas simple o un filtrado de columnas anidadas.
- Mediante TAG.

El filtrado de columnas simple solo especifica una lista de columnas para incluir o excluir. Tanto la consola Lake Formation como la API y la AWS CLI admiten un filtrado de columnas sencillo. Para ver un ejemplo, consulte Grant with Simple Column Filtering.

Las siguientes notas y restricciones corresponden al filtrado de columnas:

- AWS Glue Los trabajos de ETL admiten el filtrado de columnas únicamente mediante filtros de datos (seguridad a nivel de celda). El trabajo falla si se aplica un filtrado de columnas simple a cualquier tabla a la que haga referencia el trabajo. Si solo desea filtrar columnas, conceda acceso a las tablas mediante filtros de datos e introduzca `true` para la expresión del filtro de filas en la consola o use `AllRowsWildcard` en sus llamadas a la API.
- Para conceder `SELECT` con la opción de concesión y el filtrado de columnas, debe usar una lista de inclusión, no una lista de exclusiones. Sin la opción de conceder, puede utilizar listas de inclusión o exclusión.
- Para conceder `SELECT` en una tabla con filtrado de columnas, debe haber obtenido la autorización `SELECT` en la tabla con la opción de concesión y sin restricciones de filas. Debe tener acceso a todas las filas.
- Si concede `SELECT` con la opción de concesión y filtrado de columnas a una entidad principal de su cuenta, dicha entidad principal deberá especificar el filtrado de columnas para las mismas columnas o un subconjunto de las columnas beneficiarias cuando conceda a otra entidad principal. Si concede `SELECT` con la opción de concesión y filtrado de columnas a una cuenta externa, el administrador del lago de datos de la cuenta externa puede conceder `SELECT` en todas las columnas a otra entidad principal de su cuenta. Sin embargo, incluso con `SELECT` en todas las columnas, esa entidad principal tendrá visibilidad solo en las columnas concedidas a la cuenta externa.
- No puede aplicar el filtrado de columnas a las claves de partición.
- A una entidad principal con el permiso `SELECT` sobre un subconjunto de columnas de una tabla no se le puede conceder el permiso `ALTER`, `DROP`, `DELETE` o `INSERT` sobre esa tabla. Para una entidad principal con el permiso `ALTER`, `DROP`, `DELETE` o `INSERT` en una tabla, si concede el permiso `SELECT` con filtrado de columnas, no tiene ningún efecto.

Las siguientes notas y restricciones se aplican al filtrado de columnas anidadas:

- Puede incluir o excluir cinco niveles de campos anidados en un filtro de datos.

  Example

  Col1.Col1_1.Col1_1_1.Col1_1_1_1.Col1_1_1_1_1
- No puede aplicar el filtrado de columnas a los campos anidados dentro de las columnas de partición.
- Si el esquema de la tabla contiene un nombre de columna de nivel superior (“cliente”.”dirección”) que tiene el mismo patrón de representación de un campo anidado dentro de un filtro de datos (una columna anidada con un nombre de columna de nivel superior `customer` y un nombre de campo anidado `address` se especifica como `"customer"."address"` en un filtro de datos), no puede especificar explícitamente el acceso a la columna de nivel superior o al campo anidado, ya que ambos se representan con el mismo patrón en las listas de inclusión/exclusión. Esto es ambiguo y Lake Formation no puede resolverlo si especifica la columna de nivel superior o el campo anidado.
- Si una columna de nivel superior o un campo anidado contiene comillas dobles dentro del nombre, debe incluir una segunda comilla doble cuando especifique el acceso a un campo anidado en la lista de inclusión y exclusión de un filtro de celdas de datos.

  Example

  Ejemplo de nombre de columna anidado con comillas dobles: `a.b.double"quote`

  Example

  Ejemplo de representación de columna anidada dentro de un filtro de datos: `"a"."b"."double""quote"`

Permisos necesarios para consultar tablas con filtrado a nivel de celda

Los siguientes permisos (IAM) AWS Identity and Access Management son necesarios para ejecutar consultas en tablas con filtrado a nivel de celda.

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

```
            "Effect": "Allow",
            "Action": [
                "lakeformation:StartQueryPlanning",
                "lakeformation:GetQueryState",
                "lakeformation:GetWorkUnits",
                "lakeformation:GetWorkUnitResults"
            ],
            "Resource": "*"
        }
    ]
}
```

Para obtener más información sobre permisos de Lake Formation, consulte Personas de Lake Formation y referencia de permisos IAM.

Administrar filtros de datos

Con el fin de implementar la seguridad a nivel de columna, fila y celda, puede crear y mantener filtros de datos. Cada filtro de datos pertenece a una tabla del catálogo de datos. Puede crear varios filtros de datos para una tabla y, a continuación, utilizar uno o varios de ellos al conceder permisos sobre la tabla. También puede definir y aplicar filtros de datos en las columnas anidadas que tienen tipos de datos `struct` y que permiten a los usuarios acceder solo a las subestructuras de las columnas anidadas.

Se requiere un permiso `SELECT` con la opción de concesión para crear o ver un filtro de datos. Para que las entidades principales de su cuenta puedan ver y utilizar un filtro de datos, puede conceder el permiso `DESCRIBE` sobre este.

> **Note**
>
> Lake Formation no es compatible con la concesión de permisos `Describe` en un filtro de datos que se comparte desde otra cuenta.

Puede administrar los filtros de datos utilizando la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Para obtener más información acerca de los filtros de datos, consulte Filtros de datos en Lake Formation

Creación de un filtro de datos

Puede crear uno o varios filtros de datos para cada tabla del catálogo de datos.

Para crear un filtro de datos para una tabla del catálogo de datos (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

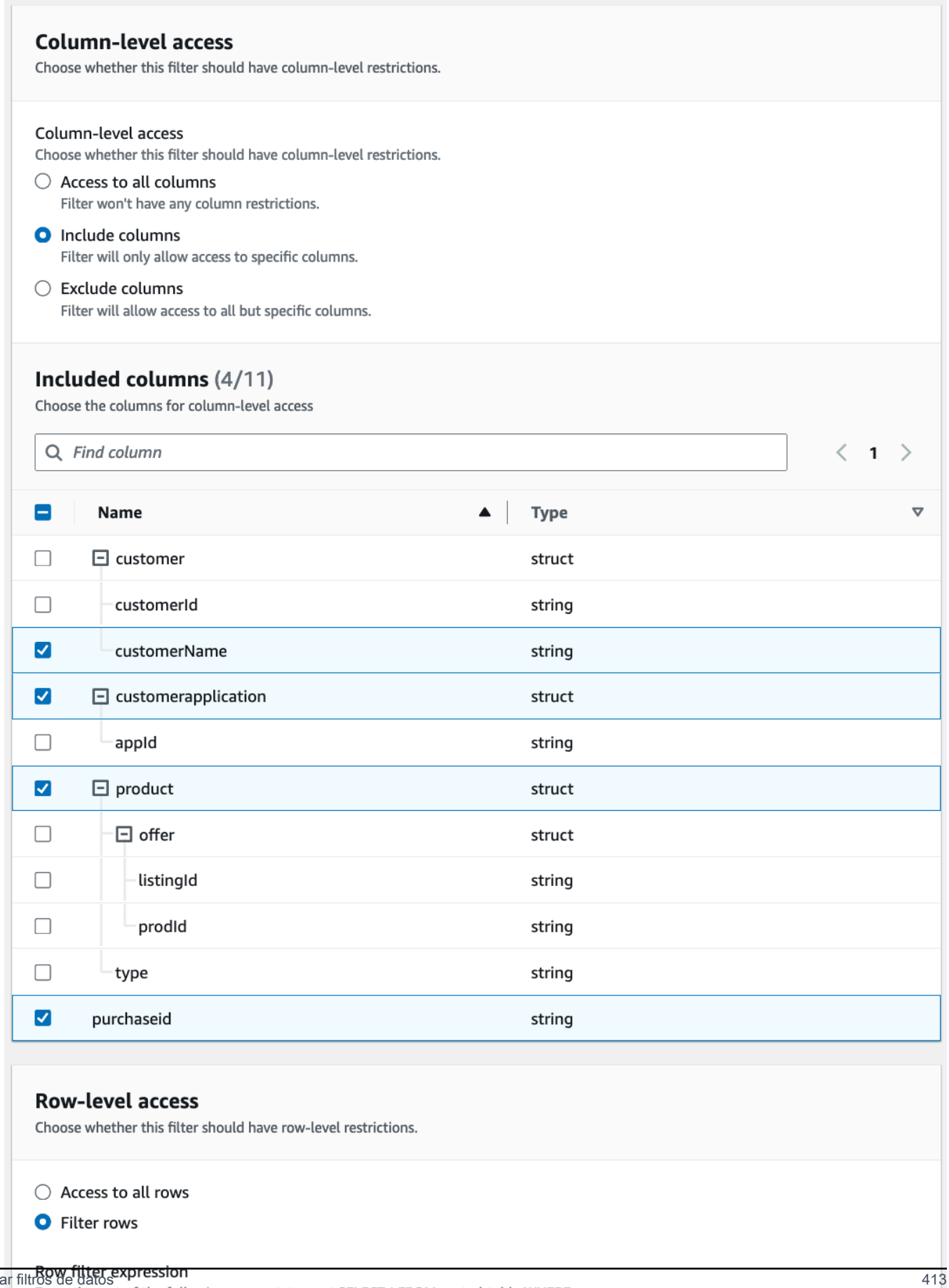
   Inicie sesión como administrador del lago de datos, como propietario de la tabla de destino o como entidad principal con permiso de Lake Formation en la tabla de destino.
2. En el panel de navegación, en Catálogo de datos, seleccione Filtros de datos.
3. En la página Filtros de datos, seleccione Crear filtro nuevo.
4. En el cuadro de diálogo Crear filtro de datos, introduzca la información siguiente:

   - Nombre del filtro de datos
   - Base de datos de destino. Indique la base de datos que contiene la tabla.
   - Tabla de destino
   - Acceso a nivel de columna. Deje esta opción como Acceso a todas las columnas para especificar únicamente el filtrado solo de filas. Elija Incluir columnas o Excluir columnas para especificar el filtrado de columnas o celdas y, a continuación, especifique las columnas que desea incluir o excluir.

     Columnas anidadas: si aplica el filtro a una tabla que contiene columnas anidadas, puede especificar explícitamente las subestructuras de las columnas de estructura anidada dentro de un filtro de datos.

     Al conceder el permiso SELECT a una entidad principal en este filtro, la entidad principal que ejecute la siguiente consulta solo verá los datos de `customer.customerName` y no de `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'

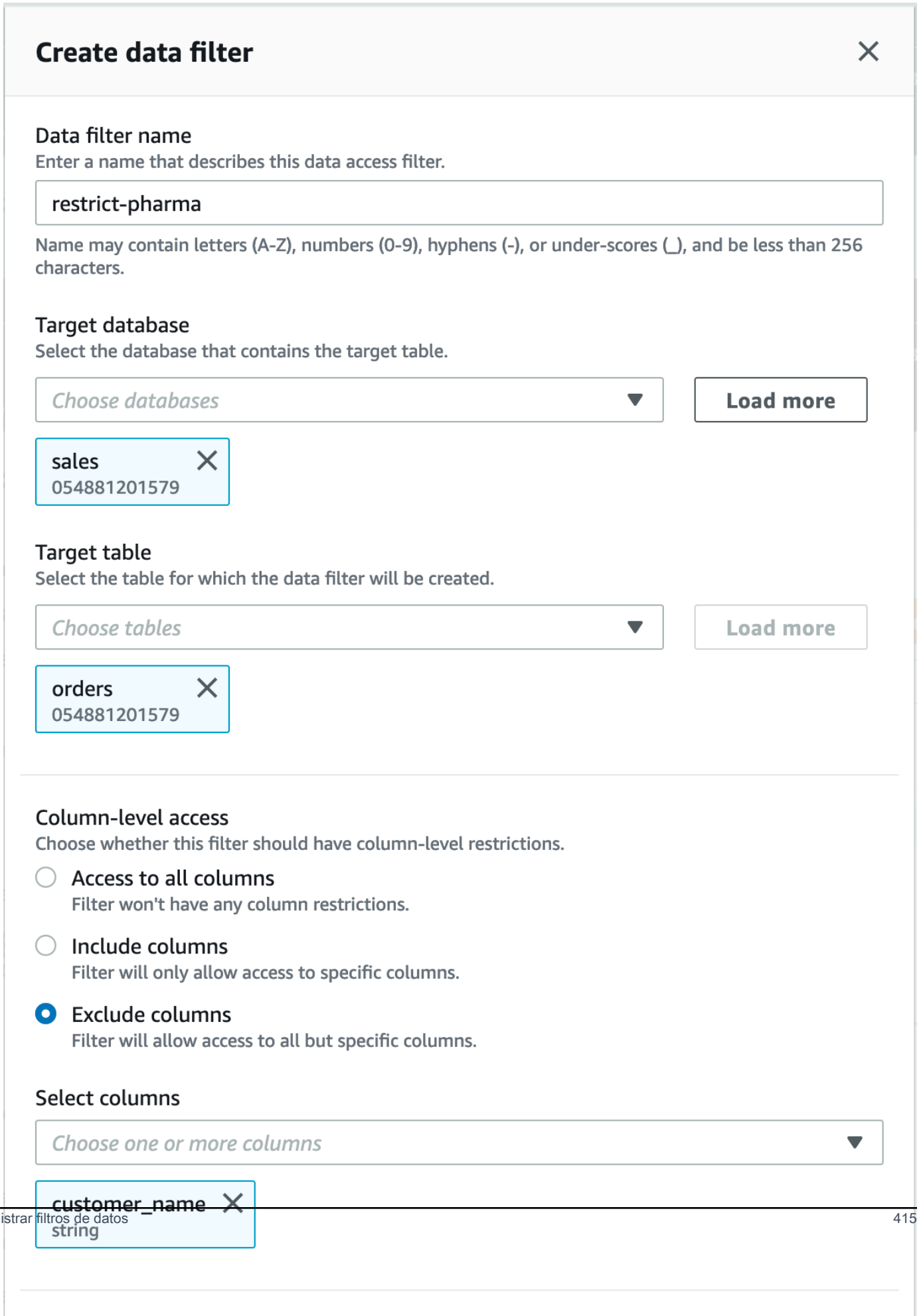
Al conceder permisos a la columna `customer`, la entidad principal recibe el acceso a la columna y a los campos anidados debajo de la columna (`customerName` y `customerID`).

- Expresión de filtro de filas. Introduzca una expresión de filtro para especificar el filtro de filas o celdas. Para conocer los tipos de datos y operadores compatibles, consulte Compatibilidad con PartiQL en expresiones de filtro de filas. Elija Acceso a todas las filas para conceder acceso a todas.

  Puede incluir estructuras de columnas parciales de columnas anidadas en una expresión de filtro de filas para filtrar las filas que contienen un valor concreto.

  Cuando a una entidad principal se le conceden permisos para una tabla con una expresión de filtro de filas `Select * from example_nestedtable where customer.customerName <>'John'` y el acceso a Nivel de columna se establece en Acceso a todas las columnas, los resultados de la consulta solo muestran las filas en las que `customerName <>'John'` se evalúa como verdadero.

La siguiente captura de pantalla muestra un filtro de datos que implementa el filtrado de celdas. En las consultas a la tabla `orders`, deniega el acceso a la columna `customer_name` y muestra solo las filas que tienen "pharma" en la columna `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. Elija Crear filtro.

Para crear un filtro de datos con políticas de filtrado de celdas en un campo anidado

En esta sección se utiliza el siguiente esquema de ejemplo para mostrar cómo crear un filtro de celdas de datos:

```
[
    { 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. En la página Crear un filtro de datos, escriba un nombre para el filtro.
2. A continuación, utilice el menú desplegable para elegir un nombre de base de datos y un nombre de tabla.
3. En la sección de Acceso a nivel de columna, elija Columnas incluidas y seleccione una columna anidada (`customer.customerName`).
4. En la sección de Acceso a nivel de fila, elija la opción Acceso a todas las filas.
5. Elija Crear filtro.

   Al conceder el permiso `SELECT` en este filtro, la entidad principal tiene acceso a todas las filas de la columna `customerName`.
6. A continuación, defina otro filtro de datos para la misma base de datos o tabla.
7. En la sección de Acceso a nivel de columna, elija Columnas incluidas y seleccione otra columna anidada (`customer.customerid`).
8. En la sección de Acceso a nivel de fila, elija Filtrar filas e introduzca una Expresión de filtro de fila (`customer.customerid <> 5`).
9. Elija Crear filtro.

   Al conceder el permiso `SELECT` en este filtro, la entidad principal recibe acceso a todas las filas de los campos `customerName` y `customerId`, excepto a la celda en la que el valor es 5 en la columna `customerId`.

Concesión de permisos de filtrado de datos

Puede conceder los permisos `SELECT`, `DESCRIBE` y `DROP` de Lake Formation sobre los filtros de datos a las entidades principales.

Al principio, solo el usuario puede ver los filtros de datos que crea para una tabla. Para que otra entidad principal pueda ver un filtro de datos y conceder permisos del catálogo de datos con el filtro de datos, debe:

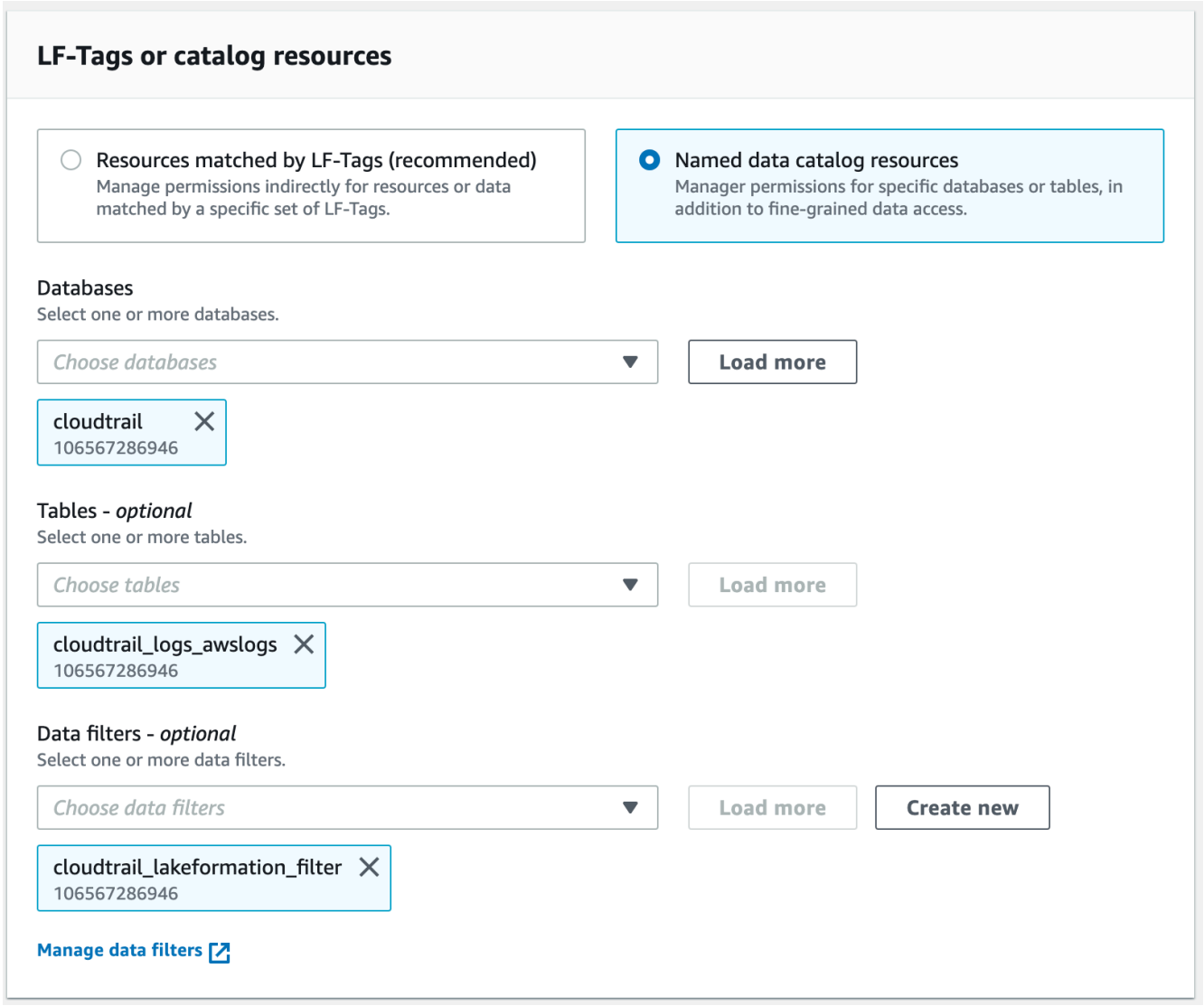
- Conceder `SELECT` sobre una tabla a la entidad principal con la opción de concesión, y aplicar el filtro de datos a la concesión.
- Conceda el permiso `DESCRIBE` o `DROP` sobre el filtro de datos a la entidad principal.

Puede conceder el permiso `SELECT` a una cuenta externa AWS. Un administrador del lago de datos de esa cuenta puede entonces conceder ese permiso a otras entidades principales de la cuenta. Cuando conceda el permiso a una cuenta externa, debe incluir la opción de concesión para que el administrador de la cuenta externa pueda extender el permiso en cascada a otros usuarios de su cuenta. Al conceder a una entidad principal de su cuenta, la concesión con la opción de concesión es opcional.

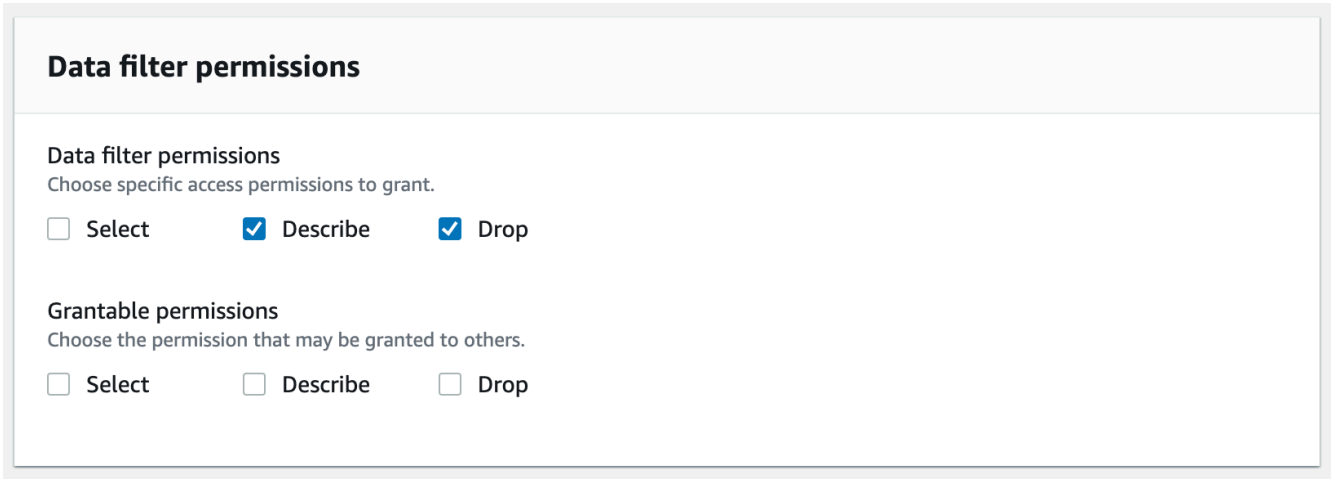
Para conceder y revocar permisos sobre los filtros de datos, puede utilizar la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Console

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, en Permisos, seleccione Permisos de lago de datos.
3. En la página Permisos, en la sección Permisos de datos, seleccione Conceder.
4. En la página Conceder permisos de datos, seleccione las entidades principales a las que conceder los permisos.
5. En la sección de las etiquetas LF o recursos del catálogo, seleccione Recursos del catálogo de datos con nombre. A continuación, elija la base de datos, la tabla y el filtro de datos para los que desea conceder permisos.



6. En la sección Permisos del filtro de datos, seleccione los que desea conceder a las entidades principales seleccionadas.



AWS CLI

- Introduzca un comando `grant-permissions`. Especifique `DataCellsFilter` para el argumento de `resource` y `DESCRIBE` o `DROP` para el argumento de `Permissions` y, opcionalmente, para el argumento de `PermissionsWithGrantOption`.

  En el siguiente ejemplo, se concede `DESCRIBE` con la opción de concesión al usuario `datalake_user1` del filtro de datos `restrict-pharma`, que pertenece a la tabla `orders` de la base de datos `sales` de la cuenta AWS 1111-2222-3333.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

  A continuación se muestra el contenido del archivo `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"]
}
```

Concesión de permisos de datos proporcionados por filtros de datos

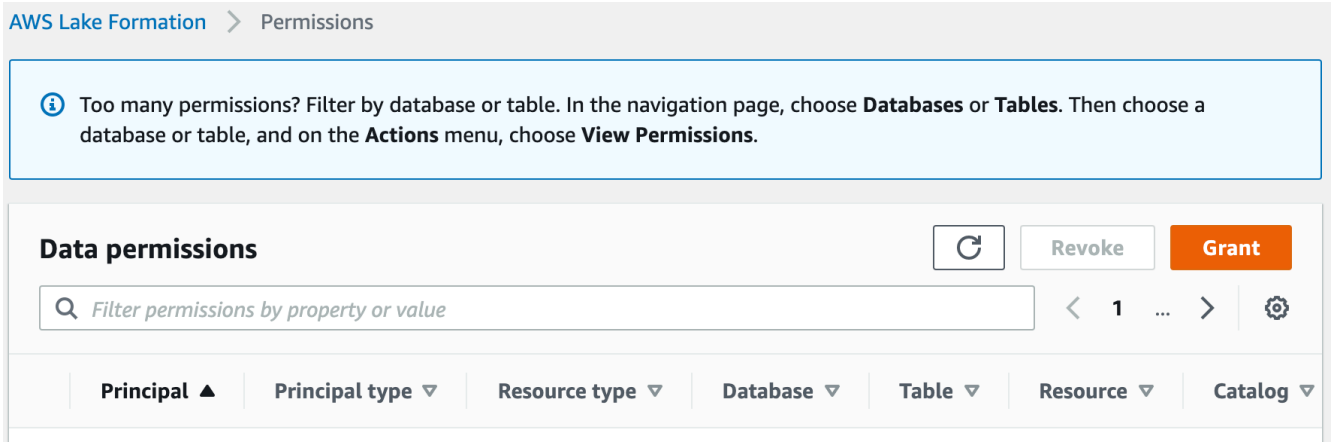
Los filtros de datos representan un subconjunto de datos dentro de una tabla. Para proporcionar acceso a los datos a las entidades principales, es necesario concederles permisos `SELECT`. Con este permiso las entidades principales pueden:

- Ver el nombre real de la tabla en la lista de tablas compartidas con su cuenta.
- Cree filtros de datos en la tabla compartida y conceda permisos a sus usuarios sobre esos filtros de datos.

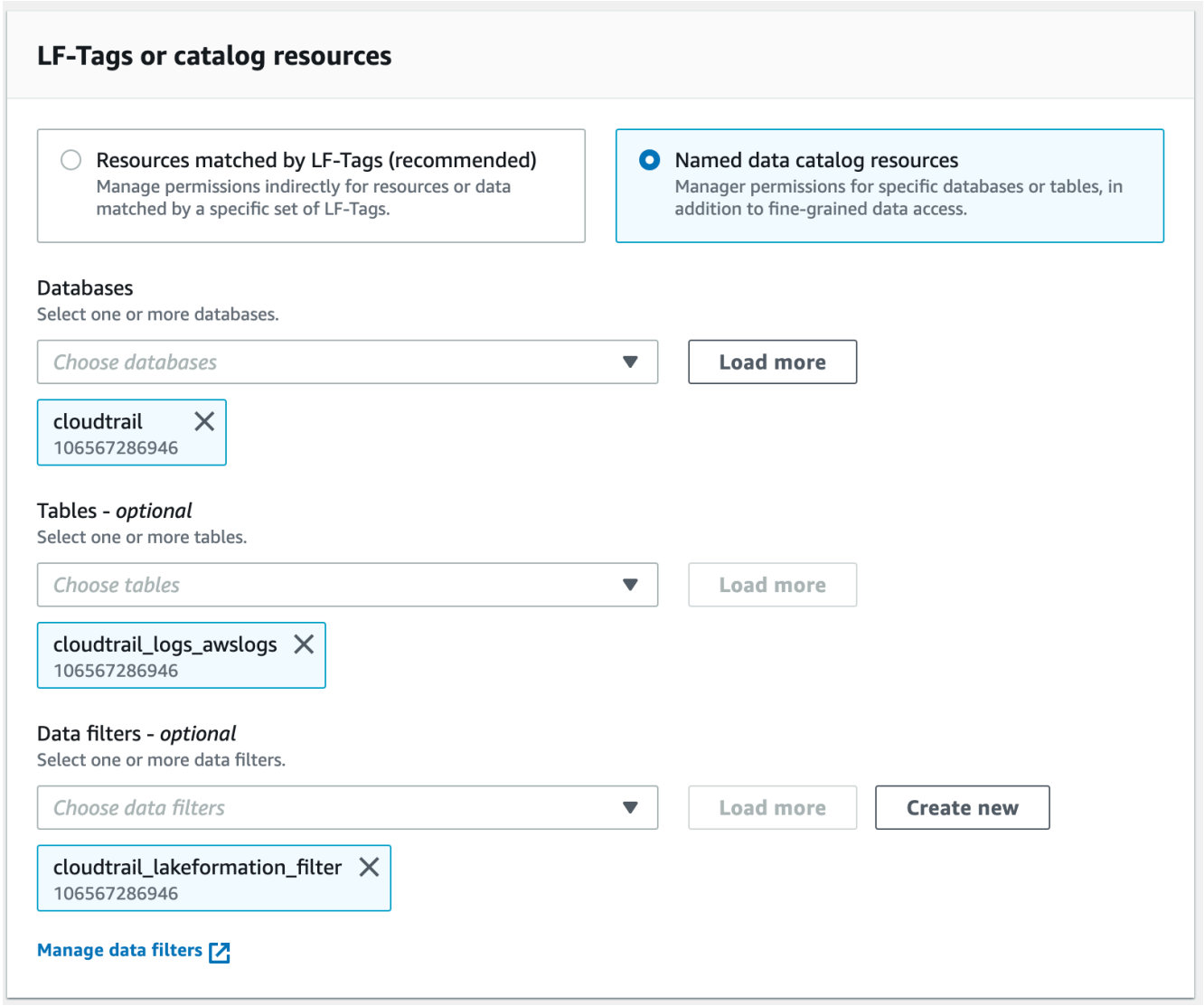
Console

Para conceder permisos SELECT

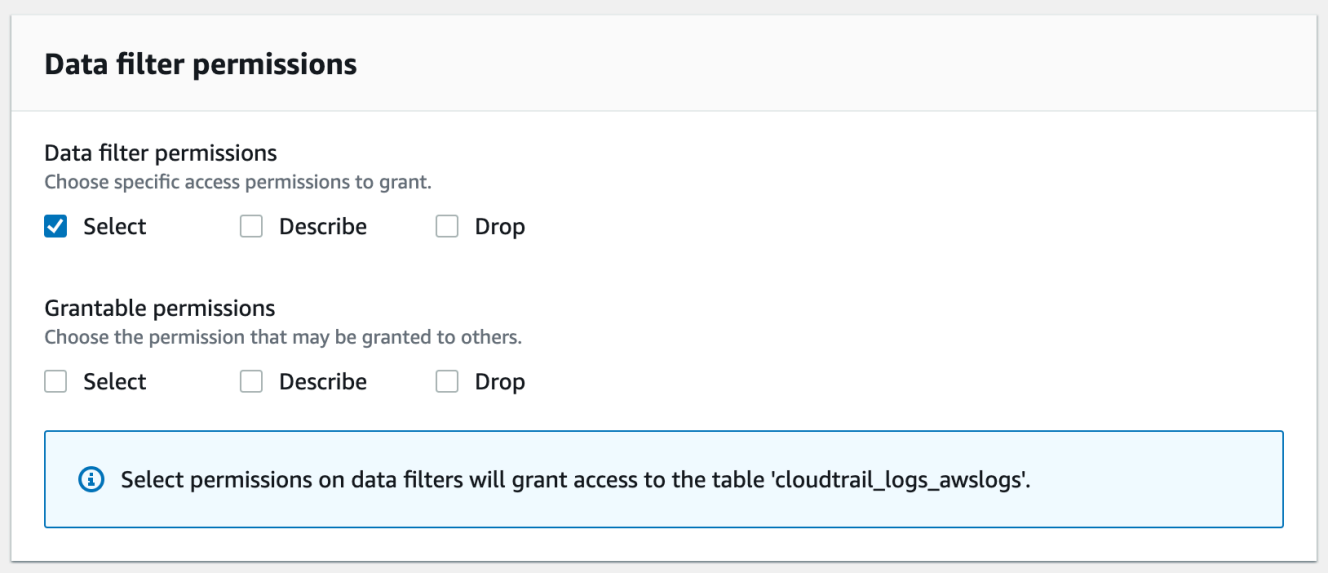
1. Vaya a la página Permisos de la consola de Lake Formation y elija Concesión.



2. Seleccione las entidades principales a las que desea proporcionar acceso y seleccione Recursos del catálogo de datos con nombre.



3. Para proporcionar acceso a los datos que representa el filtro, elija Seleccionar en Permisos de filtro de datos.



CLI

Introduzca un comando `grant-permissions`. Especifique `DataCellsFilter` el argumento de recurso y `SELECT` para el argumento de permisos.

El siguiente ejemplo concede `SELECT` con la opción grant al usuario `datalake_user1` sobre el filtro de datos `restrict-pharma`, que pertenece a la tabla `orders` de la base de datos `sales` en la cuenta de Cuenta de AWS `1111-2222-3333`.

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

A continuación, se muestra el contenido del archivo `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"]
}
```

Visualización de filtros de datos

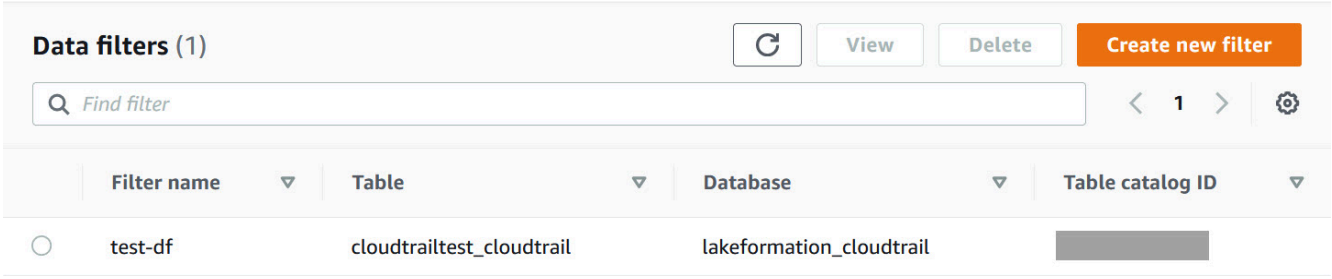
Puede utilizar la consola de Lake Formation, AWS CLI, o la API de Lake Formation para ver los filtros de datos.

Para ver los filtros de datos, debe ser administrador de lago de datos o tener los permisos necesarios sobre los filtros de datos.

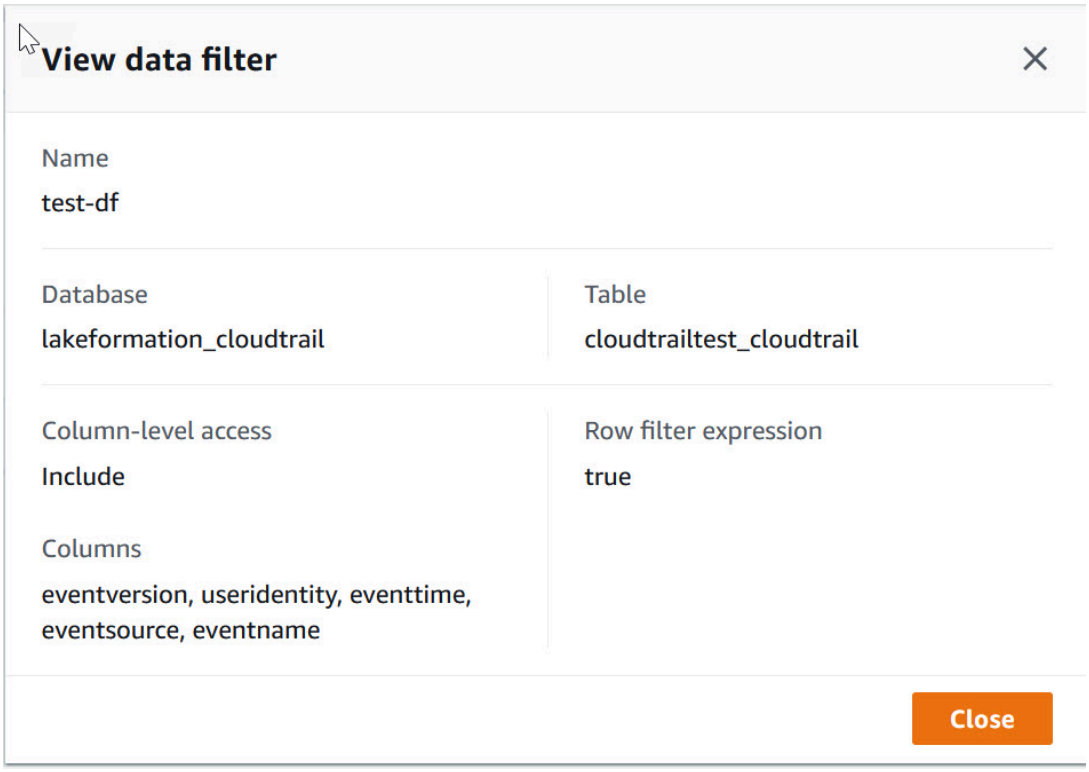
Console

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, en Catálogo de datos, seleccione Filtros de datos.

   La página muestra los filtros de datos a los que tiene acceso.



3. Para ver los detalles del filtro de datos, elija el filtro de datos y, a continuación, elija Ver. Aparece una nueva ventana con la información detallada del filtro de datos.



AWS CLI

Introduzca un comando `list-data-cells-filter` y especifique un recurso de tabla.

El siguiente ejemplo muestra los filtros de datos de la tabla `cloudtrailtest_cloudtrail`.

```
aws lakeformation list-data-cells-filter --table '{ "CatalogId":"123456789012",
"DatabaseName":"lakeformation_cloudtrail", "Name":"cloudtrailtest_cloudtrail"}'
```

API/SDK

Use la API `ListDataCellsFilter` y especifique un recurso de la tabla.

El ejemplo siguiente utiliza Python para listar los 20 primeros filtros de datos de la tabla `myTable`.

```
response = client.list_data_cells_filter(
    Table = {
        'CatalogId': '111122223333',
        'DatabaseName': 'mydb',
        'Name': 'myTable'
    },
```

```
    MaxResults=20
)
```

Lista de permisos de filtro de datos

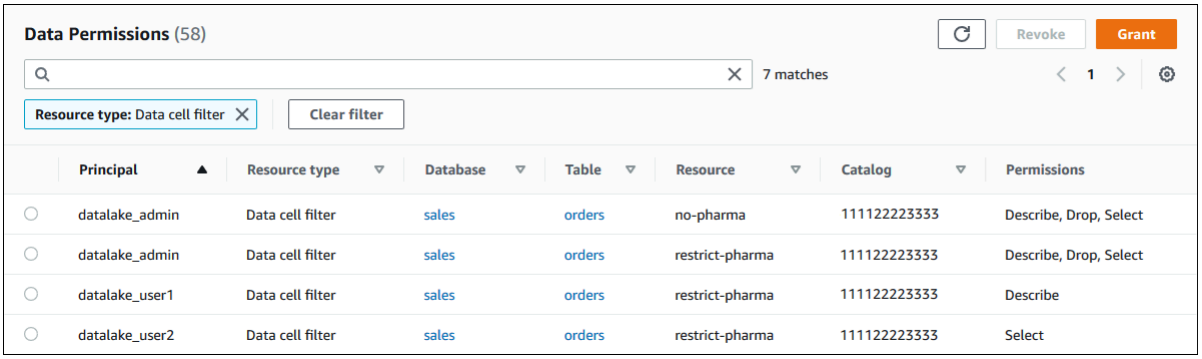
Puede utilizar la consola de Lake Formation para ver los permisos concedidos sobre los filtros de datos.

Para ver los permisos de un filtro de datos, debe ser administrador de lago de datos o tener los permisos necesarios sobre el filtro de datos.

Console

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación, en Permisos, seleccione Permisos de datos.
3. En la página Permisos de datos, haga clic o toque en el campo de búsqueda y, en el menú Propiedades, elija Tipo de recurso.
4. En el menú Tipo de recurso, seleccione Tipo de recurso: filtro de celda de datos.

   Se muestran los filtros de datos para los que tiene permisos. Es posible que tenga que desplazarse horizontalmente para ver las columnas Permisos y Concedibles.



AWS CLI

- Introduzca un comando `list-permissions`. Especifique `DataCellsFilter` para el argumento de `resource` y `DESCRIBE` o `DROP` para el argumento de `Permissions` y, opcionalmente, para el argumento de `PermissionsWithGrantOption`.

El siguiente ejemplo enumera los permisos `DESCRIBE` con la opción de concesión en el filtro de datos `restrict-pharma`. Los resultados se limitan a los permisos concedidos para la entidad principal `datalake_user1` y la tabla `orders` de la base de datos `sales` de la cuenta AWS 1111-2222-3333.

```
aws lakeformation list-permissions --cli-input-json file://list-params.json
```

A continuación, se muestra el contenido del archivo `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"]
}
```

Consulta de los permisos de bases de datos y tablas en Lake Formation

Puede consultar los permisos de Lake Formation concedidos sobre una base de datos o tabla del catálogo de datos. Para hacerlo, puede recurrir a la consola de Lake Formation, a la API o a la AWS Command Line Interface (AWS CLI).

Mediante la consola, puede ver los permisos a partir de las páginas Bases de datos o Tablas, o desde la página Permisos de datos.

> **Note**
>
> Si no es administrador de la base de datos ni propietario del recurso, podrá ver los permisos que otras entidades principales tienen sobre el recurso solo si tiene un permiso de Lake Formation sobre el recurso con la opción de concesión.
> Además de los permisos necesarios de Lake Formation, necesita los permisos de (IAM) AWS Identity and Access Management, `glue:GetDatabases`, `glue:GetDatabase`, `glue:GetTables`, `glue:GetTable` y `glue:ListPermissions`.

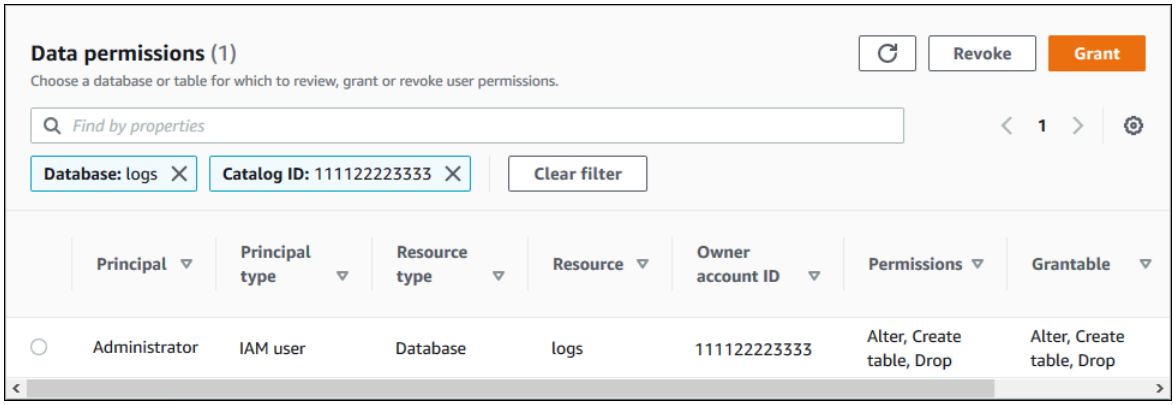
Para ver los permisos sobre una base de datos (consola, empezando por la página Bases de datos)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos, creador de la base de datos o como usuario que tenga algún permiso de Lake Formation sobre la base de datos con la opción de concesión.
2. En el panel de navegación, seleccione Databases (Bases de datos).
3. Elija una base de datos y, en el menú Acciones, seleccione Ver permisos.

   > **Note**
   >
   > Si elige un enlace de recurso de base de datos, Lake Formation muestra los permisos sobre el enlace de recursos, no sobre la base de datos de destino del enlace de recursos.

   La página Permisos de datos enumera todos los permisos de Lake Formation para la base de datos. El nombre de la base de datos y el ID de catálogo (ID de cuenta de AWS) del propietario de la base de datos aparecen como etiquetas en el cuadro de búsqueda. Los mosaicos indican que se ha aplicado un filtro para enumerar los permisos únicamente para esa base de datos. Puede ajustar el filtro cerrando un mosaico o eligiendo Borrar filtro.



Para ver los permisos sobre una base de datos (consola, empezando por la página Permisos de datos)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos, creador de la base de datos o como usuario que tenga algún permiso de Lake Formation sobre la base de datos con la opción de concesión.
2. En el panel de navegación, seleccione Permisos de datos.
3. Coloque el cursor en el cuadro de búsqueda de la parte superior de la página y, en el menú Propiedades que aparece, elija Base de datos.
4. En el menú Bases de datos que aparece, elija una.

> **Note**
>
> Si elige un enlace de recurso de base de datos, Lake Formation muestra los permisos sobre el enlace de recursos, no sobre la base de datos de destino del enlace de recursos.

La página Permisos de datos enumera todos los permisos de Lake Formation para la base de datos. El nombre de la base de datos aparece como un mosaico debajo del cuadro de búsqueda. El mosaico indica que se ha aplicado un filtro para enumerar los permisos únicamente para esa base de datos. Puede eliminar el filtro cerrando un mosaico o seleccionando Borrar filtro.

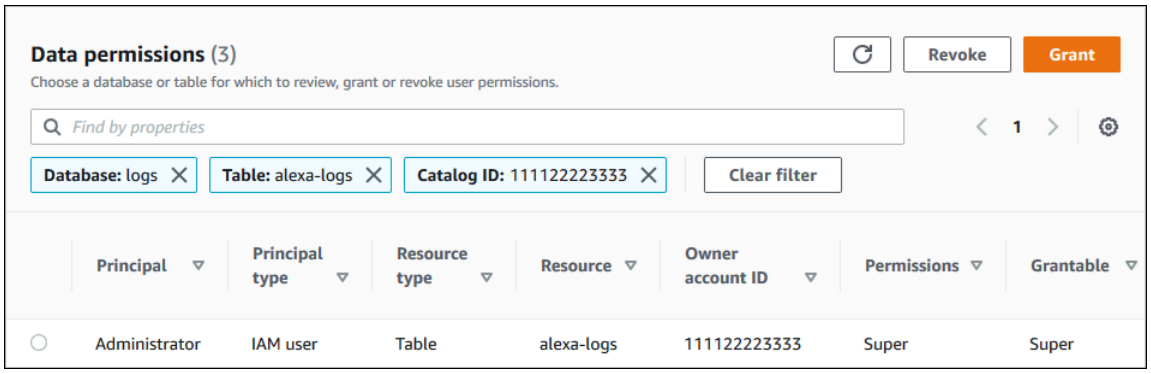
Para ver los permisos sobre una tabla (consola, empezando por la página Tablas)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos, creador de la tabla o como usuario que tenga algún permiso de Lake Formation sobre la tabla con la opción de concesión.
2. En el panel de navegación, elija Tablas.
3. Elija una tabla y, en el menú Acciones, seleccione Ver permisos.

> **Note**
>
> Si elige un enlace de recurso de tabla, Lake Formation muestra los permisos sobre el enlace de recursos, no sobre la tabla de destino del enlace de recursos.

La página Permisos de datos enumera todos los permisos de Lake Formation para la tabla. El nombre de la tabla, el nombre de la base de datos que contiene la tabla y el ID del catálogo (ID de cuenta AWS) del propietario de la tabla aparecen como etiquetas bajo el cuadro de búsqueda. Las etiquetas indican que se ha aplicado un filtro para enumerar los permisos únicamente para esa tabla. Puede ajustar el filtro cerrando una etiqueta o seleccionando Borrar filtro.



Para ver los permisos sobre una tabla (consola, empezando por la página Permisos de datos)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos, creador de la tabla o como usuario que tenga algún permiso de Lake Formation sobre la tabla con la opción de concesión.
2. En el panel de navegación, seleccione Permisos de datos.

3. Coloque el cursor en el cuadro de búsqueda de la parte superior de la página y, en el menú Propiedades que aparece, elija Base de datos.
4. En el menú Bases de datos que aparece, elija una.

   > **Important**
   >
   > Si desea ver los permisos de una tabla que se compartió con su cuenta AWS desde una cuenta externa, debe elegir la base de datos de la cuenta externa que contiene la tabla, no un enlace de recursos a la base de datos.

   La página Permisos de datos enumera todos los permisos de Lake Formation para la base de datos.
5. Coloque el cursor en el cuadro de búsqueda y, en el menú Propiedades que aparece, elija Tabla.
6. En el menú Tablas que aparece, elija una.

   La página Permisos de datos enumera todos los permisos de Lake Formation para la tabla. El nombre de la tabla y el nombre de la base de datos que la contiene aparecen como mosaicos bajo el cuadro de búsqueda. Los mosaicos indican que se ha aplicado un filtro para listar los permisos solo para esa tabla. Puede ajustar el filtro cerrando un mosaico o eligiendo Borrar filtro.

Para ver los permisos sobre una tabla (AWS CLI)

- Introduzca un comando `list-permissions`.

  En el siguiente ejemplo, se enumeran los permisos de una tabla compartida desde una cuenta externa. La propiedad `CatalogId` es el identificador de cuenta AWS de la cuenta externa y el nombre de la base de datos hace referencia a la base de datos de la cuenta externa que contiene la tabla.

```
aws lakeformation list-permissions  --resource-type TABLE --resource '{ "Table":
 {"DatabaseName":"logs", "Name":"alexa-logs", "CatalogId":"123456789012"}}'
```

Revocación de permisos mediante la consola de Lake Formation

Puede utilizar la consola para revocar todos los tipos de permisos de Lake Formation: de catálogo de datos, de etiqueta de política, de filtro de datos y de ubicación.

Para revocar los permisos de Lake Formation sobre un recurso (consola)

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos o como usuario al que se le han concedido permisos con la opción de concesión en el recurso.
2. En el panel de navegación, en Permisos, elija Permisos de lago de datos, Etiquetas de LF y permisos o Ubicaciones de datos.
3. Seleccione el permiso o la ubicación y, a continuación, elija Revocar.
4. En el cuadro de diálogo que se abre, elija Revocar.

Compartir datos entre cuentas en Lake Formation

Las capacidades multicuenta de Lake Formation permiten a los usuarios compartir de forma segura lagos de datos distribuidos entre varias AWS organizaciones o directamente con los directores de IAM en otra cuenta Cuentas de AWS, lo que proporciona un acceso detallado a los metadatos del catálogo de datos y a los datos subyacentes. Las grandes empresas suelen utilizar varias cuentas Cuentas de AWS, y es posible que muchas de esas cuentas necesiten acceder a un lago de datos gestionado por una sola persona. Cuenta de AWS Los usuarios y los trabajos de AWS Glue extracción, transformación y carga (ETL) pueden consultar y unir tablas en varias cuentas y, aun así, aprovechar las protecciones de datos a nivel de tabla y columna de Lake Formation.

Al conceder permisos de Lake Formation sobre un recurso del catálogo de datos a una cuenta externa o directamente a un director de IAM de otra cuenta, Lake Formation utiliza el servicio AWS Resource Access Manager (AWS RAM) para compartir el recurso. Si la cuenta del beneficiario está en la misma organización que la cuenta del concedente, el recurso compartido estará disponible inmediatamente para el beneficiario. Si la cuenta del concesionario no pertenece a la misma organización, AWS RAM envía una invitación a la cuenta del concesionario para que acepte o rechace la concesión de recursos. Luego, para que el recurso compartido esté disponible, el administrador del lago de datos de la cuenta del concesionario debe usar la AWS RAM consola o aceptar la invitación. AWS CLI

Lake Formation permite compartir los recursos del Catálogo de datos con cuentas externas en modo de acceso híbrido. El modo de acceso híbrido proporciona la flexibilidad de habilitar selectivamente los permisos de Lake Formation para las bases de datos y tablas de su AWS Glue Data Catalog. Con el modo de acceso híbrido, ahora tiene una ruta incremental que le permite establecer los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes.

Para obtener más información, consulte Modo de acceso híbrido.

Cómo compartir directamente entre cuentas

Las entidades principales autorizadas pueden compartir recursos de forma explícita con una entidad principal de IAM en una cuenta externa. Esta característica es útil cuando el propietario de una cuenta quiere controlar quién puede acceder a los recursos en la cuenta externa. Los permisos que reciba la entidad principal de IAM serán una combinación de concesiones directas y concesiones de cuenta que se transferirán en cascada a las entidades principales. El administrador del lago de datos de la cuenta receptora puede ver las concesiones directas entre cuentas, pero no revocar los permisos. La entidad principal que recibe el recurso compartido no puede compartir el recurso con otras entidades principales.

Métodos para compartir los recursos del Catálogo de datos

Con una sola operación de concesión de Lake Formation, puede conceder permisos entre cuentas en los siguientes recursos del Catálogo de datos.

- Una base de datos
- Una tabla individual (con filtrado de columnas opcional)
- Algunas tablas seleccionadas
- Todas las tablas de una base de datos (mediante el comodín Todas las tablas)

Existen dos opciones para compartir sus bases de datos y tablas con otra persona Cuenta de AWS o con los directores de IAM de otra cuenta.

- Control de acceso basado en etiquetas de Lake Formation (LF-TBAC) (recomendado)

  El control de acceso basado en atributos de Lake Formation es una estrategia de autorización que define permisos basados en atributos. Puede utilizar el control de acceso basado en etiquetas para compartir los recursos del catálogo de datos (bases de datos, tablas y columnas) con entidades principales, Cuentas de AWS organizaciones y unidades organizativas (OU) de IAM

externas. En Lake Formation, estos atributos se denominan etiquetas LF. Para más información, consulte Gestión de un lago de datos mediante el control de acceso basado en etiquetas de Lake Formation.

> **Note**
>
> El método LF-TBAC para conceder permisos al catálogo de datos se utiliza para conceder permisos entre cuentas. AWS Resource Access Manager
> Lake Formation ahora admite la concesión de permisos entre cuentas a organizaciones y unidades organizativas utilizando el método LF-TBAC.
> Para activar esta prestación, debe actualizar la Configuración de la versión entre cuentas a la Versión 3.
> Para obtener más información, consulte Actualización de los ajustes de la versión entre cuentas para compartir datos.

- Recursos con nombre de Lake Formation

  El intercambio de datos entre cuentas de Lake Formation mediante el método de recursos con nombre asignado le permite conceder permisos de Lake Formation con una opción de concesión en las tablas y bases de datos del catálogo de datos a directores Cuentas de AWS, organizaciones o unidades organizativas externas de IAM. La operación de concesión comparte automáticamente esos recursos.

> **Note**
>
> También puedes permitir que el AWS Glue rastreador acceda a un almacén de datos en una cuenta diferente con las credenciales de Lake Formation. Para obtener más información, consulte el rastreo entre cuentas en AWS Glue la Guía para desarrolladores.

Los servicios integrados, como Athena y Amazon Redshift Spectrum, requieren enlaces a recursos para poder incluir recursos compartidos en las consultas. Para obtener más información sobre los enlaces de recursos, consulte Cómo funcionan los enlaces de recursos en Lake Formation.

Para ver las consideraciones y limitaciones, consulte Prácticas recomendadas y consideraciones para uso compartido de datos entre cuentas.

Temas

- Requisitos previos
- Actualización de los ajustes de la versión entre cuentas para compartir datos
- Compartir tablas y bases de datos del Catálogo de datos entre Cuentas de AWS o entidades principales de IAM de cuentas externas
- Conceder permisos en una base de datos o tabla compartida con su cuenta
- Conceder permisos de enlace de recursos
- Acceso a los datos subyacentes de una tabla compartida
- Registro multicuenta CloudTrail
- Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation
- Visualización de todas las subvenciones entre cuentas mediante la operación de la API GetResourceShares

Temas relacionados de

- Descripción general de los permisos de Lake Formation
- Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos
- Creación de enlaces de recursos
- Solución de problemas de acceso entre cuentas

Requisitos previos

Antes de que su AWS cuenta pueda compartir los recursos del catálogo de datos (bases de datos y tablas) con otra cuenta o con los responsables de otra cuenta, y antes de que pueda acceder a los recursos compartidos con su cuenta, debe cumplir los siguientes requisitos previos.

Requisitos generales para compartir datos entre cuentas

- Para compartir bases de datos y tablas del Catálogo de datos en modo de acceso híbrido, debe actualizar la Configuración de la versión entre cuentas a la Versión 4.
- Antes de conceder permisos entre cuentas en un recurso del Catálogo de datos, debe revocar todos los permisos de Lake Formation del grupo `IAMAllowedPrincipals` para el recurso. Si la entidad principal que llama tiene permisos entre cuentas para acceder a un

recurso y el permiso `IAMAllowedPrincipals` existe en el recurso, Lake Formation genera `AccessDeniedException`.

Este requisito solo se aplica cuando se registra la ubicación de datos subyacente en el modo de Lake Formation. Si registra la ubicación de datos en modo híbrido, los permisos del grupo `IAMAllowedPrincipals` pueden existir en la base de datos o tabla compartida.

- En las bases de datos que contengan tablas que desee compartir, debe evitar que las nuevas tablas tengan una concesión predeterminada de `Super` a `IAMAllowedPrincipals`. En la consola de Lake Formation, edite la base de datos y desactive Usar solo el control de acceso de IAM para las nuevas tablas de esta base de datos o introduzca el siguiente AWS CLI comando, sustituyéndolo por `database` el nombre de la base de datos. Si la ubicación de datos subyacente está registrada en modo de acceso híbrido, no necesita cambiar esta configuración predeterminada. En el modo de acceso híbrido, Lake Formation le permite aplicar de forma selectiva los permisos de Lake Formation y las políticas de permisos de IAM para Amazon S3 y AWS Glue en el mismo recurso.

```
aws glue update-database --name database --database-input
 '{"Name":"database","CreateTableDefaultPermissions":[]}'
```

- Para conceder permisos entre cuentas, el otorgante debe tener los permisos AWS Identity and Access Management (IAM) necesarios para operar y prestar servicio. AWS Glue AWS RAM La política AWS gestionada `AWSLakeFormationCrossAccountManager` concede los permisos necesarios.

  Los administradores de lagos de datos de las cuentas que reciben recursos compartidos AWS RAM deben tener la siguiente política adicional. Permite al administrador aceptar invitaciones para compartir AWS RAM recursos. También permite al administrador habilitar el intercambio de recursos con las organizaciones.

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

```
                "ec2:DescribeAvailabilityZones",
                "ram:EnableSharingWithAwsOrganization"
            ],
            "Resource": "*"
        }
    ]
}
```

- Si quieres compartir los recursos del catálogo de datos con AWS Organizations o las unidades organizativas, debes habilitar el uso compartido con las organizaciones en AWS RAM.

  Para obtener información sobre cómo habilitar el uso compartido con las organizaciones, consulte Habilitar el uso compartido con AWS las organizaciones en la Guía del AWS RAM usuario.

  Debe tener el permiso `ram:EnableSharingWithAwsOrganization` para habilitar el uso compartido con las organizaciones.
- Para compartir recursos directamente con una entidad principal de IAM de otra cuenta, debe actualizar la Configuración de la versión entre cuentas a la Versión 3. Esta configuración está disponible en la página Configuración del Catálogo de datos. Si utiliza la Versión 1, consulte las instrucciones para actualizar la configuración Actualización de los ajustes de la versión entre cuentas para compartir datos.
- No puede compartir los recursos del catálogo de datos cifrados con una clave gestionada por el AWS Glue servicio con otra cuenta. Solo puede compartir los recursos del Catálogo de datos cifrados con la clave de cifrado del cliente, y la cuenta que reciba el recurso compartido debe tener permisos sobre la clave de cifrado del Catálogo de datos para descifrar los objetos.

Uso compartido de datos entre cuentas según los requisitos del LF-TBAC

- Para compartir los recursos del catálogo de datos con AWS Organizations unidades organizativas (OU), debes actualizar la configuración de la versión multicuentas a la versión 3.
- Para compartir los recursos del Catálogo de datos con la versión 3 de la Configuración de la versión entre cuentas, el concedente debe tener los permisos de IAM definidos en la política administrada AWS `AWSLakeFormationCrossAccountManager` de su cuenta.
- Si utiliza la versión 1 o la versión 2 de la Configuración de la versión entre cuentas, debe disponer de una política de recursos (`glue:PutResourcePolicy`) del Catálogo de datos que active LF-TBAC. Para obtener más información, consulte Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation.

- Si actualmente está utilizando una política de recursos del Catálogo de datos AWS Glue para compartir recursos y desea conceder permisos entre cuentas utilizando la versión 3 de la Configuración de la versión entre cuentas, debe añadir el permiso `glue:ShareResource` en la configuración del Catálogo de datos utilizando la operación de la API `glue:PutResourcePolicy` como se muestra en la sección Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation. Esta política no es obligatoria si su cuenta no ha efectuado concesiones entre cuentas utilizando la política de recursos del Catálogo de datos AWS Glue (permiso `glue:PutResourcePolicy` de uso en las versiones 1 y 2) para conceder el acceso entre cuentas.

```
{
      "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"
      ]
    }
```

- Si su cuenta ha compartido recursos entre cuentas mediante la política de recursos AWS Glue del Catálogo de datos, y actualmente está utilizando el método de recursos con nombre o LF-TBAC con la Configuración entre cuentas versión 3 para compartir recursos, que utiliza AWS RAM para compartir recursos, debe establecer el argumento `EnableHybrid` en `'true'` cuando invoque la operación API `glue:PutResourcePolicy`. Para obtener más información, consulte Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation.

Se requiere una configuración en cada cuenta que acceda al recurso compartido

- Si comparte recursos con Cuentas de AWS, al menos un usuario de la cuenta de consumidor debe ser administrador de un lago de datos para poder ver los recursos compartidos. Para obtener más información sobre cómo crear un administrador de lago de datos, consulte Crear un administrador de lago de datos.

El administrador del lago de datos puede conceder permisos de Lake Formation sobre los recursos compartidos a otras entidades principales de la cuenta. Otras entidades principales no pueden acceder a los recursos compartidos hasta que el administrador del lago de datos les conceda permisos sobre ellos.

- Los servicios integrados, como Athena y Amazon Redshift Spectrum, requieren enlaces a recursos para poder incluir recursos compartidos en las consultas. Las entidades principales deben crear un enlace de recursos en su Catálogo de datos a un recurso compartido de otra Cuenta de AWS. Para obtener más información sobre los enlaces de recursos, consulte Cómo funcionan los enlaces de recursos en Lake Formation.
- Cuando un recurso se comparte directamente con una entidad principal de IAM, para consultar la tabla utilizando Athena, la entidad principal debe crear un enlace de recursos. Para crear un enlace de recursos, la entidad principal necesita el permiso `CREATE_TABLE` o `CREATE_DATABASE` de Lake Formation y el permiso de `glue:CreateTable` o `glue:CreateDatabase` de IAM.

  Si la cuenta de productor comparte una tabla diferente en la misma base de datos con la misma entidad principal u otra, esta puede consultar la tabla inmediatamente.

> **Note**
>
> Para el administrador del lago de datos y para las entidades principales a las que el administrador del lago de datos haya concedido permisos, los recursos compartidos aparecen en el Catálogo de datos como si fueran recursos locales (de su propiedad). Las tareas de extracción, transformación y carga (ETL) pueden acceder a los datos subyacentes de los recursos compartidos.
> En el caso de los recursos compartidos, las páginas Tablas y Bases de datos de la consola de Lake Formation muestran el ID de cuenta del propietario.
> Cuando se accede a los datos subyacentes de un recurso compartido, los eventos de CloudTrail registro se generan tanto en la cuenta del destinatario del recurso compartido como en la cuenta del propietario del recurso. Los CloudTrail eventos pueden contener el ARN del principal que accedió a los datos, pero solo si la cuenta del destinatario opta por incluir el ARN principal en los registros. Para obtener más información, consulte Registro multicuenta CloudTrail .

Actualización de los ajustes de la versión entre cuentas para compartir datos

De vez en cuando, AWS Lake Formation actualiza la configuración del intercambio de datos entre cuentas para distinguir los cambios realizados en el AWS RAM uso y para admitir las actualizaciones realizadas en la función de intercambio de datos entre cuentas. Al hacer esto, Lake Formation crea una nueva versión de la Configuración de la versión entre cuentas.

Principales diferencias en las configuraciones de las versiones entre cuentas

Para obtener más información sobre cómo funciona el intercambio de datos entre cuentas en Configuraciones de la versión entre cuentas diferentes, consulte las secciones siguientes.

> **Note**
>
> Para compartir datos con otra cuenta, el concedente debe haber `AWSLakeFormationCrossAccountManager` administrado los permisos de la política de IAM. Es un requisito previo para todas las versiones.
> Actualizar la Configuración de la versión entre cuentas no afecta a los permisos que el destinatario tiene en los recursos compartidos. Esto se aplica al actualizar de la versión 1 a la versión 2, de la versión 2 a la versión 3 y de la versión 1 a la versión 3. Consulte las consideraciones que se indican a continuación al actualizar las versiones.

Versión 1

Método de recurso con nombre asignado: asigna cada permiso de Lake Formation multicuenta otorgado a un AWS RAM recurso compartido. El usuario (rol de concedente o entidad principal) no requiere permisos adicionales.

Método LF-TBAC: las concesiones de permisos entre cuentas de Lake Formation no AWS RAM se utilizan para compartir datos. Debe tener permiso de `glue:PutResourcePolicy`.

Ventajas de actualizar las versiones: versión inicial, no aplicable.

Consideraciones al actualizar las versiones: versión inicial, no aplicable

Versión 2

Método de recurso con nombre asignado: optimiza el número de AWS RAM recursos compartidos al asignar varias concesiones de permisos entre cuentas a un recurso compartido. AWS RAM El usuario no necesita permisos adicionales.

Método LF-TBAC: las concesiones de permisos entre cuentas de Lake Formation no AWS RAM se utilizan para compartir datos. Debe tener permiso de `glue:PutResourcePolicy`.

Ventajas de la actualización de las versiones: configuración escalable para varias cuentas mediante una utilización óptima de la capacidad. AWS RAM

Consideraciones a la hora de actualizar las versiones: Los usuarios que deseen conceder permisos de Lake Formation para varias cuentas deben disponer de los permisos de la política `AWSLakeFormationCrossAccountManager` AWS gestionada. De lo contrario, debe tener los permisos `ram:AssociateResourceShare` y `ram:DisassociateResourceShare` necesarios para compartir correctamente los recursos con otra cuenta.

Versión 3

Método de recurso con nombre asignado: optimiza la cantidad de AWS RAM recursos compartidos al asignar varias concesiones de permisos entre cuentas a un AWS RAM recurso compartido. El usuario no necesita permisos adicionales.

Método LF-TBAC: Lake Formation utiliza AWS RAM subvenciones entre cuentas. El usuario debe añadir la declaración glue: ShareResource al permiso. `glue:PutResourcePolicy` El destinatario debe aceptar las invitaciones para compartir recursos de AWS RAM.

Ventajas de actualizar las versiones: admite las funciones siguientes:

- Permite compartir recursos explícitamente con una entidad principal de IAM en una cuenta externa.

  Para obtener más información, consulte Concesión y revocación de permisos sobre los recursos del catálogo de datos.
- Permite compartir entre cuentas utilizando el método LF-TBAC a organizaciones o unidades organizativas (UO).
- Elimina la sobrecarga que supone mantener AWS Glue políticas adicionales para las subvenciones entre cuentas.

Consideraciones al actualizar las versiones: Si el concedente utiliza una versión anterior a la versión 3 y el destinatario utiliza la versión 3 o posterior, el concedente recibe el

siguiente mensaje de error: «Solicitud de concesión entre cuentas no válida». La cuenta de consumidor ha optado por utilizar la versión entre cuentas: v3. Actualice `DataLakeSetting` a `CrossAccountVersion` la versión mínima v3 (servicio: AmazonDataCatalog; código de estado: 400; código de error: InvalidInputException)». Sin embargo, si el concedente usa la versión 3 y el destinatario la versión 1 o 2, las concesiones entre cuentas se tramitan correctamente.

Para compartir recursos directamente con entidades principales de IAM en otra cuenta, solo el concedente debe utilizar la versión 3.

Las concesiones entre cuentas efectuadas con el método LF-TBAC requieren que los usuarios tengan una política de recursos AWS Glue Data Catalog en la cuenta. Al actualizar a la versión 3, el LF-TBAC concede los usos AWS RAM. Para permitir que AWS RAM las subvenciones multicuenta tengan éxito, debe añadir la `glue:ShareResource` declaración a las políticas de recursos actuales del catálogo de datos, tal y como se muestra en la Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation sección.

Versión 4

El concedente necesita la versión 4 o superior para compartir los recursos del Catálogo de datos en el modo de acceso híbrido.

Optimice los AWS RAM recursos compartidos

Las nuevas versiones (versión 2 y versiones posteriores) de las subvenciones multicuentas utilizan de manera óptima la AWS RAM capacidad para maximizar el uso multicuenta. Al compartir un recurso con un director externo Cuenta de AWS o de IAM, Lake Formation puede crear un nuevo recurso compartido o asociar el recurso con un recurso compartido existente. Al asociarse a los recursos compartidos existentes, Lake Formation reduce el número de invitaciones a compartir recursos que un consumidor tiene que aceptar.

Habilite el AWS RAM uso compartido a través de TBAC o comparta los recursos directamente con los directores

Para compartir recursos directamente con entidades principales de IAM en otra cuenta o para habilitar los recursos compartidos entre cuentas TBAC a organizaciones o unidades organizativas, debe actualizar la Configuración de la versión entre cuentas a la versión 3. Para obtener más información sobre los límites de AWS RAM recursos, consulte. Prácticas recomendadas y consideraciones para uso compartido de datos entre cuentas

Permisos necesarios para actualizar la configuración de las versiones entre cuentas

Si un concedente de permisos entre cuentas tiene permisos de política de IAM gestionados por `AWSLakeFormationCrossAccountManager`, no se requiere ninguna configuración de permisos adicional para el rol o la entidad principal del concedente de permisos entre cuentas. Sin embargo, si el concedente entre cuentas no está utilizando la política administrada, entonces el rol de concedente o entidad principal debe tener concedidos los siguientes permisos de IAM para que la nueva versión de la concesión entre cuentas tenga éxito.

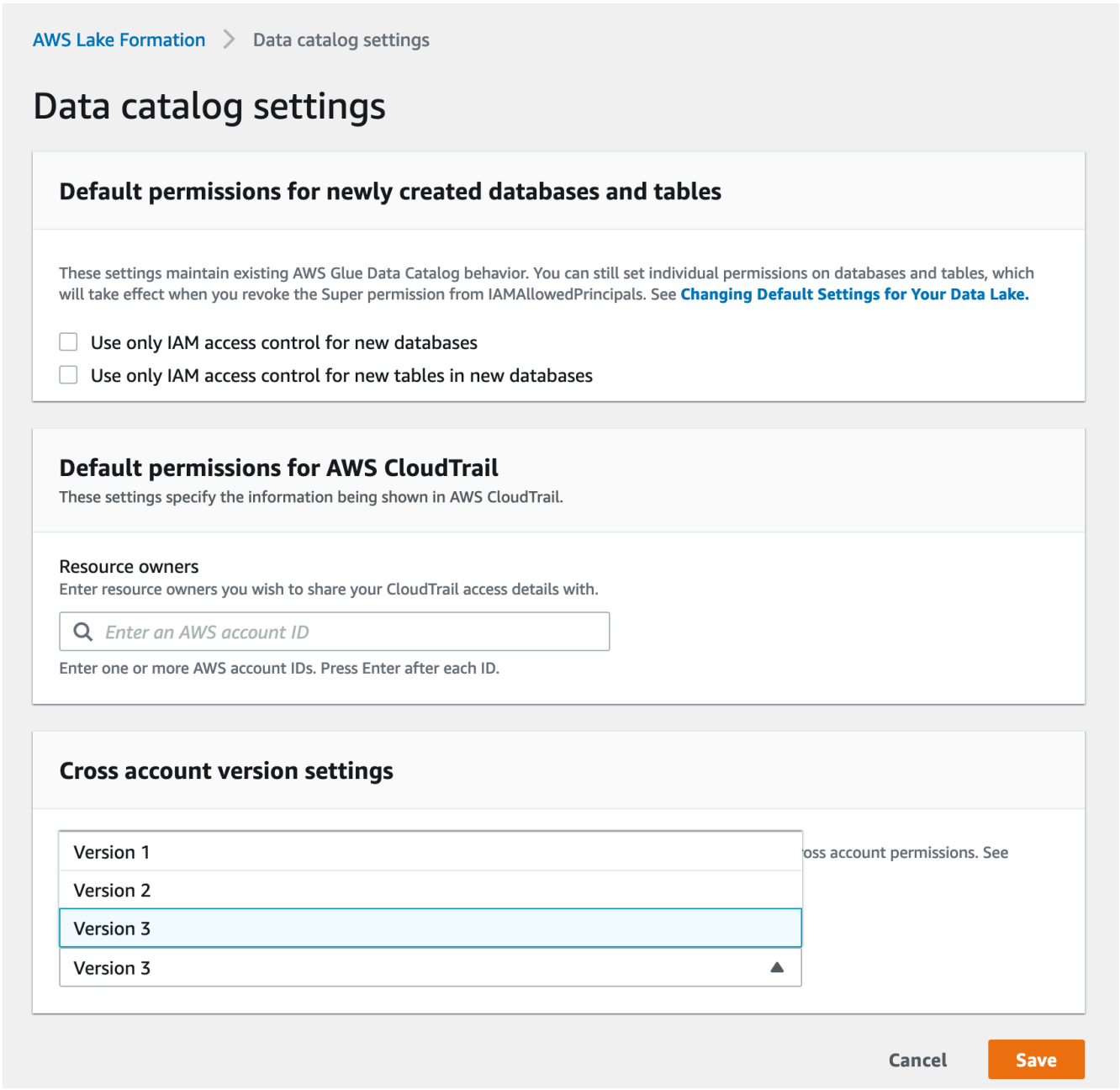
```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
         "ram:AssociateResourceShare",
         "ram:DisassociateResourceShare",
         "ram:GetResourceShares"
       ],
     "Resource": "*",
     "Condition": {
       "StringLike": {
         "ram:ResourceShareName": "LakeFormation*"
        }
      }
    }
  ]
}
```

Para habilitar la nueva versión

Siga estos pasos para actualizar la configuración de la versión multicuenta a través de la AWS Lake Formation consola o el AWS CLI.

Console

1. Elija Versión 2, Versión 3 o Versión 4 en la Configuración de la versión entre cuentas en la página Configuración del Catálogo de datos. Si selecciona Versión 1, Lake Formation utilizará el modo de recursos compartidos predeterminado.



2. Seleccione Guardar.

AWS Command Line Interface (AWS CLI)

Utilice el `put-data-lake-settings` AWS CLI comando para configurar el `CROSS_ACCOUNT_VERSION` parámetro. Los valores aceptados son 1, 2, 3 y 4.

```
aws lakeformation put-data-lake-settings --region us-east-1 --data-lake-settings
 file://settings
{
```

```
    "DataLakeAdmins": [
        {
            "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/test"
        }
    ],
    "CreateDatabaseDefaultPermissions": [],
    "CreateTableDefaultPermissions": [],
    "Parameters": {
        "CROSS_ACCOUNT_VERSION": "3"
    }
}
```

Important

En cuanto elija la Versión 2 o la Versión 3, todas las nuevas concesiones de recursos con nombre pasarán por el nuevo modo de concesión entre cuentas. Para aprovechar al máximo la AWS RAM capacidad de sus acciones multicuenta existentes, le recomendamos que revoque las concesiones que se concedieron con la versión anterior y las vuelva a conceder en el nuevo modo.

Compartir tablas y bases de datos del Catálogo de datos entre Cuentas de AWS o entidades principales de IAM de cuentas externas

En esta sección se incluyen instrucciones sobre cómo habilitar los permisos entre cuentas en las tablas y bases de datos del catálogo de datos para una AWS cuenta externa, un director de IAM, una organización o una unidad organizativa. La operación de concesión comparte automáticamente esos recursos.

Temas

- Compartir datos mediante el control de acceso basado en etiquetas
- Uso compartido de datos entre cuentas utilizando el método de recursos con nombre

Compartir datos mediante el control de acceso basado en etiquetas

Configuración necesaria en la cuenta del productor/concedente

1. Defina una etiqueta LF. Para obtener instrucciones sobre cómo crear una etiqueta LF, consulte Crear etiquetas LF.
2. Asigne la etiqueta LF al recurso de destino. Para obtener más información, consulte Asignación de varias etiquetas LF a recursos del catálogo de datos.
3. Conceda el permiso de etiqueta LF a la cuenta externa. Para obtener más información, consulte Concesión de permisos de etiqueta LF desde la consola.

   En este punto, el administrador del lago de datos del consumidor debería poder encontrar la etiqueta de política que se comparte a través de la consola de Lake Formation de la cuenta del beneficiario, en Permisos, Roles y tareas administrativas, Etiquetas LF.
4. Concede permiso de datos a la cuenta externa o del beneficiario.
   a. En el panel de navegación, en Permisos, Permisos de lago de datos, seleccione Conceder.
   b. En el caso de los principales, elija Cuentas externas e introduzca el Cuenta de AWS ID de destino o la función de IAM del principal o el nombre de recurso de Amazon (ARN) del principal (ARN principal).
   c. Para las etiquetas LF o los recursos del catálogo, elija la clave y los valores de la etiqueta LF que se va a compartir con la cuenta del consumidor (clave `Confidentiality` y valor `public`).
   d. En Permisos, bajo Recursos que coinciden con las etiquetas LF (recomendado) elija Agregar etiqueta LF.
   e. Selecciona la clave y el valor de la etiqueta que se comparte con la cuenta del beneficiario (clave `Confidentiality` y valor `public`).
   f. Para Permisos de base de datos, seleccione Describir en Permisos de bases de datos para conceder permisos de acceso a nivel de base de datos.
   g. El administrador del lago de datos del consumidor debería poder encontrar la etiqueta de política que se comparte a través de la cuenta del consumidor en la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/, en Permisos, Roles y tareas administrativas, Etiquetas LF.
   h. Seleccione Describir en Permisos otorgables para que la cuenta del consumidor pueda conceder permisos a nivel de base de datos a sus usuarios.

Como el administrador del lago de datos debe conceder permisos sobre los recursos compartidos a las entidades principales de la cuenta beneficiaria, los permisos entre cuentas siempre deben concederse con la opción de concesión.

> **Note**
>
> Las entidades principales que reciban concesiones directas entre cuentas no dispondrán de la opción Permisos concedibles.

i. En Permisos de tabla y columna, elija Seleccionar y Describir en la Permisos de tabla.

j. Elija Seleccionar y Describir en Permisos concedibles.

k. Elija Conceder.

Se requiere una configuración en la cuenta receptora o beneficiaria

1. Al compartir un recurso con otra cuenta, el recurso sigue perteneciendo a la cuenta del productor y no está visible en la consola de Athena. Para que el recurso sea visible en la consola de Athena, debe crear un enlace de recurso que apunte al recurso compartido. Para obtener instrucciones sobre cómo crear un enlace a un recurso, consulte Crear un enlace de recursos a una tabla de catálogo de datos compartida y Crear un enlace de recursos a una base de datos de catálogo de datos compartida
2. Debe crear un conjunto independiente de etiquetas LF en la cuenta del consumidor para utilizar el control de acceso basado en etiquetas LF al compartir los enlaces de recursos. Cree y asigne las etiquetas LF necesarias a las bases de datos o tablas compartidas y a los enlaces de recursos.
3. Conceda permisos sobre estas etiquetas LF a las entidades principales de IAM de la cuenta del beneficiario.

Uso compartido de datos entre cuentas utilizando el método de recursos con nombre

Puede conceder permisos directamente a los directores de la otra AWS cuenta o a una cuenta externa o. Cuentas de AWS AWS Organizations Otorgar permisos de Lake Formation a organizaciones o unidades organizativas equivale a conceder el permiso a todos los Cuenta de AWS miembros de esa organización o unidad organizativa.

Al conceder permisos a cuentas u organizaciones externas, debe incluir la opción de Permisos concedibles. Solo el administrador del lago de datos de la cuenta externa puede acceder a los

recursos compartidos hasta que el administrador conceda permisos sobre los recursos compartidos a otras entidades principales de la cuenta externa.

> **Note**
>
> La opción Permisos otorgables no es compatible cuando se conceden permisos directamente a entidades principales de IAM desde cuentas externas.

Siga las instrucciones de Concesión de permisos de base de datos mediante el método de recurso con nombre para conceder permisos entre cuentas según el método de recurso designado.

Conceder permisos en una base de datos o tabla compartida con su cuenta

Después de compartir con su AWS cuenta un recurso del catálogo de datos que pertenece a otra AWS cuenta, como administrador del lago de datos, puede conceder permisos sobre el recurso compartido a otros responsables de su cuenta. Sin embargo, no puede conceder permisos sobre el recurso a otras cuentas AWS u organizaciones.

Puede usar la AWS Lake Formation consola, la API o AWS Command Line Interface (AWS CLI) para conceder los permisos.

Para conceder permisos en una base de datos compartida (denominado método de recurso, consola)

- Siga las instrucciones en Concesión de permisos de base de datos mediante el método de recurso con nombre. En la lista Bases de datos, bajo Etiquetas LF o recursos del catálogo, asegúrese de seleccionar la base de datos en la cuenta externa y no un enlace a un recurso para la base de datos.

  Si no ve la base de datos en la lista, asegúrese de haber aceptado la invitación de la base de datos para compartir del recurso AWS Resource Access Manager (AWS RAM). Para obtener más información, consulte Aceptar una invitación para compartir recursos de AWS RAM.

  Además, para los permisos `CREATE_TABLE` y `ALTER`, siga las instrucciones de Concesión de permisos de localización de datos (misma cuenta) y asegúrese de introducir el ID de la cuenta propietaria en el campo Ubicación de la cuenta registrada.

Para conceder permisos en una tabla compartida (método de recurso con nombre, consola)

- Siga las instrucciones en Concesión de permisos de tabla mediante el método de recursos con nombre. En la lista Bases de datos, bajo Etiquetas LF o recursos del catálogo, asegúrese de seleccionar la base de datos en la cuenta externa y no un enlace a un recurso para la base de datos.

  Si no ve la tabla en la lista de tablas, compruebe que ha aceptado la invitación para compartir el recurso AWS RAM para la tabla. Para obtener más información, consulte Aceptar una invitación para compartir recursos de AWS RAM.

  Asimismo, para el permiso de `ALTER`, siga las instrucciones de Concesión de permisos de localización de datos (misma cuenta) y asegúrese de introducir el ID de la cuenta propietaria en el campo Ubicación de la cuenta registrada.

Para conceder permisos sobre recursos compartidos (método LF-TBAC, consola)

- Siga las instrucciones en Concesión de permisos del catálogo de datos . En la sección LF-Tags o recursos del catálogo, conceda la expresión LF-Tag exacta que la cuenta externa concedió a su cuenta, o un subconjunto de esa expresión.

  Por ejemplo, si una cuenta externa otorgó la expresión LF-tag `module=customers AND environment=production` a su cuenta con la opción de concesión, como administrador de un lago de datos, puede conceder esa misma expresión, `module=customers` o `environment=production` a una entidad principal de su cuenta. Puede conceder solo los mismos permisos o un subconjunto de los permisos de Lake Formation (por ejemplo, `SELECT`, `ALTER`, etc.) que se concedieron sobre los recursos mediante la expresión LF-Tag.

Para conceder permisos en una tabla compartida (método de recurso denominado AWS CLI)

- Introduzca un comando similar al siguiente. En este ejemplo:
  - El ID AWS de tu cuenta es 1111-2222-3333.
  - La cuenta propietaria de la tabla y que ha concedido el permiso a su cuenta es 1234-5678-9012.
  - El permiso `SELECT` se concede al usuario `datalake_user1` en la tabla compartida `pageviews`. Ese usuario es la entidad principal de su cuenta.

- La tabla `pageviews` se encuentra en la base de datos `analytics`, perteneciente a la cuenta 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"}}'
```

Tenga en cuenta que la cuenta propietaria debe especificarse en la propiedad `CatalogId` del argumento `resource`.

Conceder permisos de enlace de recursos

Sigue estos pasos para conceder AWS Lake Formation permisos en uno o más enlaces de recursos a una entidad principal de tu cuenta. AWS

Después de crear un enlace de recursos, solo podrá verlo y acceder a él por sí mismo. (Esto supone que Control de acceso IAM solo para nuevas tablas en esta base de datos no está activado para la base de datos). Para permitir que otras entidades principales de su cuenta accedan al enlace de recursos, conceda al menos el permiso `DESCRIBE`.

> **Important**
>
> Conceder permisos en un enlace a un recurso no otorga permisos en la tabla o base de datos (vinculada) de destino. Debe conceder los permisos en el destino por separado.

Puede conceder permisos mediante la consola de Lake Formation, la API o AWS Command Line Interface (AWS CLI).

console

Para conceder permisos de enlaces de recursos utilizando la consola de Lake Formation

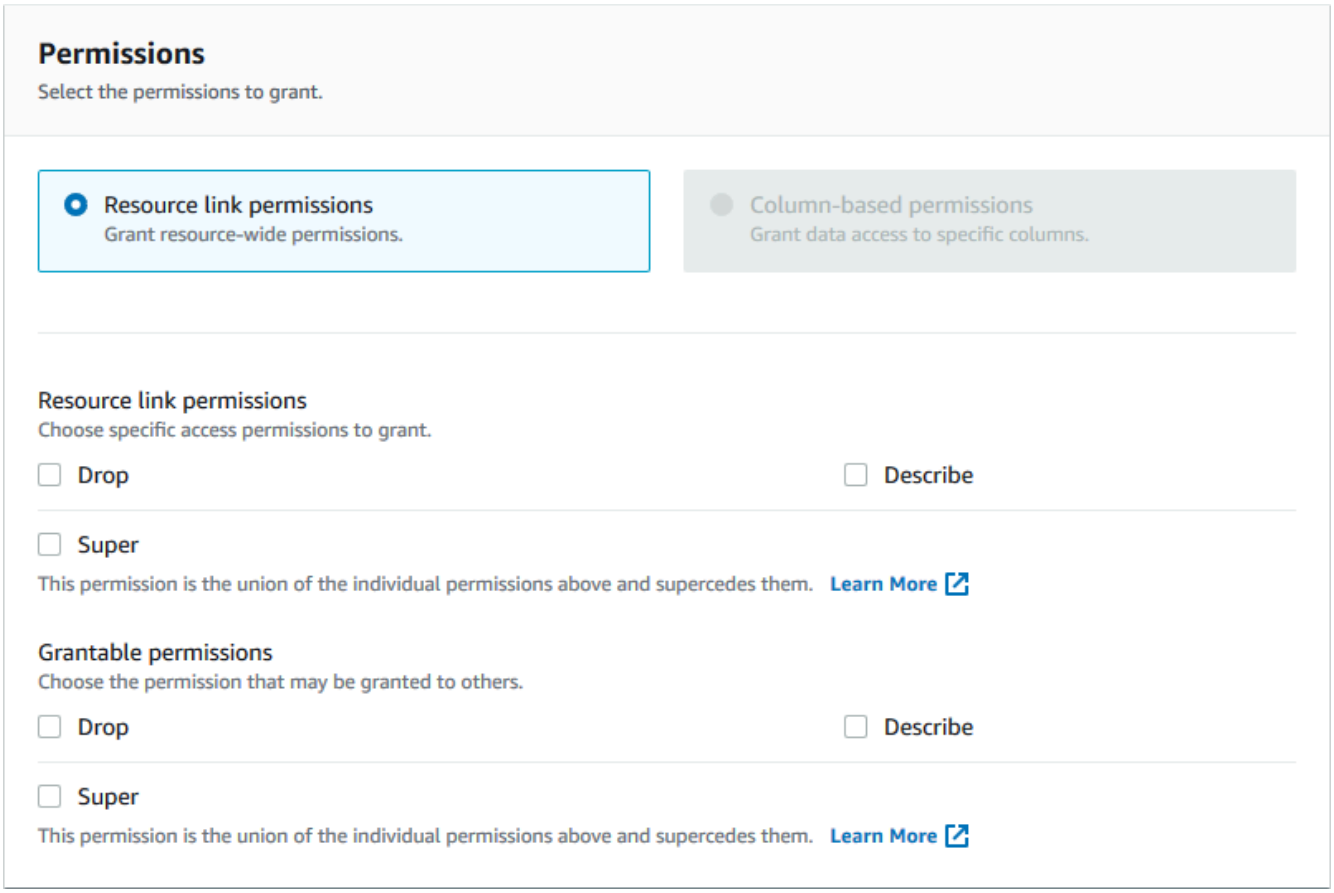
1. Realice una de las acciones siguientes:

   - Para enlaces de recursos de base de datos, siga los pasos en Concesión de permisos de base de datos mediante el método de recurso con nombre para hacer lo siguiente:

1. Abra la página Conceder permisos de lagos de datos.
2. Especifique las bases de datos. Especifique uno o varios enlaces de recursos de base de datos.
3. Especifique las entidades principales.

- Para enlaces de recursos de tabla, siga los pasos de Concesión de permisos de tabla mediante el método de recursos con nombre para hacer lo siguiente:
  1. Abra la página Conceder permisos de lagos de datos.
  2. Especifique las tablas. Especifique uno o varios enlaces de recursos de tabla.
  3. Especifique las entidades principales.

2. En Permisos, seleccione los que va a conceder. Si lo desea, seleccione los permisos que pueden concederse.



3. Elija Conceder.

AWS CLI

Para conceder permisos de enlace de recursos mediante AWS CLI

- Ejecute el comando `grant-permissions` y especifique un enlace a un recurso como recurso.

  Example

  En este ejemplo, se concede `DESCRIBE` al usuario de la tabla el `datalake_user1` enlace de recursos de la base `incidents-link` de datos de `issues` la AWS cuenta 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"}}'
```

Consulte también:

- Creación de enlaces de recursos
- Referencia de permisos de Lake Formation

Acceso a los datos subyacentes de una tabla compartida

Supongamos que la AWS cuenta A comparte una tabla del catálogo de datos con la cuenta B, por ejemplo, concediéndola `SELECT` con la opción de concesión de la tabla a la cuenta B. Para que un principal de la cuenta B pueda leer los datos subyacentes de la tabla compartida, se deben cumplir las siguientes condiciones:

- El administrador del lago de datos de la cuenta B debe aceptar el uso compartido. (Esto no es necesario si las cuentas A y B pertenecen a la misma organización o si la concesión se efectuó con el método de control de acceso basado en etiquetas de Lake Formation).
- El administrador del lago de datos debe volver a conceder a la entidad principal el permiso `SELECT` de Lake Formation que la cuenta A concedió en la tabla compartida.

- La entidad principal debe tener los siguientes permisos IAM sobre la tabla, la base de datos que la contiene y la cuenta A Catálogo de datos.

> **Note**
>
> En la política de IAM siguiente:
>
> - <account-id-A>Sustitúyala por el ID AWS de cuenta de la cuenta A.
> - Sustituya *<region>* por una región válida.
> - Sustituya *<base de datos>* por el nombre de la base de datos de la cuenta A que contiene la tabla compartida.
> - Sustituya *<table>* por el nombre de la tabla compartida.

```
{
    "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:<region>:<account-id-A>:table/<database>/<table>",
            "arn:aws:glue:<region>:<account-id-A>:database/<database>",
            "arn:aws:glue:<region>:<account-id-A>:catalog"
           ]
        },
        {
          "Effect": "Allow",
          "Action": [
            "lakeformation:GetDataAccess"
           ],
          "Resource": [
            "*"
```

```
            ],
           "Condition": {
             "StringEquals": {
               "lakeformation:GlueARN":"arn:aws:glue:<region>:<account-id-
A>:table/<database>/<table>"
             }
         }
     }
   ]
}
```

Consulte también:

- Aceptar una invitación para compartir recursos de AWS RAM

Registro multicuenta CloudTrail

Lake Formation proporciona un registro de auditoría centralizado de todos los accesos entre cuentas a los datos de su lago de datos. Cuando una AWS cuenta de destinatario accede a los datos de una tabla compartida, Lake Formation copia el CloudTrail evento en los registros de CloudTrail la cuenta propietaria. Los eventos copiados incluyen consultas a los datos por parte de servicios integrados, como Amazon Redshift Spectrum, Amazon Athena y accesos AWS Glue a los datos por parte de los trabajos.

CloudTrail Los eventos de las operaciones entre cuentas en los recursos del catálogo de datos se copian de forma similar.

Como propietario del recurso, si habilita el registro a nivel de objeto en Amazon S3, puede ejecutar consultas que unan los eventos de S3 con CloudTrail los eventos de Lake Formation CloudTrail para determinar las cuentas que han accedido a sus buckets de S3.

Temas

- Incluir las identidades principales en los registros entre cuentas CloudTrail
- Consulta de CloudTrail registros para el acceso entre cuentas de Amazon S3

Incluir las identidades principales en los registros entre cuentas CloudTrail

De forma predeterminada, los CloudTrail eventos entre cuentas que se agregan a los registros del destinatario del recurso compartido y se copian en los registros del propietario del recurso contienen solo el ID AWS principal de la cuenta externa, no el nombre de recurso de Amazon (ARN) legible para los humanos del principal (ARN principal). Al compartir recursos dentro de límites de confianza, como dentro de la misma organización o equipo, puede optar por incluir el ARN principal en los CloudTrail eventos. A continuación, las cuentas de propietarios de recursos pueden hacer un seguimiento de las entidades principales de las cuentas de destinatarios que acceden a sus recursos en propiedad.

> **Important**
>
> Como destinatario de recursos compartidos, para ver el ARN principal de los eventos en sus propios CloudTrail registros, debe optar por compartir el ARN principal con la cuenta del propietario.
> Si el acceso a los datos se produce a través de un enlace de recursos, se registran dos eventos en la cuenta del destinatario del recurso compartido: uno para el acceso al enlace de recursos y otro para el acceso al recurso de destino. El evento para el acceso al enlace de recursos sí incluye el ARN de la entidad principal. El evento de acceso al recurso de destino no incluye el ARN de la entidad principal sin la activación. El evento de acceso al enlace de recursos no se copia en la cuenta del propietario.

El siguiente es un extracto de un CloudTrail evento multicuenta predeterminado (sin suscripción). La cuenta que efectúa el acceso a los datos es 1111-2222-3333. Este es el registro que se muestra en la cuenta de llamada y en la cuenta del propietario del recurso. Lake Formation rellena los registros de ambas cuentas en el caso entre cuentas.

```
{
    "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"
    },
...
}
```

Como consumidor de recursos compartidos, cuando opta por incluir el ARN de la entidad principal, el resumen pasa a ser el siguiente. El campo `lakeFormationPrincipal` representa el rol o el usuario final que hace la consulta a través de Amazon Athena, Amazon Redshift Spectrum o trabajos de 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"
    },
...
}
```

Para optar por incluir los ARN principales en los registros de varias cuentas CloudTrail

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como usuario de `Administrator` o como usuario con la política de IAM de `Administrator Access`.
2. En el panel de navegación, seleccione Configuración.

3. En la página de configuración del catálogo de datos, en la AWS CloudTrail sección Permisos predeterminados para los propietarios de los recursos, introduzca uno o más identificadores de cuenta del propietario del AWS recurso.

   Pulse Intro después de cada ID de cuenta.
4. Seleccione Guardar.

   Ahora, CloudTrail los eventos entre cuentas almacenados en los registros tanto para el destinatario del recurso compartido como para el propietario del recurso contienen el ARN principal.

Consulta de CloudTrail registros para el acceso entre cuentas de Amazon S3

Como propietario de un recurso compartido, puede consultar CloudTrail los registros de S3 para determinar las cuentas que han accedido a sus buckets de Amazon S3 (siempre que haya habilitado el registro a nivel de objeto en Amazon S3). Esto solo se aplica a las ubicaciones S3 que haya registrado en Lake Formation. Si los consumidores de recursos compartidos optan por incluir los Rans principales en los CloudTrail registros de Lake Formation, puede determinar las funciones o los usuarios que accedieron a los depósitos.

Al ejecutar consultas con Amazon Athena, puede unir los eventos de Lake Formation y CloudTrail los eventos de S3 CloudTrail en la propiedad del nombre de la sesión. Las consultas también pueden filtrar los eventos de Lake Formation en `eventName="GetDataAccess"` y los eventos de S3 en `eventName="Get Object"` o`eventName="Put Object"`.

El siguiente es un extracto de un CloudTrail evento entre cuentas de Lake Formation en el que se accedió a los datos de una ubicación S3 registrada.

```
{
  "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"
   }
}
```

El valor `lakeFormationRoleSessionName` clave,`AWSLF-00-GL-111122223333-B8JSAjo5QA`, se puede unir al nombre de la sesión en la `principalId` clave del evento S3 CloudTrail . El siguiente es un extracto del CloudTrail evento S3. Muestra la ubicación del nombre de la sesión.

```
{
   "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"
     },
   ..............
   ..............
}
```

El nombre de la base de datos se forma de la siguiente manera:

```
AWSLF-<version-number>-<query-engine-code>-<account-id->-<suffix>
```

`version-number`

La versión de este formato, actualmente `00`. Si cambia el formato del nombre de la sesión, la siguiente versión será la `01`.

`query-engine-code`

Indica la entidad que ha accedido a los datos. Los valores actuales son:

`GL`	Trabajo ETL de AWS Glue
`AT`	Athena

RE	Amazon Redshift Spectrum

`account-id`

El identificador de la AWS cuenta que solicitó las credenciales de Lake Formation.

`suffix`

Una cadena generada aleatoriamente.

Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation

Es posible otorgar acceso entre cuentas a los recursos del Catálogo de datos y a los datos subyacentes mediante AWS Glue o AWS Lake Formation.

EnAWS Glue, se conceden permisos para varias cuentas mediante la creación o actualización de una política de recursos del catálogo de datos. En Lake Formation, concede permisos entre cuentas mediante el modelo de permisos de `GRANT/REVOKE` de Lake Formation y la operación de la API `Grant Permissions`.

> **Tip**
>
> Recomendamos confiar solo en los permisos de Lake Formation para proteger su lago de datos.

Puede ver las subvenciones entre cuentas de Lake Formation mediante la consola Lake Formation o la consola AWS Resource Access Manager (AWS RAM). Sin embargo, esas páginas de la consola no muestran los permisos entre cuentas otorgados por la política de recursos del Catálogo de datos de AWS Glue. Del mismo modo, puede ver las concesiones entre cuentas en la política de recursos del Catálogo de datos mediante la página Configuración de la consola de AWS Glue, pero esa página no muestra los permisos entre cuentas concedidos mediante Lake Formation.

Para garantizar que no pierda ninguna subvención al consultar y gestionar los permisos entre cuentas, Lake Formation y AWS Glue requieren que haga lo siguiente para confirmar que conoce y permite concesiones entre cuentas por Lake Formation y por AWS Glue.

Al conceder permisos entre cuentas usando la política de recursos del Catálogo de datos de AWS Glue

Si su cuenta (cuenta de cedente o cuenta de productor) no ha realizado concesiones multicuentas que se utilicen AWS RAM para compartir los recursos, puede guardar una política de recursos del catálogo de datos como de costumbre en. AWS Glue Sin embargo, si ya se han concedido concesiones que implican el uso compartido de AWS RAM recursos, debe realizar una de las siguientes acciones para asegurarse de que la política de recursos se ha guardado correctamente:

- Cuando guarde la política de recursos en la página Configuración de la consola AWS Glue, esta emitirá una alerta indicando que los permisos de la política se sumarán a cualquier permiso concedido mediante la consola Lake Formation. Debe elegir Continuar para guardar la política.
- Al guardar la política de recursos mediante la operación de la API `glue:PutResourcePolicy`, debe establecer el campo `EnableHybrid` en 'TRUE' (tipo = cadena). El siguiente ejemplo de código muestra cómo hacerlo en Python.

```
import boto3
import json

REGION = 'us-east-2'
PRODUCER_ACCOUNT_ID = '123456789012'
CONSUMER_ACCOUNT_IDs = ['111122223333']

glue = glue_client = boto3.client('glue')

policy = {
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Cataloguers",
            "Effect": "Allow",
            "Action": [
                "glue:*"
            ],
            "Principal": {
                "AWS": CONSUMER_ACCOUNT_IDs
            },
            "Resource": [
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:catalog",
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:database/*",
                f"arn:aws:glue:{REGION}:{PRODUCER_ACCOUNT_ID}:table/*/*"
```

```
            ]
        }
    ]
}

policy = json.dumps(policy)
glue.put_resource_policy(PolicyInJson=policy, EnableHybrid='TRUE')
```

Para obtener más información, consulte PutResourcePolicy Action (Python: put_resource_policy) en la Guía para desarrolladores.AWS Glue

Al conceder permisos entre cuentas mediante el método de recursos con nombre de Lake Formation

Si no existe una política de recursos del Catálogo de datos en su cuenta, las concesiones entre cuentas de Lake Formation que efectúe procederán como de costumbre. Sin embargo, si existe una política de recursos del Catálogo de datos, debe agregarle la siguiente instrucción para que las concesiones entre cuentas se hagan correctamente si se sigue el método de recurso indicado. <region><account-id>Sustitúyalo por un nombre de región válido y por tu ID de cuenta. AWS

```
  {
    "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"
    ]
  }
```

Sin esta declaración adicional, la subvención de Lake Formation es válida, pero queda bloqueada y la cuenta receptora no puede acceder al recurso otorgado. AWS RAM

Important

Si sigue el método de control de acceso basado en etiquetas (LF-TBAC) de Lake Formation para hacer concesiones entre cuentas, debe contar con una política de recursos del Catálogo de datos que incluya al menos los permisos especificados en Requisitos previos.

Consulte también:

- Control de acceso a los metadatos (para comparar el método de recursos con nombre frente al método de control de acceso basado en etiquetas (LF-TBAC) de Lake Formation).
- Visualización de tablas y bases de datos compartidas del catálogo de datos
- Trabajo con la configuración del Catálogo de datos en la consola de AWS Glue en la Guía para desarrolladores de AWS Glue
- Concesión de acceso entre cuentas en la Guía para desarrolladores de AWS Glue (para ver ejemplos de las políticas de recursos del Catálogo de datos)

Visualización de todas las subvenciones entre cuentas mediante la operación de la API GetResourceShares

Si su empresa concede permisos multicuenta mediante una política de AWS Glue Data Catalog recursos y subvenciones de Lake Formation, la única forma de ver todas las concesiones multicuenta en un solo lugar es mediante la operación de `glue:GetResourceShares` API.

Cuando concedes permisos a Lake Formation en todas las cuentas mediante el método de recurso designado, AWS Resource Access Manager (AWS RAM) crea una política de recursos AWS Identity and Access Management (IAM) y la almacena en tu AWS cuenta. La política otorga los permisos necesarios para acceder al recurso. AWS RAM crea una política de recursos independiente para cada concesión multicuenta. Puede ver todas estas políticas mediante la operación de la API `glue:GetResourceShares`.

Note

Esta operación también devuelve la política de recursos del Catálogo de datos. Sin embargo, si ha activado el cifrado de metadatos en la configuración del catálogo de datos y no tiene

permiso para utilizar la AWS KMS clave, la operación no devolverá la política de recursos del catálogo de datos.

Para ver todas las concesiones entre cuentas

- Introduzca el siguiente AWS CLI comando.

```
aws glue get-resource-policies
```

El siguiente es un ejemplo de política de recursos que AWS RAM crea y almacena cuando se conceden permisos sobre una tabla `t` de la base de datos `db1` a la AWS cuenta 1111-2222-3333.

```
{
  "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:<region>:111122223333:table/db1/t"
     ]
    }
  ]
}
```

Véase también:

- GetResourceShares Acción (Python: get_resource_policies) en la Guía para desarrolladores AWS Glue

Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos

Para el administrador del lago de datos y para las entidades principales a las que se han concedido permisos, los recursos que se comparten con su cuenta AWS aparecen en el catálogo de datos como si fueran recursos de su cuenta. La consola muestra la cuenta propietaria del recurso.

Puede ver los recursos compartidos con su cuenta utilizando la consola de Lake Formation. También puede utilizar la consola AWS Resource Access Manager (AWS RAM) para ver tanto los recursos que están compartidos con su cuenta como los recursos que ha compartido con otras cuentas AWS utilizando el método de recursos con nombre.

Important

Cuando alguien utiliza el método de recursos con nombre para conceder permisos entre cuentas sobre un recurso del catálogo de datos a su cuenta u organización AWS, Lake Formation utiliza el servicio AWS Resource Access Manager (AWS RAM) para compartir el recurso. Si su cuenta pertenece a la misma organización AWS que la cuenta que concede la concesión, el recurso compartido estará a su disposición inmediatamente.
Sin embargo, si su cuenta no pertenece a la misma organización, AWS RAM envía una invitación a su cuenta para que acepte o rechace el recurso compartido. A continuación, para que el recurso compartido esté disponible, el administrador del lago de datos de su cuenta deberá utilizar la consola de AWS RAM o la CLI para aceptar la invitación.
La consola de Lake Formation muestra una alerta si hay una invitación para compartir recursos de AWS RAM esperando a ser aceptada. Solo los usuarios autorizados a ver las invitaciones AWS RAM recibirán la alerta.

Consulte también:

- Uso compartido de tablas y bases de datos del Catálogo de datos entre cuentas AWS
- Compartir datos entre cuentas en Lake Formation
- Acceso a los datos subyacentes de una tabla compartida
- Control de acceso a los metadatos (para obtener información sobre el método de recursos con nombre frente al método LF-TBAC para el uso compartido de recursos).

Temas

Aceptar una invitación para compartir recursos de AWS RAM

Si un recurso del catálogo de datos se comparte con su cuenta AWS y su cuenta no pertenece a la misma organización AWS que la cuenta que lo comparte, no tendrá acceso al recurso compartido hasta que acepte una invitación para compartir recursos de AWS Resource Access Manager (AWS RAM). Como administrador del lago de datos, primero debe consultar AWS RAM si hay invitaciones pendientes y después aceptar la invitación.

Puede utilizar la consola de AWS RAM, la API o AWS Command Line Interface (AWS CLI) para ver y aceptar invitaciones.

Para ver y aceptar una invitación para compartir recursos de AWS RAM (consola)

1. Asegúrese de que dispone de los permisos de AWS Identity and Access Management (IAM) necesarios para ver y aceptar invitaciones para compartir recursos.

   Para obtener información sobre las políticas de IAM sugeridas para los administradores del lago de datos, consulte the section called “Permisos de administrador del lago de datos”.
2. Siga las instrucciones de Aceptar y rechazar invitaciones en la Guía del usuario de AWS RAM.

Para ver y aceptar una invitación para compartir recursos de AWS RAM (AWS CLI)

1. Asegúrese de que dispone de los permisos de AWS Identity and Access Management (IAM) necesarios para ver y aceptar invitaciones para compartir recursos.

   Para obtener información sobre las políticas de IAM sugeridas para los administradores del lago de datos, consulte the section called “Permisos de administrador del lago de datos”.

2. Introduzca el siguiente comando para ver las invitaciones para compartir recursos pendientes.

   ```
   aws ram get-resource-share-invitations
   ```

   El resultado debería ser similar al siguiente.

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

   Observe el estado de `PENDING`.

3. Copie el valor de la clave `resourceShareInvitationArn` en el portapapeles.

4. Pegue el valor en el siguiente comando, sustituyendo *<invitación-arn>*, e introduzca el comando.

   ```
   aws ram accept-resource-share-invitation --resource-share-invitation-
   arn <invitation-arn>
   ```

   El resultado debería ser similar al siguiente.

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

Observe el estado de `ACCEPTED`.

Visualización de tablas y bases de datos compartidas del catálogo de datos

Puede ver los recursos que se comparten con su cuenta utilizando la consola de Lake Formation o la CLI de AWS. También puede utilizar la consola AWS Resource Access Manager (AWS RAM) o la CLI para ver tanto los recursos compartidos con su cuenta como los recursos que ha compartido con otras cuentas AWS.

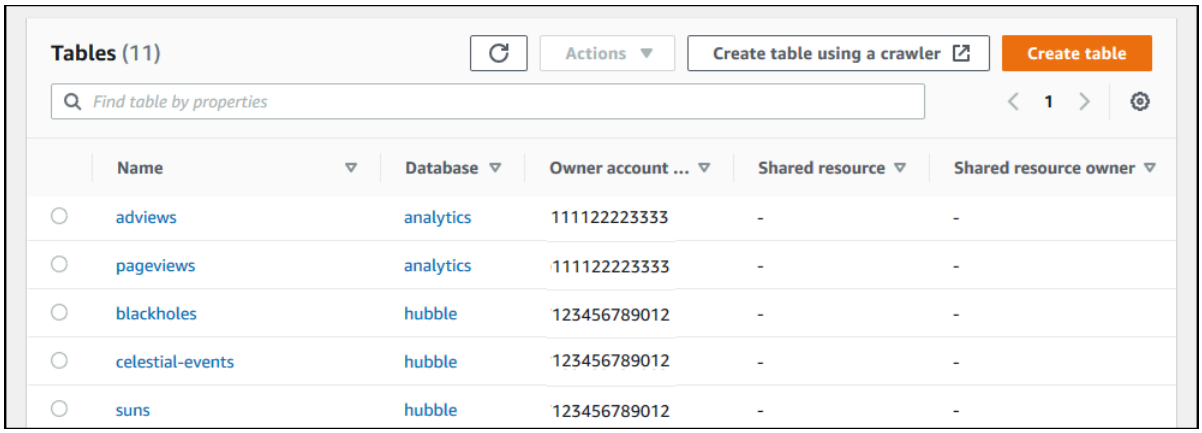
Para ver los recursos compartidos utilizando la consola de Lake Formation

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.

   Inicie sesión como administrador del lago de datos o como usuario al que se han concedido permisos sobre una tabla compartida.
2. Para ver los recursos compartidos con su cuenta AWS, lleve a cabo una de las siguientes acciones:

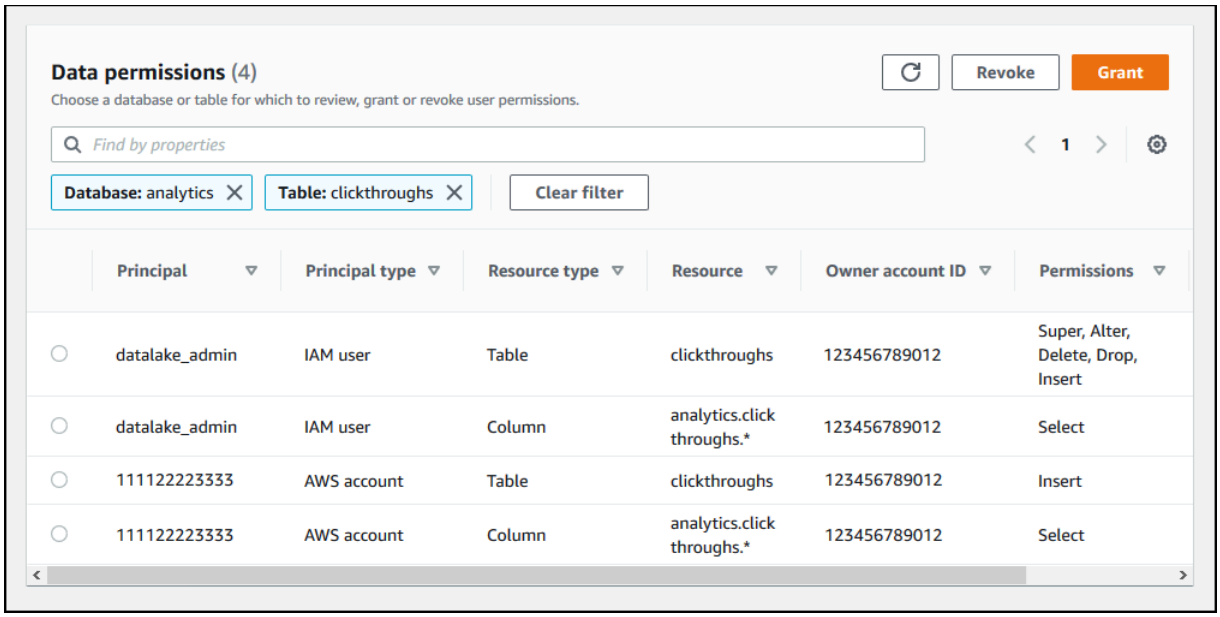
   - Para ver las tablas que se comparten con su cuenta, en el panel de navegación, seleccione Tablas.
   - Para ver las base de datos que se comparten con su cuenta, en el panel de navegación, seleccione Bases de datos.

La consola muestra una lista de las bases de datos o tablas tanto de su cuenta como compartidas con ella. En el caso de los recursos que se comparten con tu cuenta, la consola muestra el ID de la cuenta AWS del propietario en la columna ID de la cuenta del propietario (la tercera columna de la siguiente captura de pantalla).



3. Para ver los recursos compartidos con otras cuentas AWS u organizaciones, en el panel de navegación, seleccione Permisos de datos.

   Los recursos que ha compartido aparecen en la página Permisos de datos con el número de cuenta externa en la columna Entidad principal, como se ve en la imagen siguiente.



Para ver recursos compartidos con la consola de AWS RAM

1. Asegúrese de que dispone de los permisos de AWS Identity and Access Management (IAM) necesarios para ver compartir recursos usando AWS RAM.

   Como mínimo, debe tener el permiso `ram:ListResources`. Este permiso está incluido en la política administrada `AWSLakeFormationCrossAccountManager` de AWS.
2. Inicie sesión en la AWS Management Console y abra la consola de AWS RAM en https://console.aws.amazon.com/ram.
3. Haga una de las siguientes acciones:

   - Para ver los recursos que ha compartido, en el panel de navegación, en Compartidos por mí, seleccione Recursos compartidos.
   - Para ver los recursos que ha compartido, en el panel de navegación, en Compartidos por mí, seleccione Recursos compartidos.

Creación de enlaces de recursos

Los enlaces de recursos son objetos del catálogo de datos con enlaces a bases de datos y tablas de metadatos; en general, a bases de datos y tablas compartidas de otras cuentas de AWS. Ayudan a facilitar el acceso entre cuentas a los datos del lago de datos en todas las regiones de AWS.

> **Note**
>
> Lake Formation permite consultar las tablas del catálogo de datos en todas las regiones de AWS. Para acceder a las bases de datos y tablas del catálogo de datos desde cualquier región de AWS, debe crear enlaces de recursos en esas regiones que apunten a bases de datos y tablas compartidas en diferentes regiones.

Temas

- Cómo funcionan los enlaces de recursos en Lake Formation
- Crear un enlace de recursos a una tabla de catálogo de datos compartida
- Crear un enlace de recursos a una base de datos de catálogo de datos compartida
- Gestión de enlaces de recursos en las API de AWS Glue

Cómo funcionan los enlaces de recursos en Lake Formation

Un enlace de recursos es un objeto de catálogo de datos que actúa como enlace con una tabla o base de datos local o compartida. Después de crear un enlace de recursos a una tabla o una base de datos, puede emplear el nombre de dicho enlace donde vaya a utilizar el nombre de la base de datos. `glue:GetTables()` devuelve los enlaces de recursos de tablas y las tablas, ya sean suyas o compartidas, y aparecen como entradas en la página Tablas de la consola de Lake Formation. Los enlaces de recursos a bases de datos actúan de manera similar.

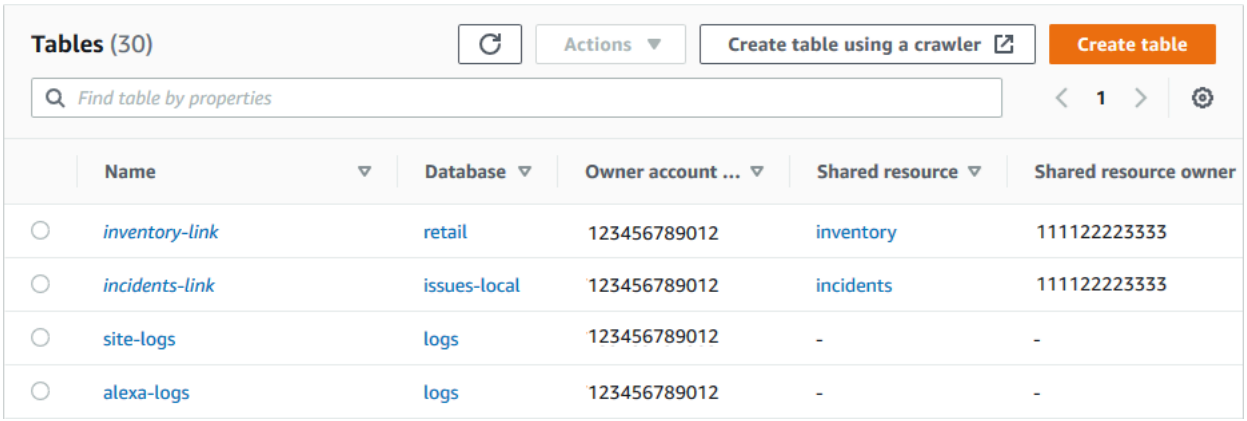
Al crear un enlace de recursos a una base de datos o tabla podrá:

- Asignar un nombre diferente a una base de datos o tabla de su catálogo de datos. Esto es especialmente útil si cuentas diferentes de AWS comparten bases de datos o tablas con el mismo nombre, o si varias bases de datos de su cuenta tienen tablas con el mismo nombre.
- Para acceder a las bases de datos y tablas del catálogo de datos desde cualquier región de AWS, cree enlaces de recursos en esas regiones que apunten a la base de datos y las tablas de otra región. Puede ejecutar consultas en cualquier región con estos enlaces de recursos mediante Athena, Amazon EMR y ejecutar trabajos de ETL Spark de AWS Glue, sin copiar los datos de origen ni los metadatos del Glue Data Catalog.
- Utilice servicios de AWS integrados, como Amazon Athena y Amazon Redshift Spectrum, para ejecutar consultas que accedan a tablas o bases de datos compartidas. Algunos servicios integrados no pueden acceder directamente a las bases de datos o tablas entre cuentas. Sin embargo, pueden acceder a los enlaces de recursos presentes en su cuenta a las bases de datos y tablas de otras cuentas.

> **Note**
>
> No es necesario crear un enlace de recursos para hacer referencia a una base de datos o tabla compartida en los scripts de extracción, transformación y carga (ETL) de AWS Glue. Sin embargo, para evitar la ambigüedad cuando varias cuentas de AWS comparten una base de datos o una tabla con el mismo nombre, puede crear y utilizar un enlace de recursos o especificar el identificador del catálogo al invocar las operaciones de ETL.

En el ejemplo siguiente, se muestra la página Tablas de la consola de Lake Formation, que muestra dos enlaces de recursos. Los nombres de los enlaces de recursos siempre se muestran en cursiva. Cada enlace de recursos se muestra junto con el nombre y el propietario del recurso

compartido vinculado. En este ejemplo, un administrador de un lago de datos de la cuenta de AWS 1111-2222-3333 compartió las tablas `inventory` y `incidents` con la cuenta 1234-5678-9012. A continuación, un usuario de esa cuenta creó enlaces de recursos a esas tablas compartidas.



Las siguientes son notas y restricciones sobre los enlaces de recursos:

- Los enlaces de recursos son necesarios para permitir que los servicios integrados, como Athena y Redshift Spectrum, consulten los datos subyacentes de las tablas compartidas. Las consultas en estos servicios integrados se crean a partir de los nombres de los enlaces de recursos.
- Si asumimos que la configuración para usar solo el control de acceso de IAM para las nuevas tablas de esta base de datos esté desactivada en la base de datos que la contiene, solo la entidad principal que creó un enlace de recursos podrá verla y acceder a ella. Para que otras entidades principales de su cuenta puedan acceder a un enlace de recursos, conceda el permiso de `DESCRIBE` correspondiente. Para que otros usuarios puedan eliminar un enlace de recursos, concédales el permiso de `DROP`. Los administradores del lago de datos pueden acceder a todos los enlaces de recursos de la cuenta. Para eliminar un enlace de recurso creado por otra entidad principal, el administrador del lago de datos primero debe concederse el permiso de `DROP` para acceder al enlace de recurso. Para obtener más información, consulte Referencia de permisos de Lake Formation.

  > **Important**
  >
  > Conceder permisos en un enlace de recursos no otorga permisos en la tabla o base de datos (vinculada) de destino. Debe conceder los permisos en el destino por separado.

- Para crear un enlace de recursos, necesita el permiso Lake Formation `CREATE_TABLE` o `CREATE_DATABASE`, así como el permiso (IAM) de `glue:CreateTable` o `glue:CreateDatabase` de AWS Identity and Access Management.

- Puede crear enlaces de recursos a los recursos del catálogo de datos local (de su propiedad), así como a los recursos compartidos con su cuenta de AWS.
- Al crear un enlace de recursos, no se comprueba si el recurso compartido de destino existe ni si tiene permisos entre cuentas sobre el recurso. Esto le permite crear el enlace de recurso y el recurso compartido en cualquier orden.
- Si elimina un enlace de recurso, el recurso compartido vinculado no se elimina. Si elimina un recurso compartido, los enlaces de recursos a ese recurso no se eliminan.
- Es posible crear cadenas de enlaces de recursos. Sin embargo, no tiene sentido hacerlo, ya que las API solo siguen el primer enlace de recursos.

Véase también:

- Concesión y revocación de permisos sobre los recursos del catálogo de datos

Crear un enlace de recursos a una tabla de catálogo de datos compartida

Puede crear un enlace de recursos a una tabla compartida en cualquier región de AWS mediante la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Para crear un enlace de recursos a una tabla compartida (consola)

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como una entidad principal que tenga el permiso de `CREATE_TABLE` de Lake Formation en la base de datos que contiene el enlace de recursos.
2. En el panel de navegación, elija Tables (Tablas) y, a continuación, seleccione Create table (Crear tabla).
3. En la página Crear tabla, seleccione el mosaico Enlace de recursos y, a continuación, proporcione la información siguiente:

   Nombre del enlace de recursos

   Introduzca un nombre que cumpla las mismas reglas que el nombre de una tabla. El nombre puede ser el mismo que el de la tabla compartida de destino.

Base de datos

La base de datos del catálogo local que contiene el enlace de recursos.

Región del propietario de la tabla

Si va a crear el enlace de recursos en una región diferente, seleccione la región de la tabla compartida de destino.

Tabla compartida

Seleccione una tabla compartida de la lista o introduzca un nombre de tabla local (propia) o compartida.

La lista contiene todas las tablas compartidas con su cuenta. Anote la base de datos y el ID de la cuenta del propietario que aparecen en cada tabla. Si no se muestra una tabla que sabe que se ha compartido con su cuenta, verifique lo siguiente:

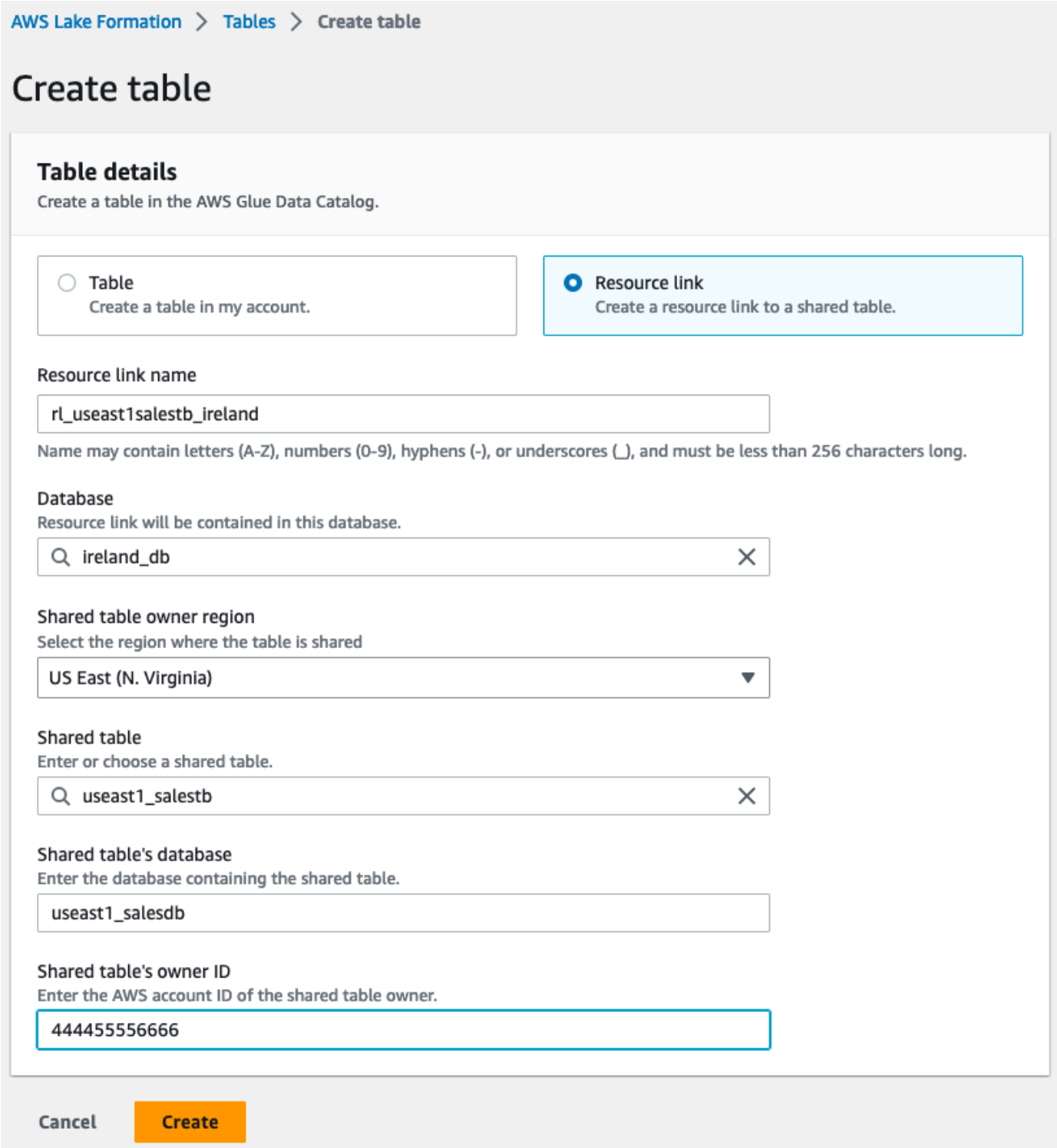
- Si no es administrador de un lago de datos, compruebe que el administrador del lago de datos le haya concedido los permisos de Lake Formation sobre la tabla.
- Si es administrador de un lago de datos y su cuenta no pertenece a la misma organización de AWS que la cuenta que los concede, asegúrese de haber aceptado la invitación para compartir recursos AWS Resource Access Manager (AWS RAM) de la tabla. Para obtener más información, consulte Aceptar una invitación para compartir recursos de AWS RAM.

Base de datos de la tabla compartida

Si ha seleccionado una tabla compartida de la lista, este campo se rellena con la base de datos de la tabla compartida en la cuenta externa. De lo contrario, introduzca una base de datos local (para un enlace de recursos a una tabla local) o la base de datos de la tabla compartida en la cuenta externa.

Propietario de tabla compartida

Si ha seleccionado una tabla compartida de la lista, este campo se rellena con el ID de la cuenta del propietario de la tabla compartida. De lo contrario, introduzca su ID de cuenta de AWS (para un enlace de recursos a una tabla local) o el ID de la cuenta de AWS que compartió la tabla.



4. Elija Crear para crear el enlace de recursos.

   A continuación, puede ver el nombre del enlace de recursos en la columna Nombre de la página Tablas.

5. (Opcional) Otorgue el permiso de DESCRIBE de Lake Formation en el enlace de recursos a las entidades principales que deben poder ver el enlace y acceder a la tabla de destino.

Para crear un enlace de recursos a una tabla compartida en la misma región (AWS CLI)

1. Escriba un comando similar al siguiente.

```
aws glue create-table --database-name myissues --table-input
 '{"Name":"my_customers","TargetTable":
{"CatalogId":"111122223333","DatabaseName":"issues","Name":"customers"}}'
```

   Este comando crea un enlace de recursos con el nombre `my_customers` a la tabla compartida `customers`, que está en la base de datos `issues` de la cuenta de AWS 1111-2222-3333. El enlace de recursos se almacena en la base de datos local `myissues`.

2. (Opcional) Otorgue el permiso de DESCRIBE de Lake Formation en el enlace de recursos a las entidades principales que deben poder ver el enlace y acceder a la tabla de destino.

Para crear un enlace de recursos a una tabla compartida en una región diferente (AWS CLI)

1. Escriba un comando similar al siguiente.

```
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"
        }
    }
}‘
```

   Este comando crea un enlace de recurso denominado `rl_useast1salestb_ireland` en la región Europa (Irlanda) a la tabla compartida `useast1_salestb`, que está en la base de datos `useast1_salesdb` de la cuenta AWS 444455556666 de la región Este de EE. UU. (Norte de Virginia). El enlace de recursos se almacena en la base de datos local `ireland_db`.

2. Conceda permiso de `DESCRIBE` de Lake Formation a las entidades principales que deben poder ver el enlace y acceder al destino del enlace a través del enlace.

> Véase también:
>
> - Cómo funcionan los enlaces de recursos en Lake Formation
> - `DESCRIBE`

Crear un enlace de recursos a una base de datos de catálogo de datos compartida

Puede crear un enlace de recursos a una base de datos compartida mediante la consola de AWS Lake Formation, la API o la AWS Command Line Interface (AWS CLI).

Para crear un enlace de recursos a una base de datos compartida (consola)

1. Abra la consola de AWS Lake Formation en https://console.aws.amazon.com/lakeformation/. Inicie sesión como administrador de un lago de datos o como creador de bases de datos.

   El creador de una base de datos es una entidad principal a la que se le ha otorgado el permiso de `CREATE_DATABASE` de Lake Formation.
2. En el panel de navegación, seleccione Bases de datos y, a continuación, Crear base de datos.
3. En la página Crear base de datos, seleccione el mosaico Enlace de recursos y, a continuación, proporcione la información siguiente:

   Nombre del enlace de recursos

   Introduzca un nombre que cumpla las mismas reglas que el nombre de una base de datos. El nombre puede ser el mismo que el de la base de datos compartida de destino.

   Región propietaria de la base de datos compartida

   Si va a crear el enlace de recursos en una región diferente, seleccione la región de la base de datos compartida de destino.

Base de datos compartida

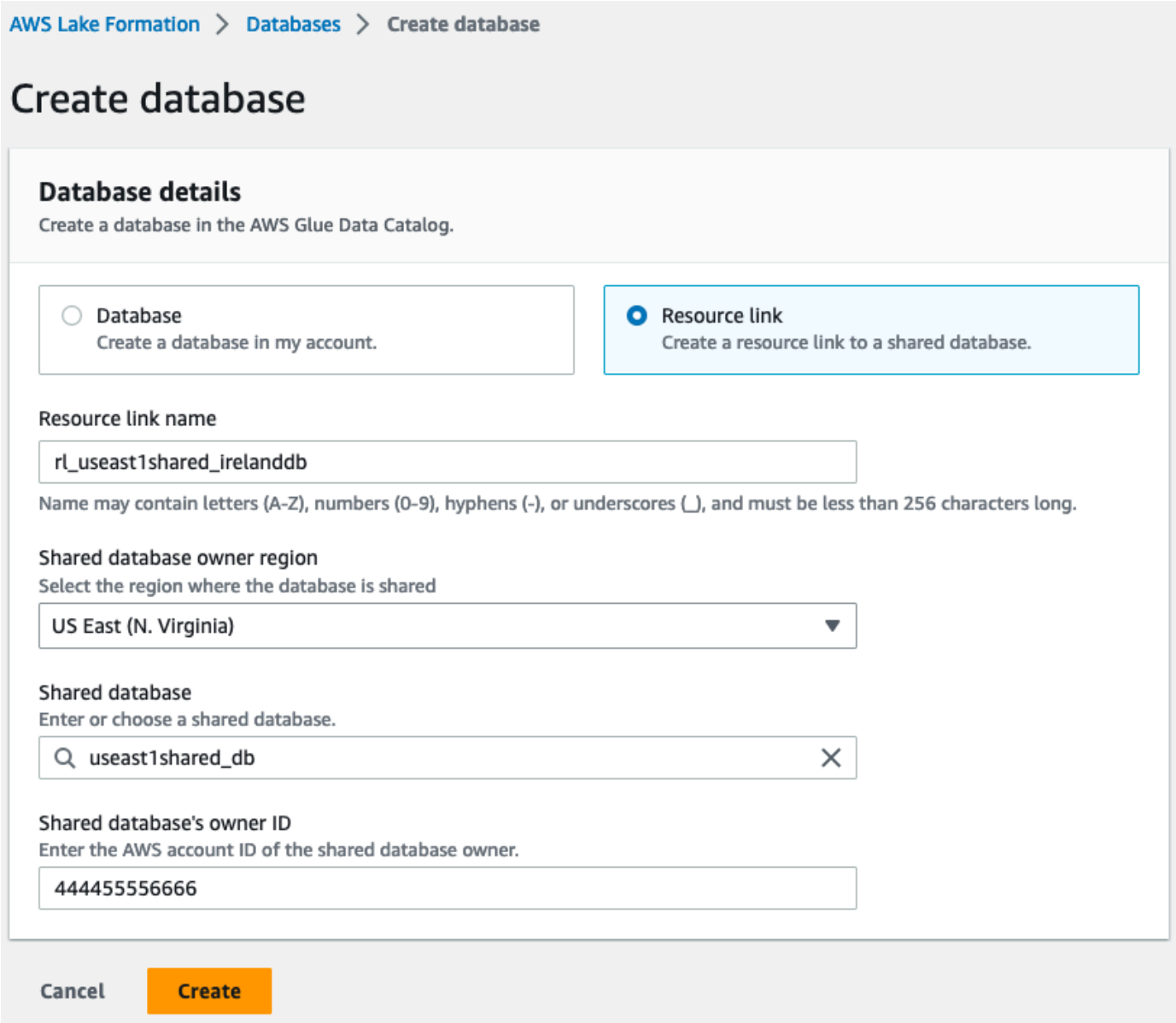
Elija una base de datos de la lista o introduzca un nombre de base de datos local (propia) o compartida.

La lista contiene todas las bases de datos compartidas con su cuenta. Observe el ID de la cuenta del propietario que aparece con cada base de datos. Si no ve una base de datos que sabe que se ha compartido con su cuenta, verifique lo siguiente:

- Si no es administrador de un lago de datos, compruebe que el administrador del lago de datos le haya concedido los permisos de Lake Formation sobre la base de datos.
- Si es administrador de un lago de datos y su cuenta no pertenece a la misma organización AWS que la cuenta concedente, asegúrese de haber aceptado la invitación para compartir recursos AWS Resource Access Manager (AWS RAM) de la base de datos. Para obtener más información, consulte Aceptar una invitación para compartir recursos de AWS RAM.

Propietario de la base de datos compartida

Si ha seleccionado una base de datos compartida de la lista, este campo se rellena con el ID de la cuenta del propietario de la base de datos compartida. De lo contrario, introduzca el ID de cuenta de AWS (para un enlace de recursos a una base de datos local) o el ID de la cuenta AWS que compartió la base de datos.



4. Elija Crear para crear el enlace de recursos.

   A continuación, puede ver el nombre del enlace de recursos en la columna Nombre de la página Bases de datos.

5. (Opcional) Conceda el permiso de `DESCRIBE` de Lake Formation en el enlace de recursos a las entidades principales de la región de Europa (Irlanda) que deben poder ver el enlace y acceder a la base de datos de destino.

Para crear un enlace de recursos a una base de datos compartida en la misma región (AWS CLI)

1. Escriba un comando similar al siguiente.

```
aws glue create-database --database-input '{"Name":"myissues","TargetDatabase":
{"CatalogId":"111122223333","DatabaseName":"issues"}}'
```

   Este comando crea un enlace de recursos con el nombre `myissues` a la base de datos compartida `issues`, que está en la cuenta de AWS 1111-2222-3333.

2. (Opcional) Otorgue el permiso de `DESCRIBE` de Lake Formation a las entidades principales del enlace de recursos para que puedan ver el enlace y acceder a la base de datos de destino.

Para crear un enlace de recursos a una base de datos compartida en una región diferente (AWS CLI)

1. Escriba un comando similar al siguiente.

```
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"
       }
    }
}'
```

   Este comando crea un enlace de recursos con el nombre `rl_useast1shared_irelanddb` de la cuenta de AWS 111122223333 de la región Europa (Irlanda) a la base de datos compartida `useast1shared_db`, que está en la cuenta AWS 444455556666 de la región Este de EE. UU. (Norte de Virginia)

2. Conceda el permiso de `DESCRIBE` de Lake Formation a las entidades principales de la región de Europa (Irlanda) que deben poder ver el enlace y acceder al destino del enlace a través del enlace.

Véase también:

- Cómo funcionan los enlaces de recursos en Lake Formation
- `DESCRIBE`

Gestión de enlaces de recursos en las API de AWS Glue

En las tablas siguientes se explica cómo las API del catálogo de datos de AWS Glue gestionan los enlaces a los recursos de bases de datos y tablas. Para todas las operaciones de la API de `Get*`, solo se devuelven las bases de datos y las tablas sobre las que la persona que llama tiene permiso. Además, al acceder a una base de datos o tabla de destino a través de un enlace de recursos, debe tener los permisos de Lake Formation y AWS Identity and Access Management (IAM) en el enlace de destino y en el de recursos. El permiso de Lake Formation que se requiere en los enlaces de recursos es `DESCRIBE`. Para obtener más información, consulte `DESCRIBE`.

Operaciones de la API de base de datos

Operación de la API	Gestión de enlaces de recursos
`CreateDatabase`	Si la base de datos es un enlace de recursos, crea el enlace de recursos a la base de datos de destino designada.
`UpdateDatabase`	Si la base de datos es un enlace de recursos, sigue el enlace y actualiza la base de datos de destino. Si el enlace de recursos se debe modificar para enlazarlo a una base de datos diferente, debe eliminarlo y crear uno nuevo.
`DeleteDatabase`	Elimina el enlace de recursos. No elimina la base de datos vinculada (de destino).
`GetDatabase`	Si la persona que llama tiene permisos sobre el destino, sigue el enlace para obtener las propiedades del destino. De lo contrario, devuelve las propiedades del enlace.
`GetDatabases`	Devuelve una lista de bases de datos, incluidos los enlaces de recursos. Para cada enlace de recursos del conjunto de resultado s, la operación sigue el enlace para obtener las propiedades del

Operación de la API	Gestión de enlaces de recursos
	destino del enlace. Debe especificar `ResourceShareType = ALL` para ver las bases de datos compartidas con su cuenta.

Operaciones de la API de tabla

Operación de la API	Gestión de enlaces de recursos
`CreateTable`	Si la base de datos es un enlace de recursos, sigue el enlace de la base de datos y crea una tabla en la base de datos de destino. Si la tabla es un enlace de recursos, la operación crea el enlace de recursos en la base de datos designada. No es compatible crear un enlace de recursos de tabla a través de un enlace de recursos de base de datos.
`UpdateTable`	Si la tabla o la base de datos designada es un enlace de recursos, actualiza la tabla de destino. Si tanto la tabla como la base de datos son enlaces de recursos, la operación falla.
`DeleteTable`	Si la base de datos designada es un enlace de recursos, sigue el enlace y borra la tabla o el enlace de recursos de tabla en la base de datos de destino. Si la tabla es un enlace de recursos, la operación borra el enlace de recursos de la tabla en la base de datos designada. La eliminación de un enlace de recursos de tabla no elimina la tabla de destino.
`BatchDeleteTable`	Igual que `DeleteTable` .
`GetTable`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y devuelve la tabla o el enlace de recursos de la tabla de la base de datos de destino. En caso contrario, si la tabla es un enlace de recursos, la operación sigue el enlace y devuelve las propiedades de la tabla de destino.
`GetTables`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y devuelve las tablas y los enlaces de recursos de tablas de la base de datos de destino. Si la base

Operación de la API	Gestión de enlaces de recursos
	de datos de destino es una base de datos compartida desde otra cuenta AWS, la operación devuelve solo las tablas compartidas de esa base de datos. No sigue los enlaces de recursos de la tabla en la base de datos de destino. De lo contrario, si la base de datos designada es una base de datos local (propia), la operación devuelve todas las tablas de la base de datos local, y sigue cada enlace de recurso de tabla para devolver las propiedades de la tabla de destino.
`SearchTables`	Devuelve enlaces a recursos de tablas y tablas. No sigue los enlaces para devolver las propiedades de la tabla de destino. Debe especificar `ResourceShareType = ALL` para ver las bases de datos compartidas con su cuenta.
`GetTableVersion`	Igual que `GetTable`.
`GetTableVersions`	Igual que `GetTable`.
`DeleteTableVersion`	Igual que `DeleteTable` .
`BatchDeleteTableVersion`	Igual que `DeleteTable` .

Operaciones de la API en una partición

Operación de la API	Gestión de enlaces de recursos
`CreatePartition`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y crea una partición en la tabla designada de la base de datos de destino. Si la tabla es un enlace de recursos, la operación sigue el enlace de recursos y crea la partición en la tabla de destino. No es compatible crear una partición mediante un enlace de recursos de tabla y un enlace de recursos de base de datos.

Operación de la API	Gestión de enlaces de recursos
`BatchCreatePartition`	Igual que `CreatePartition` .
`UpdatePartition`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y actualiza la partición en la tabla designada de la base de datos de destino. Si la tabla es un enlace de recursos, la operación sigue el enlace de recursos y actualiza la partición en la tabla de destino. No es compatible actualizar una partición mediante un enlace de recursos de tabla y un enlace de recursos de base de datos.
`DeletePartition`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y borra la partición en la tabla designada en la base de datos de destino. Si la tabla es un enlace de recursos, la operación sigue el enlace de recursos y borra la partición en la tabla de destino. No es compatible eliminar una partición mediante un enlace de recursos de tabla y un enlace de recursos de base de datos.
`BatchDeletePartition`	Igual que `DeletePartition` .
`GetPartition`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y devuelve la información de partición de la tabla designada. En caso contrario, si la tabla es un enlace de recursos, la operación sigue el enlace y devuelve información sobre la partición. Si la tabla y la base de datos son enlaces de recursos, devuelve un conjunto de resultados vacío.
`GetPartitions`	Si la base de datos designada es un enlace de recursos, sigue el enlace de la base de datos y devuelve la información de partición para todas las particiones de la tabla designada. En caso contrario , si la tabla es un enlace de recursos, la operación sigue el enlace y devuelve información sobre la partición. Si la tabla y la base de datos son enlaces de recursos, devuelve un conjunto de resultados vacío.

Operación de la API	Gestión de enlaces de recursos
`BatchGetPartition`	Igual que `GetPartition` .

Operaciones de la API con funciones definidas por el usuario

Operación de la API	Gestión de enlaces de recursos
(Todas las operaciones de la API)	Si la base de datos es un enlace de recursos, sigue el enlace de recursos y efectúa la operación en la base de datos de destino.

Véase también:

- Cómo funcionan los enlaces de recursos en Lake Formation

Acceso a las tablas entre regiones

Lake Formation permite consultar las tablas del catálogo de datos en todas las regiones de AWS. Puede acceder a los datos de una región desde otras regiones mediante Amazon Athena, Amazon EMR y ETL AWS Glue creando enlaces de recursos en otras regiones que apunten a las bases de datos y tablas de origen. Con el acceso a las tablas entre regiones, puede acceder a los datos de todas las regiones sin copiar los datos subyacentes o los metadatos en el catálogo de datos.

Por ejemplo, puede compartir una base de datos o una tabla de una cuenta de productor con una cuenta de consumidor en la Región A. Tras aceptar la invitación a compartir recursos en la Región A, el administrador del lago de datos de la cuenta de consumidor puede crear enlaces de recursos al recurso compartido en la Región B. El administrador de la cuenta de consumidor puede conceder permisos sobre el recurso compartido a las entidades principales de IAM de esa cuenta en la Región A y conceder permisos de enlace de recursos en la Región B. Con el enlace de recursos, las entidades principales de la cuenta de consumidor pueden consultar los datos compartidos de la Región B.

También puede alojar el origen de datos de Amazon S3 de la región A en una cuenta de productor y registrar la ubicación de los datos en una cuenta central de la región B. Puede crear recursos del catálogo de datos en la cuenta central, configurar los permisos de Lake Formation y compartir

datos con los consumidores de su cuenta o con cuentas externas de la región B. La característica entre regiones permite a los usuarios acceder a estas tablas del catálogo de datos desde la región C mediante enlaces de recursos.

Con esta característica, puede consultar bases de datos federadas en Apache Hive Metastores entre regiones y también unir tablas de la región local con tablas de otra región al ejecutar consultas.

Lake Formation admite las siguientes características con acceso a tablas entre regiones:

- Control de acceso basado en etiquetas LF
- Permisos de control de acceso específicos
- Escribir operaciones en la base de datos o tabla compartida con los permisos adecuados
- Intercambio de datos entre cuentas en el nivel de cuentas y directamente con el nivel de entidades principales de IAM

Los usuarios no administrativos con permisos `Create_Database` y `Create_Table` pueden crear enlaces de recursos entre regiones.

> **Note**
>
> Puede crear enlaces de recursos entre regiones en cualquier región y acceder a los datos sin aplicar los permisos de Lake Formation. Para los datos de origen en Amazon S3 que no están registrados en Lake Formation, el acceso se determina mediante las políticas de permisos de IAM para Amazon S3 y acciones de AWS Glue.

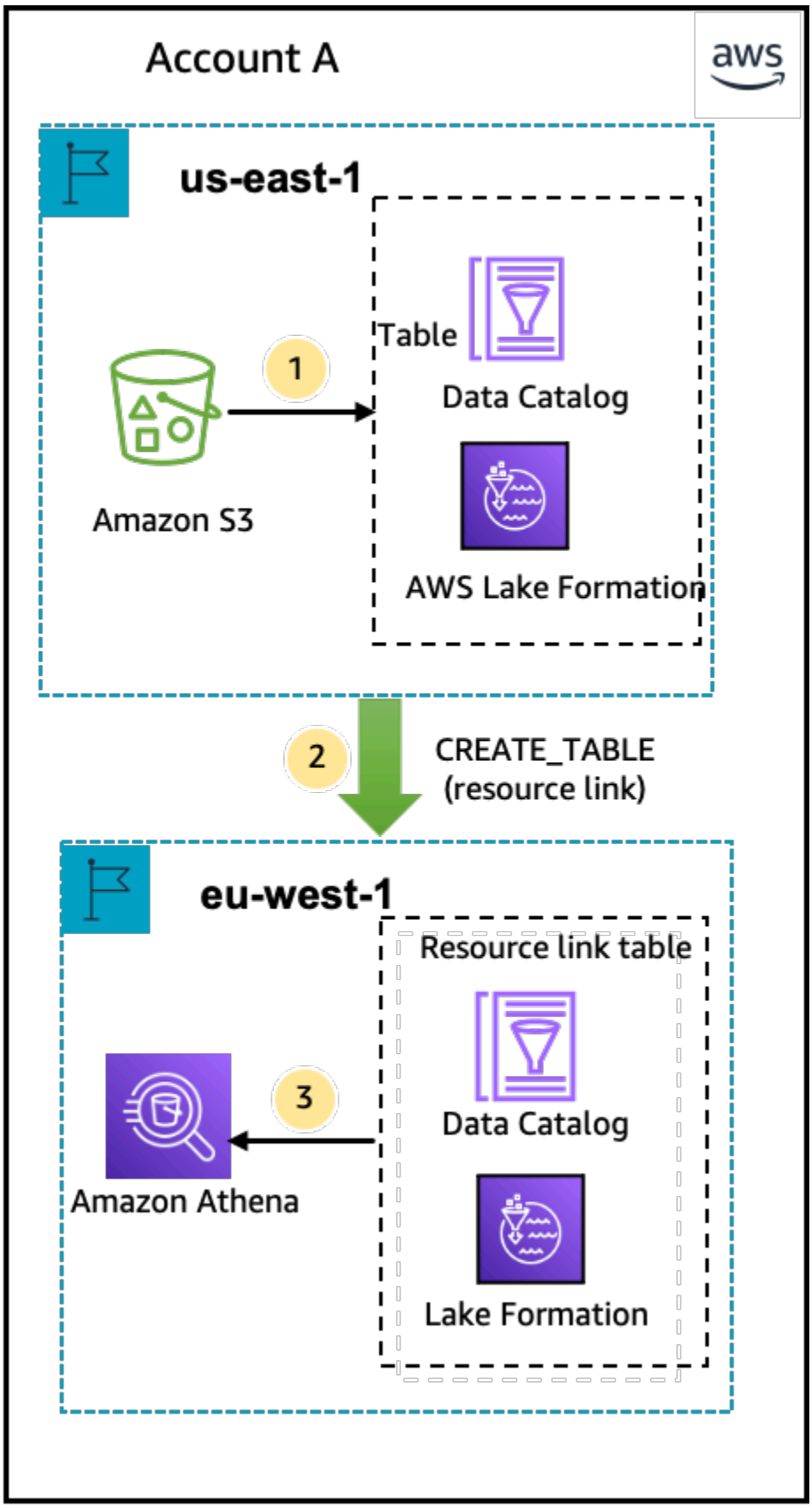
Para conocer las limitaciones, consulte Limitaciones de acceso a datos entre regiones.

Flujos de trabajo

Los siguientes diagramas muestran los flujos de trabajo para acceder a los datos de todas las regiones de AWS desde la misma cuenta AWS y desde una cuenta externa.

Flujo de trabajo para acceder a las tablas compartidas en la misma cuenta AWS

En el siguiente diagrama, los datos se comparten con un usuario de la misma cuenta AWS de la región Este de EE. UU. (Norte de Virginia) y el usuario consulta los datos compartidos desde la región Europa (Irlanda).


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

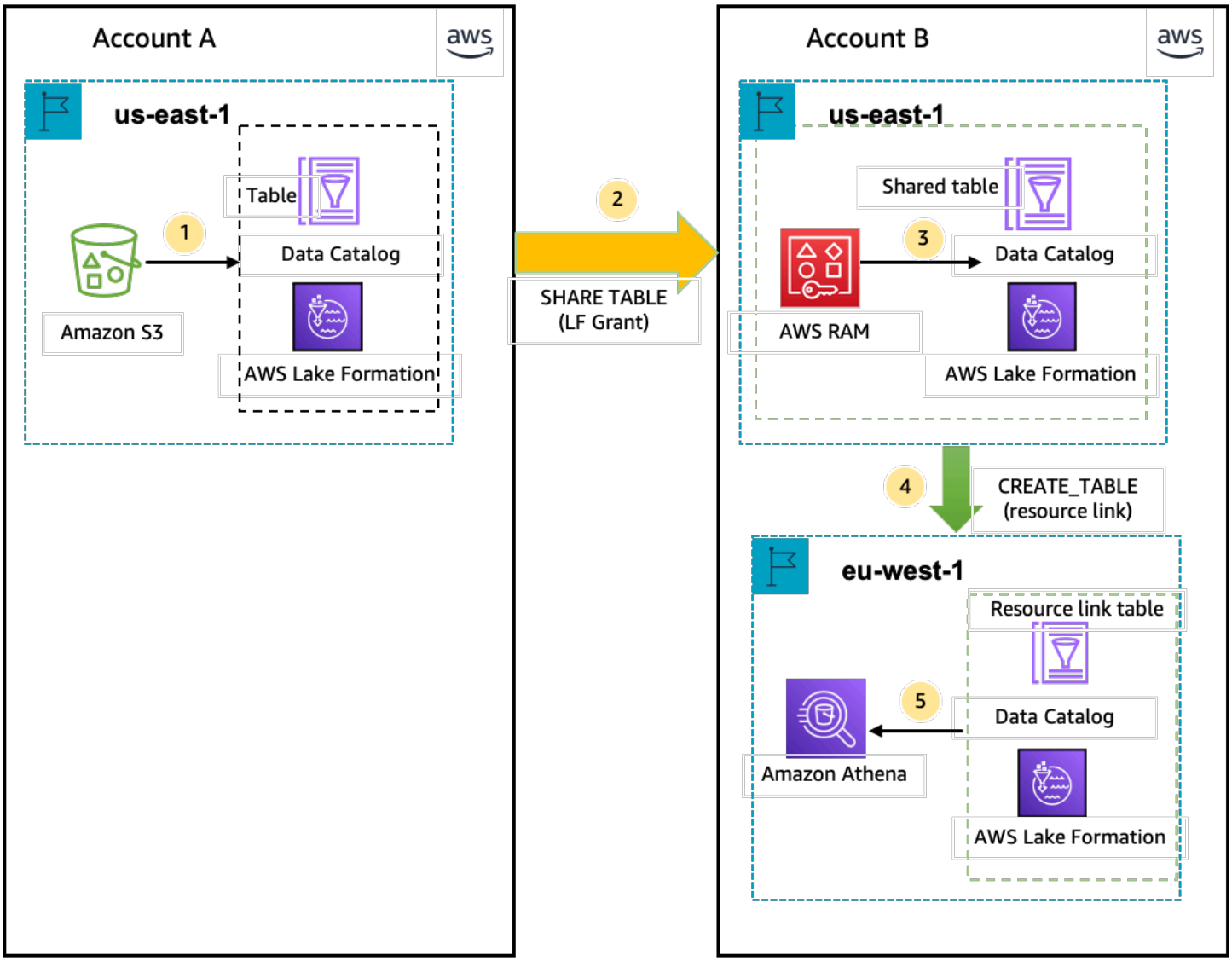
El administrador del lago de datos efectúa las actividades siguientes (pasos 1 y 2):

1. Un administrador de lago de datos configura una cuenta de AWS en las bases de datos y tablas del catálogo de datos y registra una ubicación de datos de Amazon S3 en Lake Formation en la región Este de EE. UU. (Norte de Virginia).

   Otorga permiso de `Select` sobre un recurso del catálogo de datos (tabla de productos en el diagrama) a una entidad principal (usuario) de la misma cuenta.
2. Crea un enlace de recursos en la región Europa (Irlanda) que apunta a la tabla de origen en la región Este de EE. UU. (Norte de Virginia). Otorga permiso de `DESCRIBE` a la entidad principal para utilizar el enlace de recursos de la región de Europa (Irlanda).
3. El usuario consulta la tabla desde la región Europa (Irlanda) mediante Athena.

Flujo de trabajo para acceder a las tablas compartidas con una cuenta externa de AWS

En el siguiente diagrama, la cuenta del productor (cuenta A) aloja el bucket de Amazon S3, registra la ubicación de los datos y comparte una tabla del catálogo de datos con una cuenta de consumidor (cuenta B) de la región Este de EE. UU. (Norte de Virginia) y un usuario de la cuenta de consumidor (cuenta B) consulta la tabla de la región de Europa (Irlanda).



1. Un administrador de lago de datos configura una cuenta de AWS (de productor) con los recursos del catálogo de datos y una ubicación de datos de Amazon S3 registrada en Lake Formation en la región Este de EE. UU. (Norte de Virginia).
2. El administrador del lago de datos de la cuenta del productor comparte una tabla del catálogo de datos con una cuenta de consumidor.
3. El administrador del lago de datos de la cuenta de consumidor acepta la invitación a compartir datos en la región de Este de EE. UU. (Norte de Virginia) y concede el permiso de `Select` para utilizar la tabla compartida con una entidad principal de la misma región.
4. El administrador del lago de datos de la cuenta de consumidor crea un enlace de recursos en la región de Europa (Irlanda) que apunta a la tabla compartida de destino en la región Este de EE.

UU. (Norte de Virginia) y concede al usuario permiso de `DESCRIBE` para acceder al enlace de recursos desde la región de Europa (Irlanda).

5. El usuario consulta la tabla desde la región Europa (Irlanda) mediante Athena.

Configuración del acceso a las tablas entre regiones

Para acceder a los datos de una región diferente, primero debe configurar las bases de datos y las tablas del catálogo de datos en la región en la que registra su ubicación de datos de Amazon S3. Puede compartir las bases de datos y tablas del catálogo de datos con las entidades principales de su cuenta o de otra cuenta. A continuación, debe crear administradores de lagos de datos que puedan generar enlaces de recursos que apunten a la ubicación de datos compartidos de destino en las regiones en las que los usuarios consultan los datos.

Para consultar los datos compartidos en la misma cuenta desde una región diferente

En esta sección, la región de la tabla compartida de destino se denomina Región A y los usuarios hacen consultas desde la Región B.

1. Configuración de la cuenta en la Región A (donde se crean y comparten los datos)

   El administrador de un lago de datos debe completar las acciones siguientes:

   a. Registrar una ubicación de datos de Amazon S3.

      Para obtener más información, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

   b. Crear bases de datos y tablas en la cuenta. Esto también lo puede hacer un usuario no administrativo que tenga permisos para crear bases de datos y tablas.

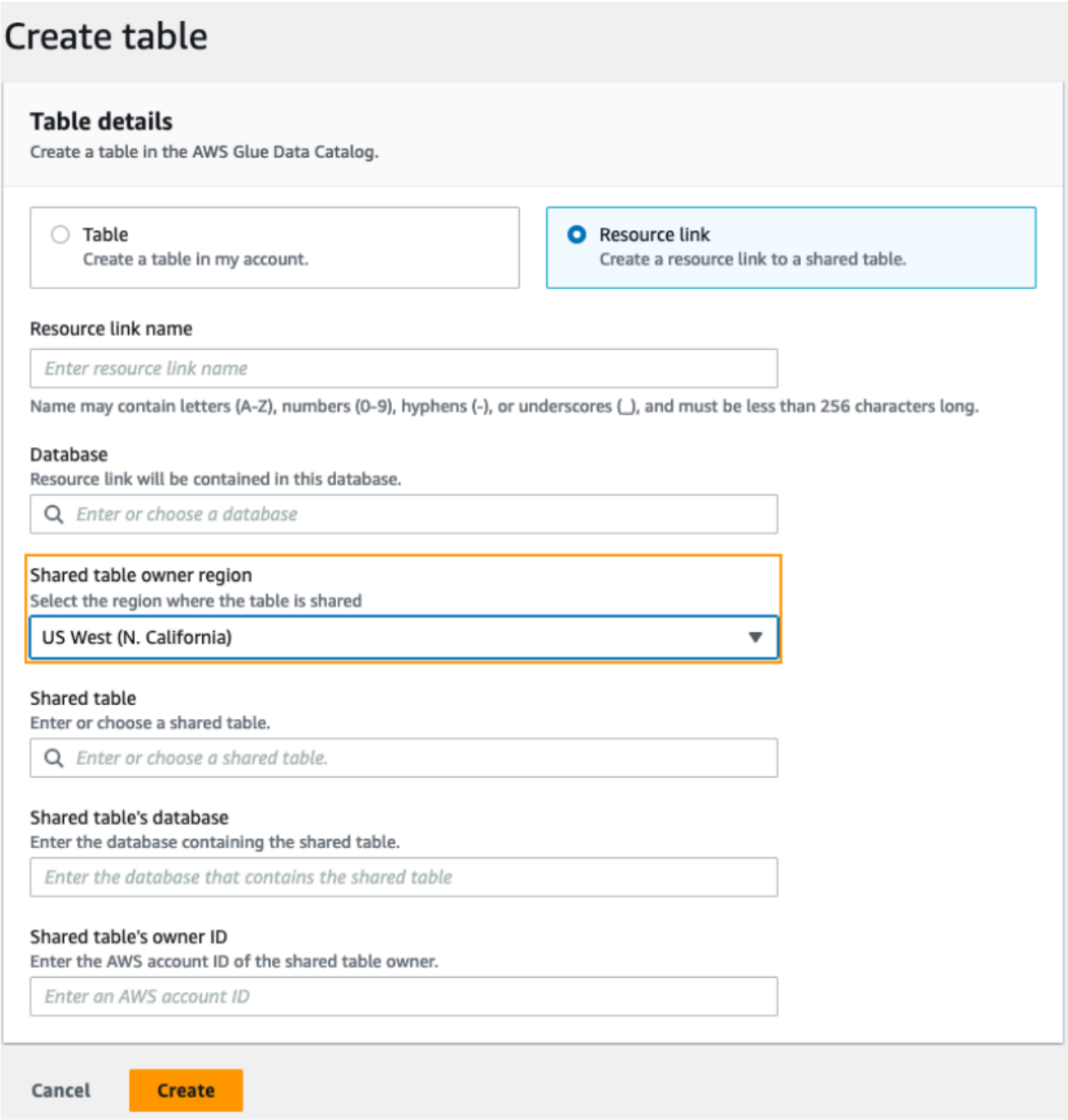
   c. Conceda permisos de datos sobre una tabla a las entidades principales con `Grantable permissions`.

      Para obtener más información, consulte, Concesión y revocación de permisos sobre los recursos del catálogo de datos.

2. Configuración de la cuenta en la Región B (donde accede a los datos)

   El administrador del lago de datos debe completar las acciones siguientes:

a. Crear un enlace de recursos en la región B que apunte a la tabla compartida de destino en la región A. Especifique la Región del propietario de la tabla compartida en la pantalla Crear tabla.



Para obtener instrucciones sobre cómo crear enlaces de recursos a bases de datos y tablas, consulte Creación de enlaces de recursos.

b. Conceda permiso de `Describe` a las entidades principales de IAM en el enlace de recursos de la Región B.

   Para obtener más información sobre conceder permisos en enlaces de recursos, consulte Conceder permisos de enlace de recursos.

   Las entidades principales de IAM de la Región B pueden consultar la tabla de destino a través del enlace con Athena.

Para acceder a los datos de varias cuentas de una región diferente

1. Configuración de la cuenta del productor/concedente

   El administrador del lago de datos debe completar las acciones siguientes:

   a. Configurar la cuenta del productor/concedente de la Región A.

   b. Registrar una ubicación de datos de Amazon S3 en la Región A.

   c. Crear bases de datos y tablas. Esto lo puede hacer un usuario no administrativo que tenga permisos para crear bases de datos y tablas.

   d. Conceda permisos de datos a la cuenta del consumidor/beneficiario de una tabla de la Región A con `Grantable permissions`.

      Para obtener más información, consulte Compartir tablas y bases de datos del Catálogo de datos entre Cuentas de AWS o entidades principales de IAM de cuentas externas.

2. Configuración de la cuenta de consumidor/beneficiario

   El administrador del lago de datos debe completar las acciones siguientes:

   a. Aceptar la invitación para compartir recursos de AWS RAM en la Región A.

   b. Crear un enlace de recursos en la región B que apunte a la tabla compartida. La Región B es donde los usuarios consultarán la tabla.

   c. Conceda permisos de datos en la tabla compartida a las entidades principales de IAM en la Región A.

> **Note**
>
> Debe conceder permisos a la tabla compartida en la misma región en la que se compartió la tabla.

d. Conceda permisos a las entidades principales en el enlace de recursos de la Región B.

   A continuación, las entidades principales de la cuenta de consumidores de la Región B consultan la tabla compartida desde la Región B utilizando Athena.

Uso compartido de datos en AWS Lake Formation

Puede utilizar la función de intercambio de AWS Lake Formation datos para conceder y gestionar permisos sobre los datos almacenados en ubicaciones distintas de Amazon S3 y los metadatos almacenados en ubicaciones distintas de laAWS Glue Data Catalog. Con la capacidad de compartir datos, puede configurar y gestionar los permisos de los conjuntos de datos en Amazon Redshift sin migrar los datos a Amazon S3. También puede utilizar la función de federación de catálogos de datos para conectarse a metaalmacenes externos.

Posteriormente, puede usar Lake Formation para administrar los datos y los permisos de acceso en un Catálogo de datos central mediante la definición de políticas de control de acceso detalladas. Los administradores del lago de datos pueden conceder permisos a otras entidades principales de IAM dentro de la cuenta o entre cuentas en los recursos del Catálogo de datos. Las entidades principales de IAM pueden consultar los datos compartidos mediante Amazon Redshift Spectrum y Amazon Athena.

Lake Formation proporciona los siguientes métodos para compartir datos y gestionar permisos sobre conjuntos de datos externos y metaalmacenes externos:

- Integración de Lake Formation con el uso compartido de datos de Amazon Redshift: utilice Lake Formation para administrar de forma centralizada los permisos de acceso a nivel de base de datos, tabla, columna y fila de los recursos compartidos de datos de Amazon Redshift y restringir el acceso de los usuarios a los objetos dentro de un recurso compartido de datos.
- Conexión AWS Glue Data Catalog a metaalmacenes externos: conecte los metaalmacenes externos AWS Glue Data Catalog para gestionar los permisos de acceso a los conjuntos de datos de Amazon S3 mediante Lake Formation. No es necesaria la migración de los metadatos a AWS Glue Data Catalog.
- Integración de Lake Formation con el intercambio de datos de AWS. Lake Formation admite la concesión de licencias de acceso a sus datos mediante AWS Data Exchange. Si está interesado en licenciar sus datos de Lake Formation, consulte Qué es AWS Data Exchange en la Guía del usuario de AWS Data Exchange.

Temas

Administración de permisos para los datos de un recurso compartido de datos de Amazon Redshift

ConAWS Lake Formation, puede gestionar los datos de forma segura en un recurso compartido de Amazon Redshift. Amazon Redshift es un servicio administrado de almacenamiento de datos a escala de petabytes en la nube AWS. Al utilizar la capacidad de compartir datos, Amazon Redshift le ayuda a compartir datos entre Cuentas de AWS. Para obtener más información sobre el uso compartido de datos, consulte Información general del uso compartido de datos en Amazon Redshift.

En Amazon Redshift, el administrador del clúster del productor crea un recurso compartido de datos y lo comparte con el administrador del lago de datos. Para obtener step-by-step instrucciones sobre cómo crear un administrador de lagos de datos, consulte. Crear un administrador de lago de datos

Una vez que el usuario (administrador del lago de datos) acepte el recurso compartido de datos, debe crear una base de datos AWS Glue Data Catalog para el recurso compartido de datos específico. Esto es para que pueda controlar el acceso a mediante los permisos de Lake Formation. Lake Formation mapea cada recurso compartido de datos a la base de datos del Catálogo de datos correspondiente. Aparecen como bases de datos federadas en Data Catalog.

Una base de datos se denomina base de datos federada cuando apunta a una entidad fuera del catálogo de datos. Las tablas y vistas del recurso compartido de datos de Amazon Redshift se muestran como tablas individuales en el Catálogo de datos. Puede compartir la base de datos federada con entidades principales de IAM y usuarios de SAML seleccionados en la misma cuenta o en otra cuenta con Lake Formation. También puede incluir expresiones de filtro de filas y columnas para restringir el acceso a determinados datos. Para obtener más información, consulte Información general del filtrado de datos.

Para proporcionar a los usuarios acceso a un recurso compartido de datos de Amazon Redshift, debe hacer lo siguiente:

1. Actualice la Configuración del Catálogo de datos para habilitar los permisos de Lake Formation.
2. Acepte la invitación para compartir datos del administrador del clúster de productores de Amazon Redshift y registre el recurso compartido de datos en Lake Formation.

   Después de completar este paso, puede administrar el intercambio de datos en el catálogo de datos de Lake Formation.
3. Cree una base de datos federada y defina los permisos en esa base de datos.

4. Otorgue permisos a los usuarios en bases de datos y tablas. Puede compartir toda la base de datos o un subconjunto de tablas con los usuarios de la misma cuenta o de otra cuenta.

Para conocer las limitaciones, consulte Limitaciones de uso compartido de datos de Amazon Redshift.

Temas

- Requisitos previos para configurar permisos en recursos compartidos de datos de Amazon Redshift
- Configurar permisos en recursos compartidos de datos de Amazon Redshift
- Consultar las bases de datos federadas

Requisitos previos para configurar permisos en recursos compartidos de datos de Amazon Redshift

Actualice la configuración predeterminada del Catálogo de datos

Para habilitar los permisos de Lake Formation para los recursos del Catálogo de datos, le recomendamos que deshabilite la configuración predeterminada del Catálogo de datos en Lake Formation. Para obtener más información, consulte Cambie el modelo de permisos predeterminado o utilice el modo de acceso híbrido.

Actualizar permisos

Además de los permisos de administrador de data lake (`AWSLakeFormationDataAdmin`), también se requieren los siguientes permisos para aceptar un datashare de Amazon Redshift en Lake Formation:

- `glue:PassConnection on aws:redshift`
- `redshift:AssociateDataShareConsumer`
- `redshift:DescribeDataSharesForConsumer`
- `redshift:DescribeDataShares`

El usuario de IAM administrador del lago de datos tiene implícitos los siguientes permisos.

- data_location_access

- create_database
- lakefomation:registerResource

Configurar permisos en recursos compartidos de datos de Amazon Redshift

En este tema se describen los pasos que debe seguir para aceptar una invitación a compartir datos, crear una base de datos federada y conceder permisos. Puede utilizar la consola de Lake Formation o AWS Command Line Interface (AWS CLI). Los ejemplos de este tema muestran el clúster de productores, el Catálogo de datos y el consumidor de datos en la misma cuenta.

Para obtener más información sobre las capacidades entre cuentas de Lake Formation, consulte Compartir datos entre cuentas en Lake Formation.

Para configurar los permisos de un recurso compartido de datos

1. Revise una invitación al recurso compartido de datos y acéptela.

   Console

   1. Inicie sesión en la consola de Lake Formation como administrador del lago de datos en https://console.aws.amazon.com/lakeformation/. Acceda a la página Intercambio de datos.
   2. Revise los datos compartidos a los que está autorizado a acceder. La columna Estado indica su estado de participación actual en el recurso compartido de datos. El estado Pendiente indica que se le ha agregado a un recurso compartido de datos, pero que aún no lo ha aceptado o ha rechazado la invitación.
   3. Para responder a una invitación a compartir datos, seleccione el nombre del recurso compartido de datos y elija Revisar la invitación. En Aceptar o rechazar el intercambio de datos, revise los detalles de la invitación. Seleccione Aceptar para aceptar la invitación o Rechazar para rechazarla. Si rechazas la invitación, no tendrás acceso al recurso compartido de datos.

   AWS CLI

   En los ejemplos siguientes se muestra cómo ver, aceptar y registrar la invitación. Sustituya el Cuenta de AWS ID por un ID válidoCuenta de AWS. Sustituya `data-share-arn` por el nombre de recurso de Amazon (ARN) real que hace referencia al recurso compartido de datos.

1. Ver una invitación pendiente.

```
aws redshift describe-data-shares \
 --data-share-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds' \
```

2. Acepte un recurso compartido de datos.

```
 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. Registre el recurso compartido de datos en la cuenta de Lake Formation. Utilice la operación RegisterResourceAPI para registrar el datashare en Lake Formation. `DataShareArn`es el parámetro de entrada de. `ResourceArn`

> **Note**
>
> Este es un paso obligatorio.

```
aws lakeformation register-resource \
 --resource-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds'
```

2. Cree una base de datos.

   Tras aceptar una invitación para un recurso compartido de datos, debe crear una base de datos que apunte a la base de datos de Amazon Redshift asociada al recurso compartido de datos. Debe ser administrador de un lago de datos para crear una base de datos.

   Console

   1. Seleccione el recurso compartido de datos en el panel Invitaciones y elija Establecer los detalles de la base de datos.

2. En Establecer los detalles de la base de datos, introduzca un nombre e identificador únicos para el recurso compartido de datos. Este identificador se utiliza para mapear el recurso compartido de datos internamente en la jerarquía de metadatos (dbName.Schema.Table).
3. Seleccione Siguiente para conceder permisos a otros usuarios en la base de datos y las tablas compartidas.

AWS CLI

Utilice el siguiente código de ejemplo para crear una base de datos que apunte a la base de datos de Amazon Redshift compartida con Lake Formation mediante. 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. Concesión de permisos.

Una vez creada la base de datos, puede conceder permisos a los usuarios de su cuenta o a organizaciones Cuentas de AWS y usuarios externos. No podrá conceder permisos de escritura de datos (insertar, eliminar) ni permisos de metadatos (modificar, eliminar, crear) en la base de datos federada que esté mapeada a un datashare de Amazon Redshift. Para obtener más información sobre cómo conceder permisos, consulte Administrar los permisos de Lake Formation.

> **Note**
>
> Como administrador de un lago de datos, solo puede ver las tablas de las bases de datos federadas. Para realizar cualquier otra acción, debe concederse más permisos en esas tablas.

Console

1. En la pantalla Conceder permisos, seleccione los usuarios a los que va a conceder los permisos.
2. Elija Conceder.

AWS CLI

Utilice los siguientes ejemplos para conceder permisos de bases de datos y tablas mediante 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"
                   ]
 }
```

Consultar las bases de datos federadas

Tras conceder los permisos, los usuarios pueden iniciar sesión y empezar a consultar la base de datos federada mediante Amazon Redshift. Los usuarios ahora pueden usar el nombre de la base de datos local para hacer referencia al recurso compartido de datos de Amazon Redshift en consultas de SQL. En Amazon Redshift, la tabla de clientes del esquema público que se comparte a través del recurso compartido de datos tendrá una tabla correspondiente creada como `public.customer` en el Catálogo de datos.

1. Antes de consultar la base de datos federada mediante Amazon Redshift, el administrador del clúster crea una base de datos a partir de la base de datos del Catálogo de datos mediante el siguiente comando:

```
CREATE DATABASE sharedcustomerdb FROM ARN
 'arn:aws:glue:<region>:111122223333:database/tahoedb' WITH DATA CATALOG SCHEMA
 tahoedb
```

2. El administrador del clúster concede permisos de uso en la base de datos.

```
GRANT USAGE ON DATABASE sharedcustomerdb TO IAM:user;
```

3. Usted (el usuario federado) ahora puede iniciar sesión en las herramientas de SQL para consultar la tabla.

```
Select * from sharedcustomerdb.public.customer limit 10;
```

Para obtener más información, visite Consulta de AWS Glue Data Catalog en la Guía de administración de Amazon Redshift.

Administración de los permisos de los conjuntos de datos que utilizan metaalmacenes externos

Con la federación de metadatos AWS Glue Data Catalog (federación de catálogos de datos), puede conectar el Catálogo de datos a metaalmacenes externos que almacenan los metadatos de sus datos de Amazon S3 y gestionar de forma segura los permisos de acceso a los datos mediante AWS Lake Formation. No tiene que migrar los metadatos del metaalmacén externo al Catálogo de datos.

El catálogo de datos proporciona un repositorio de metadatos centralizado que facilita la administración y el descubrimiento de datos en sistemas dispares. Cuando su organización administra los datos del Catálogo de datos, puede utilizar AWS Lake Formation para controlar el acceso a sus conjuntos de datos en Amazon S3.

> **Note**
>
> Actualmente, solo admitimos la federación de metaalmacenes Hive de Apache (versión 3 y superior).

Para configurar la federación de catálogos de datos, ofrecemos una aplicación AWS Serverless Application Model (AWS SAM) denominada GlueDataCatalogFederation- HiveMetastore en. AWS Serverless Application Repository

La implementación de referencia se proporciona GitHub como un proyecto de código abierto en AWS Glue Data CatalogFederation - Hive Metastore.

La aplicación AWS SAM crea e implementa los siguientes recursos necesarios para conectar el Catálogo de datos al metaalmacén de Hive:

- Una AWS Lambda función: aloja la implementación del servicio de federación que se comunica entre el catálogo de datos y el metaalmacén de Hive. AWS Glueinvoca esta función Lambda para recuperar objetos de metadatos del metabastore de Hive.
- Amazon API Gateway: el punto de conexión del metaalmacén de Hive que actúa como proxy para enrutar todas las invocaciones a la función de Lambda.
- Un rol de IAM: un rol con los permisos necesarios para crear la conexión entre el catálogo de datos y el metaalmacén de Hive.
- AWS Glueconexión: Amazon API Gateway tipo de AWS Glue conexión que almacena el Amazon API Gateway punto final y un rol de IAM para invocarlo.

Al consultar tablas, el servicio AWS Glue hace una llamada en tiempo de ejecución al metaalmacén de Hive y recupera los metadatos. La función de Lambda actúa como un traductor entre el metaalmacén de Hive y el Catálogo de datos.

Tras establecer la conexión, para sincronizar los metadatos del metaalmacén de Hive con el Catálogo de datos, debe crear una base de datos federada en el Catálogo de datos utilizando los detalles de conexión del metaalmacén de Hive y asignar esta base de datos a la base de datos de Hive. Una base de datos se denomina base de datos federada cuando apunta a una entidad ajena al Catálogo de datos.

Puede aplicar los permisos de Lake Formation mediante el control de acceso basado en etiquetas y el método de recurso con nombre en la base de datos federada, y compartirlos entre varias unidades organizativas (OU) y varias Cuentas de AWS unidades organizativas (OU). AWS Organizations También puede compartir la base de datos federada directamente con las entidades principales de IAM desde otra cuenta.

Puede definir permisos detallados a nivel de columna, nivel de fila y nivel de celda mediante los filtros de datos de Lake Formation en las tablas de Hive externas. Puede utilizar Amazon Athena, Amazon Redshift o Amazon EMR para consultar las tablas de colmenas externas gestionadas por Lake Formation.

Para obtener más información sobre el filtrado y el intercambio de datos entre cuentas, consulte:

- Compartir datos entre cuentas en Lake Formation
- Filtrado de datos y seguridad de celda en Lake Formation

Pasos básicos de la federación de metadatos del Catálogo de datos

1. Puede crear usuarios y roles de IAM que cuenten con los permisos adecuados para implementar la AWS SAM aplicación y crear bases de datos federadas.
2. Para registrar la ubicación de datos de Amazon S3 en Lake Formation, debe seleccionar la opción `Enable Data Catalog federation` para los conjuntos de datos que utilizan un metaalmacén de Hive externo.
3. Debe configurar los ajustes de la aplicación AWS SAM (nombre de la conexión AWS Glue, URL al metaalmacén de Hive y parámetros de la función de Lambda) e implementar la aplicación de AWS SAM.
4. La aplicación de AWS SAM crea e implementa los siguientes recursos necesarios para conectar el Catálogo de datos al metaalmacén de Hive.
5. Para aplicar los permisos de Lake Formation a la base de datos y las tablas de Hive, cree una base de datos en el catálogo de datos utilizando los detalles de conexión del metaalmacén de Hive y asigne esta base de datos a la base de datos de Hive.
6. Conceda permisos en las bases de datos federadas a las entidades principales de su cuenta o de otra cuenta.

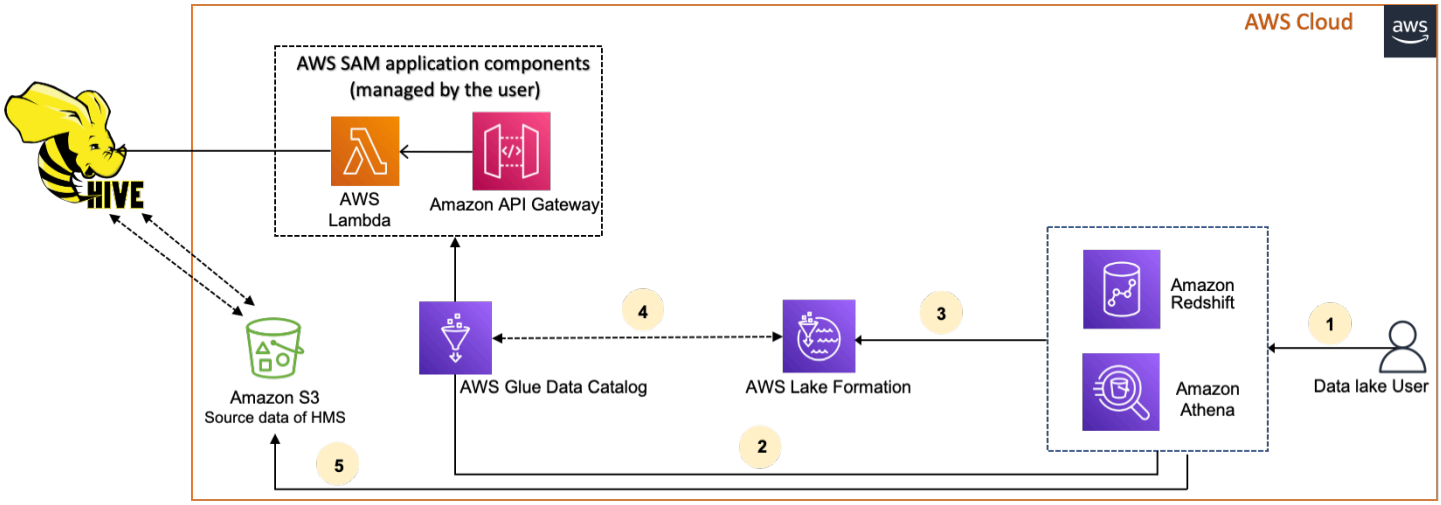
> **Note**
>
> Puede conectar el Catálogo de datos a un metaalmacén de Hive externo, crear bases de datos federadas y ejecutar consultas y scripts de ETL en bases de datos y tablas de Hive sin aplicar los permisos de Lake Formation. Para los datos de origen en Amazon S3 que no estén registrados en Lake Formation, el acceso se determina mediante las políticas de permisos de IAM para Amazon S3 y acciones de AWS Glue.

Para conocer las limitaciones, consulte Consideraciones y limitaciones del uso compartido de datos del almacén de metadatos de Hive.

Temas

Flujo de trabajo

En el siguiente diagrama, se muestra el flujo de trabajo para conectar el AWS Glue Data Catalog a un metaalmacén externo de Hive.



1. Una entidad principal envía una consulta mediante un servicio integrado como Athena o Redshift Spectrum.
2. El servicio integrado realiza una llamada al catálogo de datos para obtener los metadatos, que a su vez llama al punto final del metaalmacén de Hive disponible en la versión trasera y recibe las respuestas a las solicitudes Amazon API Gateway de metadatos.
3. El servicio integrado envía la solicitud a Lake Formation para verificar la información de la tabla y las credenciales para acceder a la tabla.
4. Lake Formation autoriza la solicitud y suministra credenciales temporales a la aplicación integrada, que permite el acceso a los datos.
5. Con las credenciales temporales recibidas de Lake Formation, el servicio integrado lee los datos de Amazon S3 y comparte los resultados con la entidad principal.

Requisitos previos para conectar el Catálogo de datos al metaalmacén de Hive

Para conectar AWS Glue Data Catalog a un metaalmacén externo de Apache Hive y configurar los permisos de acceso a los datos, debe cumplir los siguientes requisitos:

> **Note**
>
> Recomendamos que un administrador de Lake Formation despliegue la aplicación de AWS SAM y que solo un usuario privilegiado utilice la conexión del metaalmacén de Hive para crear las bases de datos federadas correspondientes.

1. Crear roles de IAM.

   Para implementar la aplicación de AWS SAM

   - Cree un rol que tenga los permisos necesarios para implementar los recursos (función de LambdaAmazon API Gateway, rol de IAM y la AWS Glue conexión) necesarios para crear una conexión al metaalmacén de Hive.

   Para crear bases de datos federadas

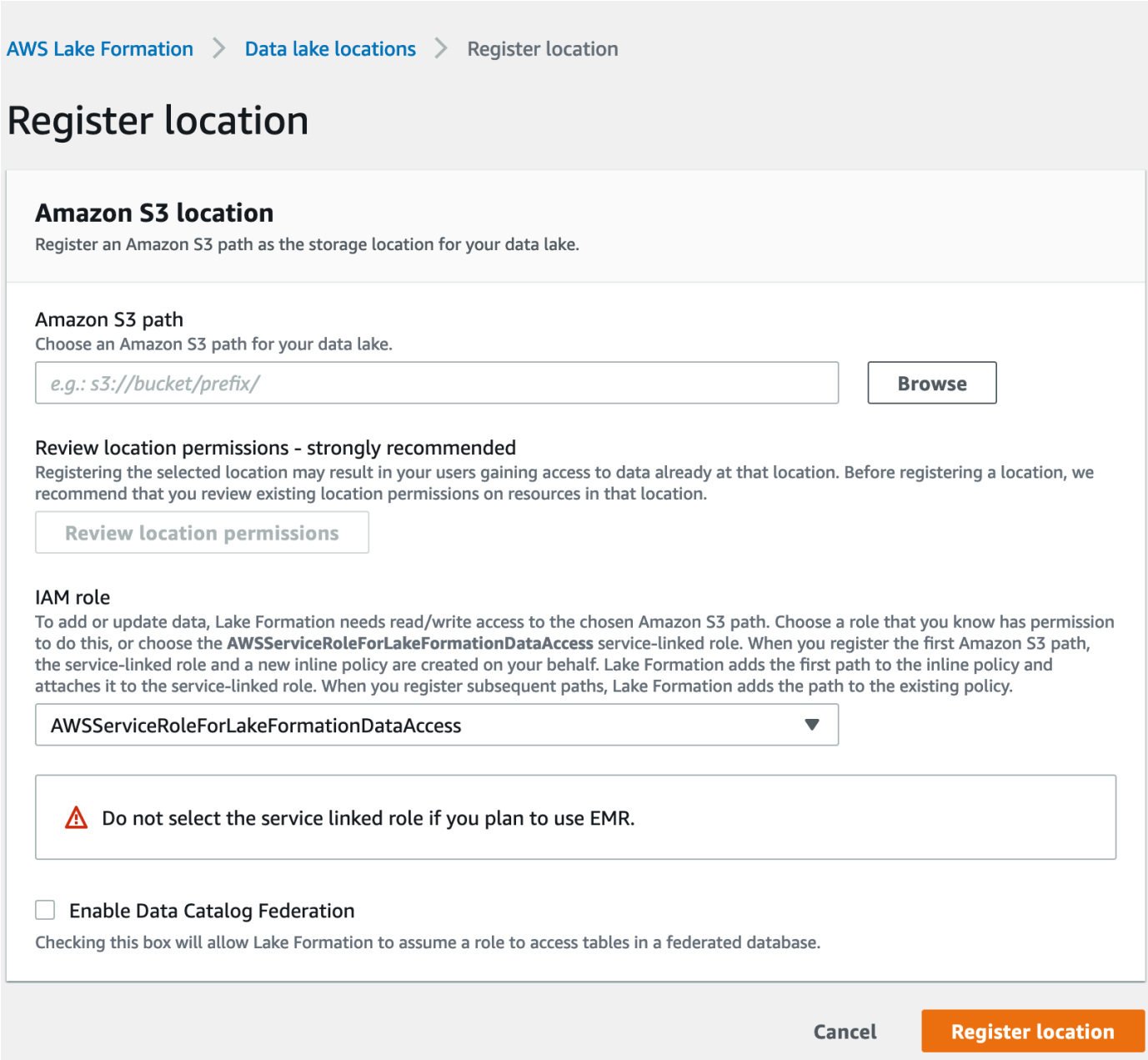
   Los recursos requieren los siguientes permisos:

   - `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. Registre la ruta de Amazon S3 en Lake Formation.

   Para utilizar Lake Formation para gestionar y proteger los datos de su lago de datos, debe registrar la ubicación de Amazon S3 que contiene los datos de las tablas en el metaalmacén de Hive con Lake Formation. De este modo, Lake Formation puede vender credenciales a servicios AWS analíticos como Athena, Redshift Spectrum y Amazon EMR.

   Para obtener más información sobre el registro de una ubicación de Amazon S3, consulte Añadir una ubicación de Amazon S3 a su lago de datos.

   Cuando registre la ubicación de Amazon S3, seleccione la casilla Habilitar la federación de catálogos de datos para permitir que Lake Formation asuma una función de acceso a las tablas de una base de datos federada.



Para obtener más información sobre el registro de una ubicación de datos en Lake Formation, consulte Configurar una ubicación de Amazon S3 para el lago de datos.

3. Utilice la versión correcta de Amazon EMR.

   Para usar Amazon EMR con las bases de datos federadas de Metastore de Hive, debe tener Hive versión 3.x o superior y Amazon EMR versión 6.x o superior.

Conexión del Catálogo de datos a un metaalmacén de Hive externo

Para conectarlos a un metaalmacén AWS Glue Data Catalog de Hive, debe implementar una aplicación llamada -. AWS SAM GlueDataCatalogFederation HiveMetastore Crea e implementa los recursos necesarios para conectar el metaalmacén de Hive al Catálogo de datos. Puede acceder a la aplicación AWS SAM en AWS Serverless Application Repository.

La aplicación AWS SAM crea la conexión para el metaalmacén de Hive detrás de Amazon API Gateway mediante una función de Lambda. La aplicación AWS SAM utiliza un identificador uniforme de recursos (URI) como entrada del usuario y conecta el metaalmacén externo de Hive al Catálogo de datos. Cuando un usuario ejecuta una consulta en las tablas de Hive, el catálogo de datos llama al punto final de API Gateway. El punto de conexión invoca la función de Lambda para recuperar los metadatos de las tablas de Hive.

Para conectar el Catálogo de datos al metaalmacén de Hive y configurar los permisos

1. Implemente la aplicación AWS SAM.
    1. Inicie sesión en la AWS Management Console y abra la AWS Serverless Application Repository.
    2. En el panel de navegación, elija Aplicaciones disponibles.
    3. Elija Aplicaciones públicas.
    4. Seleccione la opción Mostrar aplicaciones que crean roles de IAM personalizados o políticas de recursos.
    5. En el cuadro de búsqueda, introduzca el nombre GlueDataCatalogFederation- HiveMetastore.
    6. Seleccione la HiveMetastore aplicación GlueDataCatalogFederation-.
    7. En Configuración de la aplicación, introduzca la siguiente configuración mínima requerida para la función de Lambda:
        - Nombre de la aplicación: escriba el nombre de la aplicación AWS SAM.
        - GlueConnectionName- Un nombre para la conexión.
        - HiveMetastoreURIs: el URI de tu servidor de metatienda de Hive.
        - LambdaMemory- La cantidad de memoria Lambda en MB de 128 a 10240. El valor predeterminado es 1024.
        - LambdaTimeout- El tiempo de ejecución máximo de la invocación a Lambda en segundos. El valor predeterminado es 30.

- VPC y SecurityGroupIds VPCSubnetIds: información sobre la VPC en la que se encuentra el metabastore de Hive.

8. Seleccione Confirmo que esta aplicación puede crear roles de IAM y políticas de recursos personalizados. Para obtener más información, elija el enlace Info.
9. En la parte inferior derecha de la sección Configuración de aplicación, elija Implementar. Una vez finalizada la implementación, aparece la función de Lambda en la sección Recursos en la consola de Lambda.

La aplicación se implementa en Lambda. Su nombre va precedido de serverlessrepo- para indicar que la aplicación se implementó desde. AWS Serverless Application Repository Al seleccionar la aplicación, accederá a la página Recursos, donde se enumeran todos los recursos de la aplicación que se implementaron. Los recursos incluyen la función de Lambda que permite la comunicación entre el Catálogo de datos y el metaalmacén de Hive, la conexión AWS Glue y otros recursos necesarios para la federación de bases de datos.

2. Crear una base de datos federada en Data Catalog.

Después de crear una conexión al metabastore de Hive, puede crear bases de datos federadas en el catálogo de datos que apunten a las bases de datos externas del metaalmacén de Hive. Debe crear una base de datos correspondiente en el catálogo de datos para cada base de datos del metaalmacén de Hive que vaya a conectar al catálogo de datos.

Lake Formation console

1. En la página de Uso compartido de datos, seleccione la pestaña Bases de datos compartidas y, a continuación, Crear base de datos.
2. En el nombre de la conexión, elige el nombre de tu conexión a la metatienda de Hive en el menú desplegable.
3. Introduzca un nombre de base de datos único y el identificador de origen de federación para la base de datos. Este es el nombre que se utiliza en las instrucciones SQL cuando se consultan las tablas. El nombre puede tener un máximo de 255 caracteres como máximo y debe ser único en su cuenta.
4. Elija Crear base de datos.

AWS CLI

```
aws glue create-database \
```

```
'{
 "CatalogId": "<111122223333>",
  "database-input": {
    "Name":"<fed_glue_db>",
    "FederatedDatabase":{
        "Identifier":"<hive_db_on_emr>",
        "ConnectionName":"<hms_connection>"
     }
   }
 }'
```

3. Ver las tablas de la base de datos federada.

   Después de crear la base de datos federada, puede ver la lista de tablas en su metaalmacén de Hive mediante la consola de Lake Formation o la AWS CLI.

   Lake Formation console

   1. Seleccione el nombre de la base de datos en la pestaña Bases de datos compartidas.
   2. En la página Bases de datos, elija Ver tablas.

   AWS CLI

   Los siguientes ejemplos muestran cómo recuperar la definición de conexión, el nombre de la base de datos y algunas o todas las tablas de la base de datos. Sustituya el identificador del catálogo de datos por el Cuenta de AWS identificador válido que utilizó para crear la base de datos. Sustituya `hms_connection` por el nombre de la conexión.

   ```
   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. Concesión de permisos.

   Una vez creada la base de datos, puede conceder permisos a otros usuarios y roles de IAM en su cuenta o a organizaciones Cuentas de AWS y organizaciones externas. No podrá conceder permisos de escritura de datos (insertar, eliminar) ni permisos de metadatos (modificar, eliminar, crear) en las bases de datos federadas. Para obtener más información sobre cómo conceder permisos, consulte Administrar los permisos de Lake Formation.

5. Consultar las bases de datos federadas.

   Tras conceder los permisos, los usuarios pueden iniciar sesión y empezar a consultar la base de datos federada mediante Athena y Amazon Redshift. Los usuarios ahora pueden usar el nombre de la base de datos local para hacer referencia a la base de datos de Hive en las consultas de SQL.

   Ejemplo de sintaxis Amazon Athena de consulta

   `fed_glue_db`Sustitúyalo por el nombre de la base de datos local que creó anteriormente.

   ```
   Select * from fed_glue_db.customers limit 10;
   ```

Recursos adicionales de

La siguiente entrada del blog contiene instrucciones detalladas para configurar los permisos de Lake Formation en una base de datos y tablas del metaalmacén de Hive y consultarlos mediante Athena. También ilustramos un caso de uso compartido entre cuentas, en el que un director de Lake Formation en la cuenta de productor A comparte una base de datos y tablas de Hive federadas mediante la etiqueta LF con la cuenta de consumidor B.

- Consultar su metaalmacén de Apache Hive con permisos AWS Lake Formation

Seguridad en AWS Lake Formation

La seguridad en la nube de AWS es la mayor prioridad. Como cliente de AWS, se beneficia de una arquitectura de red y un centro de datos diseñados para satisfacer los requisitos de seguridad de las organizaciones más exigentes.

La seguridad es una responsabilidad compartida entre AWS y el usuario. El modelo de responsabilidad compartida la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube: AWS es responsable de proteger la infraestructura que ejecuta los servicios de AWS en la nube de AWS. AWS también proporciona servicios que puede utilizar de forma segura. Auditores externos prueban y verifican periódicamente la eficacia de nuestra seguridad en el marco de los programas de conformidad de AWS. Para obtener más información sobre los programas de conformidad que se aplican a AWS Lake Formation, consulte Servicios de AWS en el ámbito del programa de conformidad.
- Seguridad en la nube: su responsabilidad se determina según el servicio de AWS que utilice. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y los reglamentos aplicables.

Esta documentación le ayuda a comprender cómo aplicar el modelo de responsabilidad compartida cuando se utiliza Lake Formation. En los temas siguientes, se le mostrará cómo configurar Lake Formation para satisfacer sus objetivos de seguridad y conformidad. También puede aprender a utilizar otros servicios de AWS que lo ayuden a monitorear y proteger los recursos de Lake Formation.

Temas

- Protección de datos en Lake Formation
- Seguridad de infraestructuras en AWS Lake Formation
- Prevención del suplente confuso entre servicios
- Registro de eventos de seguridad en AWS Lake Formation

Protección de datos en Lake Formation

El modelo de responsabilidad compartida de AWS se aplica a la protección de datos en AWS Lake Formation. Como se describe en este modelo, AWS es responsable de proteger la infraestructura

global que ejecuta la totalidad de Nube de AWS. Usted es responsable de mantener el control sobre el contenido alojado en esta infraestructura. Usted también es responsable de las tareas de administración y configuración de seguridad para los Servicios de AWS que utiliza. Para obtener más información sobre la privacidad de los datos, consulte las Preguntas frecuentes sobre la privacidad de datos. Para obtener información sobre la protección de datos en Europa, consulte la publicación de blog Modelo de responsabilidad compartida y GDPR de AWS en el Blog de seguridad de AWS.

Con fines de protección de datos, recomendamos proteger las credenciales de la cuenta de Cuenta de AWS y configurar cuentas de usuario individuales con AWS IAM Identity Center o AWS Identity and Access Management (IAM). De esta manera, solo se conceden a cada usuario los permisos necesarios para cumplir con sus obligaciones laborales. También recomendamos proteger sus datos de la siguiente manera:

- Utilice autenticación multifactor (MFA) en cada cuenta.
- Utilice SSL/TLS para comunicarse con los recursos de AWS. reSe recomienda el uso de TLS 1.2 y recomendamos TLS 1.3.
- Configure la API y el registro de actividad del usuario con AWS CloudTrail.
- Utilice las soluciones de cifrado de AWS, junto con todos los controles de seguridad predeterminados dentro de los Servicios de AWS.
- Utilice servicios de seguridad gestionados avanzados, como Amazon Macie, que lo ayuden a detectar y proteger los datos confidenciales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados FIPS 140-2 al acceder a AWS a través de una interfaz de la línea de comandos o una API, utilice un punto de conexión de FIPS. Para obtener más información sobre los puntos de conexión de FIPS disponibles, consulte Estándar de procesamiento de la información federal (FIPS) 140-2.

Se recomienda encarecidamente no introducir nunca información confidencial o sensible, como, por ejemplo, direcciones de correo electrónico de clientes, en etiquetas o campos de formato libre, tales como el campo Nombre. Esto incluye el trabajo con Lake Formation u otros Servicios de AWS utilizando la consola, API, AWS CLI, o AWS SDK. Cualquier dato que ingrese en etiquetas o campos de formato libre utilizados para nombres se puede emplear para los registros de facturación o diagnóstico. Si proporciona una URL a un servidor externo, recomendamos encarecidamente que no incluya la información de las credenciales en la URL para validar la solicitud para ese servidor.

Cifrado en reposo

AWS Lake Formation es compatible con el cifrado de datos en las siguientes áreas:

- Datos en el lago de datos de Amazon Simple Storage Service (Amazon S3).

  Lake Formation es compatible con el cifrado de datos con AWS Key Management Service (AWS KMS). Los datos se escriben normalmente en el lago de datos mediante trabajos de extracción, transformación y carga (ETL) de AWS Glue. Para obtener información sobre cómo cifrar los datos escritos por trabajos AWS Glue, consulte Cifrado de datos escritos por rastreadores, trabajos y puntos de conexión de desarrollo en la Guía para desarrolladores de AWS Glue.
- El AWS Glue Data Catalog, que es donde Lake Formation almacena las tablas de metadatos que describen los datos del lago de datos.

  Para obtener más información, consulte Cifrado de su catálogo de datos en la Guía para desarrolladores de AWS Glue.

Para añadir una ubicación de Amazon S3 como almacenamiento en su lago de datos, debe registrar la ubicación con AWS Lake Formation. A continuación, puede utilizar los permisos de Lake Formation para un control de acceso específico a los objetos AWS Glue Data Catalog que apuntan a esta ubicación y a sus datos subyacentes.

Lake Formation es compatible con el registro de una ubicación de Amazon S3 que contenga datos cifrados. Para obtener más información, consulte Registro de una ubicación cifrada de Amazon S3.

Seguridad de infraestructuras en AWS Lake Formation

Por ser un servicio administrado, AWS Lake Formation está protegido por los procedimientos de seguridad de red globales de AWS descritos en el documento técnico Amazon Web Services: Información general sobre los procesos de seguridad.

Para acceder a Lake Formation a través de la red, se utilizan llamadas a API publicadas en AWS. Los clientes deben ser compatibles con Transport Layer Security (TLS) 1.0 o una versión posterior. Recomendamos TLS 1.2 o una versión posterior. Los clientes también deben ser compatibles con conjuntos de cifrado con confidencialidad directa total (PFS) tales como Ephemeral Diffie-Hellman (DHE) o Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad de seguridad de IAM. También puede utilizar AWS Security Token Service (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Prevención del suplente confuso entre servicios

El problema del suplente confuso es un problema de seguridad en el que una entidad que no tiene permiso para realizar una acción puede obligar a una entidad con más privilegios a realizar la acción. En AWS, la suplantación entre servicios puede dar lugar al problema del suplente confuso. La suplantación entre servicios puede producirse cuando un servicio (el servicio que lleva a cabo las llamadas) llama a otro servicio (el servicio al que se llama). El servicio que lleva a cabo las llamadas se puede manipular para utilizar sus permisos a fin de actuar en función de los recursos de otro cliente de una manera en la que no debe tener permiso para acceder. Para evitarlo, AWS proporciona herramientas que lo ayudan a proteger sus datos para todos los servicios con entidades principales de servicio a las que se les ha dado acceso a los recursos de su cuenta.

Se recomienda utilizar las claves de contexto de condición global `aws:SourceArn` y `aws:SourceAccount` en las políticas de recursos para limitar los permisos que AWS Lake Formation concede a otro servicio para el recurso. Si se utilizan ambas claves de contexto de condición global, el valor `aws:SourceAccount` y la cuenta del valor `aws:SourceArn` deben utilizar el mismo ID de cuenta cuando se utilicen en la misma declaración de política.

Actualmente, Lake Formation solo admite `aws:SourceArn` en los formatos siguientes:

```
arn:aws:lakeformation:aws-region:account-id:*
```

El ejemplo siguiente muestra cómo se pueden utilizar las claves contextuales de condición global `aws:SourceArn` y `aws:SourceAccount` en Lake Formation para evitar el problema del suplente confuso.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ConfusedDeputyPreventionExamplePolicy",
      "Effect": "Allow",
      "Principal": {
        "Service": "lakeformation.amazonaws.com"
      },
      "Action": [
        "sts:AssumeRole"
      ],
      "Condition": {
        "StringEquals": {
```

```
          "aws:SourceAccount": "account-id"
        },
        "ArnEquals": {
          "aws:SourceArn": "arn:aws:lakeformation:aws-region:account-id:*"
        }
      }
    }
  ]
}
```

Registro de eventos de seguridad en AWS Lake Formation

AWS Lake Formation se integra con AWS CloudTrail, un servicio que proporciona un registro de la actividad desarrollada por un usuario, un rol o un servicio de AWS en Lake Formation. CloudTrail captura las llamadas a la API de Lake Formation como eventos. Entre las llamadas capturadas se incluyen las efectuadas desde la consola de Lake Formation, la AWS Command Line Interface y las llamadas de código a las operaciones de la API de Lake Formation.

Para obtener más información acerca del registro de eventos de Lake Formation, consulte Registro de llamadas a la API de AWS Lake Formation mediante AWS CloudTrail.

> **Note**
>
> `GetTableObjects`, `UpdateTableObjects` y `GetWorkUnitResults` son operaciones de planos de datos de gran volumen. En la actualidad, las llamadas a estas API no se registran en CloudTrail. Para obtener más información sobre operaciones de planos de datos, consulte Registrar eventos de datos para registros de seguimiento en la Guía del usuario de AWS CloudTrail.
>
> Los cambios de Lake Formation para su compatibilidad con otros eventos de CloudTrail se documentarán en Historial de documentos de AWS Lake Formation.

Integración de servicios de terceros con Lake Formation

La integración con AWS Lake Formation permite a los servicios de terceros acceder de forma segura a los datos de sus lagos de datos basados en Amazon S3. Puede usar Lake Formation como motor de autorización para administrar o hacer cumplir los permisos de su lago de datos con servicios integrados de AWS, como Amazon Athena, Amazon EMR y Redshift Spectrum. Lake Formation ofrece dos opciones para integrar servicios:

1. Configuración de integración de la aplicación Lake Formation: Lake Formation puede suministrar credenciales temporales con alcance reducido en forma de tokens AWS STS a ubicaciones registradas de Amazon S3 según los permisos vigentes, de modo que las aplicaciones autorizadas puedan acceder a los datos en nombre de los usuarios.
2. Aplicación centralizada: las operaciones de la API de consulta de Lake Formation recuperan datos de Amazon S3 y filtran los resultados en función de los permisos efectivos. El motor o la aplicación que se integran con la operación de la API de consulta pueden depender de que Lake Formation evalúe los permisos de la identidad que efectúa la llamada y filtre los datos de forma segura en función de estos permisos. Los motores de consulta de terceros solo ven y funcionan con datos filtrados.

Temas

Uso de la integración de aplicaciones de Lake Formation

Lake Formation permite que los servicios de terceros se integren con Lake Formation y obtengan acceso temporal a los datos de Amazon S3 en nombre de sus usuarios mediante el uso GetTemporaryGlueTableCredentialsy GetTemporaryGluePartitionCredentialslas operaciones. Esto permite que los servicios de terceros utilicen la misma característica de autorización y suministro de credenciales que utilizan el resto de los servicios de análisis de AWS. En esta sección se describe cómo utilizar estas operaciones de la API para integrar un motor de consultas de terceros con Lake Formation.

Estas operaciones de la API están deshabilitadas de forma predeterminada. Existen dos opciones para autorizar a Lake Formation a integrar aplicaciones:

- Configurar las etiquetas de sesión de IAM para que se validen cada vez que se invoquen las operaciones de la API de integración de aplicaciones

  Para obtener más información, consulte Habilitar los permisos para que un motor de consultas de terceros llame a las operaciones de la API de integración de aplicaciones.

- Habilitar la opción que permite a los motores externos acceder a los datos en las ubicaciones de Amazon S3 con acceso completo a las tablas

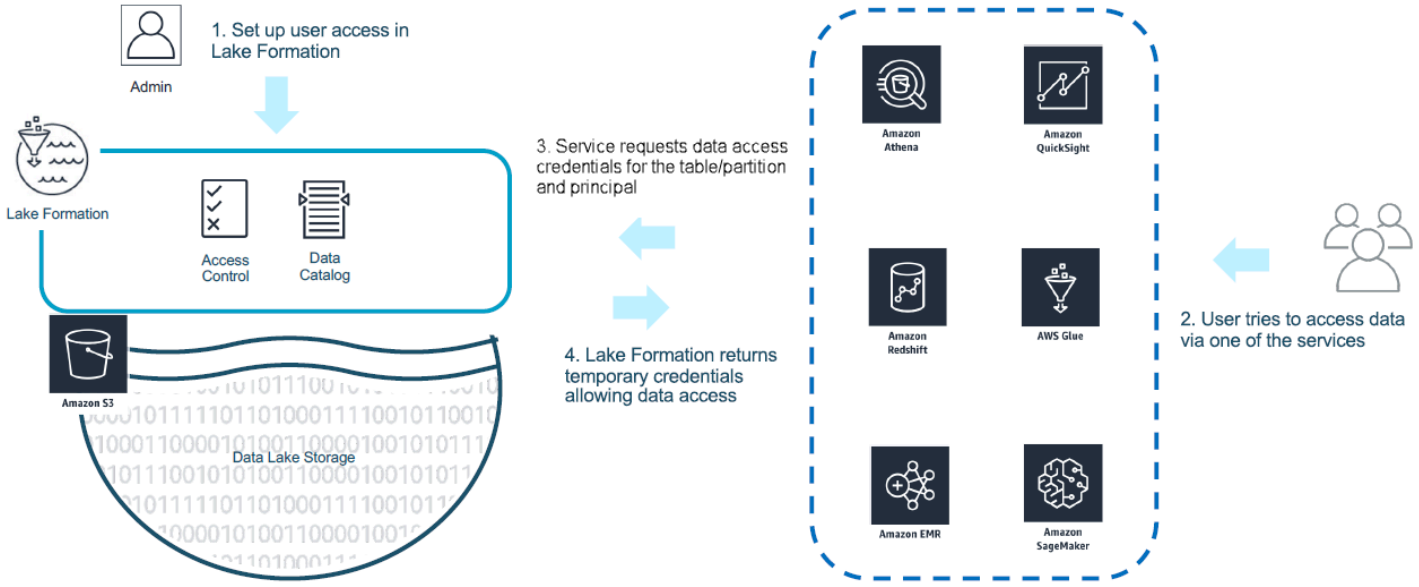
  Esta opción permite que los motores de consultas y las aplicaciones obtengan credenciales sin etiquetas de sesión de IAM si el usuario tiene acceso completo a la tabla. Proporciona beneficios de rendimiento a los motores de consulta y las aplicaciones, además de simplificar el acceso a los datos. Amazon EMR en Amazon EC2 puede aprovechar esta configuración.

  Para obtener más información, consulte Integración de aplicaciones para un acceso completo a la tabla .

Temas

Cómo funciona la integración de aplicaciones de Lake Formation

En esta sección se describe cómo utilizar las operaciones de la API de integración de aplicaciones para integrar una aplicación de terceros (motor de consultas) con Lake Formation.



1. El administrador de Lake Formation efectúa las siguientes actividades:
   - Registra una ubicación de Amazon S3 en Lake Formation proporcionando un rol de IAM (utilizado para suministrar credenciales) que tiene los permisos adecuados para acceder a los datos de la ubicación de Amazon S3
   - Registra una aplicación de terceros para hacer llamadas a las operaciones de la API de suministro de credenciales de Lake Formation. Consulte the section called “Registro de un motor de consultas de terceros”
   - Concede permisos para permitir el acceso a bases de datos y tablas

     Por ejemplo, si desea publicar un conjunto de datos de sesiones de usuario que incluya algunas columnas que contengan información de identificación personal (PII), para restringir el acceso, asigne a estas columnas una etiqueta LF-TBAC denominada «clasificación» con el valor «confidencial». A continuación, define un permiso que permita a un analista empresarial acceder a los datos de las sesiones de los usuarios, pero excluye las columnas etiquetadas con classification = sensitive.
2. Una entidad principal (usuario) envía una consulta a un servicio integrado.
3. La aplicación integrada envía la solicitud a Lake Formation solicitando la información de la tabla y las credenciales para acceder a la tabla.
4. Si la entidad principal que efectúa la consulta está autorizada a acceder a la tabla, Lake Formation devuelve las credenciales a la aplicación integrada, lo que permite el acceso a los datos.

> **Note**
>
> Lake Formation no accede a los datos subyacentes cuando vende credenciales.

5. El servicio integrado lee los datos de Amazon S3, filtra las columnas según las políticas que ha recibido y devuelve los resultados a la entidad principal.

> **Important**
>
> Las operaciones de la API de suministro de credenciales de Lake Formation permiten una aplicación distribuida con un modelo explícito de denegación en caso de fallo (cierre por error). Esto introduce un modelo de seguridad tripartito entre los clientes, los servicios de terceros y Lake Formation. Se confía en los servicios integrados para hacer cumplir adecuadamente los permisos de Lake Formation (aplicación distribuida).

El servicio integrado es responsable de filtrar los datos leídos de Amazon S3 según las políticas devueltas de Lake Formation antes de que los datos filtrados regresen al usuario. Los servicios integrados siguen un modelo de cierre por error, lo que significa que la consulta debe fallar si no pueden hacer cumplir los permisos de Lake Formation necesarios.

Roles y responsabilidades en la integración de aplicaciones de Lake Formation

Rol	Responsabilidad
El cliente	• Habilita la configuración de integración de aplicaciones de Lake Formation (consulte the section called “Registro de un motor de consultas de terceros”). • Registra explícitamente los terceros aprobados en Lake Formation (consulte the section called “Registro de un motor de consultas de terceros”). • Prueba y valida soluciones de terceros con los permisos de Lake Formation.

Rol	Responsabilidad
	• Supervisa y audita el uso de terceros de las operaciones de la API de suministro de credenciales de Lake Formation.
El tercero	• Documenta públicamente la capacidad compatible con cada revisión del software y proporciona instrucciones para habilitarla correctam ente. • Anuncia con precisión las capacidades compatibles al llamar a las operaciones de la API de suministro de credenciales de Lake Formation (según la documentación). • Almacena y gestiona de forma segura las credenciales suministradas para evitar la filtración de credenciales y el aumento de privilegios. • Aplica los permisos de acuerdo con las capacidades compatibles y devuelve solo los datos filtrados a los usuarios • No se puede llevar a cabo la consulta si no se pueden aplicar correctamente los permisos necesarios
`AWS Lake Formation`	• Deriva y devuelve correctamente los permisos efectivos para una determinada entidad principal. • Valida las capacidades compatibles con terceros en función del funcionamiento de la API. call-by-call • Devuelve las credenciales de IAM con un alcance reducido solo cuando las capacidades anunciadas del motor coinciden con las definidas en los recursos del catálogo; de lo contrario, devuelve un error.

Flujo de trabajo de Lake Formation para las operaciones de la API de integración de aplicaciones

A continuación se muestra el flujo de trabajo para las operaciones de la API de integración de aplicaciones:

1. Un usuario envía una consulta o solicitud de datos mediante un motor de consultas integrado de terceros. El motor de consultas asume un rol de IAM que representa al usuario o a un grupo de

usuarios y recupera las credenciales de confianza para utilizarlas al llamar a las operaciones de la API de integración de aplicaciones.

2. El motor de consultas invoca `GetUnfilteredTableMetadata`. Si se trata de una tabla con particiones, invoca `GetUnfilteredPartitionsMetadata` para recuperar los metadatos y la información sobre políticas del Catálogo de datos.
3. Lake Formation autoriza la solicitud. Si el usuario no tiene los permisos adecuados en la tabla, AccessDeniedExceptionse descarta.
4. Como parte de la solicitud, el motor de consultas envía el filtrado que admite. Hay dos indicadores que se pueden enviar dentro de una matriz: COLUMN_PERMISSIONS y CELL_FILTER_PERMISSION. Si el motor de consultas no admite ninguna de estas funciones y existe una política sobre la misma, se lanza una y la consulta no PermissionTypeMismatchExceptionse realiza correctamente. Esto ocurre para evitar la fuga de datos.
5. La respuesta devuelta contiene lo siguiente:
   - El esquema completo de la tabla, que los motores de consultas pueden usar para analizar los datos almacenados.
   - Una lista de columnas autorizadas a las que el usuario tiene acceso. Si la lista de columnas autorizadas está vacía, significa que el usuario tiene permisos de `DESCRIBE`, pero no tiene permisos de `SELECT`, y se produce un error en la consulta.
   - Una marca, `IsRegisteredWithLakeFormation`, que indica si Lake Formation puede vender credenciales a los datos de este recurso. Si el resultado es falso, deben usarse credenciales de clientes para acceder a Amazon S3.
   - Una lista de `CellFilters` si hay alguno que deba aplicarse a las filas de datos. Esta lista contiene columnas y una expresión para evaluar cada fila. Solo debe rellenarse si se envía CELL_FILTER_PERMISSION como parte de la solicitud y hay un filtro de datos en la tabla para el usuario que llama.
6. Una vez recuperados los metadatos, el motor de consultas invoca `GetTemporaryGlueTableCredentials` o `GetTemporaryGluePartitionCredentials` para obtener credenciales AWS y recuperar los datos de la ubicación de Amazon S3.
7. El motor de consultas lee los objetos relevantes de Amazon S3, filtra los datos según las políticas recibidas en el paso 2 y devuelve los resultados al usuario.

Las operaciones de la API de integración de aplicaciones para Lake Formation incluyen contenido adicional para configurar la integración con motores de consultas de terceros. Puede ver los detalles de la operación en la sección Operaciones de la API de suministro de credenciales.

Registro de un motor de consultas de terceros

Antes de que un motor de consultas de terceros pueda utilizar las operaciones de la API de integración de aplicaciones, debe habilitar explícitamente los permisos para que el motor de consultas llame a las operaciones de la API en su nombre. Esto se efectúa en unos pocos pasos:

1. Debe especificar las cuentas de AWS y las etiquetas de sesión de IAM que requieren permiso para acceder a las operaciones de la API de integración de aplicaciones a través de la consola de AWS Lake Formation, la AWS CLI, la API o el SDK.
2. Cuando el motor de consultas de terceros asume el rol de ejecución en su cuenta, el motor de consultas debe adjuntar una etiqueta de sesión registrada en Lake Formation que represente al motor de terceros. Lake Formation usa esta marca para validar si la solicitud proviene de un motor aprobado. Para obtener más información acerca de las marcas de sesión, consulte Etiquetas de sesión el la Guía del usuario de IAM.
3. Al configurar un rol de ejecución de un motor de consultas de terceros, debe tener el siguiente conjunto mínimo de permisos en la política de IAM:

```
{
  "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. Configure una política de confianza de roles en el rol de ejecución del motor de consultas para tener un control de acceso preciso sobre qué par clave-valor de etiqueta de sesión puede adjuntarse a este rol. En el ejemplo siguiente, a este rol solo se le permite adjuntar una clave de etiqueta de sesión `"LakeFormationAuthorizedCaller"` y un valor de etiqueta de sesión `"engine1"`, y no se permite ningún otro par de valores clave de etiqueta de sesión.

```
{
    "Sid": "AllowPassSessionTags",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/query-execution-role"
    },
    "Action": "sts:TagSession",
    "Condition": {
    "StringLike": {
        "aws:RequestTag/LakeFormationAuthorizedCaller": "engine1"        }
    }
}
```

Cuando se `LakeFormationAuthorizedCaller` llama a la operación STS: AssumeRole API para obtener credenciales para que las utilice el motor de consultas, la etiqueta de sesión debe incluirse en la AssumeRole solicitud. La credencial temporal devuelta se puede usar para efectuar solicitudes a la API de integración de aplicaciones de Lake Formation.

Las operaciones de la API de integración de aplicaciones de Lake Formation requieren que la entidad principal que llama tenga un rol de IAM. El rol de IAM debe incluir una etiqueta de sesión con un valor predeterminado que se haya registrado en Lake Formation. Esta etiqueta permite a Lake Formation verificar si el rol utilizado para llamar a las operaciones de la API de integración de aplicaciones está autorizado a hacerlo.

Habilitar los permisos para que un motor de consultas de terceros llame a las operaciones de la API de integración de aplicaciones

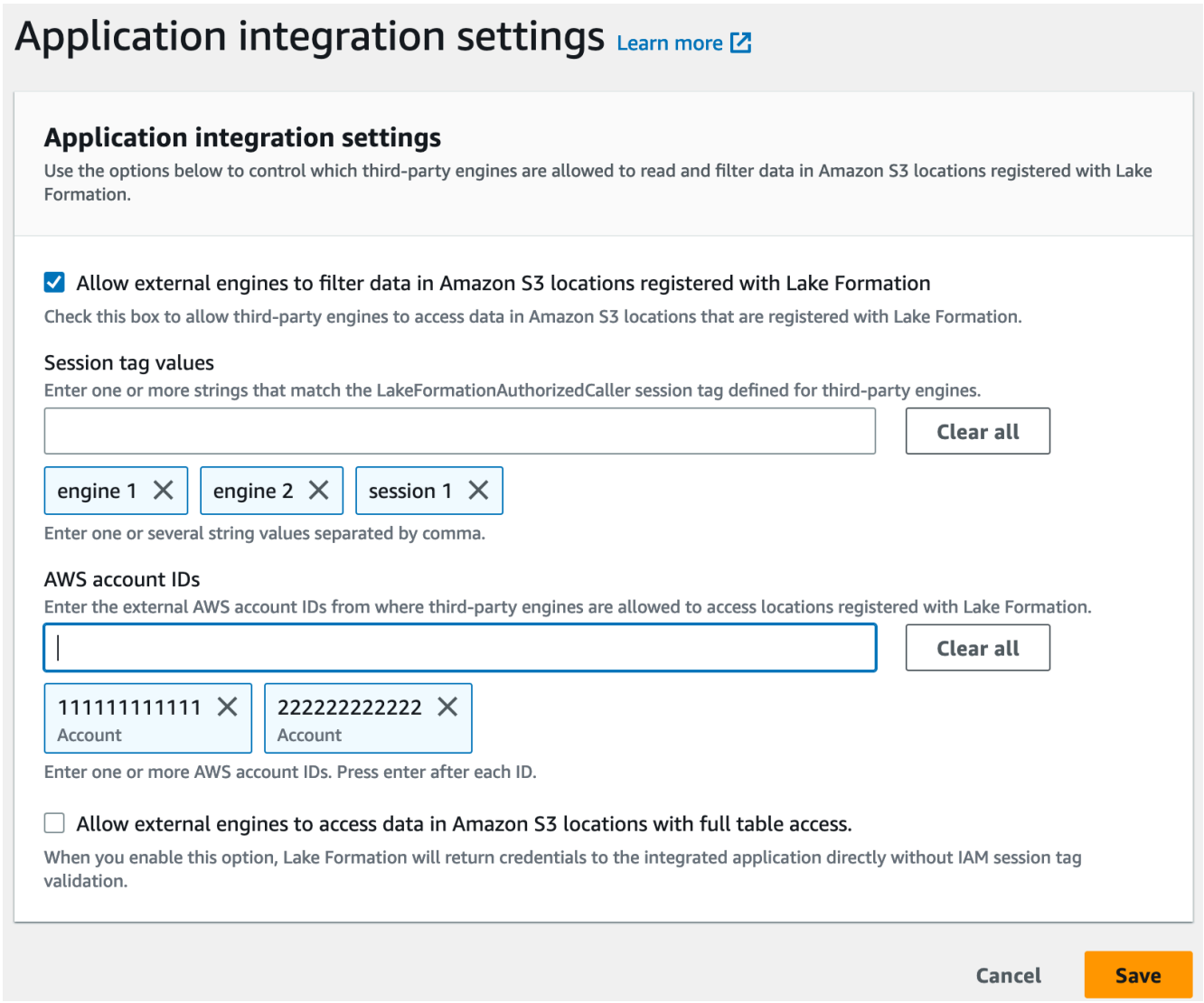
Siga estos pasos para permitir que un motor de consultas de terceros llame a las operaciones de la API de integración de aplicaciones a través de la consola AWS Lake Formation, la AWS CLI, la API o el SDK.

Console

Para registrar tu cuenta para el filtrado de datos externo:

1. Inicie sesión en la AWS Management Console y abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación de la izquierda, expanda los Permisos y, a continuación, seleccione Configuración de integración de aplicaciones.
3. En la página Configuración de integración de aplicaciones, elija la opción Permitir que motores externos filtren datos en las ubicaciones de Amazon S3 registradas en Lake Formation.
4. Introduzca las etiquetas de sesión que creó para el motor de terceros. Para obtener más información, consulte Transferencia de etiquetas de sesión en AWS STS en la Guía del usuario de AWS Identity and Access Management.
5. Introduzca los ID de cuenta de los usuarios que pueden utilizar el motor de terceros para acceder a la información de metadatos sin filtrar y a las credenciales de acceso a los datos de los recursos de la cuenta actual.

   También puede usar el campo de ID de la cuenta de AWS para configurar el acceso entre cuentas.



CLI

Utilice el siguiente comando de CLI `put-data-lake-settings` para establecer los parámetros siguientes.

Al utilizar este comando AWS CLI, se deben configurar tres campos:

- `allow-external-data-filtering` – (booleano) Indica que un motor de terceros puede acceder a la información de metadatos y a las credenciales de acceso a los datos sin filtrar de los recursos de la cuenta actual.
- `external-data-filtering-allow-list` – (matriz) Una lista de los ID de cuenta que pueden acceder a la información de metadatos sin filtrar y a las credenciales de acceso a los datos de los recursos de la cuenta actual cuando se utiliza un motor de terceros.

- `authorized-sessions-tag-value-list` – (matriz) Una lista de valores de etiquetas de sesión autorizados (cadenas). Si se ha adjuntado una credencial de rol de IAM a un par clave-valor autorizado, en caso de que la etiqueta de sesión figure en la lista, la sesión tendrá acceso a la información de metadatos y a las credenciales de acceso a los datos sin filtrar de los recursos de la cuenta configurada. La clave de etiqueta de sesión autorizada se define como `*LakeFormationAuthorizedCaller*`.
- `AllowFullTableExternalDataAccess` - (booleano) Si se debe permitir que un motor de consultas de terceros obtenga credenciales de acceso a los datos sin etiquetas de sesión cuando la persona que llama tiene todos los permisos de acceso a los datos.

Por ejemplo:

```
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

Para ello, use la operación de la API `PutDataLakeSetting` para establecer los parámetros siguientes.

Al utilizar esta operación de la API, se deben configurar tres campos:

- `AllowExternalDataFiltering` – (booleano) Indica SI un motor de terceros puede acceder a la información de metadatos y a las credenciales de acceso a los datos sin filtrar de los recursos de la cuenta actual.
- `ExternalDataFilteringAllowList` – (matriz) Una lista de los ID de cuenta que pueden acceder a la información de metadatos sin filtrar y a las credenciales de acceso a los datos de los recursos de la cuenta actual al utilizar un motor de terceros.
- `AuthorizedSectionsTagValueList` – (matriz) Una lista de valores de etiquetas autorizadas (cadenas). Si se ha adjuntado una credencial de rol de IAM a una etiqueta autorizada, la sesión recibirá acceso a la información de metadatos y a las credenciales de acceso a los datos sin filtrar de los recursos de la cuenta configurada. La clave de etiqueta de sesión autorizada se define como `*LakeFormationAuthorizedCaller*`.
- `AllowFullTableExternalDataAccess` - (booleano) Si se debe permitir que un motor de consultas de terceros obtenga credenciales de acceso a los datos sin etiquetas de sesión cuando la persona que llama tiene todos los permisos de acceso a los datos.

Por ejemplo:

```
//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));
}
```

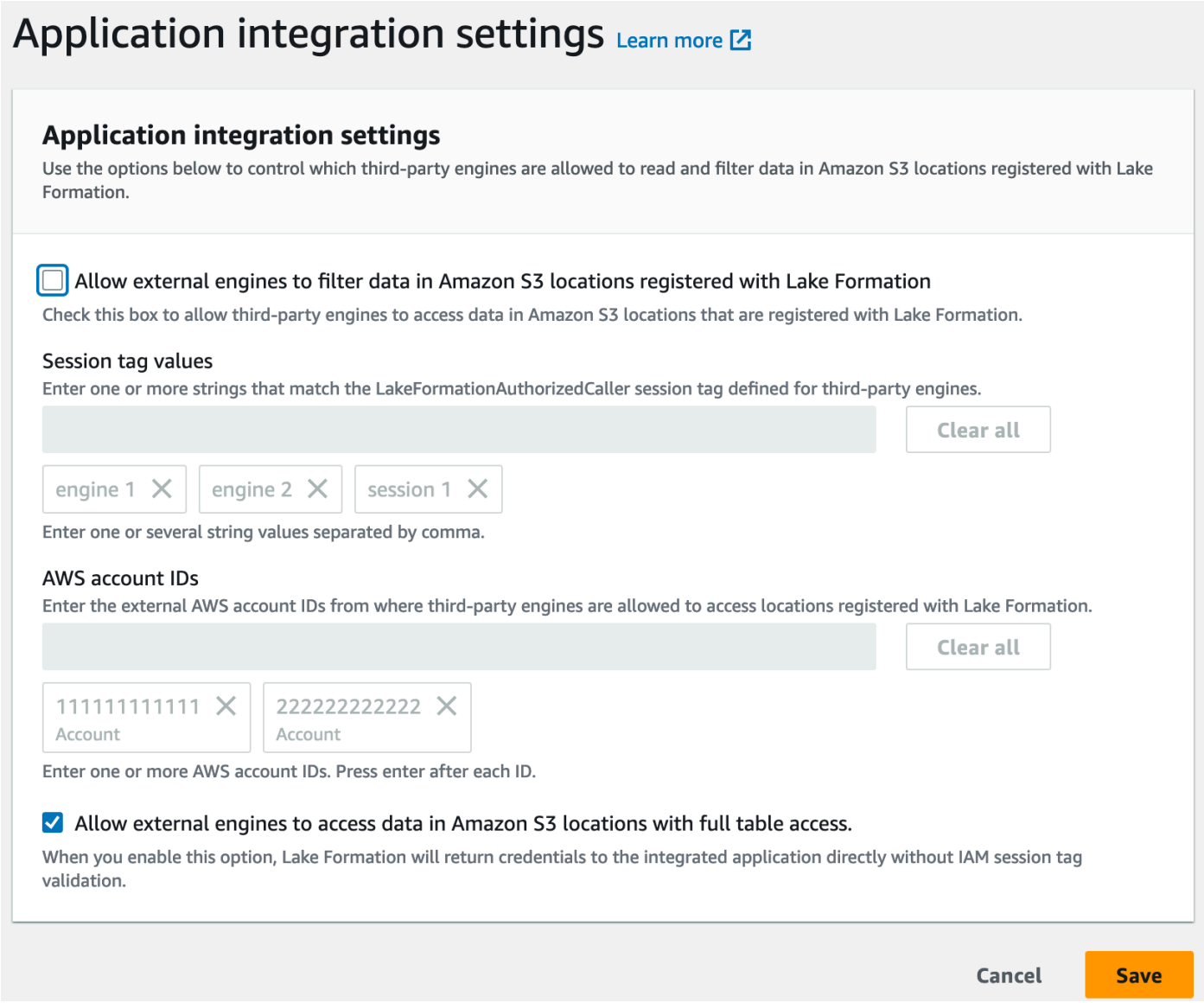
Integración de aplicaciones para un acceso completo a la tabla

Siga estos pasos para permitir que los motores de consultas de terceros accedan a los datos sin la validación de la etiqueta de sesión de IAM:

Console

1. Abra la consola de Lake Formation en https://console.aws.amazon.com/lakeformation/.
2. En el panel de navegación de la izquierda, expanda los Permisos y seleccione Configuración de la integración de aplicaciones.
3. En la página Configuración de integración de aplicaciones, marque la casilla Permitir que motores accedan a datos en las ubicaciones de Amazon S3 con acceso total a las tablas.

   Al activar esta opción, Lake Formation devolverá las credenciales a la aplicación que efectúa la consulta directamente sin validar la etiqueta de sesión de IAM.



AWS CLI

Utilice el comando CLI `put-data-lake-settings` para establecer el parámetro `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
    }
}
```

Trabajar con otros AWS servicios

AWS servicios como Amazon Athena AWS Glue, Amazon Redshift Spectrum y Amazon EMR pueden utilizar Lake Formation para acceder de forma segura a los datos de las ubicaciones de Amazon S3 registradas en Lake Formation. Con Lake Formation, puede definir y administrar permisos de control de acceso específicos (FGAC) para los datos de AWS Glue Data Catalog. Cada uno de estos AWS servicios es una persona de confianza que llama a Lake Formation, y Lake Formation proporciona acceso a los datos almacenados en Amazon S3 mediante credenciales temporales. Para obtener más información, consulte Cómo funciona la integración de aplicaciones de Lake Formation.

Para aprovechar estas prestaciones, Lake Formation requiere que antes registre la ubicación de Amazon S3 y asigne los permisos adecuados a la entidad principal de IAM para acceder a la tabla, la base de datos y la ubicación de Amazon S3. Para obtener más información, consulte Administrar los permisos de Lake Formation.

Temas

Uso AWS Lake Formation con Amazon Athena

Amazon Athena es un servicio de consultas sin servidor que le ayuda a analizar datos estructurados, semiestructurados y no estructurados almacenados en Amazon S3. Athena admite consultas de datos con formatos CSV, JSON, Parquet y Avro. Athena también admite formatos de tabla como las gobernadas por Apache Hive, Apache Hudi, Apache Iceberg y Lake Formation. Athena se integra con el AWS Glue Data Catalog para almacenar los metadatos de sus conjuntos de datos en Amazon S3. Athena puede usar Lake Formation para definir y mantener políticas de control de acceso en esos conjuntos de datos.

Estos son algunos casos de uso comunes en los que puede usar Lake Formation con Athena.

- Utilice los permisos de Lake Formation para acceder a los recursos del Catálogo de datos (bases de datos y tablas) desde Athena. Puede usar el método de recurso designado o las etiquetas LF para definir los permisos en la base de datos y las tablas. Para obtener más información, consulte:
  - Concesión de permisos de base de datos mediante el método de recurso con nombre
  - Control de acceso basado en etiquetas de Lake Formation

  > **Note**
  >
  > Los permisos de Lake Formation se aplican solo al utilizar Athena para consultar datos de origen de Amazon S3 y metadatos en el Catálogo de datos.

  Los permisos de Lake Formation admiten operaciones de lectura y escritura en bases de datos y tablas.

  > **Note**
  >
  > No puede aplicar filtros de datos cuando usa etiquetas LF para administrar los permisos en los recursos del Catálogo de datos.

- Controle los resultados de las consultas con Filtros de datos en Lake Formation para asegurar las tablas de sus lagos de datos Amazon S3 mediante permisos de columna, fila y celda. Consulte la limitación de la proyección de particiones en la Guía del usuario de Amazon Athena.
- Aplique un control de acceso detallado a los datos disponibles para el usuario de Athena basado en SAML al ejecutar consultas federadas.

  Los controladores JDBC y ODBC de Athena admiten la configuración del acceso federado al origen de datos mediante un proveedor de identidades (IdP) basado en SAML. Utilice Amazon QuickSight integrado con Lake Formation con su rol de IAM actual o con usuarios o grupos de SAML para visualizar los resultados de las consultas de Athena.

  > **Note**
  >
  > Los permisos de Lake Formation para usuarios y grupos de SAML solo se reconocen cuando se utiliza el controlador JDBC u ODBC para enviar consultas a Athena.

Para obtener más información, consulte Uso de Lake Formation y de los controladores JDBC y ODBC de Athena para el acceso federado a Athena.

> **Note**
>
> Actualmente, no se admite la autorización del acceso a las identidades de SAML en Lake Formation en las regiones siguientes:
>
> - Medio Oriente (Baréin): me-south-1
> - Asia-Pacífico (Hong Kong): ap-east-1
> - África (Ciudad del Cabo): af-south-1
> - China (Ningxia): cn-northwest-1
> - Asia-Pacífico (Osaka): ap-northeast-3

- Utilice Compartir datos entre cuentas en Lake Formation para consultar tablas en otra cuenta.

> **Note**
>
> Para obtener más información sobre las limitaciones al utilizar los permisos de Lake Formation para `Views`, consulte Consideraciones y limitaciones.

Compatibilidad con formatos de tablas transaccionales

Con los permisos de Lake Formation puede proteger sus datos transaccionales en sus lagos de datos basados en Amazon S3. La tabla siguiente muestra los formatos de tablas transaccionales compatibles con los permisos de Athena y Lake Formation. Lake Formation impone estos permisos cuando los usuarios de Athena ejecutan sus consultas.

Formato de tabla	Descripción y operaciones permitidas	Permisos de formación de lagos admitidos en Athena
Apache Hudi	Formato utilizado para simplificar el procesamiento incremental de datos y el	Utilice Filtrado de datos y seguridad de celda en Lake Formation para proteger la tabla de Hudi mediante

Formato de tabla	Descripción y operaciones permitidas	Permisos de formación de lagos admitidos en Athena
	desarrollo de canalizaciones de datos. Athena admite operaciones de creación y lectura mediante formatos de tablas Apache Hudi en conjuntos de datos de Amazon S3 para los tipos de tablas Hudi Copy on Write (CoW) y Merge On Read (MoR). Athena no admite operaciones de escritura en tablas Hudi. Utilice Athena para consultar conjuntos de datos de Hudi.	permisos de tabla, columna, fila y celda.
Apache Iceberg	Iceberg administra grandes colecciones de archivos como tablas y admite operaciones de lago de datos analíticos modernos, como las consultas de inserción, actualización y eliminación de registros, y viajes en el tiempo. Para obtener más informaci ón sobre la compatibilidad de Athena con las tablas Iceberg, consulte Uso de tablas Iceberg.	Son compatibles los permisos de tabla, columna, fila y celda. Actualmente, Lake Formation no admite la administración de permisos en operaciones de escritura como `VACUUM`, `MERGE`, `UPDATE` y `OPTIMIZE` en tablas en formatos de tabla abierta.

Formato de tabla	Descripción y operaciones permitidas	Permisos de formación de lagos admitidos en Athena
Linux Foundation Delta Lake	Delta Lake es un proyecto de código abierto que ayuda a implementar arquitecturas de lago de datos modernas basadas habitualmente en Amazon S3 o en Sistema de archivos distribuido de Hadoop (HDFS) Athena admite tablas de Delta Lake creadas mediante una definición de tabla de manifiesto basada en enlaces simbólicos a partir de una tabla AWS Glue Data Catalog de Delta Lake. Para obtener más información, consulte Rastrear las tablas de Delta Lake con rastreadores. AWS Glue Athena (versión 3 del motor) admite la lectura de tablas nativas de Delta Lake. Para obtener más información, consulte Presentamos el soporte nativo de Delta Lake para tablas con AWS Glue rastreadores.	Los permisos de tabla, columna, fila y celda son compatibles con las tablas de enlaces simbólicos y las tablas nativas de Delta Lake.

Recursos adicionales de

Publicaciones de blog, vídeos y talleres

- Consultar un conjunto de datos de Apache Hudi en un lago de datos de Amazon S3 con Amazon Athena
- Cree un lago de datos de Apache Iceberg con Amazon Athena, Amazon EMR y AWS Glue
- Insertar, actualizar y eliminar en Amazon S3 con Athena y Apache Iceberg
- Control de acceso basado en etiquetas de LF Taller de Lake Formation sobre las consultas en un lago de datos.

Uso AWS Lake Formation con Amazon Redshift Spectrum

Con Amazon Redshift Spectrum puede consultar y recuperar datos en lagos de datos de Amazon S3 sin tener que cargar datos en nodos de clústeres de Amazon Redshift.

Redshift Spectrum admite dos formas de registrar un catálogo de AWS Glue datos externo habilitado con Lake Formation.

- Uso de un rol de IAM adjunto a un clúster que tenga permiso para el Catálogo de datos

  Para crear un rol de IAM, siga los pasos descritos en el siguiente procedimiento.

  Para crear un rol de IAM para Amazon Redshift mediante AWS Glue Data Catalog una función habilitada para AWS Lake Formation
- Uso de una identidad de IAM federada configurada para gestionar el acceso a recursos AWS Glue Data Catalog externos

  Redshift Spectrum admite la consulta de tablas de Lake Formation mediante identidades de IAM federadas. Las identidades de IAM pueden ser un usuario de IAM o un rol de IAM. Para obtener más información sobre la federación de identidades IAM en Redshift Spectrum, consulte Uso de una identidad federada para administrar el acceso de Amazon Redshift a los recursos locales y a las tablas externas de Redshift Spectrum.

Con la integración de Lake Formation y Redshift Spectrum, puede definir los permisos de control de acceso de fila, columna y celda en las tablas después de registrar sus datos en Lake Formation.

Para obtener más información, consulte Uso de Redshift Spectrum con. AWS Lake Formation

Redshift Spectrum admite lecturas o consultas de `SELECT` en las tablas de esquemas externos administradas por Lake Formation.

Para obtener más información, consulte Creación de esquemas externos para Redshift Spectrum.

Compatibilidad con tipos de tablas transaccionales

La tabla siguiente muestra los formatos de tablas transaccionales compatibles con los permisos de Athena y Lake Formation.

Formatos de tabla compatibles

Formato de tabla	Descripción y operaciones permitidas	Permisos de Lake Formation compatibles con Redshift Spectrum
Apache Hudi	Formato utilizado para simplificar el procesamiento incremental de datos y el desarrollo de canalizaciones de datos. Redshift Spectrum admite operaciones de inserción , eliminación y escritura alterada mediante el formato de tabla Apache Hudi Copy on Write (CoW) en Amazon S3. Para obtener más información, consulte Creación de tablas externas para datos administrados en Apache Hudi.	Utilice Filtrado de datos y seguridad de celda en Lake Formation para proteger las tablas de Hudi mediante permisos de tabla, columna, fila y celda.
Apache Iceberg	Iceberg administra grandes colecciones de archivos como tablas y admite operaciones de lago de datos analíticos modernos, como las consultas	Redshift Spectrum admite tablas Apache Iceberg para efectuar consultas.

Formato de tabla	Descripción y operaciones permitidas	Permisos de Lake Formation compatibles con Redshift Spectrum
	de inserción, actualización y eliminación de registros, y viajes en el tiempo. Para obtener más informaci ón, consulte Uso de tablas de Apache Iceberg con Amazon Redshift.	
Linux Foundation Delta Lake	Delta Lake es un proyecto de código abierto que ayuda a implementar arquitecturas de lago de datos modernos comúnmente construidas sobre Amazon S3 o Sistema de archivos distribuido de Hadoop (HDFS). Redshift Spectrum admite la consulta de tablas Delta Lake. Para obtener más información, consulte Creación de tablas externas para datos administr ados en Delta Lake	Son compatibles los permisos de tabla, columna, fila y celda.

Recursos adicionales de

Publicaciones de blog, vídeos y talleres

- Centralice la gobernanza de su lago de datos utilizando, al AWS Lake Formation mismo tiempo, una arquitectura de datos moderna con Amazon Redshift Spectrum
- Utilice Redshift Spectrum para consultar tablas Apache HUDI Copy On Write (CoW) en el lago de datos de Amazon S3

Utilizándolo con AWS Lake FormationAWS Glue

Los ingenieros y DevOps profesionales de datos utilizan AWS Glue Extract, Transform and Load (ETL) con Apache Spark para realizar transformaciones en sus conjuntos de datos en Amazon S3 y cargar los datos transformados en lagos de datos y almacenes de datos para fines de análisis, aprendizaje automático y desarrollo de aplicaciones. Dado que diferentes equipos acceden al mismo conjunto de datos en Amazon S3, es imprescindible conceder y restringir los permisos en función de sus roles.

AWS Lake Formation se basa en AWS Glueél y los servicios interactúan de las siguientes maneras:

- Lake Formation y AWS Glue comparten el mismo Catálogo de datos.
- Las siguientes características de la consola de Lake Formation invocan la consola de AWS Glue:
  - Trabajos: para obtener más información, consulte Agregar trabajos en la Guía para desarrolladores de AWS Glue .
  - Rastreadores: para obtener más información, consulte la sección Catalogación de tablas con un rastreador en la Guía para desarrolladores de AWS Glue .
- Los flujos de trabajo que se generan cuando se utiliza un esquema AWS Gluede Lake Formation son flujos de trabajo de . Puede ver y gestionar estos flujos de trabajo tanto en la consola de Lake Formation como en la consola de AWS Glue.
- Las transformaciones de machine learning se proporcionan con Lake Formation y se basan en las operaciones de la API de AWS Glue. Puede crear y administrar transformaciones de machine learning en la consola de AWS Glue. Para obtener más información, consulte Transformaciones de machine learning en la Guía para desarrolladores de AWS Glue .

Puede utilizar el control de acceso detallado de Lake Formation para gestionar los recursos del Catálogo de datos existentes y las ubicaciones de datos de Amazon S3.

> **Note**
>
> AWS Glue ETL requiere acceso total a toda la tabla mientras recupera los datos de la ubicación subyacente de Amazon S3. AWS Glue El trabajo de ETL falla si se aplican permisos a nivel de columna a una tabla. Sin embargo, puede crear seguridad a nivel de columna y de fila mediante filtros de datos. Para obtener más información, consulte Notas y restricciones para el filtrado de nivel de columna Lake Formation evalúa el filtro de datos

definido en la tabla y recupera solo los datos filtrados de Amazon S3 necesarios para el trabajo de AWS Glue ETL.

Compatibilidad con tipos de tablas transaccionales

Con los permisos de Lake Formation puede proteger sus datos transaccionales en sus lagos de datos basados en Amazon S3. La siguiente tabla muestra los formatos de tablas transaccionales admitidos AWS Glue y los permisos de Lake Formation. Lake Formation hace cumplir estos permisos para AWS Glue las operaciones.

Formatos de tabla compatibles

Formato de tabla	Descripción y operaciones permitidas	Los permisos de Lake Formation son compatibles en AWS Glue
Apache Hudi	Formato de tabla abierta para simplificar el procesamiento incremental de datos y el desarrollo de canalizaciones de datos. Para ver ejemplos, consulte Uso del marco Hudi en AWS Glue.	Los permisos a nivel de tabla están disponibles para las tablas Hudi. Para obtener más informaci ón, consulte la sección sobre límites.
Apache Iceberg	Formato de tabla abierta que gestiona grandes colecciones de archivos como tablas. Para ver ejemplos, consulte Uso del marco Iceberg en. AWS Glue	Los permisos a nivel de tabla están disponibles para las tablas Iceberg. Para obtener más informaci ón, consulte la sección sobre límites.
Linux Foundation Delta Lake	Delta Lake es un proyecto de código abierto que ayuda a implementar arquitecturas de lago de datos modernos	Los permisos a nivel de tabla están disponibles para las tablas Delta Lake.

Formato de tabla	Descripción y operaciones permitidas	Los permisos de Lake Formation son compatibles en AWS Glue
	comúnmente construid as sobre Amazon S3 o el Sistema de archivos distribui do de Hadoop (HDFS). Para ver ejemplos, consulte Uso del marco Delta Lake en AWS Glue.	Para obtener más informaci ón, consulte la sección sobre límites.

Recursos adicionales de

Publicaciones de blog y repositorios

- Utilice el AWS Glue conector para leer y escribir tablas de Apache Iceberg con transacciones ACID y realizar viajes en el tiempo
- Escritura en tablas de Apache Hudi mediante AWS Glue un conector personalizado
- AWS repositorio de plantillas de Cloudformation y ejemplos de código de pyspark para analizar los datos de streaming mediante AWS Glue Apache Hudi y Amazon S3.

Uso AWS Lake Formation con Amazon EMR

Amazon EMR es una plataforma flexible de clústeres AWS gestionados en la que puede ejecutar cualquier código personalizado en marcos de big data compatibles, como Hadoop Map-Reduce, Spark, Hive, Presto, etc. Las organizaciones también utilizan Amazon EMR para ejecutar aplicaciones de procesamiento de datos por lotes y en streaming en un clúster altamente distribuido. Con Amazon EMR, puede ejecutar sus transformaciones de datos y código personalizado en bases de datos y tablas cuyos permisos administra Lake Formation.

Hay tres opciones para implementar Amazon EMR:

- EMR en EC2
- EMR sin servidor
- Amazon EMR en EKS

Para obtener más información, consulte Integrar Amazon EMR con Lake Formation o Uso de EMR Serverless with para un control de acceso detallado AWS Lake Formation

Compatibilidad con formatos de tablas transaccionales

Las versiones 6.15.0 y posteriores de Amazon EMR incluyen compatibilidad con los permisos de control de acceso a nivel de tabla, fila, columna y celda de Lake Formation en los formatos de tabla Apache Hudi, Apache Iceberg y Delta Lake al leer y escribir datos con Spark SQL.

Formatos de tabla compatibles

Formato de tabla	Descripción y operaciones permitidas	Permisos de Lake Formation admitidos en Amazon EMR
Apache Hudi	Formato de tabla abierta para simplificar el procesamiento incremental de datos y el desarrollo de canalizaciones de datos. Para obtener una lista de las operaciones compatibles, consulte Apache Hudi y Lake Formation.	Amazon EMR es compatibl e con el control de acceso a nivel de tabla, fila, columna y celda con Apache Hudi.
Apache Iceberg	Formato de tabla abierta que gestiona grandes colecciones de archivos como tablas. Para obtener una lista de las operaciones compatibles, consulte Apache Iceberg y Lake Formation.	Amazon EMR es compatibl e con el control de acceso a nivel de tabla, fila, columna y celda con Apache Iceberg.
Linux Foundation Delta Lake	Delta Lake es un proyecto de código abierto que ayuda a implementar arquitecturas de lago de datos modernos comúnmente construid as sobre Amazon S3 o el	Amazon EMR admite el control de acceso a nivel de tabla, fila, columna y celda con las tablas de Delta Lake.

Formato de tabla	Descripción y operaciones permitidas	Permisos de Lake Formation admitidos en Amazon EMR
	Sistema de archivos distribuido de Hadoop (HDFS). Para obtener una lista de las operaciones compatibles, consulte Delta Lake y Lake Formation.	

Recursos adicionales de

Guía del usuario, publicaciones de blog y talleres

- Integración con Amazon EMR mediante roles de tiempo de ejecución
- Introducción rápida a Apache Hudi, Apache Iceberg y Delta Lake con Amazon EMR en EKS
- Uso de OSS de Delta Lake con EMR sin servidor

Uso AWS Lake Formation con Amazon QuickSight

Amazon QuickSight admite la exploración de conjuntos de datos gestionados por los permisos de Lake Formation en Amazon S3 mediante Athena.

Tanto los usuarios de las ediciones Standard como Enterprise de Amazon se QuickSight integran con Lake Formation, pero de forma ligeramente diferente.

- Edición empresarial: otorga permisos de control de acceso (FGAC) detallados a QuickSight usuarios individuales de Amazon, grupos y roles de IAM para acceder a bases de datos y tablas.
- Edición Standard: permisos a los roles de IAM para acceder a bases de datos y tablas.

> **Note**
>
> De forma predeterminada, Amazon QuickSight usa un rol denominado`aws-quicksight-service-role-v0`. También puedes definir funciones personalizadas con los permisos necesarios que permitan QuickSight a Amazon acceder a Athena.

Para obtener más información, consulte Autorizar conexiones mediante AWS Lake Formation

Recursos adicionales de

Publicaciones de blog

- Habilite permisos detallados para los autores de Amazon en QuickSight AWS Lake Formation
- Analice sus datos de forma segura con AWS Lake Formation Amazon QuickSight

AWS Lake Formation Utilizándolo con AWS CloudTrail Lake

AWS CloudTrail Lake permite explorar los almacenes de datos de eventos utilizando Amazon Athena permisos detallados en. AWS Lake Formation

> **Note**
>
> CloudTrail Solo se puede consultar a Lake. Amazon Athena

Para registrar su almacén de datos de eventos de CloudTrail Lake en Lake Formation, consulte Federar un almacén de datos de eventos.

Registro de llamadas a la API de AWS Lake Formation mediante AWS CloudTrail

AWS Lake Formation se integra con AWS CloudTrail, un servicio que proporciona un registro de las acciones que efectúa un usuario, rol o servicio de AWS en Lake Formation. CloudTrail captura las llamadas a la API de Lake Formation como eventos. Las llamadas capturadas incluyen las efectuadas desde la consola de Lake Formation, las AWS Command Line Interface, y las llamadas de código a las acciones de la API de Lake Formation. Si crea un registro de seguimiento, puede habilitar la entrega continua de eventos de CloudTrail a un bucket de Amazon S3, incluidos los eventos para Lake Formation. Si no configura un registro de seguimiento, puede ver los eventos más recientes de la consola de CloudTrail en el Historial de eventos. Con la información recopilada por CloudTrail, puede determinar la solicitud que se hizo a Lake Formation, la dirección IP desde la que se hizo, quién y cuándo la hizo, entre otros datos.

Para obtener más información acerca de CloudTrail, consulte la Guía del usuario de AWS CloudTrail.

Información de Lake Formation en CloudTrail

CloudTrail se habilita de forma predeterminada cuando crea una cuenta nueva de AWS. Cuando se produce una actividad en Lake Formation, esta se registra en un evento de CloudTrail junto con los eventos de los demás servicios de AWS en el Historial de eventos. Un evento representa una solicitud específica realizada desde una fuente y contiene información sobre la acción solicitada, la fecha y la hora de la acción y los parámetros de la solicitud. Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad del usuario le ayuda a determinar lo siguiente:

- Si la solicitud se realizó con credenciales de usuario AWS Identity and Access Management (IAM) o credenciales de usuario raíz.
- Si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado.
- Si la solicitud la realizó otro servicio de AWS.

Para obtener más información, consulte el elemento userIdentity de CloudTrail.

Puede ver, buscar y descargar los últimos eventos de su cuenta de AWS. Para obtener más información, consulte Ver eventos con el historial de eventos de CloudTrail.

Para mantener un registro continuo de eventos en la cuenta de AWS, incluidos los eventos de Lake Formation, cree un registro de seguimiento. Un registro de seguimiento permite a CloudTrail enviar archivos de registro a un bucket de Amazon S3. De forma predeterminada, cuando se crea un registro de seguimiento en la consola, este se aplica a todas las regiones de AWS. El registro de seguimiento registra los eventos de todas las regiones de la partición de AWS y envía los archivos de registro al bucket de Amazon S3 especificado. También es posible configurar otros servicios de AWS, como Amazon Athena, para analizar en profundidad y actuar según los datos de eventos recopilados en los registros de CloudTrail. CloudTrail también puede enviar archivos de registro a registros de Amazon CloudWatch y eventos de CloudWatch.

Para obtener más información, consulte lo siguiente:

- Introducción a la creación de registros de seguimiento
- Servicios e integraciones compatibles con CloudTrail
- Configuración de notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de registro de CloudTrail de varias regiones y Recibir archivos de registro de CloudTrail de varias cuentas

Descripción de los eventos de Lake Formation

CloudTrail registra todas las acciones de la API de Lake Formation, que se documentan en la Guía para desarrolladores de AWS Lake Formation. Por ejemplo, las llamadas a las acciones `PutDataLakeSettings`, `GrantPermissions` y `RevokePermissions` generan entradas en los archivos de log de CloudTrail.

En el ejemplo siguiente se muestra una entrada de registro de CloudTrail para la acción `GrantPermissions`. La entrada incluye el usuario que concedió el permiso (`datalake_admin`), la entidad principal a la que se concedió el permiso (`datalake_user1`) y el permiso que se concedió (`CREATE_TABLE`). La entrada también muestra que la concesión ha fallado porque la base de datos de destino no estaba especificada en el argumento de `resource`.


```
{
    "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"
}
```

En el ejemplo siguiente se muestra una entrada de registro de CloudTrail para la acción `GetDataAccess`. Las entidades principales no llaman directamente a esta API. En vez de eso, se registra `GetDataAccess` cuando una entidad principal o servicio de AWS integrado solicita credenciales temporales para acceder a los datos en una ubicación de lago de datos registrada con 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"
    },
...
}
```

Véase también

- Registro multicuenta CloudTrail

Prácticas recomendadas, consideraciones y limitaciones de Lake Formation

Utilice esta sección para encontrar rápidamente las prácticas recomendadas, las consideraciones y las limitaciones de AWS Lake Formation.

Consulte Service Quotas para conocer el número máximo de recursos u operaciones de servicio de su Cuenta de AWS.

Temas

Prácticas recomendadas y consideraciones para uso compartido de datos entre cuentas

Las capacidades multicuenta de Lake Formation permiten a los usuarios compartir de forma segura lagos de datos distribuidos entre varias AWS organizaciones o directamente con los directores de IAM en otra cuenta Cuentas de AWS, lo que proporciona un acceso detallado a los metadatos del catálogo de datos y a los datos subyacentes.

Tenga en cuenta las siguientes prácticas recomendadas en el uso compartido de datos entre cuentas de Lake Formation:

- No hay límite en cuanto a la cantidad de permisos de Lake Formation que puede conceder a los directores por su propia AWS cuenta. Sin embargo, Lake Formation usa la capacidad AWS Resource Access Manager (AWS RAM) para las concesiones entre cuentas que su cuenta puede realizar con el método de recurso indicado. Para maximizar la AWS RAM capacidad, sigue estas prácticas recomendadas para el método de recurso indicado:
  - Usa el nuevo modo de concesión multicuenta (versión 3 y superior en la configuración de la versión multicuenta) para compartir un recurso con una persona externa Cuenta de AWS. Para obtener más información, consulte Actualización de los ajustes de la versión entre cuentas para compartir datos.
  - Organice las AWS cuentas en organizaciones y conceda permisos a las organizaciones o unidades organizativas. Una concesión a una organización o unidad organizativa cuenta como una concesión.

    La concesión a organizaciones o unidades organizativas también elimina la necesidad de aceptar una AWS Resource Access Manager (AWS RAM) invitación a compartir recursos para obtener la subvención. Para obtener más información, consulte Acceso y visualización de tablas y bases de datos compartidas del catálogo de datos.
  - En lugar de conceder permisos en muchas tablas individuales de una base de datos, utilice el comodín especial Todas las tablas para conceder permisos en todas las tablas de la base de datos. La concesión en Todas las tablas se considera una única concesión. Para obtener más información, consulte Concesión y revocación de permisos sobre los recursos del catálogo de datos.

  > **Note**
  >
  > Para obtener más información sobre cómo solicitar un límite superior para el número de recursos compartidos AWS RAM, consulte las cuotas de AWS servicio en el Referencia general de AWS.

- Debe crear un enlace de recursos a una base de datos compartida para que dicha base de datos aparezca en los editores de consultas Amazon Athena y en Amazon Redshift Spectrum. Del mismo modo, para poder consultar tablas compartidas con Athena y Redshift Spectrum, debe crear enlaces de recursos a las tablas. Los enlaces de recursos aparecen entonces en la lista de tablas de los editores de consultas.

  En lugar de crear enlaces de recursos para hace consultas en muchas tablas por separado, puede utilizar el comodín Todas las tablas para conceder permisos en todas las tablas de una base de

datos. A continuación, cuando cree un enlace de recursos para esa base de datos y lo seleccione en el editor de consultas, tendrá acceso a todas las tablas de esa base de datos para su consulta. Para obtener más información, consulte Creación de enlaces de recursos.

- Cuando comparte recursos directamente con entidades principales de otra cuenta, es posible que la entidad principal de IAM de la cuenta de destinatario no tenga permiso para crear enlaces de recursos para poder consultar las tablas compartidas mediante Athena y Amazon Redshift Spectrum. En lugar de crear un enlace de recursos para cada tabla compartida, el administrador del lago de datos puede crear una base de datos de marcadores de posición y conceder permisos de `CREATE_TABLE` al grupo `ALLIAMPrincipal`. A continuación, todas las entidades principales de IAM de la cuenta de destinatario pueden crear enlaces de recursos en la base de datos del marcador de posición y empezar a consultar las tablas compartidas.

  Consulte el comando CLI de ejemplo para conceder permisos a `ALLIAMPrincipals` en Concesión de permisos de base de datos mediante el método de recurso con nombre.
- Athena y Redshift Spectrum admiten el control de acceso a nivel de columna, pero solo para la inclusión, no para la exclusión. Los trabajos de ETL de AWS Glue no admiten el control de acceso a nivel de columna.
- Cuando se comparte un recurso con su AWS cuenta, solo puede conceder permisos sobre el recurso a los usuarios de su cuenta. No puedes conceder permisos sobre el recurso a otras AWS cuentas, a organizaciones (ni siquiera a la tuya) ni al `IAMAllowedPrincipals` grupo.
- No puede conceder `DROP` ni `Super` sobre una base de datos a una cuenta externa.
- Revoque los permisos entre cuentas antes de eliminar una base de datos o una tabla. De lo contrario, debe eliminar los recursos compartidos huérfanos en AWS Resource Access Manager.

Véase también

- Prácticas recomendadas y consideraciones sobre el control de acceso basado en etiquetas de Lake Formation
- `CREATE_TABLE` en Referencia de permisos de Lake Formation para más normas y limitaciones de acceso entre cuentas.

Limitaciones de acceso a datos entre regiones

Lake Formation admite las consultas a tablas de catálogo de datos entre Regiones de AWS. Puede acceder a los datos de una región desde otras regiones mediante Amazon Athena Amazon EMR y AWS Glue ETL creando enlaces de recursos en otras regiones que apunten a las bases de datos y tablas de origen. Con el acceso a las tablas entre regiones, puede acceder a los datos de todas las regiones sin copiar los datos subyacentes o los metadatos en el Catálogo de datos.

Se aplican las siguientes limitaciones al acceso a las tablas entre regiones.

- Lake Formation no admite la consulta de tablas del Catálogo de datos de otra región mediante Amazon Redshift Spectrum.
- En la consola de Lake Formation, las vistas de bases de datos y tablas no muestran los nombres de las bases de datos o tablas de la región de origen.
- Para ver la lista de tablas de una base de datos compartida de otra región, primero debe crear un enlace de recursos a la base de datos compartida, luego seleccionar el enlace de recursos y elegir Ver tablas.
- La función de acceso a las tablas entre regiones no funciona cuando se crean enlaces de recursos Regiones de AWS que apuntan a bases de datos y tablas compartidas creadas de forma opcional en Regions.

  Para obtener más información, consulta la sección Optar por regiones en la página de servicios Regiones de AWS y servicios compatibles.
- Lake Formation no admite llamadas de enlaces de recursos entre regiones realizadas por usuarios de SAML.

Vistas, consideraciones y limitaciones del catálogo de datos

En AWS Glue Data Catalog, una vista es una tabla virtual en la que el contenido se define mediante una consulta que hace referencia a una o más tablas. Puede crear una vista que haga referencia a un máximo de 10 tablas mediante editores de SQL para Amazon Athena o Amazon Redshift. Las tablas de referencia subyacentes de una vista pueden pertenecer a la misma base de datos o a bases de datos distintas dentro de la misma Cuenta de AWS.

Las siguientes consideraciones y limitaciones se aplican a las vistas del catálogo de datos.

- Amazon Redshift siempre crea vistas con columnas varchar a partir de tablas con cadenas. Debe convertir columnas de cadenas en varchar con una longitud explícita al añadir dialectos de otros motores.
- Al conceder permisos de lago de datos a `All views` dentro de una base de datos, el beneficiario tendrá permisos en todas las tablas y vistas de la base de datos.
- No se pueden crear vistas:
  - Que hagan referencia a otras vista.
  - Cuando una tabla de referencia es un enlace de recursos.
  - Cuando las tablas de referencia tienen permisos de entidad principal de `IAM_ALLOWED_GROUP`.
  - Cuando la tabla de referencia está en otra cuenta.
  - De almacenes de metadatos externos de Hive.

Limitaciones de filtrado de datos

Al conceder permisos de Lake Formation en una tabla del Catálogo de datos, puede incluir especificaciones de filtrado de datos para restringir el acceso a determinados datos en los resultados de las consultas y en los motores integrados con Lake Formation. Lake Formation utiliza el filtrado de datos para ofrecer una seguridad a nivel de columna, fila y celda. Puede definir y aplicar filtros de datos en las columnas anidadas si los datos de origen contienen estructuras anidadas.

Tenga en cuenta las siguientes notas y restricciones para el filtrado a nivel de fila y a nivel de celda:

- La seguridad a nivel de celda no se admite en las columnas anidadas.
- Todas las expresiones que se admiten en las columnas de nivel superior también se admiten en las columnas anidadas. Sin embargo, NO se debe hacer referencia a los campos anidados de las columnas de partición al definir expresiones anidadas a nivel de fila.
- La seguridad a nivel de celda está disponible en todas las regiones cuando se utiliza la versión 3 del motor Athena o Amazon Redshift Spectrum. Para otros servicios, la seguridad a nivel de celda solo está disponible en las regiones mencionadas en la Regiones admitidas.
- No se admiten las instrucciones `SELECT INTO`.
- Los tipos de datos `array` y `map` no se admiten en las expresiones de filtro de filas. Solo se admite el tipo `struct`.
- Para ejecutar operaciones de consulta en tablas que utilizan filtros a nivel de fila y de celda, debe utilizar un grupo de trabajo especial llamado `AmazonAthenaLakeFormation`. Si desea obtener

más información sobre grupos de trabajo en Athena, consulte Uso de grupos de trabajo para la ejecución de consultas en la Guía del usuario de Amazon Athena.

- No hay límite en el número de filtros de datos que se pueden definir en una tabla, pero hay un límite de 100 permisos `SELECT` de filtro de datos para una sola entidad principal en una tabla.
- El número máximo de filtros de datos que se pueden incluir en una concesión de una tabla es 10.
- Para aplicar un filtro de datos con una expresión de filtro de filas, debe tener `SELECT` con la opción de conceder en todas las columnas de la tabla. Esta restricción no se aplica a los administradores de cuentas externas cuando la concesión se hizo a la cuenta externa.
- Si una entidad principal es miembro de un grupo y tanto la entidad principal como el grupo tienen concedidos permisos sobre un subconjunto de filas, los permisos efectivos de fila de la entidad principal son la unión de los permisos de la entidad principal y los permisos del grupo.
- Los siguientes nombres de columna están restringidos en una tabla para el filtrado de nivel de fila y de celda:
  - ctid
  - oid
  - xmin
  - cmin
  - xmax
  - cmax
  - tableoid
  - insertxid
  - deletexid
  - importoid
  - redcatuniqueid
- Si aplica la expresión de filtro de todas las filas a una tabla de forma simultánea con otras expresiones de filtro con predicados, la expresión de todas las filas prevalecerá sobre todas las demás expresiones de filtro.
- Cuando se conceden permisos en un subconjunto de filas a una AWS cuenta externa y el administrador del lago de datos de la cuenta externa concede esos permisos a un director de esa cuenta, el predicado de filtro efectivo del principal es la intersección del predicado de la cuenta y cualquier predicado que se haya otorgado directamente al principal.

Por ejemplo, si la cuenta tiene permisos de fila con el predicado `dept='hr'` y a la entidad principal se le concedió por separado permiso para `country='us'`, la entidad principal solo tiene acceso a las filas con `dept='hr'` y `country='us'`.

Para obtener más información sobre el filtrado de celdas, consulte Filtrado de datos y seguridad de celda en Lake Formation.

Consideraciones y limitaciones del modo de acceso híbrido

El modo de acceso híbrido proporciona la flexibilidad de habilitar selectivamente los permisos de Lake Formation para las bases de datos y tablas de su AWS Glue Data Catalog.
Con el modo de acceso híbrido, ahora tiene una ruta incremental que le permite establecer los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes.

Las consideraciones y limitaciones indicadas a continuación se aplican al modo de acceso híbrido.

Limitaciones

- Actualizar el registro de ubicaciones de Amazon S3: no es posible editar los parámetros de una ubicación registrada en Lake Formation mediante un rol vinculado a un servicio.
- Opción de incluir al utilizar etiquetas LF: cuando puede conceder permisos de Lake Formation utilizando etiquetas LF, puede incluir entidades principales para aplicar los permisos de Lake Formation como paso consecutivo eligiendo bases de datos y tablas que tengan etiquetas LF adjuntas.
- Incluir entidades principales: actualmente, solo un rol de administrador del lago de datos puede incluir a las entidades principales en los recursos.
- Incluir todas las tablas de una base de datos: en concesiones entre cuentas, al conceder permisos e incluir todas las tablas de una base de datos, también es necesario incluir la base de datos para que los permisos funcionen.

Consideraciones

- Actualización de la ubicación de Amazon S3 registrada en Lake Formation al modo de acceso híbrido: no recomendamos convertir una ubicación de datos de Amazon S3 que ya esté registrada en Lake Formation a un modo de acceso híbrido, aunque se puede hacer.

- Comportamientos de la API cuando se registra una ubicación de datos en modo de acceso híbrido
  - CreateTable — La ubicación se considera registrada en Lake Formation independientemente del indicador del modo de acceso híbrido y del estado de suscripción. Por lo tanto, el usuario necesita el permiso de ubicación de datos para crear una tabla.
  - CreatePartition/BatchCreatePartitions/UpdatePartitions (cuando la ubicación de la partición se actualiza para que apunte a la ubicación registrada en el modo híbrido): la ubicación de Amazon S3 se considera registrada en Lake Formation, independientemente del indicador del modo de acceso híbrido y del estado de activación. Así, el usuario necesita el permiso de localización de datos para crear o actualizar una base de datos.
  - CreateDatabase/UpdateDatabase (cuando la ubicación de la base de datos se actualiza para que apunte a la ubicación registrada en el modo de acceso híbrido): la ubicación se considera registrada en Lake Formation independientemente del indicador del modo de acceso híbrido y del estado de activación. Así, el usuario necesita el permiso de localización de datos para crear o actualizar una base de datos.
  - UpdateTable (cuando la ubicación de una tabla se actualiza para que apunte a la ubicación registrada en el modo de acceso híbrido): la ubicación se considera registrada en Lake Formation independientemente del indicador del modo de acceso híbrido y del estado de activación. Por lo tanto, el usuario necesita permiso de localización de datos para actualizar la tabla. Si la ubicación de la tabla no está actualizada o apunta a una ubicación que no está registrada en Lake Formation, el usuario no necesitará permiso de ubicación de datos para actualizar la tabla.

Consideraciones y limitaciones del uso compartido de datos del almacén de metadatos de Hive

Con la federación de AWS Glue Data Catalog metadatos (federación de catálogos de datos), puede conectar el catálogo de datos a metaalmacenes externos que almacenan los metadatos de sus datos de Amazon S3 y administrar de forma segura los permisos de acceso a los datos mediante AWS Lake Formation.

Las siguientes consideraciones y limitaciones se aplican a las bases de datos federadas que se crean a partir de bases de datos de Hive:

Consideraciones

- AWS SAM soporte de aplicaciones: usted es responsable de la disponibilidad de los recursos de la aplicación que se AWS SAM despliega (Amazon API Gateway y de la función Lambda). Asegúrese de que la conexión entre el metabastore de Hive AWS Glue Data Catalog y el metabastore de Hive funcione cuando los usuarios ejecuten consultas.
- Requisito de la versión de almacén de metadatos de Hive: puede crear bases de datos federadas con Apache Hive versión 3 o posterior.
- Requisito de base de datos mapeada: cada base de datos de Hive debe estar mapeada a una nueva base de datos en Lake Formation.
- Soporte de federación a nivel de base de datos: puede conectarse al metaalmacén de Hive a nivel de base de datos.
- Permisos en bases de datos federadas: los permisos aplicados a una base de datos federada o a las tablas de una base de datos federada persisten incluso cuando se elimina una tabla de origen o una base de datos. Cuando se vuelve a crear la base de datos o tabla de origen, no es necesario volver a conceder los permisos. Cuando se elimina una tabla federada con permisos de Lake Formation en el origen, los permisos de Lake Formation siguen siendo visibles y puede revocarlos si es necesario.

  Si un usuario elimina una base de datos federada, se pierden todos sus permisos correspondientes. Si se vuelve a crear la misma base de datos con el mismo nombre, no se recuperarán los permisos de Lake Formation. Los usuarios deberán volver a configurar los permisos nuevos.
- Permisos de AllowedPrincipal grupos de IAM en bases de datos federadas: según la`DataLakeSettings`, Lake Formation podría establecer permisos para todas las bases de datos y tablas de un grupo virtual denominado. `IAMAllowedPrincipal` `IAMAllowedPrincipal`Se refiere a todos los directores de IAM que tienen acceso a los recursos del catálogo de datos a través de las políticas principales y las políticas de recursos de IAM. AWS Glue Si estos permisos existen en una base de datos o una tabla, todas las entidades principales tienen acceso a la base de datos o tabla.

  Sin embargo, Lake Formation no permite permisos `IAMAllowedPrincipal` en las tablas de bases de datos federadas. Al crear bases de datos federadas, asegúrese de pasar el parámetro `CreateTableDefaultPermissions` como una lista vacía.

  Para obtener más información, consulte Cambiar la configuración predeterminada de su lago de datos.

- Unir tablas en las consultas: puede unir las tablas del metaalmacén de Hive con las tablas nativas del Catálogo de datos para ejecutar consultas.

Limitaciones

- Limitación en la sincronización de metadatos entre el metaalmacén de Hive AWS Glue Data Catalog y el metabastore de Hive: después de establecer la conexión con el metabastore de Hive, debe crear una base de datos federada para sincronizar los metadatos del metaalmacén de Hive con el. AWS Glue Data Catalog Las tablas de la base de datos federada se sincronizan en tiempo de ejecución cuando los usuarios ejecutan consultas.
- Limitación al crear tablas nuevas en una base de datos federada: no podrá crear tablas nuevas en bases de datos federadas.
- Limitación de permisos de datos: la compatibilidad con los permisos en las vistas de tabla del metaalmacén de Hive no está disponible.

Limitaciones de uso compartido de datos de Amazon Redshift

AWS Lake Formation le permite gestionar de forma segura los datos de un recurso compartido de Amazon Redshift. Amazon Redshift es un servicio de almacenamiento de datos en la nube totalmente gestionado y a escala de petabytes. AWS Al utilizar la capacidad de compartir datos, Amazon Redshift le ayuda a compartir datos entre Cuentas de AWS. Para obtener más información sobre el uso compartido de datos, consulte Información general del uso compartido de datos en Amazon Redshift.

Las notas y restricciones siguientes se aplican a las bases de datos federadas que se crean a partir de recursos compartidos de datos de Amazon Redshift:

- Requisito de base de datos mapeada: cada recurso compartido de datos de Amazon Redshift debe estar mapeado a una nueva base de datos en Lake Formation. Esto es necesario para mantener nombres de tabla únicos cuando la representación de los objetos de recurso compartido de datos está aplanada en la base de datos del Catálogo de datos.
- Limitación al crear tablas nuevas en una base de datos federada: no podrá crear tablas nuevas en bases de datos federadas.
- Permisos en bases de datos federadas: los permisos aplicados a una base de datos federada o a las tablas de una base de datos federada persisten incluso cuando se elimina una tabla de origen o una base de datos. Cuando se vuelve a crear la base de datos o tabla de origen, no es

necesario volver a conceder los permisos. Cuando se elimina una tabla federada con permisos de Lake Formation en el origen, los permisos de Lake Formation siguen siendo visibles y puede revocarlos si es necesario.

Si un usuario elimina una base de datos federada, se pierden todos sus permisos correspondientes. Si se vuelve a crear la misma base de datos con el mismo nombre, no se recuperarán los permisos de Lake Formation. Los usuarios deberán volver a configurar los permisos nuevos.

- Permisos de AllowedPrincipal grupos de IAM en bases de datos federadas: según la`DataLakeSettings`, Lake Formation podría establecer permisos para todas las bases de datos y tablas de un grupo virtual denominado. `IAMAllowedPrincipal` `IAMAllowedPrincipal`Se refiere a todos los directores de IAM que tienen acceso a los recursos del catálogo de datos a través de las políticas principales y las políticas de recursos de IAM. AWS Glue Si estos permisos existen en una base de datos o una tabla, todas las entidades principales tienen acceso a la base de datos o tabla.

  Sin embargo, Lake Formation no permite permisos `IAMAllowedPrincipal` en las tablas de bases de datos federadas. Al crear bases de datos federadas, asegúrese de pasar el parámetro `CreateTableDefaultPermissions` como una lista vacía.

  Para obtener más información, consulte Cambiar la configuración predeterminada de su lago de datos.
- Filtrado de datos: en Lake Formation, puede conceder permisos en una tabla de una base de datos federada con filtrado a nivel de columna y de fila. Sin embargo, no puede combinar el filtrado a nivel de columna y de fila para restringir el acceso con una especificidad a nivel de celda a las tablas de bases de datos federadas.
- Identificador de distinción entre mayúsculas y minúsculas: los objetos de recursos compartidos de datos de Amazon Redshift administrados por Lake Formation solo admitirán nombres de tablas y columnas en minúsculas. No active el identificador de distinción entre mayúsculas y minúsculas en las bases de datos, tablas y columnas de los recurso compartido de datos de Amazon Redshift, si se van a compartir y gestionar mediante Lake Formation.

Para obtener más información acerca de las limitaciones del recurso compartido de datos en Amazon Redshift, consulte Limitaciones del uso compartido de datos en la Guía para desarrolladores de bases de datos Amazon Redshift.

Limitaciones de la integración de IAM Identity Center

Con él AWS IAM Identity Center, puede conectarse a los proveedores de identidad (IdPs) y administrar de forma centralizada el acceso de los usuarios y grupos a todos AWS los servicios de análisis. Puede configurarlo AWS Lake Formation como una aplicación habilitada en el Centro de identidades de IAM, y los administradores del lago de datos pueden conceder permisos detallados sobre los recursos a los usuarios y grupos autorizados. AWS Glue Data Catalog

Se aplican las siguientes limitaciones a la integración de Lake Formation con IAM Identity Center:

- No puede asignar usuarios y grupos de IAM Identity Center como administradores de lagos de datos o administradores de solo lectura en Lake Formation.
- Los usuarios y grupos del IAM Identity Center no pueden consultar las tablas del catálogo de datos que estén cifradas mediante claves (). AWS Key Management Service AWS KMS AWS KMS no admite la propagación de identidades fiables.
- Los usuarios y grupos de IAM Identity Center solo pueden invocar las operaciones de API que figuran en la política `AWSIAMIdentityCenterAllowListForIdentityContext` proporcionada por el IAM Identity Center.

Prácticas recomendadas y consideraciones sobre el control de acceso basado en etiquetas de Lake Formation

Puede crear, mantener y asignar etiquetas LF para controlar el acceso a las bases de datos, tablas y columnas del catálogo de datos.

Tenga en cuenta las siguientes prácticas recomendadas al usar el control de acceso basado en etiquetas de Lake Formation:

- Todas las etiquetas LF deben estar predefinidas antes de poder asignarlas a los recursos del Catálogo de datos o concederse a las entidades principales.

  El administrador del lago de datos puede delegar las tareas de administración de etiquetas generando creadores de etiquetas LF con los permisos de IAM necesarios. Los ingenieros y analistas de datos deciden las características y las relaciones de las etiquetas LF. Luego, los creadores de las etiquetas LF crean y mantienen las etiquetas LF en Lake Formation.
- Puede asignar varias etiquetas LF a los recursos del Catálogo de datos. Solo se puede asignar un valor para una clave en particular a un recurso concreto.

Por ejemplo, puede asignar `module=Orders`, `region=West`, `division=Consumer`, etc. a una base de datos, tabla o columna. No puede asignar `module=Orders,Customers`.

- No puede asignar etiquetas LF a los recursos al crearlos. Solo puede añadir etiquetas LF a los recursos existentes.
- Puede conceder expresiones de etiquetas LF, no solo etiquetas LF individuales, a una entidad principal.

  Una expresión LF-Tag tiene el aspecto siguiente (en pseudocódigo).

  ```
  module=sales AND division=(consumer OR commercial)
  ```

  Una entidad principal a la que se le conceda esta expresión de etiqueta LF solo puede acceder a los recursos del Catálogo de datos (bases de datos, tablas y columnas) que tienen asignado `module=sales` y `division=consumer` o `division=commercial`. Si desea que la entidad principal pueda acceder a los recursos que tienen `module=sales` o `division=commercial`, no incluya ambos en la misma concesión. Haga dos concesiones, una para `module=sales` y otra para `division=commercial`.

  La expresión de etiqueta LF más sencilla consiste en una sola etiqueta LF, como `module=sales`.
- Una entidad principal que reciba permisos sobre una etiqueta LF con varios valores puede acceder a los recursos del Catálogo de datos con cualquiera de esos valores. Por ejemplo, si a un usuario se le concede una etiqueta LF con clave= `module` y valores= `orders,customers`, el usuario tiene acceso a los recursos que están asignados o bien a `module=orders` o `module=customers`.
- Debe tener permiso `Grant with LF-Tag expressions` para conceder permisos de datos sobre los recursos del Catálogo de datos mediante el método LF-TBAC. El administrador del lago de datos y el creador de la etiqueta LF reciben este permiso de forma implícita. La entidad principal que tenga el permiso `Grant with LFTag expressions` puede conceder permisos de datos sobre los recursos mediante:
  - el método de recurso con nombre
  - el método LF-TBAC, pero solo usando la misma expresión de etiqueta LF

    Por ejemplo, supongamos que el administrador del lago de datos hace la siguiente concesión (en pseudocódigo).

```
GRANT (SELECT ON TABLES) ON TAGS module=customers, region=west,south TO user1 WITH
 GRANT OPTION
```

En este caso, el `user1` puede conceder `SELECT` en tablas a otras entidades principales mediante el método LF-TBAC, pero solo con la expresión de etiqueta `module=customers, region=west,south` LF completa.

- Si a una entidad principal recibe permisos sobre un recurso con el método LF-TBAC y con el método de recurso con nombre, los permisos que dicha entidad tiene sobre el recurso son la unión de los permisos concedidos por ambos métodos.
- Lake Formation admite conceder `DESCRIBE` y `ASSOCIATE` en etiquetas LF entre cuentas, y conceder permisos sobre los recursos del Catálogo de datos en todas las cuentas mediante el método LF-TBAC. En ambos casos, el principal es un identificador de AWS cuenta.

> **Note**
>
> Lake Formation ahora es compatible con la concesión de permisos entre cuentas a organizaciones y unidades organizativas utilizando el método LF-TBAC. Para utilizar esta prestación, debe actualizar la configuración de la versión entre cuentas a la Versión 3.

Para obtener más información, consulte Compartir datos entre cuentas en Lake Formation.

- Los recursos del Catálogo de datos creados en una cuenta solo se pueden etiquetar con etiquetas LF creadas en la misma cuenta. Las etiquetas LF creadas en una cuenta no se pueden asociar a los recursos compartidos de otra cuenta.
- El uso del control de acceso basado en etiquetas (LF-TBAC) de Lake Formation para conceder acceso entre cuentas a los recursos del catálogo de datos requiere adiciones a la política de recursos del catálogo de datos de su cuenta. AWS Para obtener más información, consulte Requisitos previos.
- Las claves y los valores de las etiquetas LF no pueden superar los 50 caracteres.
- La cantidad máxima de etiquetas LF que puede asignar a un recurso del Catálogo de datos es 50.
- Los siguientes límites son flexibles:
  - El número máximo de etiquetas LF que se pueden crear es 1000.
  - El número máximo de valores que se pueden definir para una etiqueta LF es 1000.
- Las etiquetas, las claves y los valores se convierten en minúsculas cuando se almacenan.

- Solo se puede asignar un valor para una etiqueta LF a un recurso concreto.
- Si se conceden varias etiquetas LF a una entidad principal con una sola concesión, la entidad principal solo podrá acceder a los recursos del Catálogo de datos que tengan todas las etiquetas LF.
- Los trabajos AWS Glue ETL requieren acceso completo a la tabla. Los trabajos fallarán si el rol AWS Glue ETL no tiene acceso a todas las columnas de una tabla. Es posible aplicar etiquetas LF a nivel de columna, pero esto puede provocar que los roles de AWS Glue ETL pierdan el acceso total a la tabla y que se produzcan errores en los trabajos. El uso de filtros de datos para filtrar columnas o filas no se ve afectado por esta limitación.
- Si la evaluación de una expresión de etiqueta LF da como resultado el acceso solo a un subconjunto de columnas de la tabla, pero el permiso Lake Formation que se concede cuando hay una coincidencia es uno de los permisos que requiere acceso completo a las columnas, es decir `Alter`, `Drop`, `Insert` o `Delete`, entonces no se concede ninguno de esos permisos. En su lugar, solo se concede `Describe`. Si el permiso concedido es `All` (`Super`), solo se conceden `Select` y `Describe`.
- No se utilizan caracteres comodín con las etiquetas LF. Para asignar una etiqueta LF a todas las columnas de una tabla, debe asignar la etiqueta LF a la tabla y todas las columnas de la tabla heredarán la etiqueta LF. Para asignar una etiqueta LF a todas las tablas de una base de datos, debe asignar la etiqueta LF a la base de datos y todas las tablas de la base de datos heredarán esa etiqueta LF.

Formatos compatibles y limitaciones de la compactación de datos administrada

Para mejorar el rendimiento de lectura de los servicios de AWS análisis, como Amazon Athena, Amazon EMR y AWS Glue ETL Jobs, AWS Glue Data Catalog proporciona compactación gestionada (un proceso que compacta objetos pequeños de Amazon S3 en objetos más grandes) para las tablas Iceberg del catálogo de datos.

La compactación de datos admite una variedad de tipos de datos y formatos de compresión para leer y escribir datos, incluida la lectura de datos de tablas cifradas.

La compactación de datos admite:

- Tipos de datos: booleano, entero, largo, flotante, doble, cadena, decimal, fecha, hora, marca de tiempo, cadena, UUID, binario

- Compresión: zstd, gzip, snappy, sin comprimir
- Cifrado: la compactación de datos solo admite el cifrado Amazon S3 (SSE-S3) y el cifrado KMS del lado del servidor (SSE-KMS).
- Compactación de bin pack
- Evolución del esquema
- Tablas con el tamaño de archivo objetivo (escriba. target-file-size-bytes propiedad en configuración iceberg) dentro del rango inclusivo de 128 MB a 512 MB.
- Regiones
  - Asia-Pacífico (Tokio)
  - Asia-Pacífico (Seúl)
  - Asia-Pacífico (Bombay)
  - Europa (Irlanda)
  - Europa (Fráncfort)
  - Este de EE. UU. (Norte de Virginia)
  - Este de EE. UU. (Ohio)
  - Oeste de EE.UU. (Norte de California)
  - América del Sur (São Paulo)
- Puede ejecutar la compactación desde la cuenta en la que reside el catálogo de datos cuando el bucket de Amazon S3 que almacena los datos subyacentes esté en otra cuenta. Para ello, el rol de compactación requiere acceso al bucket de Amazon S3.

La compactación de datos actualmente no admite:

- Tipos de datos: fijos
- Compresión: brotli, lz4
- Compactación de archivos a medida que evoluciona la especificación de la partición.
- Clasificación regular o clasificación en orden Z
- Combinar o eliminar archivos: el proceso de compactación omite los archivos de datos que tienen archivos de eliminación asociados.
- Compactación en tablas con varias cuentas: no se puede ejecutar la compactación en tablas con varias cuentas.

- Compactación en tablas en varias regiones: no se puede ejecutar la compactación en tablas en varias regiones.
- Habilitar la compactación en los enlaces de recursos
- Puntos de conexión para los buckets de Amazon S3

Solución de problemas de Lake Formation

Si le surgen problemas al trabajar con AWS Lake Formation, consulte los temas de esta sección.

Temas

Solución de problemas generales

Utilice esta información para diagnosticar y solucionar diversos problemas de Lake Formation.

Error: permisos insuficientes para Lake Formation en <ubicación de Amazon S3>.

Se ha intentado crear o modificar un recurso del catálogo de datos sin permisos de ubicación de datos en la ubicación de Amazon S3 a la que apunta el recurso.

Si una base de datos o tabla del catálogo de datos apunta a una ubicación de Amazon S3, cuando conceda los permisos Lake Formation `CREATE_TABLE` o `ALTER`, también deberá conceder el permiso `DATA_LOCATION_ACCESS` en la ubicación. Si concede estos permisos a cuentas externas o a organizaciones, debe incluir la opción de concesión.

Una vez concedidos estos permisos a una cuenta externa, el administrador del lago de datos de dicha cuenta deberá conceder los permisos a las entidades principales (usuarios o roles) de la cuenta. Al conceder el permiso que `DATA_LOCATION_ACCESS` se recibió de otra cuenta, debe especificar el ID del catálogo (ID de cuenta AWS) de la cuenta propietaria. La cuenta propietaria es la cuenta que registró la ubicación.

Para más información, consulte Control de acceso a los datos subyacentes y Conceder permisos de ubicación de datos.

Error: "Permisos de clave de cifrado insuficientes para la API Glue"

Se ha intentado conceder permisos de Lake Formation sin permisos AWS Identity and Access Management (IAM) sobre la clave de cifrado AWS KMS para un catálogo de datos cifrado.

Mi consulta Amazon Athena o la de Amazon Redshift que usa manifiestos está fallando

Lake Formation no es compatible con las consultas que utilizan manifiestos.

Error: "Permiso(s) de Lake Formation insuficientes: se requiere crear etiqueta en el catálogo"

El usuario o rol debe ser administrador del lago de datos.

Error al eliminar administradores de lagos de datos no válidos

Debe eliminar todos los administradores de lagos de datos no válidos (funciones de IAM eliminadas que se definen como administradores de lagos de datos) simultáneamente. Si intenta eliminar administradores de lagos de datos no válidos por separado, Lake Formation da un error de entidad principal no válida.

Solución de problemas de acceso entre cuentas

Utilice esta información para diagnosticar y solucionar los problemas de acceso entre cuentas.

Temas

- ConcurrentModificationException en solicitudes de concesión/revocación a cuentas externas
- Error al utilizar Amazon EMR para acceder a datos compartidos entre cuentas

He concedido un permiso entre cuentas Lake Formation pero el destinatario no puede ver el recurso

- ¿Es el usuario de la cuenta de destinatario administrador del lago de datos? Solo los administradores del lago de datos pueden ver el recurso en el momento del uso compartido.
- ¿Está compartiendo con una cuenta externa a su organización utilizando el método de recursos con nombre? Si es así, el administrador del lago de datos de la cuenta receptora debe aceptar una invitación para compartir recursos en AWS Resource Access Manager (AWS RAM).

  Para obtener más información, consulte the section called “Aceptar una invitación para compartir recursos de AWS RAM”.
- ¿Está utilizando políticas de recursos a nivel de cuenta (catálogo de datos) en AWS Glue? En caso afirmativo, si utiliza el método de recursos con nombre, deberá incluir una instrucción especial en la política que autorice a AWS RAM a compartir políticas en su nombre.

  Para obtener más información, consulte the section called “Administración de los permisos entre cuentas mediante AWS Glue y Lake Formation”.
- ¿Dispone de los permisos AWS Identity and Access Management (IAM) necesarios para conceder el acceso entre cuentas?

  Para obtener más información, consulte the section called “Requisitos previos”.
- El recurso sobre el que ha concedido permisos no debe tener ningún permiso de Lake Formation concedido al grupo `IAMAllowedPrincipals`.
- ¿Existe alguna instrucción `deny` sobre el recurso en la política a nivel de cuenta?

Las entidades principales de la cuenta de destinatario pueden ver el recurso del catálogo de datos pero no pueden acceder a los datos subyacentes

Las entidades principales de la cuenta de destinatario deben tener los permisos requeridos AWS Identity and Access Management (IAM). Para más información, consulte Acceso a los datos subyacentes de una tabla compartida.

Error: "Ha fallado la asociación porque el comunicante no estaba autorizado" al aceptar una invitación para compartir un recurso AWS RAM

Tras conceder acceso a un recurso a una cuenta diferente, cuando la cuenta receptora intenta aceptar la invitación a compartir el recurso, la acción falla.

```
$ 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
        }
    ]
}
```

El error se produce porque el `glue:PutResourcePolicy` es invocado por AWS Glue cuando la cuenta receptora acepta la invitación a compartir recursos. Para resolver el problema, permita la acción `glue:PutResourcePolicy` mediante el rol asumido utilizado por la cuenta del productor/concedente.

Error: “No está autorizado a conceder permisos para el recurso”

Se ha intentado conceder permisos entre cuentas en una base de datos o tabla que es propiedad de otra cuenta. Cuando una base de datos o una tabla se comparte con su cuenta, como administrador del lago de datos, puede conceder permisos sobre ella solo a los usuarios de su cuenta.

Error: “Acceso denegado para recuperar la información de la organización AWS“

Su cuenta es una cuenta de administración de organizaciones AWS y no dispone de los permisos necesarios para recuperar información de la organización, como las unidades organizativas de la cuenta.

Para obtener más información, consulte Required permissions for cross-account grants.

Error: "Organización <organization-ID> no encontrada"

Se ha intentado compartir un recurso con una organización, pero el uso compartido con organizaciones no está habilitado. Habilite el uso compartido de recursos con organizaciones.

Para obtener más información, consulte Habilitar el uso compartido con AWS Organizations en la Guía del usuario de AWS RAM.

Error: "Permisos de Lake Formation insuficientes: combinación ilegal"

Un usuario compartió un recurso del catálogo de datos mientras que los permisos de Lake Formation se concedieron al grupo `IAMAllowedPrincipals` para el recurso. El usuario debe revocar todos los permisos de Lake Formation de `IAMAllowedPrincipals` antes de compartir el recurso.

ConcurrentModificationException en solicitudes de concesión/revocación a cuentas externas

Cuando los usuarios realizan múltiples solicitudes concurrentes de concesión y/o revocación de permisos para una entidad principal en las políticas de etiquetas LF, Lake Formation lanza ConcurrentModificationException. Los usuarios deben captar la excepción y volver a intentar la solicitud de concesión/revocación fallida. El uso de las versiones por lotes de las operaciones de la API `GrantPermissions`/`RevokePermissions` - BatchGrantPermissionsy BatchRevokePermissions alivia en cierta medida este problema al reducir el número de solicitudes concurrentes de concesión/revocación.

Error al utilizar Amazon EMR para acceder a datos compartidos entre cuentas

Al usar Amazon EMR para acceder a los datos compartidos desde otra cuenta, algunas bibliotecas de Spark intentarán llamar a la operación de la API `Glue:GetUserDefinedFunctions`. Como

las versiones 1 y 2 de los permisos administrados de AWS RAM no admiten esta acción, obtiene el siguiente mensaje de error:

```
"ERROR: User: arn:aws:sts::012345678901:assumed-role/my-
spark-role/i-06ab8c2b59299508a is not authorized to perform:
glue:GetUserDefinedFunctions on resource: arn:exampleCatalogResource
because no resource-based policy allows the glue:GetUserDefinedFunctions
action"
```

Para resolver este error, el administrador del lago de datos que creó el recurso compartido debe actualizar los permisos administrados de AWS RAM asociados al recurso compartido. La versión 3 de los permisos administrados de AWS RAM permite a las entidades principales llevar a cabo la acción `glue:GetUserDefinedFunctions`.

Si crea un nuevo recurso compartido, Lake Formation aplica la última versión del permiso administrado por AWS RAM de forma predeterminada y no es necesario que haga nada. Para habilitar el acceso a los datos entre cuentas para los recursos compartidos existentes, debe actualizar los permisos administrados de AWS RAM a la versión 3.

Puede ver los permisos de AWS RAM asignados a los recursos compartidos en AWS RAM. Los siguientes permisos se incluyen en la versión 3:

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
  AWSRAMPermissionGlueAllTablesReadWriteForCatalog
  AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

Para actualizar la versión de los permisos administrados de AWS RAM de los recursos compartidos existentes

Como administrador del lago de datos, puede actualizar los permisos administrados de AWS RAM a una versión más reciente. Para ello, siga las instrucciones de la Guía del usuario de AWS RAM

o revoque todos los permisos existentes para el tipo de recurso y vuelva a concederlos. Si revoca los permisos, AWS RAM elimina el recurso compartido de AWS RAM asociado al tipo de recurso. Al volver a conceder los permisos, AWS RAM adjunta la versión más reciente de los permisos administrados de AWS RAM para crear nuevos recursos compartidos.

Solución de problemas de esquemas y flujos de trabajo

Utilice esta información para diagnosticar y solucionar problemas de esquemas y flujos de trabajo.

Temas

Mi esquema falló con "Usuario: <usuario-ARN> no está autorizado para: iam:PassRole en recurso: <rol-ARN>"

Se ha intentado crear un esquema por parte de un usuario que no tiene permisos suficientes para pasar el rol elegido.

Actualice la política de IAM del usuario para poder pasar el rol, o pídale que elija un rol diferente con los permisos necesarios para pasar el rol.

Para obtener más información, consulte the section called “Personas de Lake Formation y referencia de permisos IAM”.

Mi flujo de trabajo ha fallado con " Usuario: <usuario-ARN> no está autorizado para ejecutar: iam:PassRole en el recurso: <rol-ARN>"

El rol que especificó para el flujo de trabajo no tenía una política en línea que permitiera que el rol se pasara a sí mismo.

Para obtener más información, consulte the section called “(Opcional) Cree un rol de IAM para los flujos de trabajo”.

Un rastreador en mi flujo de trabajo falló con "El recurso no existe o el solicitante no está autorizado para acceder a los permisos solicitados"

Una posible causa es que el rol pasado no tenía permisos suficientes para crear una tabla en la base de datos de destino. Concede al rol el permiso `CREATE_TABLE` sobre la base de datos.

Un rastreador en mi flujo de trabajo falló con "Se produjo un error (AccessDeniedException) al llamar a la operación CreateTable..."

Una posible causa es que el rol de flujo de trabajo no tuviera permisos de ubicación de datos en el almacén de destino. Conceda permisos de ubicación de datos al rol.

Para obtener más información, consulte the section called “`DATA_LOCATION_ACCESS`”.

Problemas conocidos de AWS Lake Formation

Revise estos problemas conocidos de AWS Lake Formation.

Temas

- Limitación del filtrado de metadatos de las tablas
- Problema al renombrar una columna excluida
- Problema con la eliminación de columnas en tablas CSV
- Las particiones de tabla deben añadirse bajo una ruta común
- Problema con la creación de una base de datos durante la creación del flujo de trabajo
- Problema al eliminar un usuario y, a continuación, volver a crearlo
- Las API GetTables y SearchTables no actualizan el valor del parámetro IsRegisteredWithLakeFormation
- Las operaciones de la API del catálogo de datos no actualizan el valor del parámetro IsRegisteredWithLakeFormation
- Las operaciones de Lake Formation no admiten el registro de esquemas AWS Glue

Limitación del filtrado de metadatos de las tablas

Los permisos a nivel de columna de AWS Lake Formation pueden utilizarse para restringir el acceso a columnas específicas de una tabla. Cuando un usuario recupera metadatos sobre la tabla utilizando la consola o una API como `glue:GetTable`, la lista de columnas del objeto de tabla contiene solo los campos a los que tiene acceso. Es importante comprender las limitaciones de este filtrado de metadatos.

Aunque Lake Formation pone a disposición de los servicios integrados metadatos sobre los permisos de las columnas, el filtrado real de las columnas en las respuestas a las consultas es responsabilidad del servicio integrado. Los clientes de Lake Formation compatibles con el filtrado a nivel de columna, incluidos Amazon Athena, Amazon Redshift Spectrum y Amazon EMR, filtran los datos en función de los permisos de columna registrados en Lake Formation. Los usuarios no podrán leer datos a los que no deberían tener acceso. Actualmente, AWS Glue ETL no es compatible con el filtrado de columnas.

> **Note**
>
> Los clústeres de EMR no se administran completamente mediante AWS. Por lo tanto, es responsabilidad de los administradores de los EMR asegurar adecuadamente los clústeres para evitar el acceso no autorizado a los datos.

Algunas aplicaciones o formatos pueden almacenar metadatos adicionales, incluidos los nombres y tipos de las columnas, en el mapa de `Parameters` como propiedades de la tabla. Estas propiedades se devuelven sin modificar y son accesibles por cualquier usuario con permiso de `SELECT` sobre cualquier columna.

Por ejemplo, el Avro SerDe almacena una representación JSON del esquema de la tabla en una propiedad de la tabla llamada `avro.schema.literal`, que está disponible para todos los usuarios con acceso a la tabla. Le recomendamos que evite almacenar información sensible en las propiedades de las tablas y que sea consciente de que los usuarios pueden conocer el esquema completo de las tablas con formato Avro. Esta limitación es específica de los metadatos sobre una tabla.

AWS Lake Formation elimina cualquier propiedad de la tabla que empiece por `spark.sql.sources.schema` al responder a una petición `glue:GetTable` o similar si el autor de la llamada no tiene permisos `SELECT` sobre todas las columnas de la tabla. Esto impide a los

usuarios acceder a metadatos adicionales sobre tablas creadas con Apache Spark. Cuando se ejecutan en Amazon EMR, las aplicaciones Apache Spark siguen pudiendo leer estas tablas, pero es posible que no se apliquen ciertas optimizaciones y que no se admitan los nombres de columnas que distinguen entre mayúsculas y minúsculas. Si el usuario tiene acceso a todas las columnas de la tabla, Lake Formation devuelve la tabla sin modificar con todas las propiedades de la tabla.

Problema al renombrar una columna excluida

Si utiliza permisos a nivel de columna para excluir una columna y, a continuación, cambia el nombre de la columna, esta dejará de estar excluida de las consultas, como `SELECT *`.

Problema con la eliminación de columnas en tablas CSV

Si crea una tabla del catálogo de datos con el formato CSV y luego elimina una columna del esquema, las consultas podrían devolver datos erróneos y es posible que no se respeten los permisos a nivel de columna.

Solución alternativa: cree una tabla nueva en su lugar.

Las particiones de tabla deben añadirse bajo una ruta común

Lake Formation espera que todas las particiones de una tabla estén bajo una ruta común que se establece en el campo de ubicación de la tabla. Cuando utilice el rastreador para añadir particiones a un catálogo, esto funcionará sin problemas. Pero si añade particiones manualmente, y estas particiones no están bajo la ubicación establecida en la tabla principal, el acceso a los datos no funciona.

Problema con la creación de una base de datos durante la creación del flujo de trabajo

Al crear un flujo de trabajo a partir de un esquema utilizando la consola de Lake Formation, puede crear la base de datos de destino si no existe. Al hacerlo, el usuario que ha iniciado sesión obtiene el permiso de `CREATE_TABLE` sobre la base de datos creada. Sin embargo, el rastreador que genera el flujo de trabajo asume el papel de este cuando intenta crear una tabla. Esto produce un error porque el rol no tiene el permiso `CREATE_TABLE` en la base de datos.

Solución: si crea la base de datos a través de la consola durante la configuración del flujo de trabajo, antes de ejecutar el flujo de trabajo, debe dar al rol asociado al flujo de trabajo el permiso `CREATE_TABLE` sobre la base de datos que acaba de crear.

Problema al eliminar un usuario y, a continuación, volver a crearlo

El siguiente escenario da como resultado permisos erróneos de Lake Formation devueltos por `lakeformation:ListPermissions`:

1. Cree un usuario y conceda permisos a Lake Formation.
2. Elimine el usuario.
3. Vuelva a crear el usuario con el mismo nombre.

`ListPermissions` devuelve dos entradas, una para el usuario anterior y otra para el usuario nuevo. Si intenta revocar los permisos concedidos al usuario anterior, se revocan los permisos del nuevo usuario.

Las API **`GetTables`** y **`SearchTables`** no actualizan el valor del parámetro **`IsRegisteredWithLakeFormation`**

Existe una limitación conocida por la que las operaciones de la API del catálogo de datos como `GetTables` y `SearchTables` no actualizan el valor de `IsRegisteredWithLakeFormation parameter`, y devuelven el valor predeterminado, que es falso. Se recomienda utilizar la API `GetTable` para ver el valor correcto del `IsRegisteredWithLakeFormation parameter`.

Las operaciones de la API del catálogo de datos no actualizan el valor del parámetro **`IsRegisteredWithLakeFormation`**

Existe una limitación conocida por la que las operaciones de la API del catálogo de datos como `GetTables` y `SearchTables` no actualizan el valor del parámetro `IsRegisteredWithLakeFormation` y devuelven el valor predeterminado, que es falso. Se recomienda utilizar la API `GetTable` para ver el valor correcto del parámetro `IsRegisteredWithLakeFormation`.

Las operaciones de Lake Formation no admiten el registro de esquemas AWS Glue

Las operaciones de Lake Formation no admiten tablas AWS Glue que contengan un `SchemaReference` en el `StorageDescriptor` para ser utilizadas en el Registro de Esquemas.

Mensaje de error actualizado

AWS Lake Formation ha actualizado las excepciones específicas de los recursos al mensaje de error general `EntityNotFound` para las siguientes operaciones de la API con el fin de cumplir los objetivos de seguridad y conformidad.

- RevokePermissions
- GrantPermissions
- GetResourceLFTags
- GetTable
- GetDatabase

API de AWS Lake Formation

> **Note**
>
> Ya está disponible la referencia actualizada la API de servicio de AWS Lake Formation.

Contenido

- Patrones de cadena

API de permisos

En la sección API de permisos se describen las operaciones y los tipos de datos necesarios para conceder y revocar permisos en AWS Lake Formation. Consulte la Guía de referencia de la API de Lake Formation para ver todas las operaciones y tipos de datos de la API de AWS Lake Formation.

Operaciones

- GrantPermissions
- RevokePermissions
- BatchGrantPermissions
- BatchRevokePermissions
- GetEffectivePermissionsForPath
- ListPermissions

Data Types

- Resource
- DatabaseResource
- TableResource
- TableWithColumnsResource
- DataCellsFilterResourcee
- DataLocationResource
- DataLakePrincipal
- PrincipalPermissions
- PrincipalResourcePermissions
- DetailsMap
- ColumnWildcard
- BatchPermissionsRequestEntry
- BatchPermissionsFailureEntry

API de configuración de lagos de datos

En esta sección figuran las operaciones de la API de configuración del lago de datos y los tipos de datos para gestionar los administradores del lago de datos.

Operaciones

- GetDataLakeSettings
- PutDataLakeSettings

Data Types

- DataLakeSettings

API de integración de IAM Identity Center

Esta contiene las operaciones para crear y administrar la integración de Lake Formation con IAM Identity Center.

Operaciones

- CreateLakeFormationIdentityCenterConfiguration
- DeleteLakeFormationIdentityCenterConfiguration
- DescribeLakeFormationIdentityCenterConfiguration
- UpdateLakeFormationIdentityCenterConfiguration

Data Types

- ExternalFilteringConfiguration

API de modo de acceso híbrido

En la sección API de modo de acceso híbrido se describen las operaciones y los tipos de datos necesarios para configurar el modo de acceso híbrido en AWS Lake Formation. Consulte la Guía de

referencia de la API de Lake Formation para ver todas las operaciones y tipos de datos de la API de AWS Lake Formation.

Operaciones

- CreateLakeFormationOptIn
- DeleteLakeFormationOptIn
- ListLakeFormationOptIns

Data Types

- Resource
- DatabaseResource
- TableResource
- Resource Info
- LakeFormationOptInsInfo
- DataLocationResource

API de expedición de credenciales

La sección API de expedición de credenciales describe las operaciones y los tipos de datos relacionados con el uso del servicio de AWS Lake Formation para la expedición de credenciales y para registrar y administrar un recurso del lago de datos.

Operaciones

- RegisterResource
- DeregisterResource
- ListResources
- GetUnfilteredTableMetadata
- GetUnfilteredPartitionsMetadata
- GetTemporaryGluePartitionCredentials
- GetTemporaryGlueTableCredentials

- UpdateResource

Data Types

- FilterCondition
- RowFilter
- ResourceInfo

API de etiquetado

La sección API de etiquetado describe las operaciones y los tipos de datos relacionados con una estrategia de autorización que define un modelo de permisos sobre los atributos o las etiquetas de pares clave-valor.

Operaciones

- AddLFTagsToResource
- RemoveLFTagsFromResource
- GetResourceLFTags
- ListLFTags
- CreateLFTag
- GetLFTag
- UpdateLFTag
- DeleteLFTag
- SearchTablesByLFTags
- SearchDatabasesByLFTags

Data Types

- LFTagKeyResource
- LFTagPolicyResource
- TaggedTable
- TaggedDatabase

- LFTag
- LFTagPair
- LFTagError
- ColumnLFTag

API de filtrado de datos

En la sección API de filtro de datos se describe cómo administrar los filtros de celdas de datos en AWS Lake Formation.

Operaciones

- CreateDataCellsFilter
- DeleteDataCellsFilter
- ListDataCellsFilter
- GetDataCellsFilter
- UpdateDataCellsFilter

Tipos de datos

- DataCellsFilter
- RowFilter

Tipos de datos comunes

Los tipos de datos comunes describen diversos tipos de datos comunes en AWS Lake Formation.

Estructura ErrorDetail

Contiene detalles sobre un error.

Campos

- `ErrorCode`: cadena UTF-8, con 1 byte de largo como mínimo y 255 bytes de largo como máximo, que coincide con el Single-line string pattern.

El código asociado a este error.

- `ErrorMessage`: cadena de descripción de un máximo de 2048 bytes de largo, que coincide con el URI address multi-line string pattern.

  Mensaje que describe el error.

Patrones de cadena

La API utiliza las siguientes expresiones regulares para definir qué es un contenido válido para diversos miembros y parámetros de cadena:

- Patrón de cadena de línea única: "`[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*`"
- Patrón de cadena de varias líneas de la dirección URI: "`[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`"
- Patrón de cadena personalizado n.º 3: “`^\w+\.\w+\.\w+$`”
- Patrón de cadena personalizado n.º 4: “`^\w+\.\w+$`”
- Patrón de cadena personalizado n.º 5: “`arn:aws:iam::[0-9]*:role/.*`”
- Patrón de cadena personalizado n.º 6: “`arn:aws:iam::[0-9]*:user/.*`”
- Patrón de cadena personalizado n.º 7: “`arn:aws:iam::[0-9]*:group/.*`”
- Patrón de cadena personalizado n.º 8: "`arn:aws:iam::[0-9]*:saml-provider/.*`"
- Patrón de cadena personalizado n.º 9: "`^([\p{L}\p{Z}\p{N}_.:\/=+\-@%]*)$`"
- Patrón de cadena personalizado n.º 10: “`^([\p{L}\p{Z}\p{N}_.:\*\/=+\-@%]*)$`”
- Patrón de cadena personalizado n.º 11: “`[\p{L}\p{N}\p{P}]*`”

Regiones admitidas

Esta sección contiene información sobre el soporte Regiones de AWS y la funcionalidad de Lake Formation.

Disponibilidad general

Para conocer los servicios Regiones de AWS compatibles AWS Lake Formation, consulte la lista de AWS servicios disponibles por región.

Para obtener una lista de los puntos de conexión de Lake Formation de cada región y las Service Quotas de Lake Formation, consulte puntos de conexión y cuotas de AWS Lake Formation.

AWS GovCloud (US)

Para obtener una descripción general de las diferencias entre AWS GovCloud (US) la región y el estándar Regiones de AWS, consulte En qué AWS Lake Formation se diferencia AWS GovCloud (US).

Optimización de transacciones y almacenamiento

Las funciones de tablas gobernadas, soporte de transacciones y optimización del almacenamiento de Lake Formation están disponibles en las siguientes ubicaciones: Regiones de AWS

Nombre de la región	Parámetro de la región	Punto de conexión
Este de EE. UU. (Norte de Virginia)	us-east-1	lakeformation.us-east-1.amazonaws.com lakeformation-fips.us-east-1.amazonaws.com
Este de EE. UU. (Ohio)	us-east-2	lakeformation.us-east-2.amazonaws.com lakeformation-fips.us-east-2.amazonaws.com

Nombre de la región	Parámetro de la región	Punto de conexión
Oeste de EE. UU. (Oregón)	us-west-2	lakeformation.us-west-2.amazonaws.com lakeformation-fips.us-west-2.amazonaws.com
Asia Pacífico (Mumbai)	ap-south-1	lakeformation.ap-south-1.amazonaws.com
Asia Pacífico (Seúl)	ap-northeast-2	lakeformation.ap-northeast-2.amazonaws.com
Asia Pacífico (Singapur)	ap-southeast-1	lakeformation.ap-southeast-1.amazonaws.com
Asia Pacífico (Sídney)	ap-southeast-2	lakeformation.ap-southeast-2.amazonaws.com
Asia Pacífico (Tokio)	ap-northeast-1	lakeformation.ap-northeast-1.amazonaws.com
Europa (Frankfurt)	eu-central-1	lakeformation.eu-central-1.amazonaws.com
Europa (Irlanda)	eu-west-1	lakeformation.eu-west-1.amazonaws.com
Europa (Londres)	eu-west-2	lakeformation.eu-west-2.amazonaws.com
Europa (Estocolmo)	eu-north-1	lakeformation.eu-north-1.amazonaws.com

Nombre de la región	Parámetro de la región	Punto de conexión
Canadá (Central)	ca-central-1	lakeformation.ca-central-1.amazonaws.com
América del Sur (São Paulo)	sa-east-1	lakeformation.sa-east-1.amazonaws.com

Historial de documentos de AWS Lake Formation

En la siguiente tabla se describen cambios importantes en la documentación para AWS Lake Formation.

Cambio	Descripción	Fecha
Configuración actualizada de Lake Formation	Se actualizaron los pasos de la AWS Lake Formation sección de configuración.	7 de febrero de 2024
Cambio de política actualizado	Se han añadido nuevos permisos a la política en línea del rol vinculado al servicio. Para obtener más información, consulte Uso de funciones vinculadas a servicios para Lake Formation.	7 de febrero de 2024
Cambio de política actualizado	Documentó el cambio en la LakeFormationDataAccessServiceRolePolicypolítica.	2 de febrero de 2024
Limitaciones de Lake Formation consolidadas	Se ha creado una sección unificada de las limitaciones y consideraciones de Lake Formation. Para obtener más información, consulte Limitaciones de Lake Formation.	15 de diciembre de 2023
Se ha añadido documentación para la compactación de Iceberg	Para mejorar el rendimiento de lectura de los servicios de análisis de AWS, como Athena y Amazon EMR, y los trabajos de ETL AWS Glue, AWS Glue Data Catalog	25 de noviembre de 2023

	proporciona una compactación administrada (un proceso que compacta objetos pequeños de Amazon S3 para convertir los en objetos más grandes) para las tablas de Iceberg del catálogo de datos. Para obtener más información, consulte Optimización de las tablas de Iceberg.	
Se ha añadido documentación para la integración de IAM Identity Center	Las integraciones de IAM Identity Center permiten a los usuarios y grupos acceder a los recursos del catálogo de datos haciendo cumplir los permisos de Lake Formation. Para obtener más información consulte Integración de IAM Identity Center.	25 de noviembre de 2023
Se ha añadido documentación para las vistas del catálogo de datos	Puede crear vistas en AWS Glue Data Catalog que hagan referencia a un máximo de 10 tablas mediante editores de SQL para Amazon Athena o Amazon Redshift. Para obtener más información, consulte Creación de vistas.	25 de noviembre de 2023
Actualizado el cambio de política	Documentó el cambio en la AWSLakeFormationCrossAccountManagerpolítica.	25 de octubre de 2023

Agregada documentación para el modo de acceso híbrido	El modo de acceso híbrido proporciona la flexibilidad de habilitar selectivamente los permisos de Lake Formation para las bases de datos y tablas de su AWS Glue Data Catalog. Con el modo de acceso híbrido, ahora tiene una ruta incremental que le permite establecer los permisos de Lake Formation para un conjunto específico de usuarios sin interrumpir las políticas de permisos de otros usuarios o cargas de trabajo existentes. Para obtener más información, consulte Modo de acceso híbrido.	26 de septiembre de 2023
Agregada documentación para crear tablas de Apache Iceberg	Ahora puede crear tablas Iceberg de Apache que utilicen el formato de datos de Apache Parquet en el AWS Glue Data Catalog con datos que residan en Amazon S3. Para obtener más información, consulte Creación de tablas Iceberg.	16 de agosto de 2023

Agregada documentación para el acceso a datos entre regiones	Lake Formation es compatible con las consultas a las tablas del Catálogo de datos entre regiones de AWS. Puede acceder a los datos de una región desde otras regiones mediante Athena, Amazon EMR y ejecutar AWS Glue ETL creando enlaces de recursos en otras regiones que apunten a las bases de datos y tablas de origen. Puede conectar el Catálogo de datos a metaalmacenes externos que almacenan los metadatos de sus datos de Amazon S3 y gestionar de forma segura los permisos de acceso a los datos mediante AWS Lake Formation. Para obtener más información, consulte Acceso a tablas entre regiones.	30 de junio de 2023
Contenido reorganizado	Reorganizados los capítulos de la guía para adaptarlos a la experiencia del usuario de Lake Formation.	15 de mayo de 2023

Agregada documentación para la federación de HMS	Puede conectar el Catálogo de datos a metaalmacenes externos que almacenan los metadatos de sus datos de Amazon S3 y gestionar de forma segura los permisos de acceso a los datos mediante AWS Lake Formation. Para obtener más información, consulte Administración de los permisos de los conjuntos de datos que utilizan metaalmacenes externos.	15 de abril de 2023
Agregada documentación sobre el uso compartido de datos de Amazon Redshift	Ahora puede gestionar de forma segura los datos de un recurso compartido de datos de Amazon Redshift con los permisos de Lake Formation. Lake Formation es compatible con licencias de acceso a sus datos mediante AWS Data Exchange. Para obtener más información, consulte Uso compartido de datos en AWS Lake Formation.	30 de noviembre de 2022
Compatibilidad para compartir datos entre cuentas directamente con las entidades principales	Agregada información sobre cómo compartir datos directamente con las entidades principales de IAM en otra cuenta. Para obtener más información, consulte Uso compartido de datos entre cuentas en AWS Lake Formation.	10 de noviembre de 2022

Compatibilidad con el uso compartido de datos habilitado para AWS RAM mediante TBAC	Agregada información sobre el método LF-TBAC de concesión de permisos del catálogo, uso de AWS Resource Access Manager para concesiones entre cuentas.	10 de noviembre de 2022
Agregada una sección sobre cómo trabajar con otros servicios	Agregada información sobre cómo los servicios de AWS como Amazon Athena, AWS Glue, Redshift Spectrum y Amazon EMR pueden utilizar Lake Formation para acceder de forma segura a los datos de Amazon S3 registradas en Lake Formation. Para más información, consulte Trabajar con otros AWS servicios.	10 de noviembre de 2022
???	Agregada información sobre cómo solucionar un error al utilizar Amazon EMR para acceder a los datos entre cuentas. Para obtener más información, consulte Error al utilizar Amazon EMR para acceder a datos compartidos entre cuentas.	7 de noviembre de 2022
Actualizaciones del uso compartido de recursos entre cuentas	Agregada una descripción sobre cómo funcionan los recursos compartidos entre cuentas en Lake Formation . Documentó el cambio en la AWSLakeFormationCrossAccountManagerpolítica.	6 de mayo de 2022

Nuevos tutoriales	Agregados nuevos tutoriales para crear tablas gobernadas, proteger los lagos de datos y compartir lagos de datos. Para obtener más detalles, consulte la sección Comenzar.	20 de abril de 2022
Nueva página de inicio de Lake Formation	Se actualizó la página de inicio de Lake Formation para incluir enlaces a tutoriales que proporcionan step-by-s tep instrucciones sobre cómo construir un lago de datos, ingerir datos, compartir y proteger lagos de datos con Lake Formation.	20 de abril de 2022
Compatibilidad para el suministro de credenciales	Agregada información sobre el suministro de credencia les, compatible con Lake Formation para permitir que los servicios de terceros se integren con Lake Formation mediante las operaciones de la API de suministro de credenciales. Para obtener más información, consulte Cómo funciona el suministr o de credenciales en Lake Formation.	28 de febrero de 2022

Compatibilidad para tablas gobernadas y filtrado de datos avanzado	Agregada información acerca de tablas gobernadas por Lake Formation, que admiten transacciones ACID, compactación automática de datos y consultas de viaje en el tiempo. Agregada información sobre la creación de filtros de datos para admitir la seguridad a nivel de columna, fila y celda. Para obtener más información, consulte Tablas gobernadas en Lake Formation y Filtrado de datos y seguridad a nivel de celda en Lake Formation.	30 de noviembre de 2021
Compatibilidad con puntos de conexión de interfaz de VPC	Agregada información sobre la creación de un punto de conexión de interfaz de nube privada virtual (VPC) para Lake Formation, de forma que la comunicación entre su VPC y Lake Formation se haga de forma completa y segura dentro de la red AWS. Para obtener más información, consulte Uso de Lake Formation con puntos de conexión de VPC.	11 de octubre de 2021

Compatibilidad con las políticas de punto de conexión de VPC	Agregada información acerca de la compatibilidad con las políticas de punto de conexión de nube privada virtual (VPC) en Lake Formation. Para obtener más información, consulte Uso de Lake Formation con puntos de conexión de VPC.	11 de octubre de 2021
Compatibilidad con control de acceso basado en etiquetas.	El control de acceso basado en etiquetas de Lake Formation proporciona una forma nueva y más escalable de administrar el acceso a los recursos del Catálogo de datos y a los datos subyacentes mediante etiquetas LF. Para obtener más información, consulte Control de acceso basado en etiquetas de Lake Formation.	7 de mayo de 2021
Nuevo requisito de suscripción para el filtrado de datos en Amazon EMR.	Agregada información sobre el requisito de optar por permitir que Amazon EMR filtre los datos gestionados por Lake Formation. Para obtener más información, consulte Permitir el filtrado de datos en Amazon EMR.	9 de octubre de 2020

Compatibilidad para conceder permisos entre cuentas completos en las bases de datos del Catálogo de datos	Agregada información sobre la concesión de permisos completos de Lake Formation en bases de datos del Catálogo de datos entre cuentas de AWS, incluido `CREATE_TABLE` . Para obtener más información, consulte Compartir bases de datos del Catálogo de datos.	1 de octubre de 2020
Compatibilidad con la autenticación de usuarios Amazon Athena a través de SAML.	Agregada información sobre la compatibilidad para los usuarios de Athena que se conectan a través del controlador JDBC u ODBC y se autentican a través de proveedores de identidad SAML, como Okta y Microsoft Active Directory Federation Service (AD FS). Para más información, consulte Integraciones de servicio AWS con Lake Formation.	30 de septiembre de 2020
Compatibilidad para acceso entre cuentas con un Catálogo de datos cifrado	Agregada información sobre la concesión de permisos entre cuentas cuando el Catálogo de datos está cifrado. Para obtener más información, consulte Requisitos previos para el acceso entre cuentas.	30 de julio de 2020

Compatibilidad para el acceso entre cuentas al lago de datos	Agregada información sobre la concesión de permisos de AWS Lake Formation en las bases de datos y tablas del Catálogo de datos a cuentas AWS y organizaciones externas, y sobre el acceso a los objetos del Catálogo de datos compartidos desde cuentas externas. Para obtener más información, consulte Acceso entre cuentas.	7 de julio de 2020
Integración con Amazon QuickSight	Se agregó información sobre cómo conceder permisos de Lake Formation a los usuarios de Amazon QuickSight Enterprise Edition para que puedan acceder a los conjuntos de datos que residen en ubicaciones registradas de Amazon S3. Para obtener más información, consulte Concesión de permisos de ubicación de datos.	29 de junio de 2020
Actualizaciones de los capítulos de configuración y de introducción	Reorganizados y mejorados los capítulos de configuración e introducción. Actualizados los permisos de AWS Identity and Access Management (IAM) recomendados para el administrador del lago de datos.	27 de febrero de 2020

Compatibilidad con AWS Key Management Service	Agregada información sobre cómo la compatibilidad de Lake Formation para AWS Key Management Service (AWS KMS) simplifica la configuración de servicios integrados para leer y escribir datos cifrados en ubicaciones registradas de Amazon Simple Storage Service (Amazon S3). Agregada información sobre cómo registrar las ubicacion es de Amazon S3 que están cifradas con AWS KMS keys. Para obtener más informaci ón, consulte the section called “Añadir una ubicación de Amazon S3 a su lago de datos”.	27 de febrero de 2020
Actualizaciones de los esquemas y las políticas de IAM de los administradores de lagos de datos	Aclarados los parámetros de entrada para los esquemas de bases de datos increment ales. Actualizadas las políticas de IAM necesarias para el administrador de un lago de datos.	20 de diciembre de 2019
Reescritura y actualización de los capítulos de seguridad	Mejorados los capítulos de seguridad y actualización.	29 de octubre de 2019
El superpermiso reemplaza a todos los permisos	Actualizados los capítulos de seguridad y actualización para reflejar la sustitución del permiso `All` por `Super`.	10 de octubre de 2019

Adiciones, correcciones y aclaraciones	Efectuadas adiciones, correcciones y aclaraciones de acuerdo con los comentarios. Revisado el capítulo de seguridad. Actualizados los capítulos de seguridad y actualización para reflejar la sustitución del grupo `Everyone` por `IAMAllowedPrincipals` .	11 de septiembre de 2019
Nueva guía	Esta es la versión inicial de la Guía para desarrolladores de AWS Lake Formation.	8 de agosto de 2019

Glosario de AWS

Para ver la terminología más reciente de AWS, consulte el Glosario de AWS en la Referencia de Glosario de AWS.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.