開發人員指南

AWS Lake Formation

AWS Lake Formation: 開發人員指南

Table of Contents

什麼是 AWS Lake Formation？

歡迎使用開 AWS Lake Formation 發人員指南。

AWS Lake Formation 協助您集中控管、保護並全球共用資料，以進行分析和機器學習。使用 Lake Formation，您可以在 Amazon Simple Storage Service (Amazon S3) 及其中的中 AWS Glue Data Catalog繼資料上管理資料湖資料的精細存取控制。

Lake Formation 提供了自己的許可模型，用於增強 IAM 許可模型。Lake Formation 權限模型可透過簡單的授權或撤銷機制，對儲存在資料湖中的資料進行細粒度存取，就像關聯式資料庫管理系統 (RDBMS) 一樣。使用跨 AWS 分析和機器學習服務 (包括亞馬遜雅典娜、Amazon Amazon Redshift Spectrum、Amazon EMR 和) 的欄、列和儲存格層級的精細控制來強制執行 Lake Formation 許可。Amazon QuickSight AWS Glue

Lake Formation 混合存取模式 AWS Glue Data Catalog 可讓您使用 Amazon S3 的 Lake Formation 許可和 IAM 許可政策和 AWS Glue 動作來保護和存取已編目的資料。透過混合式存取模式，資料管理員可以選擇性地逐步上載 Lake Formation 權限，一次專注於一個資料湖使用案例。

Lake Formation 也可讓您在內部和外部跨多個 AWS 組織共用資料 AWS 帳戶，或直接與另一個帳戶中的 IAM 主體共用資料，以提供對中 AWS Glue Data Catalog 繼資料和基礎資料的精細存取權。

主題

Lake Formation 功能

Lake Formation 可幫助您打破資料孤島，並將不同類型的結構化和非結構化資料合併到集中式儲存庫中。首先，識別 Amazon S3 中的現有資料存放區或關聯式和 NoSQL 資料庫，然後將資料移到資料湖

中。然後編目、編目並準備資料以進行分析。接下來，透過選擇的分析服務，為您的使用者提供對資料的安全自助存取。

主題

資料擷取與管理

從已存在的資料庫匯入資料 AWS

一旦您指定現有資料庫的位置並提供存取認證，Lake Formation 就會讀取資料及其中繼資料 (結構描述) 以瞭解資料來源的內容。然後，它會將資料匯入新的資料湖，並將中繼資料記錄在中央目錄中。使用 Lake Formation，您可以從在 Amazon RDS 中執行或託管在 Amazon EC2 中執行的 MySQL、PostgreSQL、SQL 伺服器、MariaDB 和甲骨文資料庫匯入資料。支援大量和增量資料載入。

從其他外部來源匯入資料

您可以使用 Lake Formation，透過與 Java 資料庫連線 (JDBC) 連線，從內部部署資料庫移動資料。識別您的目標來源並在主控台中提供存取認證，Lake Formation 會讀取您的資料並將其載入資料湖。若要從上述資料庫以外的資料庫匯入資料，您可以使 AWS Glue用建立自訂 ETL 工作。

編目和標記您的資料

您可以使用 AWS Glue 檢索器讀取 Amazon S3 中的資料，並擷取資料庫和表格結構描述，並將該資料存放在可搜尋 AWS Glue Data Catalog範圍內。然後，使用 Lake Formation 基於 Lake Formation 標籤的訪問控制 (TBAC) 來管理資料庫、資料表和資料行的權限。若要取得有關將表格加入至資料目錄的更多資訊，請參閱建立資料目錄表格和資料庫。

安全性管理

定義和管理存取控制

Lake Formation 提供單一位置來管理資料湖中資料的存取控制。您可以定義安全性原則，以限制資料庫、表格、欄、列和儲存格層級的資料存取權。這些政策適用於 IAM 使用者和角色，以及透過外部身分識別提供者聯合時的使用者和群組。您可以使用精細的控制來存取由 Amazon Redshift Spectrum、Athena、 AWS Glue ETL 和 Amazon EMR 中的 Lake Formation 保護的資料 (適用於

Apache Spark)。每當您建立 IAM 身分時，請務必遵循 IAM 最佳實務。如需詳細資訊，請參閱 IAM 使用者指南中的安全性最佳做法。

混合存取模式

Lake Formation m 混合存取模式提供了靈活性，可選擇性地啟用 AWS Glue Data Catalog. 透過混合式存取模式，您現在擁有一個增量路徑，可讓您為一組特定使用者設定 Lake Formation 權限，而不會中斷其他現有使用者或工作負載的權限原則。如需詳細資訊，請參閱 混合存取模式。

實作稽核記錄

Lake Formation 提供全面的稽核記錄， CloudTrail 以監控存取並顯示是否符合集中定義的政策。您可以跨分析和機器學習服務稽核資料存取歷史記錄，這些服務會透過 Lake Formation 讀取資料湖中的資料。這可讓您查看哪些使用者或角色嘗試存取哪些資料、使用哪些服務以及何時存取。您可以使用 CloudTrail API 和主控台存取任何其他 CloudTrail 記錄檔的方式來存取稽核記錄。如需 CloudTrail 記錄檔的詳細資訊，請參閱使用記錄 AWS Lake Formation API 調用 AWS CloudTrail。

列與儲存格層級安全性

Lake Formation 提供資料篩選器，可讓您限制對欄和列組合的存取。使用資料列和儲存格層級安全性來保護敏感資料，例如個人識別資訊 (PII)。如需列層級安全性的詳細資訊，請參閱。資料篩選概觀

標籤式存取控制

透過建立稱為 Lf-tags 的自訂標籤，使用 Lake Formation 標籤型存取控制來管理數百甚至數千個資料權限。您現在可以定義 LF 標籤，並將它們附加到資料庫、資料表或資料行。然後，跨分析、機器學習 (ML) 共用受控存取，以及擷取、轉換和載入 (ETL) 服務以供使用。LF 標籤可確保資料控管能夠輕鬆調整規模，方法是將數千個資源的原則定義取代為幾個邏輯標籤。Lake Formation 針對此中繼資料提供以文字為基礎的搜尋，因此您的使用者可以快速找到他們需要分析的資料。

跨帳戶存取

Lake Formation 權限管理功能可透過集中式方法簡化跨多個 AWS 帳戶的分散式資料湖的保護和管理，為資料目錄和 Amazon S3 位置提供精細的存取控制。如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。

資料共用

資料共用功能可讓您對存放在 Amazon Redshift 等不同資料來源中的資料集設定許可，而無需將資料或中繼資料移轉至 Amazon S3 或 AWS Glue Data Catalog。您可以使用以下方法在 Lake Formation 中共享數據：

有關更多信息，請參閱 Lake Formation 中的數據共享。

- 將湖泊形成與 Amazon Redshift 資料共用整合 — 使用 Lake Formation 集中管理 Amazon Redshift 資料庫的資料庫、表格、欄和資料列層級存取權限，並限制使用者存取資料清單中的物件。
- 連線 AWS Glue Data Catalog 至外部中繼存放區 — Connect AWS Glue Data Catalog 至外部中繼存放區，以使用 Lake Formation 管理 Amazon S3 中資料集的存取許可。不需要將中繼資料 AWS Glue Data Catalog 移轉至。

  如需更多資訊，請參閱管理使用外部中繼存放區之資料集的權限
- 將湖泊形成與 AWS Data Exchange 整合 — Lake Formation 支援透過授權存取您的資料 AWS Data Exchange。如果您對 Lake Formation 資料的授權有興趣，請參閱AWS Data Exchange 使用者指南AWS Data Exchange中的內容。

AWS Lake Formation：運作方式

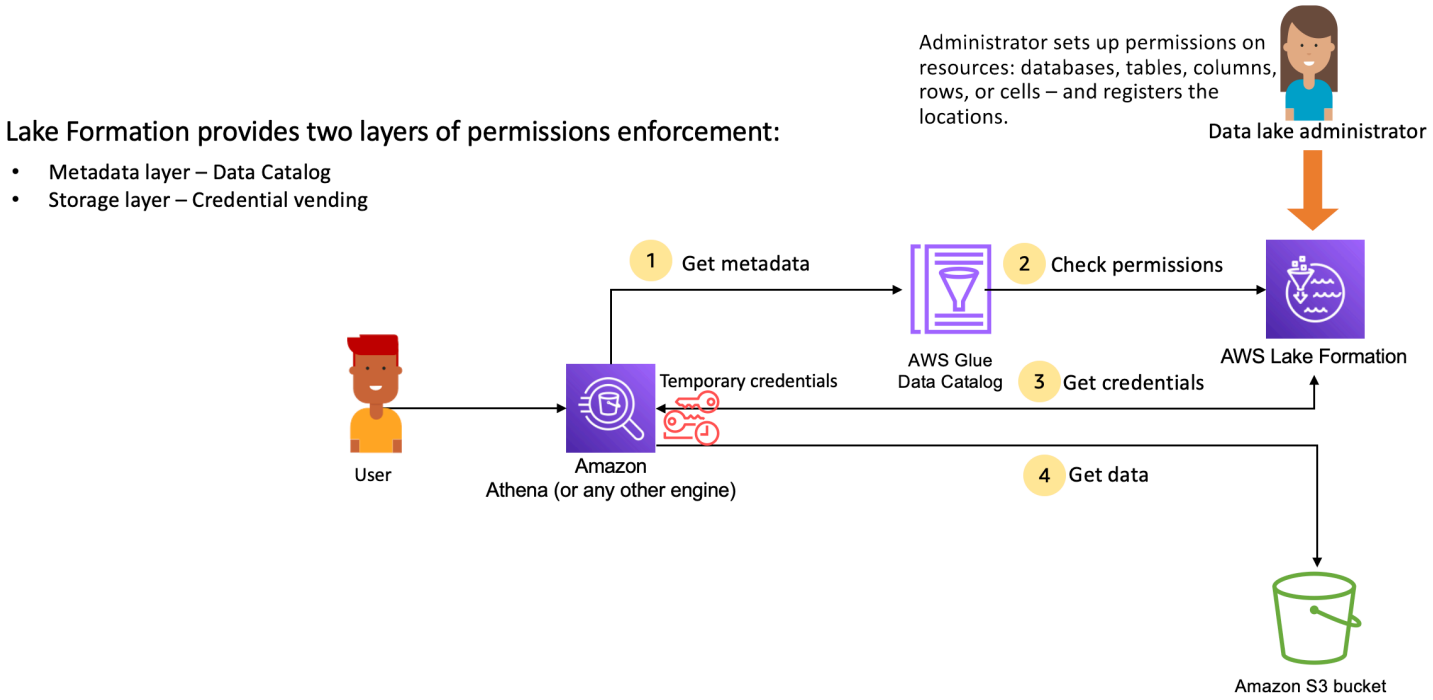
AWS Lake Formation 提供關聯式資料庫管理系統 (RDBMS) 許可模型，以授與或撤銷資料目錄資源的存取權，例如資料庫、表格和 Amazon S3 中含有基礎資料的資料行。易於管理的 Lake Formation 許可取代複雜的 Amazon S3 儲存貯體政策和對應的 IAM 政策。

在 Lake Formation 中，您可以在兩個級別上實現權限：

- 對資料目錄資源 (例如資料庫和表格) 強制執行中繼資料層級權限
- 代表整合引擎管理 Amazon S3 中存放的基礎資料的儲存存取權限

Lake Formation 權限管理工作流程

Lake Formation 與分析引擎整合，以查詢已在 Lake Formation 註冊的 Amazon S3 資料存放區和中繼資料物件。下圖說明了許可管理在 Lake Formation 中的運作方式。



Lake Formation 許可管理高級步驟

在 Lake Formation 可以為資料湖中的資料提供存取控制之前，資料湖管理員或具有管理權限的使用者會設定個別的「資料目錄」表格使用者原則，以允許或拒絕使用 Lake Formation 權限存取「資料目錄」表格。

然後，資料湖管理員或管理員委派的使用者將 Lake Formation 許可授與資料目錄資料庫和表格上的使用者，並將表格的 Amazon S3 位置註冊到 Lake Formation。

1. 取得中繼資料 — 主體 (使用者) 將查詢或 ETL 指令碼提交至整合式分析引擎，例如 Amazon Athena AWS Glue、Amazon EMR 或 Amazon Redshift Spectrum。整合式分析引擎可識別要求的資料表，並將中繼資料要求傳送至資料目錄。
2. 檢查權限 —「資料目錄」會透過 Lake Formation 檢查使用者的權限，如果使用者獲得存取資料表的授權，則會將允許使用者查看的中繼資料傳回引擎。
3. 取得認證 — 資料目錄可讓引擎知道表格是否由 Lake Formation 管理。如果基礎資料已向 Lake Formation 註冊，則分析引擎會要求 Lake Formation 透過授予臨時存取權來提供資料存取權。
4. 取得資料 — 如果使用者獲得存取資料表的授權，Lake Formation 會提供對整合式分析引擎的暫時存取權。分析引擎會使用暫時存取從 Amazon S3 擷取資料，並執行必要的篩選，例如欄、列或儲存格篩選。當引擎完成執行作業時，會將結果傳回給使用者。此過程稱為憑證自動售貨機。

   如果表格不是由 Lake Formation 管理，則會直接向 Amazon S3 進行分析引擎的第二個呼叫。系統會針對資料存取評估相關的 Amazon S3 儲存貯體政策和 IAM 使用者政策。

每當您使用 IAM 政策時，請務必遵循 IAM 最佳實務。如需詳細資訊，請參閱《IAM 使用者指南》中的 IAM 中的安全性最佳實務。

主題

元數據權限

Lake Formation 為資料目錄提供授權和存取控制。當 IAM 角色從任何系統進行資料目錄 API 呼叫時，資料目錄會驗證使用者的資料許可，並僅傳回使用者具有存取權限的中繼資料。例如，如果 IAM 角色只能存取資料庫中的一個資料表，而且擔任該角色的服務或使用者執行`GetTables`作業，則無論資料庫中的資料表數目為何，回應都只會包含一個資料表。

預設設定-**`IAMAllowedPrincipal`** 群組權限

AWS Lake Formation依預設，會將所有資料庫和資料表的權限設定為名為的虛擬群組`IAMAllowedPrincipal`。這個群體是獨一無二的，只有在 Lake Formation 中可見。該`IAMAllowedPrincipal`群組包括可透過 IAM 主體政策和 AWS Glue 資源政策存取資料目錄資源的所有 IAM 主體。如果此權限存在於資料庫或資料表上，則會授與所有主體存取資料庫或資料表的存取權。

如果您想要在資料庫或資料表上提供更精細的權限，請移除`IAMAllowedPrincipal`權限，並且 Lake Formation 會強制執行與該資料庫或資料表相關聯的所有其他原則。例如，如果有一個原則允許使用者 A `IAMAllowedPrincipal` 存取具有`DESCRIBE`權限的資料庫 A，且具有所有權限，則使用者 A 將繼續執行所有其他動作，直到`IAMAllowedPrincipal`權限被撤銷為止。

此外，根據預設，`IAMAllowedPrincipal`群組在建立時擁有所有新資料庫和資料表的權限。有兩種配置可以控制此行為。第一個是在新建立的資料庫啟用此功能的帳戶和區域層級，第二個是在資料庫層級。若要修改預設設定，請參閱變更預設權限模型或使用混合存取模式。

授予許可

資料湖管理員可以將「資料目錄」權限授與主體，以便主參與者可以建立和管理資料庫和表格，以及存取基礎資料。

資料庫和資料表層級權限

當您在 Lake Formation 內授與權限時，授與者必須指定要授與權限的主參與者、要授與權限的資源，以及受權者應具有存取權限的動作。對於 Lake Formation 中的大多數資源，要授與權限的主參與者清單和資源類似，但受權者可以執行的動作會根據資源類型而有所不同。例如，SELECT資料表可以讀取資料表的SELECT權限，但資料庫不允許使用權限。資料庫允CREATE_TABLE許使用權限，但不允許資料表。

您可以使用兩種方法授予 AWS Lake Formation 權限：

- 具名資源方法 — 可讓您選擇資料庫和表格名稱，同時將權限授與使用者。
- LF 標籤型存取控制 (LF-TBAC) — 使用者建立 LF 標籤、將它們與資料目錄資源產生關聯、授與 LF 標籤的權限、將Describe權限關聯至個別使用者，以及使用 LF 標籤將 LF 權限原則寫入不同的使用者。這類以 LF 標籤為基礎的原則會套用至與這些 LF 標籤值相關聯的所有資料目錄資源。

> **Note**
>
> LF-標籤是 Lake Formation 獨有的。它們僅在 Lake Formation 中可見，不應與 AWS 資源標籤混淆。

LF-TBAC 是一項功能，可讓使用者將資源分組到 LF 標籤的使用者定義類別，並對這些資源群組套用權限。因此，這是在大量資料目錄資源之間擴展權限的最佳方式。

如需詳細資訊，請參閱 基於 Lake Formation 標籤的訪問控制。

當您授與權限給主參與者時，Lake Formation 會將權限評估為該使用者所有策略的聯集。例如，如果主參與者的資料表上有兩個原則，其中一個原則會透過具名的資源方法授與資料行 col1、col2 和 col3 的權限，而另一個原則會授與相同資料表和主體的權限給 col5，以及 col6 至 LF 標籤，有效權限的聯集將是 col1、col2、col3、col5 和 col6。這也包括資料篩選器和列。

資料位置權限

資料位置許可讓非管理使用者能夠在特定 Amazon S3 位置建立資料庫和表格。如果使用者嘗試在沒有建立權限的位置建立資料庫或資料表，則建立工作會失敗。這是為了防止使用者在資料湖中的任意位置建立資料表，並提供控制這些使用者可以讀取和寫入資料的位置。在建立資料庫內的 Amazon S3 位置建立資料表時，會有隱含權限。如需詳細資訊，請參閱 授與資料位置權限。

建立資料表和資料庫權限

默認情況下，非管理用戶沒有在數據庫中創建數據庫或表的權限。使用 Lake Formation 設定在帳戶層級控制資料庫建立，因此只有授權的主體才能建立資料庫。如需詳細資訊，請參閱 建立資料庫。若要建立資料表，主體需要建立資料表所在資料庫的CREATE_TABLE權限。如需詳細資訊，請參閱 建立資料表。

隱含和明確的權限

Lake Formation 根據人物角色和人物角色執行的動作提供隱含的權限。例如，資料湖管理員會自動取得資料目錄中所有資源的DESCRIBE權限、所有位置的資料位置權限、在所有位置建立資料庫和表格的權限，以Grant及任何資源的Revoke權限和權限。資料庫建立者會自動取得他們所建立之資料庫的所有資料庫權限，而資料表建立者則會取得他們所建立之資料表的所有權限。如需詳細資訊，請參閱 隱含 Lake Formation 權限。

可授予的權限

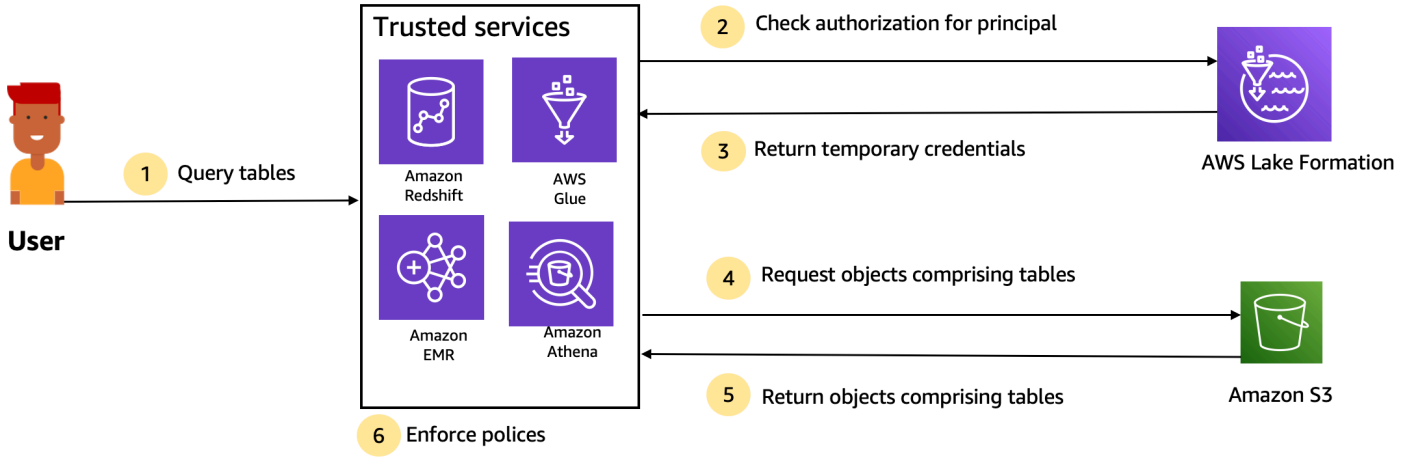
資料湖管理員可以透過提供可授與的權限，將權限管理委派給非系統管理使用者。當提供資源的可授與權限集的主參與者權限時，該主參與者就可以將權限授與該資源上其他主參與者。

儲存存取權管理

Lake Formation 使用憑證自動售貨功能來提供對 Amazon S3 資料的臨時存取。憑據自動售貨或令牌自動售貨是一種常見的模式，它為用戶，服務或某些其他實體提供臨時憑據，用於授予對資源的短期訪問權限。

Lake Formation 利用這種模式來提供短期存取 AWS 分析服務 (例如 Athena)，以代表呼叫主體存取資料。授予許可時，使用者不需要更新其 Amazon S3 儲存貯體政策或 IAM 政策，也不需要直接存取 Amazon S3。

下圖顯示了 Lake Formation 如何提供對註冊位置的臨時訪問權限：



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

1. 主體 (使用者) 透過受信任的整合服務 (例如 Athena、Amazon EMR、Redshift 頻譜或) 輸入資料表的查詢或請求。 AWS Glue
2. 整合式服務會針對表格和要求的資料行檢查 Lake Formation 的授權，並做出授權決定。如果使用者未獲得授權，則 Lake Formation 會拒絕存取資料，且查詢會失敗。
3. 在授權成功並開啟資料表和使用者的儲存區授權之後，整合式服務會從 Lake Formation 擷取暫時認證以存取資料。
4. 整合式服務會使用來自 Lake Formation 的臨時登入資料，向 Amazon S3 請求物件。
5. Amazon S3 提供 Amazon S3 對象的集成服務。Amazon S3 對象包含表中的所有數據。
6. 整合式服務會執行 Lake Formation 原則的必要強制執行，例如欄層級、資料列層級和/或儲存格層級篩選。整合式服務會處理查詢，並將結果傳回給使用者。

啟用「資料目錄」表格的儲存層級權限強制執行

依預設，「資料目錄」中的資料表不會啟用儲存層級強制執行。若要啟用儲存層級強制執行，您必須向 Lake Formation 註冊來源資料的 Amazon S3 位置，並提供 IAM 角色。對於具有相同表格位置路徑或 Amazon S3 位置前綴的所有表格，都會啟用儲存層級許可。

當整合式服務代表使用者要求存取資料位置時，Lake Formation 服務會擔任此角色，並將認證傳回至具有降低資源權限的要求服務，以便進行資料存取。已註冊的 IAM 角色必須具有對 Amazon S3 位置 (包括 AWS KMS 金鑰) 的所有必要存取權。

如需詳細資訊，請參閱 註冊 Amazon S3 位置。

支援的 AWS 服務

AWS 分析服務，例如 Athena，Redshift 頻譜，Amazon EMR AWS Glue Amazon QuickSight，並使用 AWS Lake Formation 憑證自動售貨 API 操作與 Lake Formation Amazon SageMaker 集成。若要查看與 Lake Formation 整合的完整 AWS 服務清單，以及它們支援的粒度和表格格式等級，請參閱與其他 AWS 服務合作。

Lake Formation 的跨賬戶數據共享

使用 Lake Formation，您可以使用具名的資源方法或 LF-tag，在簡單的設定中，在帳 AWS 戶內和跨帳戶共用資料目錄資源 (資料庫和表格)。您可以共用整個資料庫，也可以從資料庫中選取資料表給帳戶中的任何 IAM 主體 (IAM 角色和使用者)、帳戶層級的其他 AWS 帳戶，或直接向其他帳戶中的 IAM 主體選取資料表。

您也可以使用資料篩選器共用「資料目錄」表格，以限制對資料列層級和儲存格層級詳細資訊的詳細資訊的存取。Lake Formation 使用 AWS Resource Access Manager （AWS RAM）來促進帳戶之間授予權限。在兩個帳號之間共用資源時， AWS RAM 會將邀請傳送至收件者帳戶。當使用者接受 AWS RAM 共用邀請時，會向 Lake Formation AWS RAM 提供必要的權限，讓資料目錄資源可用，並啟用儲存層級強制執行。如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。

當收件者帳戶的資料湖管理員接受共用時，收件者帳戶中即可使用共用資源。 AWS RAM 如果管理員具有共用資源的權限，則資料湖管理員會將共用資源的進一步 Lake Formation `GRANTABLE` 權限授與收件者帳戶中的其他 IAM 主體。

不過，如果沒有資源連結，主體就無法使用 Athena 或 Redshift 頻譜查詢共用資源。資源連結是資料目錄中的實體，類似於 Linux 符號連結概念。

收件者帳號的資料湖管理員會在共用資源上建立資源連結。管理員會將資源連結的`Describe`權限以及原始共用資源的必要權限授與其他使用者。接著，收件者帳號中的使用者可以使用資源連結，使用 Athena 和 Redshift 頻譜查詢共用資源。如需有關資源連結的詳細資訊，請參閱建立資源連結。

Lake Formation 部分

AWS Lake Formation 依賴數個元件的互動來建立和管理資料湖。

Lake Formation 控制台

您可以使用 Lake Formation 主控台來定義和管理資料湖，並授予和撤銷 Lake Formation 權限。您可以使用主控台上的藍圖來探索、清理、轉換和擷取資料。您也可以啟用或停用個別 Lake Formation 使用者對主控台的存取。

Lake Formation API 和命令行界面

Lake Formation 通過幾個特定於語言的 SDK 和（ ）提供 API 操作。 AWS Command Line Interface AWS CLI Lake Formation API 與 AWS Glue API 一起工作。Lake Formation API 主要著重於管理 Lake Formation 權限，而 AWS Glue API 則提供資料目錄 API 和受管理的基礎架構，用於在您的資料上定義、排程和執行 ETL 作業。

如需 AWS Glue API 的相關資訊，請參閱AWS Glue 開發人員指南。若要取得有關使用的資訊 AWS CLI，請參閱《AWS CLI 指令參考》。

其他 AWS 服務

Lake Formation 使用以下服務：

- AWS Glue以協調工作和編目器，以使用轉換來轉換資料。AWS Glue
- IAM 將許可政策授予 Lake Formation 校長。Lake Formation 許可模型增強了 IAM 許可模型，以保護您的資料湖。

Lake Formation 術語

以下是您在本指南中會遇到的一些重要術語。

資料湖

資料湖是您存放在 Amazon S3 中並由 Lake Formation 使用資料目錄管理的持續性資料。資料湖通常會儲存下列項目：

- 結構化和非結構化資料
- 原始資料和轉換資料

若要讓 Amazon S3 路徑位於資料湖內，必須向 Lake Formation 註冊。

資料存取

Lake Formation 透過擴充 AWS Identity and Access Management (IAM) 政策的新授權/撤銷許可模型，提供安全且精細的資料存取。

分析師和資料科學家可以使用完整的 AWS 分析和機器學習服務產品組合 (例如 Amazon Athena) 來存取資料。設定的 Lake Formation 安全性原則有助於確保使用者只能存取其授權存取的資料。

混合存取模式

Hyrbid 存取模式可讓您使用 Lake Formation 許可和 IAM 和 Amazon S3 許可來保護和存取已編目的資料。混合式存取模式可讓資料管理員選擇性地逐步上載 Lake Formation 權限，一次專注於一個資料湖使用案例。

藍圖

藍圖是一種資料管理範本，可讓您輕鬆地將資料內嵌到資料湖中。Lake Formation 提供數個藍圖，每個藍圖用於預先定義的來源類型，例如關聯式資料庫或 AWS CloudTrail 記錄檔。您可以從藍圖建立工作流程。工作流程包含 AWS Glue 編目器、工作和觸發器，這些工作是為了協調資料的載入和更新而產生的。藍圖會將資料來源、資料目標和排程視為輸入，以設定工作流程。

工作流程

工作流程是一組相關AWS Glue工作、編目器和觸發器的容器。您可以在 Lake Formation 中建立工作流程，並在AWS Glue服務中執行。Lake Formation 可以將工作流程作為單一實體進行追蹤。

定義工作流程時，您可以選取工作流程所依據的藍圖。然後，您可以根據需要或按排程執行工作流程。

您在 Lake Formation 中建立的工作流程在AWS Glue主控台中會顯示為有向無環圖 (DAG)。使用 DAG，您可以追蹤工作流程的進度並執行疑難排解。

Data Catalog

資料目錄是您的永久性中繼資料存放區。這是一項託管服務，可讓您以在 Apache Hive 中繼存放區中相同的方式在 AWS 雲端中儲存、註解和共用中繼資料。它提供了一個統一的存儲庫，其中不同的系統可以存儲和查找元數據以跟踪數據孤島中的數據，然後使用該元數據查詢和轉換數據。Lake Formation 使用資AWS Glue料目錄來儲存有關資料湖、資料來源、轉換和目標的中繼資料。

有關資料來源和目標的中繼資料採用資料庫和表格的形式。表格儲存結構定義資訊、位置資訊等。數據庫是表的集合。Lake Formation 提供權限階層，以控制對資料目錄中資料庫和表格的存取。

每個 AWS 帳戶每個 AWS 區域都有一個資料目錄。

基礎資料

基礎資料是指「資料目錄」表格所指向的資料湖中的來源資料或資料。

Principal

主體是 AWS Identity and Access Management (IAM) 使用者或角色，或是作用中目錄使用者。

資料湖管理員

資料湖管理員是可以授與任何資料目錄資源或資料位置的任何權限的任何主參與者 (包括自己) 的主參與者。指定資料湖管理員做為「資料目錄」的第一個使用者。然後，此使用者可以將更精細的資源權限授與其他主體。

> **Note**
>
> IAM 管理使用者 (具有`AdministratorAccess` AWS 受管政策的使用者) 不會自動成為資料湖管理員。例如，除非已授與目錄物件的 Lake Formation 權限，否則他們無法授與 Lake Formation 權限。但是，他們可以使用 Lake Formation 主控台或 API 將自己指定為資料湖管理員。

如需有關資料湖管理員權能的資訊，請參閱隱含 Lake Formation 權限。如需有關將使用者指定為資料湖管理員的資訊，請參閱建立資料湖管理員。

AWS 與 Lake Formation 的服務集成

您可以使用 Lake Formation 對存放在 Amazon S3 中的資料管理資料庫、表格和欄層級存取許可。您的數據在 Lake Formation 註冊後，您可以使用 AWS 分析服務 AWS Glue，如 Amazon Athena，Amazon Redshift Spectrum，Amazon EMR 來查詢數據。以下 AWS 服務與 Lake Formation 權限整合 AWS Lake Formation 並獲得榮譽。

AWS 服務	整合細節
AWS Glue	參考主題：AWS Lake Formation 搭配使用 AWS Glue AWS Glue和 Lake Formation 共享相同的數據目錄。對於控制台操作（例如查看表格列表）和所有 API 操作，AWS Glue用戶只能訪問他們具有 Lake Formation 權限的數據庫和表格。
Amazon Athena	參考主題：使 AWS Lake Formation 用 Amazon Athena

AWS 服務	整合細節
	使用 Lake Formation 允許或拒絕在 Amazon S3 中讀取資料的許可。當 Amazon Athena 使用者在查詢編輯器中選取AWS Glue目錄時，只能查詢他們擁有 Lake Formation 權限的資料庫、資料表和欄。不支援使用資訊清單的查詢。 目前，Lake Formation 不支援管理寫入作業的權限，例如 `VACUUM`、`MERGE`，以`UPDATE`及開`OPTIMIZE`啟資料表格式中的資料表。 除了透過 AWS Identity and Access Management (IAM) 向 Athena 進行驗證的主體之外，Lake Formation 還支援透過 JDBC 或 ODBC 驅動程式連線並透過 SAML 進行驗證的 Athena 使用者。支援的 SAML 提供者包括 Okta 和 Microsoft 作用中目錄同盟服務 (AD FS)。
Amazon Redshift Spectrum	參考主題：AWS Lake Formation 與 Amazon Redshift Spectrum 一起使用 當 Amazon Redshift 使用者在中的資料庫上建立外部結構描述時 AWS Glue Data Catalog，他們只能查詢他們擁有 Lake Formation 權限的結構描述中的資料表和欄。
Amazon QuickSight 企業版	參考：使 AWS Lake Formation 用 Amazon QuickSight 當 Amazon QuickSight 企業版使用者查詢 Amazon S3 位置中的資料集時，使用者必須擁有該資料的 Lake Formation `SELECT` 權限。
Amazon EMR	參考：AWS Lake Formation 與 Amazon EMR 一起使用 當您建立具有執行階段角色的 Amazon EMR 叢集時，您可以整合 Lake Formation 許可。 執行階段角色是您與 Amazon EMR 任務或查詢相關聯的 IAM 角色，然後 Amazon EMR 會使用此角色存取 AWS 資源。

Lake Formation 也與 AWS Key Management Service(AWS KMS) 搭配使用，讓您能夠更輕鬆地設定這些整合服務，以加密和解密 Amazon Simple Storage Service (Amazon S3) 位置中的資料。

其他 Lake Formation 資源

主題

- 部落格
- 技術講座和網絡研討
- 現代, 天, 建築
- 資料網格資源
- 最佳做法指南

部落格

- AWS Lake Formation 2022 年度回顧
- 高度彈性的多區域現代資料架構
- 使用 LF 標籤跨帳戶共用以指導 IAM 主體
- Lake Formation 權限庫存儀表板
- 事件驅動資料網格

技術講座和網絡研討

- RE: 創新 2020 — 資料湖：輕鬆建置、保護並與之共用 AWS Lake Formation
- RE: 2022 年發明 — 在 Amazon S3 上構建和運營數據庫
- AWS 峰會 SF 2022 — 了解並實現現代數據架構
- AWS 2022 年 ATL 峰會 — 現代資料湖 AWS Lake Formation，Amazon Redshift 和 AWS Glue
- AWS 澳新銀行 2022 年峰會 — 資料湖、湖泊房屋和資料網：什麼、為什麼以及如何？
- AWS 線上技術講座 — 簡化資料湖中的權限和治理

現代, 天, 建築

- 現代日, 建築學, 圖

資料網格資源

- 使用 AWS Lake Formation 以標籤為基礎的存取控制，大規模建置現代資料架構和資料網格模式
- 摩根大通如何構建數據網格架構以推動重要價值以增強其企業數據平台
- 建置資料網格 AWS

最佳做法指南

- AWS Lake Formation 最佳做法指南

開始使用 Lake Formation

我們建議您從下列各節開始著手：

- AWS Lake Formation：運作方式— 瞭解基本術語以及各種元件的互動方式。
- 開始使用 Lake Formation— 取得必要條件的相關資訊，並完成重要的設定工作。
- 教學課程— 按照 step-by-step 教程學習如何使用 Lake Formation。
- 中的安全性 AWS Lake Formation— 瞭解如何協助安全地存取 Lake Formation 中的資料。

開始使用 Lake Formation

如果您尚未註冊 AWS 或需要開始使用的協助，請務必完成下列工作。

主題

完成初始 AWS 設定工作

若要使用 AWS Lake Formation，您必須先完成以下任務：

主題

註冊一個 AWS 帳戶

如果您沒有 AWS 帳戶，請完成以下步驟來建立一個。

若要註冊成為 AWS 帳戶

1. 開啟 https://portal.aws.amazon.com/billing/signup。
2. 請遵循線上指示進行。

   部分註冊程序需接收來電，並在電話鍵盤輸入驗證碼。

   當您註冊時 AWS 帳戶，會建立AWS 帳戶根使用者一個。根使用者有權存取該帳戶中的所有 AWS 服務 和資源。安全性最佳做法是將管理存取權指派給使用者，並僅使用 root 使用者來執行需要 root 使用者存取權的工作。

AWS 註冊過程完成後，會向您發送確認電子郵件。您可以隨時登錄 https://aws.amazon.com/ 並選擇我的帳戶，以檢視您目前的帳戶活動並管理帳戶。

建立具有管理權限的使用者

註冊後，請保護 AWS 帳戶 AWS 帳戶根使用者、啟用和建立系統管理使用者 AWS IAM Identity Center，這樣您就不會將 root 使用者用於日常工作。

保護您的 AWS 帳戶根使用者

1. 選擇 Root 使用者並輸入您的 AWS 帳戶 電子郵件地址，以帳戶擁有者身分登入。AWS Management Console在下一頁中，輸入您的密碼。

   如需使用根使用者登入的說明，請參閱 AWS 登入 使用者指南中的以根使用者身分登入。
2. 若要在您的根使用者帳戶上啟用多重要素驗證 (MFA)。

   如需指示，請參閱《IAM 使用者指南》中的為 AWS 帳戶 根使用者啟用虛擬 MFA 裝置 (主控台)。

建立具有管理權限的使用者

1. 啟用 IAM Identity Center。

   如需指示，請參閱 AWS IAM Identity Center 使用者指南中的啟用 AWS IAM Identity Center。
2. 在 IAM 身分中心中，將管理存取權授予使用者。

   若要取得有關使用 IAM Identity Center 目錄 做為身分識別來源的自學課程，請參閱《使用指南》IAM Identity Center 目錄中的「以預設值設定使用AWS IAM Identity Center 者存取」。

以具有管理權限的使用者身分登入

- 若要使用您的 IAM Identity Center 使用者簽署，請使用建立 IAM Identity Center 使用者時傳送至您電子郵件地址的簽署 URL。

  如需使用 IAM 身分中心使用者登入的說明，請參閱使用AWS 登入 者指南中的登入 AWS 存取入口網站。

指派存取權給其他使用者

1. 在 IAM 身分中心中，建立遵循套用最低權限許可的最佳做法的權限集。

如需指示，請參閱《AWS IAM Identity Center 使用指南》中的「建立權限集」。

2. 將使用者指派給群組，然後將單一登入存取權指派給群組。

   如需指示，請參閱《AWS IAM Identity Center 使用指南》中的「新增群組」。

授與程式設計存取權

如果使用者想要與 AWS 之外的 AWS Management Console. 授與程式設計存 AWS取權的方式取決於正在存取的使用者類型。

若要授與使用者程式設計存取權，請選擇下列其中一個選項。

哪個使用者需要程式設計存取權？	到	By
人力身分 (IAM Identity Center 中管理的使用者)	使用臨時登入資料來簽署對 AWS CLI、 AWS SDK 或 AWS API 的程式設計要求。	請依照您要使用的介面所提供的指示操作。 • 如需詳細資訊 AWS CLI，請參閱《使 AWS CLI 用AWS Command Line Interface者指南》 AWS IAM Identity Center中的〈配置使用〉。 • 如需 AWS SDK、工具和 AWS API，請參閱 AWS SDK 和工具參考指南中的 IAM 身分中心身分驗證。
IAM	使用臨時登入資料來簽署對 AWS CLI、 AWS SDK 或 AWS API 的程式設計要求。	遵循《IAM 使用者指南》中的〈將臨時登入資料搭配 AWS 資源使用〉中的指示
IAM	(不建議使用) 使用長期認證來簽署對 AWS CLI、 AWS SDK 或 AWS API 的程式設計要求。	請依照您要使用的介面所提供的指示操作。 • 如需相關資訊 AWS CLI，請參閱使用指南中的使用

哪個使用者需要程式設計存取權？	到	By
		IAM 使用者登入資料進行驗證。AWS Command Line Interface • 對於 AWS SDK 和工具，請參閱 AWS SDK 和工具參考指南中的使用長期憑據進行身份驗證。 • 如需 AWS API，請參閱 IAM 使用者指南中的管理 IAM 使用者的存取金鑰。

設定 AWS Lake Formation

以下各節提供了有關首次設置 Lake Formation 的信息。並非本節中的所有主題都需要開始使用 Lake Formation。您可以使用指示設定 Lake Formation 許可模型，以管理 Amazon Simple Storage Service (Amazon S3) 中的現有 AWS Glue Data Catalog 物件和資料位置。

1. 建立資料湖管理員
2. 變更預設權限模型或使用混合存取模式
3. the section called “為您的資料湖設定 Amazon S3 位置”
4. the section called “將權限分配給 Lake Formation 用戶”
5. the section called “整合 IAM 身分識別中心”
6. the section called “(選擇性) 外部資料篩選設定”
7. the section called “(選擇性) 授與資料目錄加密金鑰的存取權”
8. (選擇性) 為工作流程建立 IAM 角色

本節介紹如何以兩種不同的方式設置 Lake Formation 資源：

- 使用 AWS CloudFormation 範本
- 使用 Lake Formation 控制台

要使用 AWS 控制台設置 Lake Formation，請轉到建立資料湖管理員。

使用 AWS CloudFormation 模板設置 Lake Formation 資源

> **Note**
>
> AWS CloudFormation 堆疊會執行上述步驟 1 到 6，但步驟 2 和 5 除外。從 Lake Formation 控制台進行變更預設權限模型或使用混合存取模式並the section called “整合 IAM 身分識別中心”手動執行。

1. 以美國東部 (維吉尼亞北部) 區域的 IAM 管理員身分登入 AWS CloudFormation 主控台，網址為 https://console.aws.amazon.com/cloudformation。
2. 選擇「啟動堆疊」。
3. 在「建立堆疊」畫面中選擇「下一步」
4. 輸入堆疊名稱。
5. 在DatalakeAdminName和中 DatalakeAdminPassword，輸入資料湖管理員使用者的使用者名稱和密碼。
6. 對於 DatalakeUser1Name 和 DatalakeUser1Password，請輸入資料湖分析師使用者的使用者名稱和密碼。
7. 針對 DataLakeBucketName，輸入要建立的新值區名稱。
8. 選擇下一步。
9. 在下一頁上，選擇 [下一步]。
10. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
11. 選擇建立。

    堆疊建立最多可能需要兩分鐘的時間。

清除資源

如果您想清理堆 AWS CloudFormation 棧資源：

1. 取消註冊堆疊建立並註冊為資料湖位置的 Amazon S3 儲存貯體。
2. 刪除 AWS CloudFormation 堆疊。這將刪除堆棧創建的所有資源。

建立資料湖管理員

資料湖管理員最初是唯 AWS Identity and Access Management 一可以將資料位置和資料目錄資源的 Lake Formation 權限授與任何主體 (包括自己) 的 (IAM) 使用者或角色。如需有關資料湖管理員權能的更多資訊，請參閱隱含 Lake Formation 權限。根據預設，Lake Formation 允許您建立多達 30 個資料湖管理員。

您可以使用湖泊形成控制台或湖泊形成 API 的`PutDataLakeSettings`操作來創建數據 Lake Formation 管理員。

建立資料湖管理員需要下列權限。`Administrator`使用者具有隱含的這些權限。

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

如果您授與使用者`AWSLakeFormationDataAdmin`原則，該使用者將無法建立其他 Lake Formation 管理員使用者。

若要建立資料湖管理員 (主控台)

1. 如果要成為資料湖管理員的使用者尚不存在，請使用 IAM 主控台建立該使用者。否則，請選擇要成為資料湖管理員的現有使用者。

   > **Note**
   >
   > 建議您不要選取 IAM 管理使用者 (具有`AdministratorAccess` AWS 受管政策的使用者) 做為資料湖管理員。

   將下列 AWS 受管理的策略附加到使用者：

政策	強制性？	備註
`AWSLakeFormationDataAdmin`	強制性	基本資料湖管理員權限。此 AWS 受管政策包含 Lake Formation API 作業的明確拒絕，`PutDataLakeSetting` 限制使用者建立新的資料湖管理員。

政策	強制性？	備註
AWSGlueConsoleFullAccess ，CloudWatchLogsReadOnlyAccess	選用	如果資料湖管理員將對從 Lake Formation 藍圖建立的工作流程進行疑難排解，請附加這些原則。這些原則可讓資料湖管理員在AWS Glue主控台和主控台中檢視疑難排解資訊。 Amazon CloudWatch Logs 如需工作流程的資訊，請參閱the section called “使用工作流程匯入資”。
AWSLakeFormationCrossAccountManager	選用	附加此原則可讓資料湖管理員授與和撤銷資料目錄資源的跨帳戶權限。如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。
AmazonAthenaFullAccess	選用	如果資料湖管理員將在中執行查詢，請附加此原則 Amazon Athena。

2. 附加下列內嵌政策，該原則會授與資料湖管理員建立 Lake Formation 服務連結角色的權限。策略的建議名稱為LakeFormationSLR。

   服務連結角色可讓資料湖管理員更輕鬆地向 Lake Formation 註冊 Amazon S3 位置。有關 Lake Formation 服務鏈接角色的更多信息，請參閱the section called “使用服務連結角色”。

   > **Important**
   >
   > 在下列所有策略中，請以<account-id>有效的 AWS 帳號取代。

```
{
    "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. (選擇性) 將下列PassRole內嵌原則附加至使用者。此原則可讓資料湖管理員建立和執行工作流程。此iam:PassRole權限可讓工作流程擔任建立爬行者程式和工作的角色LakeFormationWorkflowRole，以及將角色附加至建立的爬行者程式和工作。策略的建議名稱為UserPassRole。

> **Important**
>
> 以<account-id>有效的 AWS 帳號取代。

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

4. （可選）如果您的帳戶將授予或接收跨帳戶 Lake Formation 權限，請附加此額外的內嵌政策。此原則可讓資料湖管理員檢視並接受 AWS Resource Access Manager (AWS RAM) 資源共用邀請。此外，對於 AWS Organizations 管理帳戶中的資料湖管理員，此原則包含啟用跨帳戶授與組織的權限。如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。

   策略的建議名稱為RAMAccess。

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

5. 在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以您在其中建立的系統管理員使用者身分登入，建立具有管理權限的使用者或以具有使用者 AWS 管理策略的AdministratorAccess使用者身分登入。
6. 如果出現「歡迎使用 Lake Formation」視窗，請選擇您在步驟 1 中建立或選取的 IAM 使用者，然後選擇 [開始使用]。
7. 如果您沒有看到「歡迎使用 Lake Formation」視窗，請執行以下步驟來配置 Lake Formation 管理員。

   a. 在瀏覽窗格的 [系統管理員] 下，選擇 [系統管理角色和工作]。在主控台頁面的 [資料湖管理員] 區段中，選擇 [新增]。

   b. 在 [新增管理員] 對話方塊的 [存取類型] 下，選擇 [資料湖管理員]。

   c. 對於 IAM 使用者和角色，請選擇您在步驟 1 中建立或選取的 IAM 使用者，然後選擇 [儲存]。

變更預設權限模型或使用混合存取模式

Lake Formation 從啟用與現有 AWS Glue Data Catalog 行為兼容的「僅使用 IAM 訪問控制」設置開始。此設定可讓您透過 IAM 政策和 Amazon S3 儲存貯體政策，管理對資料湖及其中繼資料中資料的存取。

為了簡化將資料湖許可從 IAM 和 Amazon S3 模型轉換為 Lake Formation 許可，我們建議您對資料目錄使用混合存取模式。使用混合式存取模式，您將擁有一個增量路徑，您可以在其中為一組特定使用者啟用 Lake Formation 權限，而不會中斷其他現有使用者或工作負載。

如需詳細資訊，請參閱 混合存取模式。

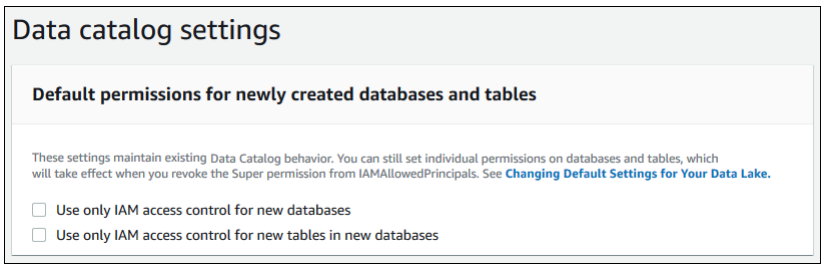
停用預設設定，只需一個步驟，即可將表格的所有現有使用者移至 Lake Formation。

> **Important**
>
> 如果您有現有的AWS Glue Data Catalog資料庫和資料表，請勿遵循本節中的指示。或者，請遵循the section called “將AWS Glue資料權限升級至 Lake Formation 型模型”中的說明進行。

> **Warning**
>
> 如果您已在資料目錄中建立資料庫和表格的自動化，則下列步驟可能會導致自動化和下游擷取、轉換和載入 (ETL) 工作失敗。只有在修改現有程序或將明確的 Lake Formation 權限授與必要主參與者之後，才繼續執行。如需 Lake Formation 權限的相關資訊，請參閱the section called “Lake Formation 權限參考”。

變更預設資料目錄設定的步驟

1. 繼續在 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/. 請確定您以您在其中建立的系統管理員使用者身分登入，建立具有管理權限的使用者或以具有`AdministratorAccess` AWS 受管理策略的使用者身分登入。
2. 修改「資料目錄」設定：
   a. 在導覽窗格的 [系統管理] 下，選擇 [資料目錄設定]。
   b. 清除這兩個核取方塊並選擇 [儲存]。



3. 撤銷資料庫建立者的`IAMAllowedPrincipals`權限。

   a. 在功能窗格的 [系統管理] 下，選擇 [系統管理角色和工作]。

   b. 在 [管理角色和工作] 主控台頁面的 [資料庫建立者] 區段中，選取`IAMAllowedPrincipals`群組，然後選擇 [撤銷]。

      [撤銷權限] 對話方塊隨即出現，並顯示`IAMAllowedPrincipals`具有 [建立資料庫] 權限。

   c. 選擇「撤銷」。

將權限分配給 Lake Formation 用戶

建立可存取中資料湖的使用者 AWS Lake Formation。此使用者具有查詢資料湖的最低權限。

如需建立使用者或群組的詳細資訊，請參閱 IAM 使用者指南中的 IAM 身分。

將權限附加至非系統管理員使用者以存取 Lake Formation 資料

1. 在以下位置開啟 IAM 主控台，https://console.aws.amazon.com/iam並以您在其中建立的管理員使用者身分登入，建立具有管理權限的使用者或以`AdministratorAccess` AWS 受管政策的使用者身分登入。
2. 選擇使用者或使用者群組。
3. 在清單中，選擇要內嵌政策的使用者或群組名稱。

   選擇許可。
4. 選擇 [新增權限]，並選擇 [直接附加原則]。`Athena`在篩選策略文字欄位中輸入。在結果清單中，勾選的方塊`AmazonAthenaFullAccess`。
5. 選擇 [建立原則] 按鈕。在建立政策頁面上，選擇 JSON 標籤。將下列程式碼複製並貼到原則編輯器中。

```
{
    "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. 選擇底部的「下一步」按鈕，直到看到「檢閱政策」頁面為止。輸入策略的名稱，例如`DatalakeUserBasic`。選擇「建立策略」，然後關閉「策略」標籤或瀏覽器視窗。

為您的資料湖設定 Amazon S3 位置

若要使用 Lake Formation 管理和保護資料湖中的資料，您必須先註冊 Amazon S3 位置。當您註冊一個位置時，Amazon S3 路徑和該路徑下的所有資料夾都會註冊，這可讓 Lake Formation 強制執行儲存層級許可。當使用者向 Amazon Athena 等整合式引擎請求資料時，Lake Formation 會提供資料存取，而不是使用使用者許可。

註冊位置時，您可以指定 IAM 角色，以授與該位置的讀取/寫入許可。Lake Formation 在提供臨時登入資料給要求存取已註冊 Amazon S3 位置資料的整合 AWS 服務時擔任該角色。您可以指定 Lake Formation 服務連結角色 (SLR) 或建立自己的角色。

在下列情況下使用自訂角色：

- 您計劃在 Amazon CloudWatch 日誌中發布指標。除了 SLR 權限之外，使用者定義角色還必須包含用於在 CloudWatch 記錄檔中新增記錄檔和發佈指標的原則。如需授與必要 CloudWatch 權限的範例內嵌政策，請參閱用於註冊地點的角色需求。

- Amazon S3 位置存在於不同的帳戶中。如需詳細資訊，請參閱 the section called “在另一個 AWS 帳戶中註冊 Amazon S3 位置”。
- Amazon S3 位置包含使用 AWS 受管金鑰. 如需詳細資訊，請參閱 註冊加密的 Amazon S3 位置 和 跨 AWS 帳戶註冊加密的 Amazon S3 位置。
- 您計劃使用 Amazon EMR 訪問 Amazon S3 位置。如需有關角色需求的詳細資訊，請參閱《Amazon EMR 管理指南》中的 Lake Formation 的 IAM 角色。

您選擇的角色必須具有必要的權限，如中所述用於註冊地點的角色需求。如需有關如何註冊 Amazon S3 位置的指示，請參閱將 Amazon S3 位置新增至您的資料湖。

(選擇性) 外部資料篩選設定

如果您打算使用第三方查詢引擎分析和處理資料湖中的資料，則必須選擇加入以允許外部引擎存取由 Lake Formation 管理的資料。如果您沒有選擇加入，外部引擎將無法存取在 Lake Formation 註冊的 Amazon S3 位置中的資料。

Lake Formation 支持列級權限，以限制對表中特定列的訪問。整合式分析服務 (例如 Amazon Athena Amazon Redshift Spectrum 和 Amazon EMR) 會從中擷取未篩選的表格中繼資料。 AWS Glue Data Catalog查詢回應中欄的實際篩選是整合式服務的責任。協力廠商管理員有責任妥善處理權限，以避免未經授權存取資料。

選擇加入以允許第三方引擎存取和篩選資料（主控台）

1. 繼續在 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/. 確保您以主體的身分登入，且該主體具有 Lake Formation `PutDataLakeSettings` API 作業的 IAM 許可。您在中建立的 IAM 管理員使用者註冊一個 AWS 帳戶具有此權限。
2. 在功能窗格的 [系統管理] 下，選擇 [應用程式整合設定]。
3. 在 [應用程式整合設定] 頁面上，執行下列動作：
   a. 勾選「允許外部引擎篩選在 Lake Formation 註冊之 Amazon S3 位置中的資料」方塊。
   b. 輸入為第三方引擎定義的工作階段標籤值。
   c. 對於AWS 帳號 ID，請輸入允許第三方引擎存取在 Lake Formation 註冊地點的帳號 ID。在每個帳戶 ID 之後按 Enter 鍵。
   d. 選擇儲存。

若要允許外部引擎在沒有工作階段標籤驗證的情況下存取資料 完整表格存取的應用程式整合

(選擇性) 授與資料目錄加密金鑰的存取權

如果 AWS Glue Data Catalog 已加密，請將 AWS KMS 金鑰 AWS Identity and Access Management (IAM) 權限授與任何需要在資料目錄資料庫和表格上授與 Lake Formation 權限的主體。

如需詳細資訊，請參閱《AWS Key Management Service 開發人員指南》。

(選擇性) 為工作流程建立 IAM 角色

使用 AWS Lake Formation，您可以使用 AWS Glue 檢索器執行的工作流程來匯入資料。工作流程定義了將資料匯入資料湖的資料來源和排程。您可以使用藍圖或 Lake Formation 提供的範本輕鬆定義工作流程。

建立工作流程時，您必須為其指派一個 AWS Identity and Access Management (IAM) 角色，以授予 Lake Formation 擷取資料的必要權限。

下列程序假設熟悉 IAM。

若要為工作流程建立 IAM 角色

1. 在的位置開啟 IAM 主控台，https://console.aws.amazon.com/iam並以您在其中建立的管理員使用者身分登入，建立具有管理權限的使用者或以`AdministratorAccess` AWS 受管政策的使用者身分登入。
2. 在導覽窗格中，選擇角色，然後選擇建立角色。
3. 在 [建立角色] 頁面上，選擇 [AWS 服務]，然後選擇 [Glue]。選擇下一步。
4. 在 [新增權限] 頁面上，搜尋受AWSGlueServiceRole管理的原則，然後選取清單中原則名稱旁邊的核取方塊。然後完成 [建立角色] 精靈，命名角色`LFWorkflowRole`。若要完成，請選擇 [建立角色]。
5. 返回「角色」頁面，搜尋`LFflowRole`並選擇角色名稱。
6. 在角色 [摘要] 頁面的 [權限] 索引標籤下，選擇 [建立內嵌原則]。在「建立政策」畫面上，瀏覽至 JSON 索引標籤，然後新增下列內嵌政策。策略的建議名稱為`LakeFormationWorkflow`。

> **Important**
>
> 在下列原則中，請以<account-id>有效的 AWS 帳戶 數字取代。

```
{
```

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

以下是此原則中權限的簡短說明：

- `lakeformation:GetDataAccess`可讓工作流程建立的工作寫入目標位置。
- `lakeformation:GrantPermissions`可讓工作流程授與目標資料表的`SELECT`權限。
- `iam:PassRole`可讓服務擔任建立爬行者程式和工作 (工作流程執行個體) 的角色，並將角色附加至建立的爬行者程式和工作。`LakeFormationWorkflowRole`

7. 確認角色`LakeFormationWorkflowRole`已附加兩個原則。
8. 如果您要擷取資料湖位置之外的資料，請新增內嵌政策，授與讀取來源資料的權限。

將AWS Glue資料權限升級至 AWS Lake Formation 模型

AWS Lake Formation 權限可為資料湖中的資料提供精細的存取控制。您可以使用 Lake Formation 許可模型來管理 Amazon Simple Storage Service (Amazon S3) 中的現有 AWS Glue Data Catalog 物件和資料位置。

Lake Formation 許可模型使用粗粒度 AWS Identity and Access Management （ IAM ）許可進行 API 服務訪問。它限制了您的用戶和這些服務可以通過 Lake Formation 功能訪問的數據。相較之下，該AWS Glue模型會透過精細的存取控制 IAM 許可授予資料存取權。要進行切換，請按照本指南中的步驟進行操作。

如需詳細資訊，請參閱 Lake Formation 許可權概述 。

主題

關於升級到 Lake Formation 型權限模型

為了維持向下相容性AWS Glue，預設情況下`Super`， AWS Lake Formation 會將所有現有資AWS Glue料目錄資源的`Super`權限授與`IAMAllowedPrincipals`群組，並在啟用僅使用 IAM 存取控制設定時授與新資料目錄資源的權限。這有效地導致資料目錄資源和 Amazon S3 位置的存取僅由 AWS Identity and Access Management (IAM) 政策控制。該`IAMAllowedPrincipals`群組包括您的 IAM 政策允許存取資料目錄物件的所有 IAM 使用者和角色。此`Super`權限可讓主體在授與該作業的資料庫或表格上執行每個受支援的 Lake Formation 作業。

您可以透過在 Lake Formation 中註冊現有資料目錄資源的位置，或使用混合存取模式，開始使用 Lake Formation 來管理對資料的存取。當您以混合式存取模式註冊 Amazon S3 位置時，可以透過選擇該位置下資料庫和表格的主體來啟用 Lake Formation 許可。

為了簡化將資料湖許可從 IAM 和 Amazon S3 模型轉換為 Lake Formation 許可的過渡，我們建議您對資料目錄使用混合存取模式。使用混合式存取模式，您將擁有一個增量路徑，您可以在其中為一組特定使用者啟用 Lake Formation 權限，而不會中斷其他現有使用者或工作負載。

如需詳細資訊，請參閱 混合存取模式。

停用預設「資料目錄」設定，以在單一步驟中將表格的所有現有使用者移至 Lake Formation。

若要開始對現有AWS Glue資料目錄資料庫和表格使用 Lake Formation 權限，您必須執行下列動作：

1. 針對每個資料庫和資料表判斷使用者現有的 IAM 許可。
2. 在 Lake Formation 中複製這些權限。

3. 對於包含資料的每個 Amazon S3 位置：
   a. 撤銷參考該位置之每個「資料目錄」資源上之`IAMAllowedPrincipals`群組的`Super`權限。
   b. 向 Lake Formation 註冊位置。
4. 清理現有的精細存取控制 IAM 政策。

> **Important**
>
> 若要在轉換資料目錄的過程中新增使用者，您必須像以前一樣在 AWS Glue IAM 中設定精細許可。您還必須按照本節中的說明複製 Lake Formation 中的這些權限。如果新使用者具有本指南中所述的粗略 IAM 政策，他們可以列出任何具有授與`Super`權限的資料庫或表格。`IAMAllowedPrincipals`他們也可以檢視這些資源的中繼資料。

請按照本節中的步驟升級到 Lake Formation 權限模型。請從the section called “步驟 1：列出現有權限”開始。

步驟 1：列出使用者和角色的現有權限

若要開始對現有AWS Glue資料庫和資料表使用 AWS Lake Formation 權限，您必須先決定使用者現有的權限。

> **Important**
>
> 在開始之前，請確定您已完成中的工作開始使用。

主題

使用 API 作業

使用 AWS Identity and Access Management (IAM) ListPoliciesGrantingServiceAccessAPI 操作來決定附加到每個主體 (使用者或角色) 的 IAM 政策。從結果中傳回的政策中，您可以決定授與主體的 IAM 許可。您必須個別呼叫每個主體的 API。

Example

下列 AWS CLI 範例會傳回附加至使用者的策略`glue_user1`。

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

命令會傳回類似下列內容的結果。

```
{
    "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
}
```

使用 AWS Management Console

您也可以在 AWS Identity and Access Management (IAM) 主控台、使用者或角色「摘要」頁面的「存取顧問」索引標籤中查看此資訊：

1. 前往 https://console.aws.amazon.com/iam/ 開啟 IAM 主控台。
2. 在服務導覽窗格中，選擇 Users (使用者) 或者 Roles (角色)。
3. 在清單中選擇名稱以開啟其「摘要」頁面，然後選擇「存取建議程式」頁籤。
4. 檢查每個原則，以判斷每個使用者都有權限的資料庫、資料表和動作的組合。

在此過程中，除了使用者之外，還要檢查角色，因為您的資料處理工作可能會假定角色來存取資料。

使用 AWS CloudTrail

確定現有權限的另一種方法是查找 AWS Glue API 調 AWS CloudTrail 用，其中日誌的`additionaleventdata`字段包含一個`insufficientLakeFormationPermissions`條目。此項目列出使用者需要 Lake Formation 權限才能採取相同動作的資料庫和表格。

這些是資料存取記錄，因此不保證會產生完整的使用者及其權限清單。我們建議您選擇較寬的時間範圍，以擷取大部分使用者的資料存取模式，例如數週或數月。

如需詳細資訊，請參閱AWS CloudTrail 使用指南中的檢視具有 CloudTrail 事件歷程記錄的事件。

接下來，您可以設置 Lake Formation 權限以匹配AWS Glue權限。請參閱第 2 步：設置等效的 Lake Formation 權限。

第 2 步：設置等效的 Lake Formation 權限

使用中收集的資訊步驟 1：列出使用者和角色的現有權限，授與與 AWS Lake Formation 權限相符的 AWS Glue權限。使用下列任一方法來執行授權：

- 使用 Lake Formation 控制台或 AWS CLI.

  請參閱the section called “授與和撤銷資料目錄權限”。
- 使用`GrantPermissions`或 `BatchGrantPermissions` API 作業。

  請參閱權限 API。

如需詳細資訊，請參閱 Lake Formation 許可權概述 。

設置 Lake Formation 許可權後，繼續步驟 3：授予使用者 IAM 許可以使用 Lake Formation。

步驟 3：授予使用者 IAM 許可以使用 Lake Formation

若要使用 AWS Lake Formation 權限模型，主體必須具有 Lake Formation API 的 AWS Identity and Access Management (IAM) 許可。

在 IAM 中建立以下政策，並將其附加到需要存取資料湖的每個使用者。將政策命名為 `LakeFormationDataAccess`。

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

接下來，升級到 Lake Formation 權限一次一個數據位置。請參閱步驟 4：將資料存放區切換至 Lake Formation 型權限模型。

步驟 4：將資料存放區切換至 Lake Formation 型權限模型

每次升級至 Lake Formation 權限一個資料位置。若要這麼做，請重複此整節，直到您註冊資料目錄參考的所有 Amazon 簡單儲存服務 (Amazon S3) 路徑為止。

主題

- 驗證 Lake Formation 權限
- 保護現有資料目錄資源
- 為您的 Amazon S3 位置開啟 Lake Formation 許可

驗證 Lake Formation 權限

在註冊位置之前，請執行驗證步驟，以確保正確的主參與者具有必要的 Lake Formation 權限，並且不會將 Lake Formation 權限授與不應擁有這些權限的主參與者。使用 Lake Formation `GetEffectivePermissionsForPath` API 作業，識別參考 Amazon S3 位置的資料目錄資源，以及對這些資源具有許可的主體。

下列 AWS CLI 範例會傳回參照 Amazon S3 儲存貯體的資料目錄資料庫和表格`products`。

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

請注意profile選項。建議您以資料湖管理員身分執行命令。

以下是傳回結果的摘錄。

```
{
        "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**
>
> 如果您的AWS Glue資料目錄已加密，則只會GetEffectivePermissionsForPath傳回在 Lake Formation 正式運作之後建立或修改的資料庫和表格。

保護現有資料目錄資源

接下來IAMAllowedPrincipals，撤銷您為該位置識別的每個資料表和資料庫的Super權限。

> **Warning**
>
> 如果您已在資料目錄中建立資料庫和表格的自動化，則下列步驟可能會導致自動化和下游擷取、轉換和載入 (ETL) 工作失敗。只有在修改現有程序或將明確的 Lake Formation 權限授與

必要主參與者之後，才繼續執行。如需 Lake Formation 權限的相關資訊，請參閱the section called “Lake Formation 權限參考”。

若要在表格**IAMAllowedPrincipals**上撤銷 **Super**

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入。
2. 在導覽窗格中，選擇 Tables (資料表)。
3. 在「表格」頁面上，選取所需表格旁邊的圓鈕。
4. 在 [動作] 功能表上，選擇 [撤銷]。
5. 在 [撤銷權限] 對話方塊的 [IAM 使用者和角色] 清單中，向下捲動至 [群組] 標題，然後選擇 [IAM] AllowedPrincipals。
6. 在 [資料表權限] 底下，確定已選取 [超級]，然後選擇 [撤銷]。

若要撤**Super**銷**IAMAllowedPrincipals**資料庫

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入。
2. 在導覽窗格中，選擇 Databases (資料庫)。
3. 在「資料庫」頁面上，選取所需資料庫旁邊的圓鈕。
4. 在 Actions (動作) 功能表上，選擇 Edit (編輯)。
5. 在 [編輯資料庫] 頁面上，清除對此資料庫中的新資料表僅使用 IAM 存取控制，然後選擇 [儲存]。
6. 返回 [資料庫] 頁面，確定資料庫仍處於選取狀態，然後在 [動作] 功能表上選擇 [撤銷]。
7. 在 [撤銷權限] 對話方塊的 [IAM 使用者和角色] 清單中，向下捲動至 [群組] 標題，然後選擇 [IAM] AllowedPrincipals。
8. 在 [資料庫權限] 底下，確定已選取 [超級]，然後選擇 [撤銷]。

為您的 Amazon S3 位置開啟 Lake Formation 許可

接下來，向 Lake Formation 註冊 Amazon S3 位置。若要這麼做，您可以使用中所述的程序將 Amazon S3 位置新增至您的資料湖。或者，如中所述使用 `RegisterResource` API 操作憑證自動販賣 API。

> **Note**
>
> 如果已註冊父地點，則不需要註冊子位置。

在您完成這些步驟並測試使用者是否可以存取其資料之後，您已成功升級至 Lake Formation 權限。繼續執行下一個步驟，步驟 5：保護新的資料目錄資源。

步驟 5：保護新的資料目錄資源

接下來，透過變更預設資料目錄設定來保護所有新資料目錄資源的安全。關閉僅對新資料庫和資料表使用 AWS Identity and Access Management (IAM) 存取控制的選項。

> **Warning**
>
> 如果您已在資料目錄中建立資料庫和表格的自動化，則下列步驟可能會導致自動化和下游擷取、轉換和載入 (ETL) 工作失敗。只有在修改現有程序或將明確的 Lake Formation 權限授與必要主參與者之後，才繼續執行。如需 Lake Formation 權限的相關資訊，請參閱the section called “Lake Formation 權限參考”。

變更預設資料目錄設定的步驟

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以 IAM 管理使用者 (具有`AdministratorAccess` AWS 受管政策的使用者`Administrator`或其他使用者) 身分登入。
2. 在導覽窗格中，選擇設定。
3. 在 [資料目錄設定] 頁面上，清除這兩個核取方塊，然後選擇 [儲存]。

下一個步驟是將來授與使用者存取其他資料庫或資料表 future 權限。請參閱步驟 6：為使用者提供新的 IAM 政策，以便 future 存取資料湖。

步驟 6：為使用者提供新的 IAM 政策，以便 future 存取資料湖

若要在 future 授與使用者存取其他「資料目錄」資料庫或表格，您必須向他們提供粗粒度 AWS Identity and Access Management (IAM) 內嵌政策。將政策命名為 `GlueFullReadAccess`。

Important

如果您在資料目錄Super中IAMAllowedPrincipals的每個資料庫和表格撤銷之前將此原則附加至使用者，則該使用者可以檢視授與之任何資源的所有中Super繼資料。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

在此步驟和先前步驟中指定的內嵌政策包含最少的 IAM 許可。如需資料湖管理員、資料分析師和其他角色角色的建議政策，請參閱the section called “Lake Formation 角色和 IAM 許可參考”。

接下來，繼續步驟 7：清理現有的 IAM 政策。

步驟 7：清理現有的 IAM 政策

設定 AWS Lake Formation 許可並建立並附加粗粒度存取控制 AWS Identity and Access Management (IAM) 政策之後，請完成以下最後一個步驟：

- 從使用者、群組和角色中移除您在 Lake Formation 中複製的舊有精細存取控制 IAM 政策。

這樣可以確保這些主體不再能夠直接存取 Amazon Simple Storage Service (Amazon S3) 中的資料。然後，您可以完全透過 Lake Formation 管理這些主體的資料湖存取。

AWS Lake Formation 和介面 VPC 端端點 ()AWS PrivateLink

Amazon VPC 是一項 AWS 服務，可用於在您定義的虛擬網路中啟動 AWS 資源。您可利用 VPC 來控制您的網路設定，例如 IP 地址範圍、子網路、路由表和網路閘道。

如果您使用 Amazon Virtual Private Cloud (Amazon VPC) 託管資 AWS 源，則可以在 VPC 和 Lake Formation 之間建立私有連接。您可以使用此連接，以便 Lake Formation 可以在不通過公共互聯網的情況下與 VPC 中的資源進行通信。

您可以在 VPC 和 AWS Lake Formation 建立介面 VPC 端點之間建立私人連線。介面端點採用這項技術 AWS PrivateLink，可讓您在沒有網際網路閘道、NAT 裝置、VPN 連線或 AWS Direct Connect 連線的情況下私有存取 Lake Formation API。VPC 中的執行個體不需要公用 IP 位址即可與 Lake Formation API 進行通訊。您的 VPC 和 Lake Formation 之間的流量不會離開 Amazon 網絡。

每個介面端點都是由您子網路中的一或多個彈性網路介面表示。

如需詳細資訊，請參閱《Amazon VPC 使用者指南》中的介面 VPC 端點 (AWS PrivateLink)。

Lake Formation VPC 端點的考量

在為 Lake Formation 設定介面 VPC 端點之前，請務必先檢閱 Amazon VPC 使用者指南中的介面端點屬性和限制。

Lake Formation 支持從您的 VPC 調用其所有 API 操作。您可以在所有支援 Lake Formation 和 Amazon VPC 端點 AWS 區域 的 VPC 端點中使用 Lake Formation。

為 Lake Formation 創建接口 VPC 端點

您可以使用 Amazon VPC 主控台或 AWS Command Line Interface (AWS CLI) 為 Lake Formation 服務建立 VPC 端點。如需詳細資訊，請參閱《Amazon VPC 使用者指南》中的建立介面端點。

使用以下服務名稱為 Lake Formation 建立 VPC 端點：

- com.amazonaws.*region*.lakeformation

如果您為端點啟用私有 DNS，則可以使用該區域的預設 DNS 名稱向 Lake Formation 發出 API 請求，例如`lakeformation.us-east-1.amazonaws.com`。

如需詳細資訊，請參閱《Amazon VPC 使用者指南》中的透過介面端點存取服務。

為 Lake Formation 創建 VPC 端點策略

Lake Formation 支持 VPC 端點策略。VPC 端點政策是您在建立或修改端點時附加到端點的 AWS Identity and Access Management (IAM) 資源政策。

您可以將端點政策附加到 VPC 端點，以控制對 Lake Formation 的存取。此政策會指定下列資訊：

- 可執行動作的主體。
- 可執行的動作。
- 可供執行動作的資源。

如需詳細資訊，請參閱 Amazon VPC 使用者指南中的使用 VPC 端點控制對服務的存取。

範例：Lake Formation 行動的 VPC 端點政策

下列 Lake Formation 的 VPC 端點原則範例允許使用 Lake Formation 權限進行憑證自動販賣。您可以使用此政策，透過 Amazon Redshift 叢集或位於私有子網路中的叢集使用 Amazon EMR Lake Formation 許可執行查詢。

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

> **Note**
>
> 如果您在建立端點時未附加原則，則會附加允許完整存取服務的預設原則。

如需詳細資訊，請參閱 Amazon VPC 文件中的以下主題：

- 什麼是 Amazon VPC？
- 建立介面端點
- 使用 VPC 端點原則

教學課程

以下教學課程分為三個軌道，並提供有 step-by-step 關如何使用建置資料湖、擷取資料、共用 AWS Lake Formation和保護資料湖的指示：

1. 建置資料湖並擷取資料：學習如何建置資料湖，並使用藍圖來移動、儲存、編目、清理和組織資料。您還將學習如何設置受管理的表格。受管理的資料表是一種新的 Amazon S3 表格類型，支援原子、一致、隔離和持久 (ACID) 交易。

   在開始之前，請確定您已完成中的步驟開始使用 Lake Formation。

   - 從 AWS CloudTrail 來源建立資料湖

     使用您自己的 CloudTrail 記錄做為資料來源，建立並載入您的第一個資料湖。

   - 從 Lake Formation 的 JDBC 源創建數據湖

     使用其中一個 JDB 可存取的資料倉庫 (例如關聯式資料庫) 做為資料來源來建立資料湖。

2. 保護資料湖：瞭解如何使用標籤式和資料列層級存取控制，有效地保護和管理資料湖的存取。

   - 在 Lake Formation 中設置開放表格式存儲格式的權限

     本教程演示如何設置在 Lake Formation 開源事務表格式（阿帕奇冰山，阿帕奇胡迪和 Linux 基金會三角洲湖表）的權限。

   - 使用以 Lake Formation 標籤為基礎的存取控制來管理資料湖

     了解如何在 Lake Formation 中使用以標籤為基礎的存取控制來管理資料湖中資料的存取。

   - 使用資料列層級存取控制來保護資料湖

     瞭解如何設定列層級權限，以便您根據 Lake Formation 中的資料合規性和治理原則限制對特定資料列的存取。

3. 共用資料：瞭解如何 AWS 帳戶 使用標籤式存取控制 (TBAC) 安全地共用資料，並管理之間共用資料集的精細權限。 AWS 帳戶

   - 使用 Lake Formation 標籤式存取控制和具名資源共用資料湖

     在本教程中，您將學習如何 AWS 帳戶 使用 Lake Formation 安全地共享數據。

   - 使用 Lake Formation 精細存取控制共用資料湖

     在本教學課程中，您將學習如何在使用管理多個 AWS 帳戶 項目時，使用 Lake Formation 快速輕鬆地共用資料集 AWS Organizations。

主題

從 AWS CloudTrail 來源建立資料湖

本教學課程會引導您完成在 Lake Formation 主控台上執行的動作，以便從 AWS CloudTrail 來源建立和載入您的第一個資料湖。

建立資料湖的高階步驟

1. 將 Amazon Simple Storage Service (Amazon S3) 路徑註冊為資料湖。
2. 授與 Lake Formation 許可，以寫入資料目錄和資料湖中的 Amazon S3 位置。
3. 建立資料庫以組織「資料目錄」中的詮釋資料表格。
4. 使用藍圖建立工作流程。執行工作流程以從資料來源擷取資料。
5. 設定您的 Lake Formation 權限，以允許其他人管理「資料目錄」和「資料湖」中的資料。
6. 設定 Amazon Athena 以查詢您匯入 Amazon S3 資料湖的資料。
7. 對於某些資料存放區類型，請設定 Amazon Redshift Spectrum 以查詢您匯入 Amazon S3 資料湖的資料。

主題

- 步驟 5：授予資料位置權限
- 步驟 6：在「資料目錄」中建立資料庫
- 步驟 7：授予資料權限
- 步驟 8：使用藍圖建立工作流程
- 步驟 9：執行工作流程
- 第 10 步：在表格上授予選擇
- 步驟 11：使用查詢資料湖 Amazon Athena

目標對象

下表列示了本自學課程中用於建立資料湖的角色。

目標對象

角色	描述
IAM 管理員	具有受 AWS 管策略：AdministratorAccess 。可以建立 IAM 角色和 Amazon S3 儲存貯體。
資料湖管理員	可存取資料目錄、建立資料庫以及將 Lake Formation 權限授與其他使用者的使用者。IAM 許可比 IAM 管理員少，但足以管理資料湖。
資料分析	可以對資料湖執行查詢的使用者。只有足夠的權限來執行查詢。
工作流角色	具有執行工作流程所需 IAM 政策的角色。如需詳細資訊，請參閱 (選擇性) 為工作流程建立 IAM 角色。

必要條件

開始之前：

- 請確定您已完成中的工作設定 AWS Lake Formation。

- 知道 CloudTrail 日誌的位置。
- 在使用 Athena 之前，Athena 要求資料分析師角色建立 Amazon S3 儲存貯體來存放查詢結果。

假設熟悉 AWS Identity and Access Management (IAM)。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

步驟 1：建立資料分析師使用者

此使用者具有查詢資料湖的最低權限集。

1. 在 https://console.aws.amazon.com/iam 開啟 IAM 主控台。以您在中建立的系統管理員使用者身分登入，建立具有管理權限的使用者或以具有AdministratorAccess AWS 受管理策略的使用者身分登入。
2. 使用下列設定建立datalake_user名為的使用者：

   - 啟用 AWS Management Console 存取權。
   - 設定密碼，不需要重設密碼。
   - 附加受AmazonAthenaFullAccess AWS 管理的策略。
   - 附加下列內嵌政策。將政策命名為 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": "*"
```

```
        }
    ]
}
```

步驟 2：將讀取 AWS CloudTrail 記錄檔的權限新增至工作流程角色

1. 將下列內嵌原則附加至角色LakeFormationWorkflowRole。該策略授予讀取 AWS CloudTrail 日誌的權限。將政策命名為 DatalakeGetCloudTrail。

   若要建立 LakeFormationWorkflowRole 角色，請參閱(選擇性) 為工作流程建立 IAM 角色。

   > **Important**
   >
   > 以<your-s3-cloudtrail-bucket> CloudTrail 資料的 Amazon S3 位置取代。

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

2. 確認有三個原則附加至角色。

步驟 3：為資料湖建立 Amazon S3 儲存貯體

建立做為資料湖根位置的 Amazon S3 儲存貯體。

1. 在 https://console.aws.amazon.com/s3/ 開啟 Amazon S3 主控台，然後以您建立的管理員使用者身分登入建立具有管理權限的使用者。
2. 選擇 [建立值區]，然後透過精靈建立名為的值區<*yourName*>-datalake-cloudtrail，其中<yourName>是您的名字和姓氏。例如：jdoe-datalake-cloudtrail。

   如需建立 Amazon S3 儲存貯體的詳細指示，請參閱建立儲存貯體。

步驟 4：註冊 Amazon S3 路徑

將 Amazon S3 路徑註冊為資料湖的根位置。

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。以資料湖管理員身分登入。
2. 在導覽窗格中的 [註冊並擷取] 下，選擇 [資料湖位置]。
3. 選擇註冊位置，然後選擇瀏覽。
4. 選取您先前建立的*<yourName>*-datalake-cloudtrail值區，接受預設的 IAM 角色AWSServiceRoleForLakeFormationDataAccess，然後選擇 [註冊位置]。

   如需註冊位置的詳細資訊，請參閱將 Amazon S3 位置新增至您的資料湖。

步驟 5：授予資料位置權限

主參與者必須具有資料湖位置的資料位置權限，才能建立指向該位置的資料目錄表格或資料庫。您必須將資料位置許可授與工作流程的 IAM 角色，以便工作流程可以寫入資料擷取目的地。

1. 在功能窗格的 [權限] 下，選擇 [資料位置]。
2. 選擇授權，然後在授與權限對話方塊中，進行下列選擇：
   a. 對於 IAM 使用者和角色，請選擇LakeFormationWorkflowRole。
   b. 對於存儲位置，請選擇您的存儲*<yourName>*-datalake-cloudtrail桶。
3. 選擇 Grant (授予)。

如需資料位置權限的詳細資訊，請參閱Underlying data access control。

步驟 6：在「資料目錄」中建立資料庫

Lake Formation 資料目錄中的中繼資料表儲存在資料庫中。

1. 在導覽窗格的 [資料目錄] 下，選擇 [資料庫]。
2. 選擇建立資料庫，然後在資料庫詳細資訊下輸入名稱lakeformation_cloudtrail。
3. 將其他欄位保留空白，然後選擇 [建立資料庫]。

步驟 7：授予資料權限

您必須授與權限才能在「資料目錄」中建立中繼資料表。因為工作流程將與角色一起執行`LakeFormationWorkflowRole`，因此您必須將這些權限授與角色。

1. 在 Lake Formation 主控台的導覽窗格的 [資料目錄] 下，選擇 [資料庫]。
2. 選擇資`lakeformation_cloudtrail`料庫，然後從「動作」下拉式清單中選擇「權限」標題下的「授與」。
3. 在 [授與資料權限] 對話方塊中，進行下列選項：

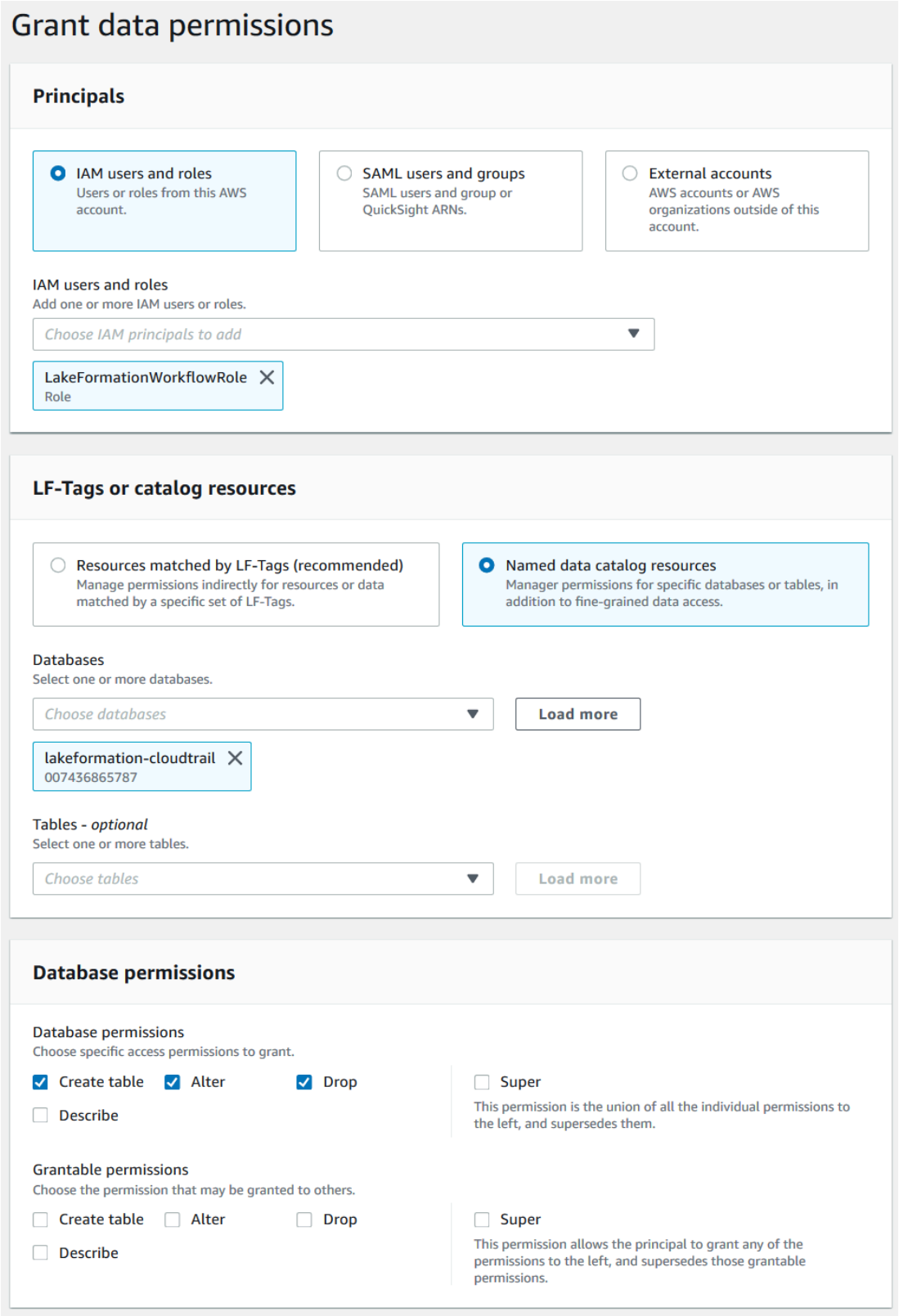
   a. 在主體下，對於 IAM 使用者和角色，選擇`LakeFormationWorkflowRole`。

   b. 在 LF 標籤或目錄資源下，選擇具名資料目錄資源。

   c. 對於數據庫，您應該看到`lakeformation_cloudtrail`數據庫已添加。

   d. 在 [資料庫權限] 底下，選取 [建立資料表]、[變更] 和 [刪除]，如果已選取，則清除超級。

   您的授予數據權限對話框現在應該看起來像這個屏幕截圖。


Grant data permissions
Principals
IAM users and roles
Users or roles from this AWS account.
SAML users and groups
SAML users and group or QuickSight ARNs.
External accounts
AWS accounts or AWS organizations outside of this account.
IAM users and roles
Add one or more IAM users or roles.
Choose IAM principals to add
LakeFormationWorkflowRole
Role
LF-Tags or catalog resources
Resources matched by LF-Tags (recommended)
Manage permissions indirectly for resources or data matched by a specific set of LF-Tags.
Named data catalog resources
Manager permissions for specific databases or tables, in addition to fine-grained data access.
Databases
Select one or more databases.
Choose databases
Load more
lakeformation-cloudtrail
007436865787
Tables - optional
Select one or more tables.
Choose tables
Load more
Database permissions
Database permissions
Choose specific access permissions to grant.
Create table
Alter
Drop
Describe
Super
This permission is the union of all the individual permissions to the left, and supersedes them.
Grantable permissions
Choose the permission that may be granted to others.
Create table
Alter
Drop
Describe
Super
This permission allows the principal to grant any of the permissions to the left, and supersedes those grantable permissions.

4. 選擇 Grant (授予)。

有關授予 Lake Formation 權限的更多信息，請參閱管理 Lake Formation 權限。

步驟 8：使用藍圖建立工作流程

為了讀取 CloudTrail 日誌，了解其結構，在數據目錄中創建適當的表，我們需要設置一個包含AWS Glue編目器，作業，觸發器和工作流程的工作流程。湖形成的藍圖簡化了這一過程。

工作流程會產生探索資料並將資料擷取到資料湖中的工作、搜尋器和觸發器。您可以根據其中一個預先定義的 Lake Formation 藍圖來建立工作流程。

1. 在「Lake Formation」主控台的導覽窗格中，選擇「藍圖」，然後選擇「使用藍圖」。
2. 在 [使用藍圖] 頁面上的 [藍圖類型] 下，選擇AWS CloudTrail。
3. 在「匯入來源」下，選擇 CloudTrail 來源和開始日期。
4. 在「匯入目標」下，指定下列參數：

目標資料庫	`lakeformation_cloudtrail`
目標儲存位置	`s3://<yourName> -datalake-cloudtrail`
資料格式	Parquet

5. 針對匯入頻率，請選擇「視需求執行」。
6. 在「匯入選項」下，指定下列參數：

工作流程名	`lakeformationcloudtrailtest`
IAM 角色	`LakeFormationWorkflowRole`
表前綴	`cloudtrailtest` **Note** 必須是小寫。

7. 選擇 [建立]，然後等待主控台回報工作流程已成功建立。

> **Tip**
>
> 您是否收到下列錯誤訊息？
> `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...`
> 如果是這樣，請檢查您是否已<account-id>在資料湖管理員使用者的內嵌原則中取代為有效的 AWS 帳戶號碼。

步驟 9：執行工作流程

因為您已指定工作流程 run-on-demand，因此您必須手動啟動工作流程。

- 在 [藍圖] 頁面上，選取工作流程`lakeformationcloudtrailtest`，然後在 [動作] 功能表上選擇 [開始]。

  當工作流程執行時，您可以在「上次執行狀態」欄中檢視其進度。偶爾選擇刷新按鈕。

  狀態會從「執行中」、「探查」、「匯入」變為「已完成」。

  工作流程完成時：

  - 資料目錄將有新的中繼資料表格。
  - 您的 CloudTrail 記錄會擷取到資料湖中。

  如果工作流程失敗，請執行下列動作：

  a. 選取工作流程，然後在 [動作] 功能表上選擇 [檢視圖表]。

     工作流程會在AWS Glue主控台中開啟。

  b. 確認已選取工作流程，然後選擇 History (歷史記錄) 標籤。

  c. 在歷史記錄下，選擇最近的運行，然後選擇查看運行詳細信息。

  d. 在動態 (程式實際執行) 圖形中選取失敗的工作或爬行者程式，然後複查錯誤訊息。失敗的節點可能是紅色或黃色。

第 10 步：在表格上授予選擇

您必須授與新「資料目錄」表格的SELECT權限，以便資料分析師可以查詢資料表所指向的資料。

> **Note**
>
> 工作流程會自動將其建立之資料表的SELECT權限授與執行該工作流程的使用者。由於資料湖管理員執行此工作流程，因此您必SELECT須授與資料分析師。

1. 在 Lake Formation 主控台的導覽窗格的 [資料目錄] 下，選擇 [資料庫]。
2. 選擇資lakeformation_cloudtrail料庫，然後從「動作」下拉式清單中選擇「權限」標題下的「授與」。
3. 在 [授與資料權限] 對話方塊中，進行下列選項：
   a. 在主體下，對於 IAM 使用者和角色，選擇datalake_user。
   b. 在 LF 標籤或目錄資源下，選擇具名資料目錄資源。
   c. 對於「資料lakeformation_cloudtrail庫」，應該已選取資料庫。
   d. 對於「表格」，請選擇cloudtrailtest-cloudtrail。
   e. 在 [資料表和欄權限] 下，選擇 [選取]。
4. 選擇 Grant (授予)。

下一個步驟是以資料分析師的身分執行。

步驟 11：使用查詢資料湖 Amazon Athena

使用 Amazon Athena 主控台查詢資 CloudTrail 料湖中的資料。

1. 在 https://console.aws.amazon.com/athena/ 開啟 Athena 主控台，然後以資料分析師和使用者身分登入datalake_user。
2. 如有必要，請選擇「開始使用」以繼續使用 Athena 查詢編輯器。
3. 對於 Data source (資料來源)，請選擇 AwsDataCatalog。
4. 針對 Database (資料庫)，輸入 lakeformation_cloudtrail。

   「表格」清單會填入。

5. 在表格旁邊的溢位功能表（3 個點水平排列）中`cloudtrailtest-cloudtrail`，選擇「預覽表格」，然後選擇「執行」。

   查詢會執行並顯示 10 列資料。

   如果您之前沒有使用過 Athena，則必須先在 Athena 主控台中設定 Amazon S3 位置以存放查詢結果。`datalake_user`必須具有必要的許可才能存取您選擇的 Amazon S3 儲存貯體。

> **Note**
>
> 現在您已經完成教學課程，請將資料權限和資料位置權限授與組織中的主參與者。

從 Lake Formation 的 JDBC 源創建數據湖

本教學課程將引導您完成在 AWS Lake Formation 主控台上執行的步驟，以使用 Lake Formation 從 JDBC 來源建立和載入第一個資料湖。

主題

目標對象

下表列出了在本 AWS Lake Formation JDBC 教程中使用的角色。

角色	描述
IAM 管理員	可建立 AWS Identity and Access Management (IAM) 使用者和角色以及 Amazon Simple Storage Service (Amazon S3) 貯體的使用者。具有受`AdministratorAccess` AWS 管策略。
資料湖管理員	可存取「資料目錄」、建立資料庫以及將 Lake Formation 權限授與其他使用者的使用者。IAM 許可比 IAM 管理員少，但足以管理資料湖。
資料分析	可對資料湖執行查詢的使用者。只有足夠的權限來執行查詢。
工作流角色	具有執行工作流程所需 IAM 政策的角色。

如需完成自學課程之先決條件的相關資訊，請參閱JDBC 教程的先決條件。

JDBC 教程的先決條件

在您開始 AWS Lake Formation JDBC 教學課程之前，請確定您已完成下列工作：

- 完成 開始使用 Lake Formation 中的任務。
- 決定要用於自學課程的 JDB 可存取資料倉庫。
- 收集建立 JDBC 類型之AWS Glue連線所需的資訊。此資料目錄物件包括資料存放區的 URL、登入登入資料，以及如果資料存放區是在 Amazon 虛擬私人雲端 (Amazon VPC) 中建立的，則需要額外的 VPC 特定組態資訊。如需詳細資訊，請參閱AWS Glue 開發人員指南中的〈資AWS Glue料目錄〉中的〈定義連線〉。

本教學課程假設您熟悉 AWS Identity and Access Management (IAM)。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

若要開始使用，請繼續執行the section called “步驟 1：建立資料分析師使用者”。

步驟 1：建立資料分析師使用者

在此步驟中，您會建立一個 AWS Identity and Access Management (IAM) 使用者，做為中資料湖的資料分析師 AWS Lake Formation。

此使用者具有查詢資料湖的最低權限集。

1. 在 https://console.aws.amazon.com/iam 開啟 IAM 主控台。以您在中建立的系統管理員使用者身分登入，建立具有管理權限的使用者或以具有AdministratorAccess AWS 受管理策略的使用者身分登入。
2. 使用下列設定建立datalake_user名為的使用者：

   - 啟用 AWS Management Console 存取權。
   - 設定密碼，不需要重設密碼。
   - 附加受AmazonAthenaFullAccess AWS 管理的策略。
   - 附加下列內嵌政策。將政策命名為 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": "*"
        }
    ]
}
```

步驟 2：建立連線 AWS Glue

> **Note**
>
> 如果您已經與 JDBC 資料來源建立AWS Glue連線，請略過此步驟。

AWS Lake Formation 透過AWS Glue連線存取 JDBC 資料來源。連線是資料目錄物件，其中包含連線至資料來源所需的所有資訊。您可以使用AWS Glue控制台建立連線。

建立連線

1. 在開啟主控台https://console.aws.amazon.com/glue/，然AWS Glue後以您在中建立的系統管理員使用者身分登入建立具有管理權限的使用者。
2. 在導覽窗格中，於 Data catalog (Data Catalog) 下選擇 Connections (連線)。
3. 在 Connectors (連接器) 頁面上，選擇 Create custom connector (建立自訂連接器)。
4. 在 [連線器內容] 頁面上**`datalake-tutorial`**，輸入連線名稱，然後選擇 JDBC 做為連線類型。然後選擇下一步。
5. 繼續執行連線精靈並儲存連線。

   如需有關建立連線的資訊，請參閱AWS Glue 開發人員指南中的 AWS Glue JDBC 連線屬性。

步驟 3：為資料湖建立 Amazon S3 儲存貯體

在此步驟中，您將建立做為資料湖根位置的 Amazon 簡單儲存服務 (Amazon S3) 儲存貯體。

1. 在 https://console.aws.amazon.com/s3/ 開啟 Amazon S3 主控台，然後以您建立的管理員使用者身分登入建立具有管理權限的使用者。
2. 選擇 [建立值區]，然後透過精靈建立名為的值區`<yourName>-datalake-tutorial`，其中<yourName>是您的名字和姓氏。例如：`jdoe-datalake-tutorial`。

   如需建立 Amazon S3 儲存貯體的詳細指示，請參閱如何建立 S3 儲存貯體？ 在 Amazon 簡單存儲服務用戶指南。

步驟 4：註冊 Amazon S3 路徑

在此步驟中，您將 Amazon Simple Storage Service (Amazon S3) 路徑註冊為資料湖的根位置。

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。以資料湖管理員身分登入。
2. 在導覽窗格中的 [註冊並擷取] 下，選擇 [資料湖位置]。
3. 選擇 [註冊位置]，然後選擇 [瀏覽]。
4. 選取您先前建立的*<yourName>*-datalake-tutorial值區，接受預設的 IAM 角色AWSServiceRoleForLakeFormationDataAccess，然後選擇 [註冊位置]。

   如需註冊位置的詳細資訊，請參閱將 Amazon S3 位置新增至您的資料湖。

步驟 5：授予資料位置權限

主參與者必須具有資料湖位置的資料位置權限，才能建立指向該位置的資料目錄表格或資料庫。您必須將資料位置許可授與工作流程的 IAM 角色，以便工作流程可以寫入資料擷取目的地。

1. 在 Lake Formation 主控台的導覽窗格的 [權限] 下，選擇 [資料位置]。
2. 選擇 [授權]，然後在 [授與權限] 對話方塊中執行下列動作：
   a. 對於 IAM 使用者和角色，請選擇LakeFormationWorkflowRole。
   b. 對於存儲位置，請選擇您的存儲*<yourName>*-datalake-tutorial桶。
3. 選擇 Grant (授予)。

如需資料位置權限的詳細資訊，請參閱Underlying data access control。

步驟 6：在「資料目錄」中建立資料庫

Lake Formation 資料目錄中的中繼資料表儲存在資料庫中。

1. 在 Lake Formation 主控台的導覽窗格的 [資料目錄] 下，選擇 [資料庫]。
2. 選擇建立資料庫，然後在資料庫詳細資訊下輸入名稱lakeformation_tutorial。
3. 將其他欄位保留空白，然後選擇 [建立資料庫]。

步驟 7：授予資料權限

您必須授與權限才能在「資料目錄」中建立中繼資料表。因為工作流程會與角色一起執行LakeFormationWorkflowRole，因此您必須將這些權限授與角色。

1. 在 Lake Formation 主控台的導覽窗格的 [權限] 下，選擇 [資料湖權限]。
2. 選擇 [授權]，然後在 [授與資料權限] 對話方塊中執行下列動作：
   a. 在主體下，對於 IAM 使用者和角色，選擇`LakeFormationWorkflowRole`。
   b. 在 LF 標籤或目錄資源下，選擇具名資料目錄資源。
   c. 對於「資料庫」，請選擇您先前建立的資料庫`lakeformation_tutorial`。
   d. 在 [資料庫權限] 底下，選取 [建立資料表]、[變更] 和 [刪除]，如果已選取，則清除超級。
3. 選擇 Grant (授予)。

有關授予 Lake Formation 權限的更多信息，請參閱Lake Formation 許可權概述 。

步驟 8：使用藍圖建立工作流程

AWS Glue工作 AWS Lake Formation 流程會產生探索資料並將資料擷取到資料湖中的工作、編目程式和觸發器。您可以根據其中一個預先定義的 Lake Formation 藍圖來建立工作流程。

1. 在「Lake Formation」主控台的導覽窗格中，選擇「藍圖」，然後選擇「使用藍圖」。
2. 在 [使用藍圖] 頁面的 [藍圖類型] 下，選擇 [資料庫快照集]。
3. 在 [匯入來源] 下，對於 [資料庫連線]，選擇您剛建立的連線`datalake-tutorial`，或選擇資料來源的現有連線。
4. 對於來源資料路徑，請在表單*`<database>`*/*`<schema>`*/*`<table>`*中輸入要擷取資料的來源路徑。

   您可以用百分比 (%) 萬用字元取代結構描述或資料表。<database>對於支援結構描述的資料庫，請輸入<database>/<schema>/% 以符合<schema>中的所有表格。Oracle 資料庫和 MySQL 不支援路徑中的結構描述，而是輸入 <database>/%。如果是「Oracle 資料庫」，<database>則是系統識別碼 (SID)。

   例如，如果 Oracle 資料庫具有 `orcl` SID，請輸入`orcl/%`以符合 JDCB 連線中指定之使用者可存取的所有表格。

   > **Important**
   > 此欄位會區分大小寫。

5. 在「匯入目標」下，指定下列參數：

目標資料庫	`lakeformation_tutorial`
目標儲存位置	`s3://<yourName> -datalake-tutorial`
資料格式	（選擇實木複合地板或 CSV）

6. 針對匯入頻率，請選擇「視需求執行」。
7. 在「匯入選項」下，指定下列參數：

工作流程名	`lakeformationjdbctest`
IAM 角色	`LakeFormationWorkflowRole`
表前綴	`jdbctest` **Note** 必須是小寫。

8. 選擇 [建立]，然後等待主控台回報工作流程已成功建立。

> **Tip**
> 您是否收到下列錯誤訊息？
> `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...`
> 如果是這樣，請檢查您是否已<account-id>在資料湖管理員使用者的內嵌原則中取代為有效的 AWS 帳戶號碼。

步驟 9：執行工作流程

因為您指定了工作流程 run-on-demand，因此您必須手動啟動中的工作流程 AWS Lake Formation。

1. 在「Lake Formation」主控台的「藍圖」頁面上，選取工作流lakeformationjdbctest程。
2. 選擇 [動作]，然後選擇 [開始]。
3. 工作流程執行時，請在「上次執行狀態」欄中檢視其進度。偶爾選擇刷新按鈕。

   狀態會從「執行中」、「探查」、「匯入」變為「已完成」。

   當工作流程完成時：

   - 資料目錄具有新的中繼資料表格。
   - 您的資料會擷取到資料湖中。

   如果工作流程失敗，請執行下列動作：

   a. 選取工作流程。選擇「動作」，然後選擇「檢視圖表」。

      工作流程會在AWS Glue主控台中開啟。
   b. 選擇工作流程並選擇「歷史記錄」標簽。
   c. 選取最近的執行，然後選擇 [檢視執行詳細資料]。
   d. 在動態 (程式實際執行) 圖形中選取失敗的工作或爬行者程式，然後複查錯誤訊息。失敗的節點可能是紅色或黃色。

第 10 步：在表格上授予選擇

您必須授與中新「資料目錄」表格的SELECT權限，以 AWS Lake Formation 便資料分析師可以查詢表格所指向的資料。

> **Note**
>
> 工作流程會自動將其建立之資料表的SELECT權限授與執行該工作流程的使用者。由於資料湖管理員執行此工作流程，因此您必SELECT須授與資料分析師。

1. 在 Lake Formation 主控台的導覽窗格的 [權限] 下，選擇 [資料湖權限]。
2. 選擇 [授權]，然後在 [授與資料權限] 對話方塊中執行下列動作：

   a. 在主體下，對於 IAM 使用者和角色，選擇datalake_user。
   b. 在 LF 標籤或目錄資源下，選擇具名資料目錄資源。

c. 對於「資料庫」，選擇lakeformation_tutorial。

「表格」清單會填入。

d. 對於「表格」，請從資料來源中選擇一或多個表格。

e. 在 [資料表和欄權限] 下，選擇 [選取]。

3. 選擇 Grant (授予)。

下一個步驟是以資料分析師的身分執行。

步驟 11：使用查詢資料湖 Amazon Athena

使用 Amazon Athena 主控台查詢資料湖中的資料。

1. 在 https://console.aws.amazon.com/athena/ 開啟 Athena 主控台，然後以資料分析師和使用者身分登入datalake_user。
2. 如有必要，請選擇「開始使用」以繼續使用 Athena 查詢編輯器。
3. 對於 Data source (資料來源)，請選擇 AwsDataCatalog。
4. 針對 Database (資料庫)，輸入 lakeformation_tutorial。

「表格」清單會填入。

5. 在其中一個表格旁的彈出式選單中，選擇「預覽表格」。

查詢會執行並顯示 10 列資料。

步驟 12：使用 Amazon Redshift Spectrum 查詢資料湖中的資料

您可以設定 Amazon Redshift Spectrum 來查詢匯入亞馬遜簡單儲存服務 (Amazon S3) 資料湖的資料。首先，建立用於啟動 Amazon Redshift 叢集和查詢 Amazon S3 資料的 AWS Identity and Access Management (IAM) 角色。然後，將您要查詢之資料表的Select權限授與此角色。然後，授予使用者使用 Amazon Redshift 查詢編輯器的權限。最後，建立一個 Amazon Redshift 叢集並執行查詢。

您可以以系統管理員身分建立叢集，並以資料分析師的身分查詢叢集。

如需有關 Amazon Redshift Spectrum 的詳細資訊，請參閱 Amazon Redshift 資料庫開發人員指南中的使用 Amazon Redshift Spectrum 查詢外部資料。

若要設定執行 Amazon Redshift 查詢的許可

1. 前往 https://console.aws.amazon.com/iam/ 開啟 IAM 主控台。以您在中建立的系統管理員使用者身分 建立具有管理權限的使用者 (使用者名稱Administrator) 或使用AdministratorAccess AWS 受管理策略的使用者身分登入。
2. 在導覽窗格中，選擇政策。

   如果這是您第一次選擇 Policies (政策)，將會顯示 Welcome to Managed Policies (歡迎使用受管政策) 頁面。選擇 Get Started (開始使用)。
3. 選擇 Create policy (建立政策)。
4. 請選擇 JSON 標籤。
5. 貼上下列 JSON 政策文件。

```
{
    "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. 完成時，選擇 Review (檢閱) 以檢閱該政策。政策驗證程式會回報任何語法錯誤。

7. 在 [檢閱原則] 頁面上，輸入您正**RedshiftLakeFormationPolicy**在建立之原則的 [名稱]。輸入 Description (說明) (選用)。檢閱政策 Summary (摘要) 來查看您的政策所授予的許可。然後選擇 Create policy (建立政策) 來儲存您的工作。
8. 在 IAM 主控台的導覽窗格中，選擇 Roles (角色)，然後選擇 Create role (建立角色)。
9. 對於 Select trusted entity (選取信任的實體) 區段，選擇 AWS service (AWS 服務)。
10. 選擇 Amazon Redshift 服務以擔任此角色。
11. 選擇服務的 Redshift Customizable (Redshift 可自訂) 使用案例。然後選擇下一步：許可。
12. 搜尋您建立的權限原則`RedshiftLakeFormationPolicy`，然後選取清單中原則名稱旁邊的核取方塊。
13. 選擇下一步：標籤。
14. 選擇下一步：檢閱。
15. 在 Role name (角色名稱) 中，輸入名稱 **`RedshiftLakeFormationRole`**。
16. (選用) 在 Role description (角色說明) 中，輸入新角色的說明。
17. 檢閱角色，然後選擇 Create role (建立角色)。

`Select`授與要在 Lake Formation 資料庫中查詢之表格的權限

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。以資料湖管理員身分登入。
2. 在功能窗格的 [權限] 下，選擇 [資料湖權限]，然後選擇 [授與]。
3. 請提供下列資訊：

   - 對於 IAM 使用者和角色，請選擇您建立的 IAM 角色`RedshiftLakeFormationRole`。當您執行 Amazon Redshift 查詢編輯器時，其會使用此 IAM 角色來取得資料的許可。
   - 針對 Database (資料庫)，輸入 `lakeformation_tutorial`。

     表格會列出填入。
   - 對於「表格」，請在資料來源中選擇要查詢的表格。
   - 選擇 「選取表格」 權限。
4. 選擇 Grant (授予)。

若要設定 Amazon Redshift Spectrum 並執行查詢

1. 在https://console.aws.amazon.com/redshift打開 Amazon Redshift 控制台。以使用者身分登入Administrator。
2. 選擇建立叢集。
3. 在 [建立叢集] 頁面上，輸入redshift-lakeformation-demo叢集識別碼。
4. 針對「節點」類型，選取 dc2.large。
5. 向下捲動，然後在 [資料庫組態] 下，輸入或接受下列參數：
   - 管理員使用者名稱：awsuser
   - 管理員使用者密碼：(*Choose a password*)
6. 展開叢集許可，對於可用的 IAM 角色，選擇RedshiftLakeFormationRole。接著選擇 Add IAM role (新增 IAM 角色)。
7. 如果您必須使用預設值 5439 以外的連接埠，請關閉 [其他組態] 旁邊的 [使用預設值] 選項。展開 [資料庫組態] 區段，然後輸入新的資料庫連接埠號碼。
8. 選擇建立叢集。

   「叢集」頁面會載入。
9. 等到叢集狀態變成 [可用]。請定期選擇重新整理圖示。
10. 授與資料分析師對叢集執行查詢的權限。若要這樣做，請完成下列步驟。
    a. 在 https://console.aws.amazon.com/iam/ 開啟 IAM 主控台，然後以Administrator使用者身分登入。
    b. 在瀏覽窗格中，選擇 [使用者]，然後將下列受管理的策略附加至使用者datalake_user。
       - AmazonRedshiftQueryEditor
       - AmazonRedshiftReadOnlyAccess
11. 登出 Amazon Redshift 主控台，然後以使用者datalake_user身分重新登入。
12. 在左側垂直工具列中，選擇 EDITOR 圖示以開啟查詢編輯器並連線至叢集。如果出現 [Connect 到資料庫] 對話方塊，請選擇叢集名稱redshift-lakeformation-demo，然後輸入資料庫名**dev**稱**awsuser**、使用者名稱和您建立的密碼。選擇 Connect to database (連線至資料庫)。

> **Note**
>
> 如果系統未提示您輸入連線參數，而查詢編輯器中已選取另一個叢集，請選擇 [變更連線] 以開啟 [Connect 線到資料庫] 對話方塊。

13. 在新查詢 1 文字方塊中，輸入並執行下列陳述式，以將 Lake Formation `lakeformation_tutorial` 中的資料庫對應至 Amazon Redshift 結構描述名稱`redshift_jdbc`：

> **Important**
>
> 以<account-id>有效的 AWS 帳戶編號取代，並<region>使用有效的 AWS 區域名稱 (例如，`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. 在「選取綱要」下的「架構」清單中，選擇「redshi ft_jdbc」。

    表格會列出填入。查詢編輯器只會顯示您被授與 Lake Formation 資料湖權限的表格。

15. 在表格名稱旁的彈出式選單上，選擇「預覽資料」。

    Amazon Redshift 返回前 10 行。

    您現在可以針對擁有權限的資料表和資料行執行查詢。

步驟 13：使用 Amazon Redshift Spectrum 授予或撤銷 Lake Formation 許可

Amazon Redshift 支援使用修改後的 SQL 陳述式授與和撤銷資料庫和資料表上的 Lake Formation 許可的功能。這些陳述式類似於現有的 Amazon Redshift 語句。如需詳細資訊，請參閱 Amazon Redshift 資料庫開發人員指南中的授權和撤銷。

在 Lake Formation 中設置開放表格式存儲格式的權限

AWS Lake Formation 支援管理開放資料表格式 (OTF) 的存取權限，例如阿帕奇冰山、阿帕奇胡迪和 Linux 基礎三角洲湖。在本教學課程中，您將學習如何在使用中使用符號連結資訊清單表建立冰山、Hudi 和 Delta 湖 AWS Glue、 AWS Glue Data Catalog 使用 Lake Formation 設定精細的許可，以及使用 Amazon Athena 查詢資料。

> **Note**
>
> AWS 分析服務並不支援所有交易表格格式。如需詳細資訊，請參閱 與其他 AWS 服務合作。本自學課程手動介紹僅使用 AWS Glue 工作在「資料目錄」中建立新資料庫和表格。

本教程包括用於快速設置的 AWS CloudFormation 模板。您可以查看和自定義它以滿足您的需求。

主題

目標對象

本教學課程適用於 IAM 管理員、資料湖管理員和商業分析師。下表列出本教學課程中使用的角色，用於使用 Lake Formation 建立受控資料表。

角色	描述
IAM 管理員	可以建立 IAM 使用者和角色以及 Amazon S3 儲存貯體的使用者。具有受`AdministratorAccess` AWS 管策略。

角色	描述
資料湖管理員	可存取「資料目錄」、建立資料庫以及將 Lake Formation 權限授與其他使用者的使用者。IAM 許可比 IAM 管理員少，但足以管理資料湖。
業務分析師	可對資料湖執行查詢的使用者。具有執行查詢的權限。

必要條件

在開始本教學課程之前，您必須擁有一 AWS 帳戶 個可以使用正確權限的使用者身分登入。如需詳細資訊，請參閱 註冊一個 AWS 帳戶 及 建立具有管理權限的使用者。

本教學課程假設您熟悉 IAM 角色和政策。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

您必須設定下列 AWS 資源，才能完成本教學課程：

- 資料湖管理員使用者
- Lake Formation, 數據, 湖, 設置
- Amazon Athena 引擎版本 3

建立資料湖管理員的步驟

1. 以系統管理員使用者身分登入 Lake Formation 主控台 https://console.aws.amazon.com/lakeformation/。您將在本教學課程的美國東部 (維吉尼亞北部) 區域建立資源。
2. 在 Lake Formation 主控台的導覽窗格的 [權限] 下，選擇 [系統管理角色和工作]。
3. 選取資料湖管理員下的選擇管理員。
4. 在快顯視窗的「管理資料湖管理員」中，選擇「IAM 使用者和角色」下的「IAM 管理員使用者」。
5. 選擇儲存。

啟用資料湖設定

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。在導覽窗格的 [資料目錄] 下，選擇 [設定]。取消勾選下列項目：

- 僅對新資料庫使用 IAM 存取控制。
- 只對新資料庫中的新資料表使用 IAM 存取控制。

2. 在「跨帳戶版本設置」 下，選擇「版本 3」作為跨帳戶版本。
3. 選擇儲存。

將 Amazon Athena 引擎升級到版本 3

1. 在 https://console.aws.amazon.com/athena/ 打開 Athena 控制台。
2. 選取工作群組並選取主要工作群組。
3. 確定工作群組的最低版本為 3。如果不是，請編輯工作群組，為升級查詢引擎選擇手動，然後選取版本 3。
4. 選擇儲存變更。

步驟 1：佈建資源

本節說明如何使用 AWS CloudFormation 範本設定 AWS 資源。

若要使用 AWS CloudFormation 範本建立資源

1. 以美國東部 (維吉尼亞北部) 區域的 IAM 管理員身分登入 AWS CloudFormation 主控台，網址為 https://console.aws.amazon.com/cloudformation。
2. 選擇「啟動堆疊」。
3. 在「建立堆疊」畫面中選擇「下一步」
4. 輸入堆疊名稱。
5. 選擇下一步。
6. 在下一頁上，選擇 [下一步]。
7. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
8. 選擇建立。

   堆疊建立最多可能需要兩分鐘的時間。

啟動雲形成堆棧創建以下資源：

- lf-otf-datalake-123456789012 — 用於存放資料的 Amazon S3 儲存貯體

> **Note**
>
> Amazon S3 儲存貯體名稱附加的帳戶識別碼會取代為您的帳戶識別碼。

- lf-otf-tutorial-123456789012 — 用於存放查詢結果和任務指令碼的 Amazon S3 儲存貯體 AWS Glue
- 冰山數據庫 — 冰山數據庫 AWS Glue
- 數據庫 — 胡迪數據庫 AWS Glue
- 四角洲 — 三角洲數據庫 AWS Glue
- native-iceberg-create — 在數據目錄中創建冰山表的 AWS Glue 工作
- native-hudi-create — 在資料目錄中建立 Hudi 資料表的 AWS Glue 工作
- native-delta-create — 在資料目錄中建立 Delta 表的 AWS Glue 工作
- LF-OTF-GlueServiceRole — 您傳遞給執行任務的 AWS Glue IAM 角色。此角色具有所需的政策附加以存取資源，例如資料目錄、Amazon S3 儲存貯體等。
- LF-OTF-RegisterRole IAM 角色，可向 Lake Formation 成註冊 Amazon S3 位置。此角色已`LF-Data-Lake-Storage-Policy`附加至該角色。
- lf-consumer-analystuser — IAM 使用者可使用 Athena 查詢資料
- lf-consumer-analystuser-credentials — 儲存在中的資料分析師使用者的密碼 AWS Secrets Manager

堆疊建立完成後，瀏覽至「輸出」索引標籤並記下下列項目的值：

- AthenaQueryResultLocation — Athena 查詢輸出的 Amazon S3 位置
- BusinessAnalystUserCredentials — 資料分析師使用者的密碼

  若要擷取密碼值：

  1. 瀏覽至 Secrets Manager 主控台以選擇`lf-consumer-analystuser-credentials`值。
  2. 在 Secret value (秘密值) 區段，選擇 Retrieve secret value (擷取秘密值)。
  3. 記下密碼的密碼值。

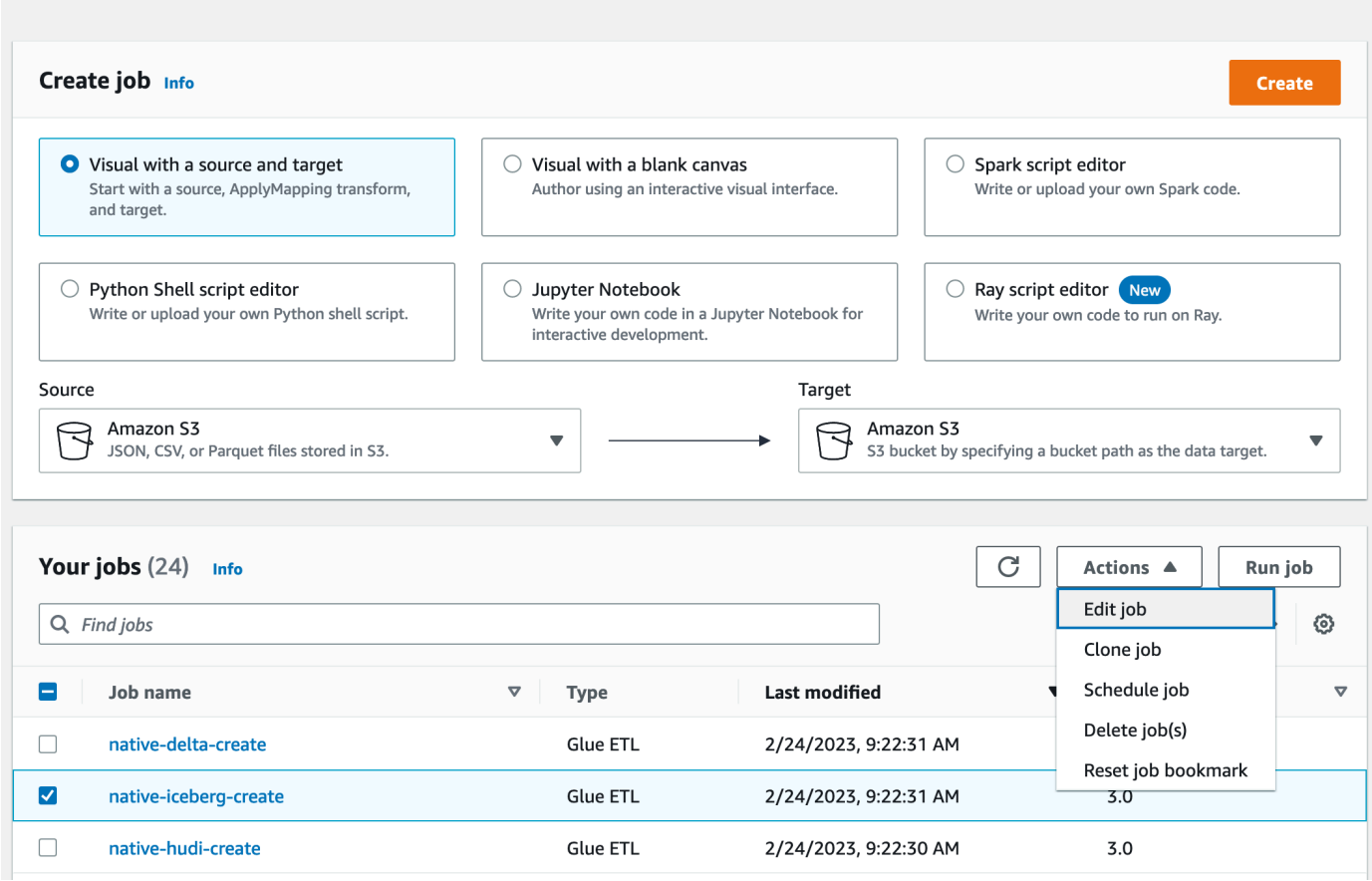
步驟 2：設置冰山表的權限

在本節中，您將學習如何在中建立 Iceberg 表格、在中設定資料許可 AWS Glue Data Catalog，以及如何使用 Amazon Athena 查詢資料。 AWS Lake Formation

要創建一個冰山表

在此步驟中，您將執行在資料目錄中建立 Iceberg 交易資料表的 AWS Glue 工作。

1. 以資料湖管理員使用者身分開啟美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/glue/ AWS Glue 主控台。
2. 從左側導覽窗格中選擇工作。
3. 選取 `native-iceberg-create`。



4. 在「動作」 下選擇「編輯工作」。
5. 在 [Job 詳細資料] 下，展開 [進階屬性]，然後核取 [用 AWS Glue Data Catalog 作 Hive 中繼資料庫] 旁邊的方塊，以在中新增表格中 AWS Glue Data Catalog繼資料。這會指定 AWS Glue Data Catalog 為工作中使用之資料目錄資源的中繼儲存區，並允許稍後在目錄資源上套用 Lake Formation 權限。
6. 選擇 儲存。
7. 選擇執行。您可以在工作執行時檢視工作的狀態。

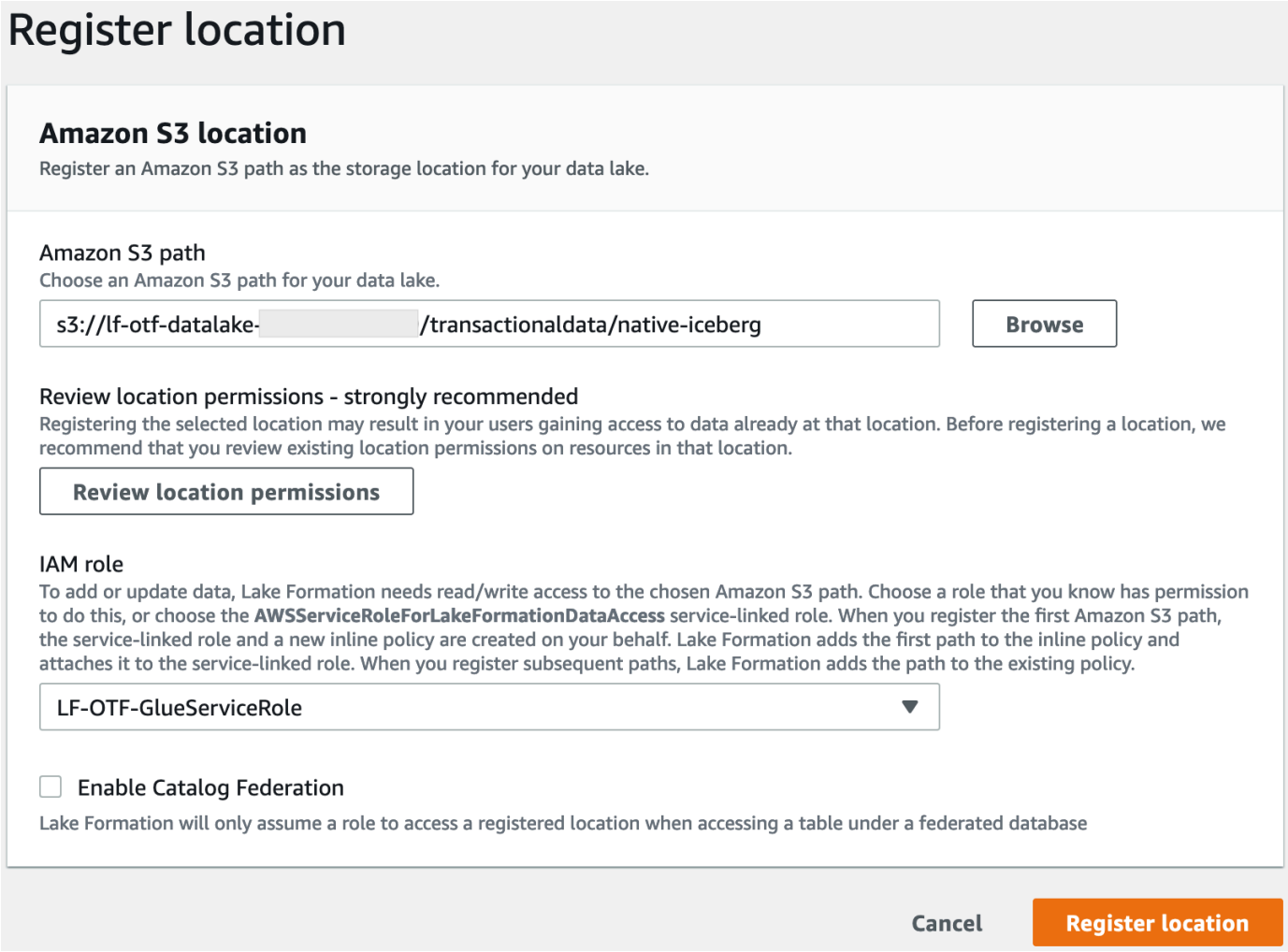
如需工作的 AWS Glue 詳細資訊，請參閱AWS Glue 開發人員指南中的在 AWS Glue 主控台上使用工作。

這項工作會在資料庫中建立名為product的冰山lficebergdb資料表。驗證 Lake Formation 控制台中的產品表。

向 Lake Formation 註冊資料位置

接下來，將 Amazon S3 路徑註冊為資料湖的位置。

1. 以資料湖管理員使用者的身分開啟 https://console.aws.amazon.com/lakeformation/ 的「湖泊形成」主控台。
2. 在功能窗格的 [註冊並擷取] 下，選擇 [資料位置]。
3. 選擇主機右上角的 [註冊位置]。
4. 在 [註冊位置] 頁面上，輸入下列資訊：
   - Amazon S3 路徑 — 選擇瀏覽並選取lf-otf-datalake-123456789012。按一下 Amazon S3 根位置旁的向右箭頭 (>) 以導覽至該s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-iceberg位置。
   - IAM 角色 — 選擇LF-OTF-RegisterRole做為 IAM 角色。
   - 選擇註冊地點。

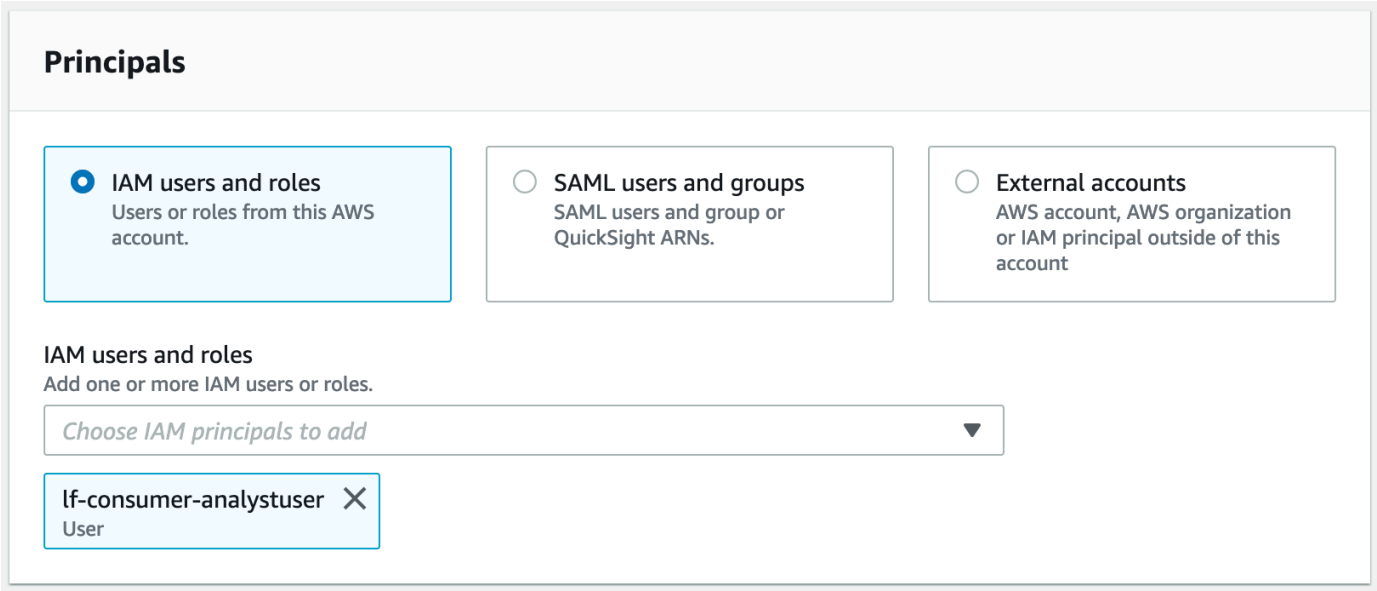


如需向 Lake Formation 註冊資料位置的詳細資訊，請參閱[將 Amazon S3 位置新增至您的資料湖]。

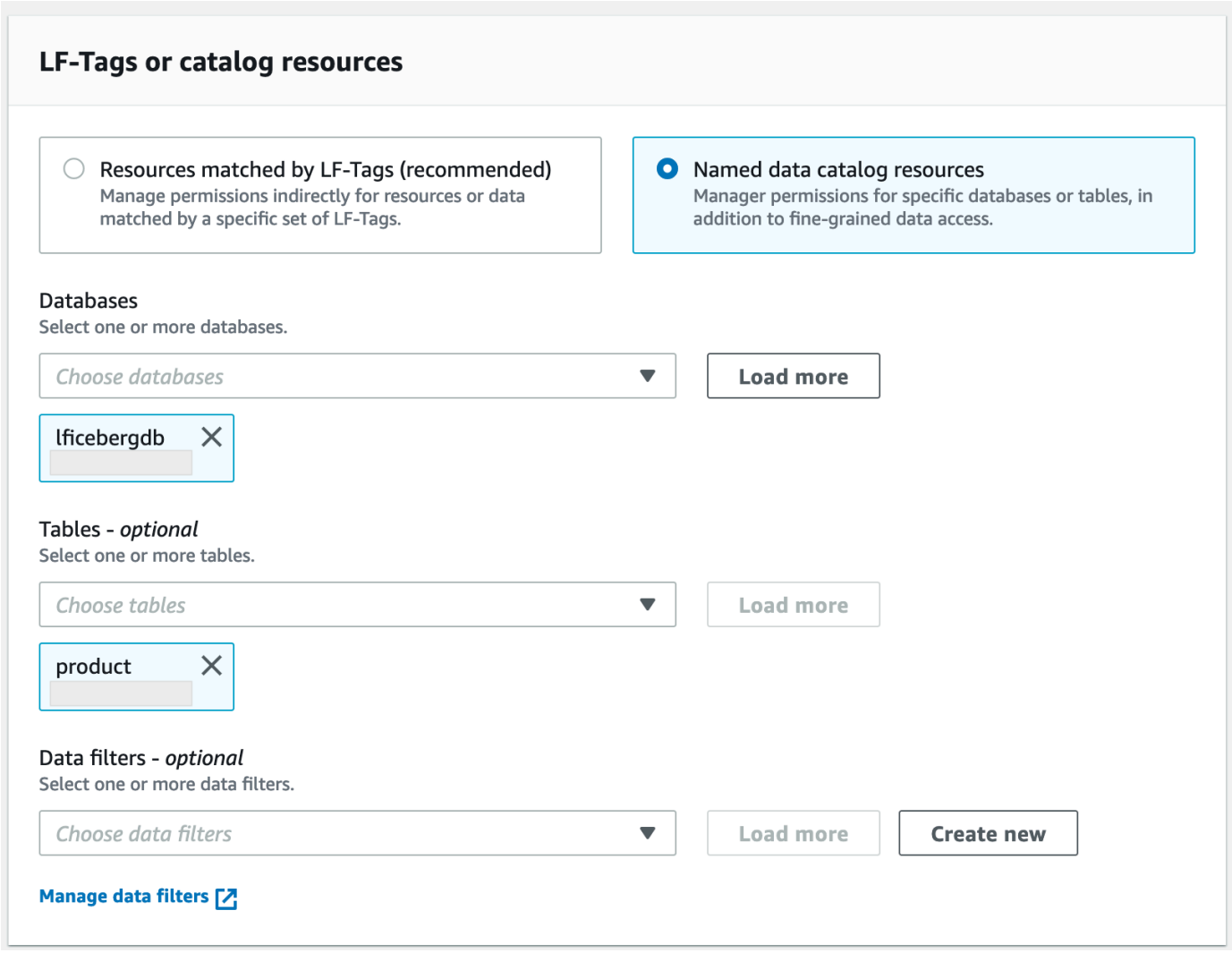
在冰山桌上授予 Lake Formation 許可權

在此步驟中，我們將授與資料湖權限給業務分析師使用者。

1. 在 [資料湖權限] 下，選擇 [授與]。
2. 在「授予資料權限」畫面上，選擇 IAM 使用者和角色。
3. `lf-consumer-analystuser`從下拉式清單中選擇。



4. 選擇具名資料目錄資源。
5. 對於數據庫，選擇`lficebergdb`。
6. 對於「表格」，請選擇`product`。



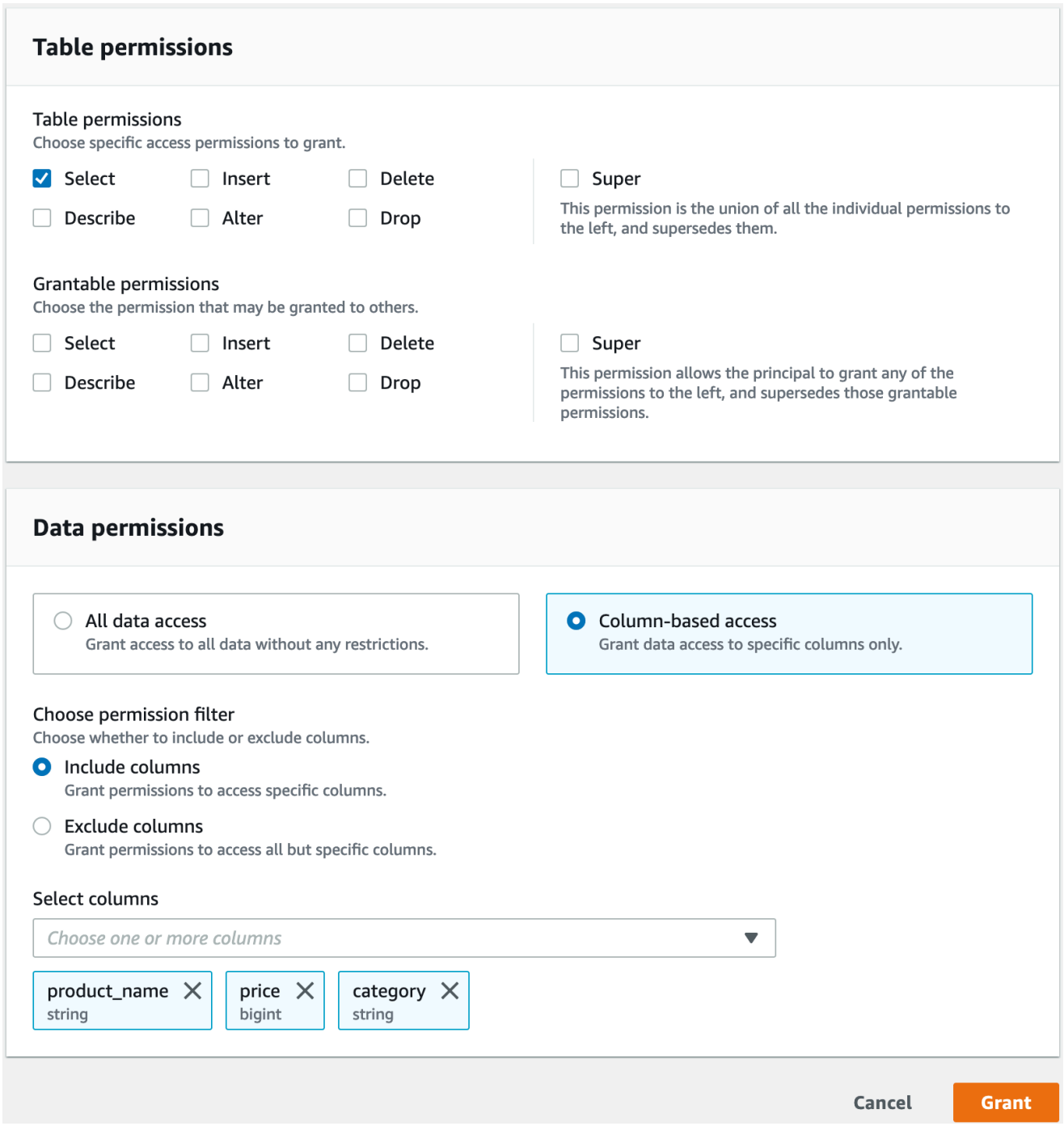
7. 接下來，您可以指定資料行來授與以資料行為基礎的存取權。

   a. 在 [資料表權限] 下，選擇 [選取

   b. 在 [資料權限] 底下，選擇 [以欄為基礎的存取]，選擇 [包含欄

   c. 選擇`product_name`、`price`和`category`欄。

   d. 選擇 Grant (授予)。



使用 Athena 查詢冰山表

現在，您可以開始查詢您使用 Athena 建立的冰山資料表了。如果這是您第一次在 Athena 執行查詢，則需要設定查詢結果位置。如需詳細資訊，請參閱指定查詢結果位置。

1. 以資料湖管理員使用者身分登出，並以美國東部 (維吉尼亞北部) 區域的身分登入，使用先前 AWS CloudFormation 輸出`lf-consumer-analystuser`中提到的密碼登入。

2. 前往 https://console.aws.amazon.com/athena/ 開啟 Athena 主控台。
3. 選擇設定，然後選取管理。
4. 在查詢結果的位置方塊中，輸入您在 AWS CloudFormation 輸出中建立的值區的路徑。複製**AthenaQueryResultLocation**（3：//lf-otf-tutorial-123456789012 /雅典音結果/）的值，然後選擇「保存」。
5. 運行以下查詢以預覽存儲在冰山表中的 10 條記錄：

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

如需使用 Athena 查詢冰山資料表的詳細資訊，請參閱 Amazon Athena 使用者指南中的查詢冰山資料表。

步驟 3：設置 Hudi 表的權限

在本節中，您將學習如何在中建立 Hudi 表格、在中設定資料許可 AWS Glue Data Catalog，以及如何使用 Amazon Athena 查詢資料。 AWS Lake Formation

建立胡迪表格的步驟

在此步驟中，您將執行在資料目錄中建立 Hudi 交易資料表的 AWS Glue 工作。

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/glue/ 登入AWS Glue主控台

   以資料湖管理員使用者身分。
2. 從左側導覽窗格中選擇工作。
3. 選取 `native-hudi-create`。
4. 在「動作」 下選擇「編輯工作」。
5. 在 [Job 詳細資料] 下，展開 [進階屬性]，然後核取 [用 AWS Glue Data Catalog 作 Hive 中繼資料庫] 旁邊的方塊，以在中新增表格中 AWS Glue Data Catalog繼資料。這會指定 AWS Glue Data Catalog 為工作中使用之資料目錄資源的中繼儲存區，並允許稍後在目錄資源上套用 Lake Formation 權限。
6. 選擇 儲存。
7. 選擇執行。您可以在工作執行時檢視工作的狀態。

   如需工作的 AWS Glue 詳細資訊，請參閱AWS Glue 開發人員指南中的在 AWS Glue 主控台上使用工作。

此工作會在資料庫中建立一個胡迪 (牛) 資料表：lfhudidb。驗證 Lake Formation 控制台中的`product`表格。

向 Lake Formation 註冊資料位置

接下來，將 Amazon S3 路徑註冊為資料湖的根位置。

1. 以資料湖管理員使用者身分登入 https://console.aws.amazon.com/lakeformation/ 的湖泊形成主控台。
2. 在功能窗格的 [註冊並擷取] 下，選擇 [資料位置]。
3. 選擇主機右上角的 [註冊位置]。
4. 在 [註冊位置] 頁面上，輸入下列資訊：
   - Amazon S3 路徑 — 選擇瀏覽並選取`lf-otf-datalake-123456789012`。按一下 Amazon S3 根位置旁的向右箭頭 (>) 以導覽至該`s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-hudi`位置。
   - IAM 角色 — 選擇`LF-OTF-RegisterRole`做為 IAM 角色。
   - 選擇註冊地點。

授與 Hudi 表格上的資料湖權限

在此步驟中，我們將授與資料湖權限給業務分析師使用者。

1. 在 [資料湖權限] 下，選擇 [授與]。
2. 在「授予資料權限」畫面上，選擇 IAM 使用者和角色。
3. `lf-consumer-analystuser`從下拉菜單。
4. 選擇具名資料目錄資源。
5. 對於數據庫，選擇`lfhudidb`。
6. 對於「表格」，請選擇`product`。
7. 接下來，您可以指定資料行來授與以資料行為基礎的存取權。
   a. 在 [資料表權限] 下，選擇 [選取
   b. 在 [資料權限] 底下，選擇 [以欄為基礎的存取]，選擇 [包含欄
   c. 選擇`product_name`、`price`、和`category`欄。

d. 選擇 Grant (授予)。

使用 Athena 查詢胡迪表

現在開始查詢您使用 Athena 建立的 Hudi 資料表。如果這是您第一次在 Athena 執行查詢，則需要設定查詢結果位置。如需詳細資訊，請參閱指定查詢結果位置。

1. 以資料湖管理員使用者身分登出，並以美國東部 (維吉尼亞北部) 區域的身分登入，使用先前 AWS CloudFormation 輸出`lf-consumer-analystuser`中提到的密碼登入。
2. 前往 https://console.aws.amazon.com/athena/ 開啟 Athena 主控台。
3. 選擇設定，然後選取管理。
4. 在查詢結果的位置方塊中，輸入您在 AWS CloudFormation 輸出中建立的值區的路徑。複製的值**`AthenaQueryResultLocation`**（3：//lf-otf-tutorial-123456789012 /雅典音結果/）並保存。
5. 執行下列查詢以預覽 Hudi 資料表中儲存的 10 筆記錄：

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

如需查詢 Hudi 資料表的詳細資訊，請參閱 Amazon Athena 使用者指南中的查詢 Hudi 資料表一節。

步驟 4：設定三角洲湖資料表的權限

在本節中，您將學習如何使用 Amazon Athena 建立具有符號連結資訊清單檔案的 Delta Lake 資料表 AWS Glue Data Catalog、在中設定資料許可，以 AWS Lake Formation 及如何使用 Amazon Athena 查詢資料。

建立三角洲湖表格的步驟

在此步驟中，您將執行在「資料目錄」中建立 Delta Lake 交易資料表的 AWS Glue 工作。

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/glue/ 登入AWS Glue主控台

   以資料湖管理員使用者身分。
2. 從左側導覽窗格中選擇工作。
3. 選取 `native-delta-create`。
4. 在「動作」 下選擇「編輯工作」。

5. 在 [Job 詳細資料] 下，展開 [進階屬性]，然後核取 [用 AWS Glue Data Catalog 作 Hive 中繼資料庫] 旁邊的方塊，以在中新增表格中 AWS Glue Data Catalog繼資料。這會指定 AWS Glue Data Catalog 為工作中使用之資料目錄資源的中繼儲存區，並允許稍後在目錄資源上套用 Lake Formation 權限。
6. 選擇儲存。
7. 選擇動作下的執行。

   此工作會建立資料庫`product`中名為的 Delta Lake 資`lfdeltadb`料表。驗證 Lake Formation 控制台中的`product`表格。

向 Lake Formation 註冊資料位置

接下來，將 Amazon S3 路徑註冊為資料湖的根位置。

1. 在 https://console.aws.amazon.com/lakeformation/ 開啟資料湖管理員使用者的湖泊形成主控台。
2. 在功能窗格的 [註冊並擷取] 下，選擇 [資料位置]。
3. 選擇主機右上角的 [註冊位置]。
4. 在 [註冊位置] 頁面上，輸入下列資訊：

   - Amazon S3 路徑 — 選擇瀏覽並選取`lf-otf-datalake-123456789012`。按一下 Amazon S3 根位置旁的向右箭頭 (>) 以導覽至該`s3/buckets/lf-otf-datalake-123456789012/transactionaldata/native-delta`位置。
   - IAM 角色 — 選擇`LF-OTF-RegisterRole`做為 IAM 角色。
   - 選擇註冊地點。

授與 Delta 湖表格上的資料湖權限

在此步驟中，我們將授與資料湖權限給業務分析師使用者。

1. 在 [資料湖權限] 下，選擇 [授與]。
2. 在「授予資料權限」畫面上，選擇 IAM 使用者和角色。
3. `lf-consumer-analystuser`從下拉菜單。
4. 選擇具名資料目錄資源。
5. 對於數據庫，選擇`lfdeltadb`。
6. 對於「表格」，請選擇`product`。

7. 接下來，您可以指定資料行來授與以資料行為基礎的存取權。

   a. 在 [資料表權限] 下，選擇 [選取

   b. 在 [資料權限] 底下，選擇 [以欄為基礎的存取]，選擇 [包含欄

   c. 選擇`product_name`、`price`和`category`欄。

   d. 選擇 Grant (授予)。

使用 Athena 查詢三角洲湖表

現在開始查詢您使用 Athena 建立的三角洲湖資料表。如果這是您第一次在 Athena 執行查詢，則需要設定查詢結果位置。如需詳細資訊，請參閱指定查詢結果位置。

1. 以資料湖管理員使用者身分登出，並使用先前 AWS CloudFormation 輸出`BusinessAnalystUser`中提到的密碼在美國東部 (維吉尼亞北部) 區域登入。
2. 前往 https://console.aws.amazon.com/athena/ 開啟 Athena 主控台。
3. 選擇設定，然後選取管理。
4. 在查詢結果的位置方塊中，輸入您在 AWS CloudFormation 輸出中建立的值區的路徑。複製的值**`AthenaQueryResultLocation`**（3：//lf-otf-tutorial-123456789012 /雅典音結果/）並保存。
5. 執行下列查詢以預覽 Delta Lake 資料表中儲存的 10 筆記錄：

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

   如需查詢三角洲湖資料表的詳細資訊，請參閱 Amazon Athena 使用者指南中的查詢三角洲湖資料表一節。

步驟 5：清理 AWS 資源

清理資源

若要避免不必要的費用 AWS 帳戶，請刪除您在本教學課程中使用的 AWS 資源。

1. 以 IAM 管理員身分登入 AWS CloudFormation 主控台 https://console.aws.amazon.com/cloudformation。
2. 刪除雲形成堆棧。您建立的資料表會隨堆疊自動刪除。

使用以 Lake Formation 標籤為基礎的存取控制來管理資料湖

成千上萬的客戶正在建置 PB 規模的資料湖。 AWS這些客戶中有許多用 AWS Lake Formation 於在整個組織中輕鬆建置和共用其資料湖。隨著資料表和使用者的數量增加，資料管理員和系統管理員正在尋找輕鬆大規模管理資料湖權限的方法。基於 Lake Formation 標籤的訪問控制（LF-TBAC）通過允許數據管理員創建 LF 標籤（基於其數據分類和本體論），然後可以附加到資源來解決此問題。

LF-TBAC 是一種授權策略，可根據屬性定義權限。在 Lake Formation，這些屬性被稱為 LF-標籤。您可以將 LF 標籤附加至資料目錄資源和 Lake Formation 主體。資料湖管理員可以使用 LF 標籤來指派和撤銷 Lake Formation 資源的權限。若要取得更多資訊，請參閱〈〉基於 Lake Formation 標籤的訪問控制。

本教學課程示範如何使用 AWS 公開資料集建立以 Lake Formation 標籤為基礎的存取控制原則。此外，它還顯示如何查詢具有與 Lake Formation 標籤相關聯之存取原則的資料表、資料庫和資料行。

您可以在下列使用案例中使用 LF-TBAC：

- 您有大量資料湖管理員必須授與存取權的表格和主參與者
- 您想要根據本體論對資料進行分類，並根據分類授予權限
- 資料湖管理員想要以鬆散耦合的方式動態指派權限

以下是使用 LF-TBAC 設定權限的高階步驟：

1. 數據管理員定義標籤本體有兩個 LF-標籤：和。`Confidential` `SensitiveConfidential=True`具有更嚴格的訪問控制的數據。具有的資料`Sensitive=True`需要分析師進行特定的分析。
2. 資料管理員會指派不同的權限層級給資料工程師，以建立具有不同 LF 標籤的資料表。
3. 資料工程師建置兩個資料庫：`tag_database`和`col_tag_database`. 中的所有表格`tag_database`都配置`Confidential=True`了。中的所有表格`col_tag_database`都使用配置`Confidential=False`。中表格的某些欄`col_tag_database`會`Sensitive=True`針對特定的分析需求加上標籤。
4. 資料工程師會將具有特定運算式條件`Confidential=True`和`Confidential=False`、的資料表的讀取權限授與分析師`Sensitive=True`。
5. 透過此配置，資料分析師可以專注於使用正確的資料執行分析。

主題

目標對象

本教學課程適用於資料管理員、資料工程師和資料分析師。在 Lake Formation 中管理 AWS Glue Data Catalog 和管理權限時，生產帳戶中的資料管理員會根據其支援的功能擁有功能所有權，並且可以授予各種消費者、外部組織和帳戶的存取權。

下表列出了本教學課程中使用的角色：

角色	描述
資料管理員 (管理員)	`lf-data-steward` 使用者具有下列存取權限： • 對資料目錄中所有資源的讀取存取權 • 可以建立 LF 標籤並與資料工程師角色相關聯，以取得其他主參與者的權限
數據工程師	`lf-data-engineer` 用戶具有以下訪問權限： • 對資料目錄中所有資源的完整讀取、寫入和更新存取 • 資料湖中的資料位置權限 • 可以關聯 LF 標籤並與資料目錄相關聯 • 可以將 LF 標籤附加至資源，此資源可根據資料管理員建立的任何原則，提供對主參與者的存取

角色	描述
資料分析	`lf-data-analyst` 使用者具有下列存取權限： • 對基於 Lake Formation 標籤的訪問策略共享的資源進行細粒度訪問

必要條件

在開始本教學課程之前，您必須擁有一 AWS 帳戶 個可用來以具有正確權限的系統管理使用者身分登入。如需詳細資訊，請參閱 完成初始 AWS 設定工作。

本教學課程假設您熟悉 IAM。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

步驟 1：佈建資源

本自學課程包括用於快速設置的 AWS CloudFormation 樣板。您可以查看和自定義它以滿足您的需求。範本會建立三個不同的角色 (列於中目標對象) 來執行此練習，並將 nyc-taxi-data 資料集複製到您的本機 Amazon S3 儲存貯體。

- Amazon S3 儲存貯體
- 適當的 Lake Formation 設置
- 適當的 Amazon EC2 資源
- 三個 IAM 角色與登入資料

建立您的資源

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/cloudformation 登入 AWS CloudFormation 主控台。
2. 選擇「啟動堆疊」。
3. 選擇下一步。
4. 在「使用者組態」段落中，輸入三個角色的密碼：`DataStewardUserPassword`、`DataEngineerUserPassword`和`DataAnalystUserPassword`
5. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
6. 選擇建立。

堆疊建立最多可能需要五分鐘的時間。

> **Note**
>
> 完成教學課程後，您可能想要刪除中的堆疊， AWS CloudFormation 以避免繼續產生費用。確認已成功刪除堆疊事件狀態中的資源。

步驟 2：註冊您的數據位置，創建 LF 標籤本體論並授予權限

在此步驟中，資料管理員使用者使用兩個 LF 標籤定義標籤本體論：`Confidential`並`Sensitive`提供特定 IAM 主體將新建立的 LF 標籤附加至資源的能力。

註冊資料位置並定義 LF 標籤本體

1. 以資料管理員使用者身分執行第一個步驟 (`lf-data-steward`)，以驗證 Amazon S3 中的資料和 Lake Formation 中的資料目錄。

   a. 使用部署 AWS CloudFormation 堆棧時使用的密碼登錄到 https://console.aws.amazon.com/lakeformation/ `lf-data-steward` 的 Lake Formation 控制台。

   b. 在導覽窗格的 [權限] 下，選擇 [系統管理角色和工作]。

   c. 在 [資料湖管理員] 區段中選擇 [新增]。

   d. 在 [新增管理員] 頁面上，對於 IAM 使用者和角色，選擇使用者`lf-data-steward`。

   e. 選擇 [儲存] 以新增`lf-data-steward`為 Lake Formation 管理員。

2. 接下來，更新「資料目錄」設定以使用 Lake Formation 權限來控制目錄資源，而不是以 IAM 為基礎的存取控制。

   a. 在導覽窗格的 [系統管理] 下，選擇 [資料目錄設定]。

   b. 取消勾選僅對新資料庫使用 IAM 存取控制。

   c. 取消勾選僅對新資料庫中的新資料表使用 IAM 存取控制。

   d. 按一下 Save (儲存)。

3. 接下來，註冊資料湖的資料位置。

   a. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。

   b. 選擇註冊地點。

c. 在 [註冊位置] 頁面上，對於 Amazon S3 路徑，輸入`s3://lf-tagbased-demo-`*`Account-ID`*。

d. 對於 IAM 角色，請保留默`AWSServiceRoleForLakeFormationDataAccess`認值。

e. 選擇 Lake Formation 作為權限模式。

f. 選擇註冊地點。

4. 接下來，通過定義 LF 標籤創建本體。

a. 在功能窗格中的 [權限] 下，選擇 [LF-標籤和權限]。 。

b. 選擇「新增 LF 標籤」。

c. 在 Key (索引鍵) 欄位，輸入 `Confidential`。

d. 對於 「值」，加入`True`和`False`。

e. 選擇「新增 LF 標籤」。

f. 重複上述步驟以建立含有值的 LF **`Sensitive`** 標籤。`True`

您已為此練習建立了所有必要的 LF 標籤。

授與權限給 IAM 使用者

1. 接下來，讓特定 IAM 主體能夠將新建立的 LF 標籤附加至資源。

a. 在功能窗格中的 [權限] 下，選擇 [LF-標籤和權限]。

b. 在 [LF 標籤權限] 區段中，選擇 [授與權限]。

c. 對於權限類型，請選擇 LF-標籤鍵值對權限。

d. 選取 IAM 使用者和角色。

e. 對於 IAM 使用者和角色，請搜尋並選擇`lf-data-engineer`角色。

f. 在 LF-標籤部分，添加`Confidential`帶有值的鍵`True`和`False`，和`key``Sensitive`與值。`True`

g. 在「權限」下，選取「權限與可授與權限」的「描述及關聯」。

h. 選擇 Grant (授予)。

2. 接下來，授予許可`lf-data-engineer`以在我們的資料目錄和建立的基礎 Amazon S3 儲存貯體上建立資料庫 AWS CloudFormation。

a. 在導覽窗格中的 [系統管理] 下，選擇 [系統管理角色和工作

b. 在「資料庫建立者」區段中，選擇授權。

c. 對於 IAM 使用者和角色，請選擇`lf-data-engineer`角色。

d. 對於目錄權限，請選取建立資料庫。

e. 選擇 Grant (授予)。

3. 接下來，將 Amazon S3 儲存貯體上的許可授與(`s3://lf-tagbased-demo-`*`Account-ID`*)使`lf-data-engineer`用者。

a. 在功能窗格的 [權限] 下，選擇 [資料位置]。

b. 選擇 Grant (授予)。

c. 選取 [我的帳戶]。

d. 對於 IAM 使用者和角色，請選擇`lf-data-engineer`角色。

e. 對於儲存位置，請輸入由 AWS CloudFormation 範本建立的 Amazon S3 儲存貯體(`s3://lf-tagbased-demo-`*`Account-ID`*)。

f. 選擇 Grant (授予)。

4. 接下來，`lf-data-engineer`授與 LF 標籤運算式關聯的資源可授與權限。`Confidential=True`

a. 在功能窗格的 [權限] 下，選擇 [資料湖權限]。

b. 選擇 Grant (授予)。

c. 選取 IAM 使用者和角色。

d. 選擇角色`lf-data-engineer`。

e. 在 LF 標籤或目錄資源區段中，選取「LF 標籤符合的資源」。

f. 選擇添加 LF-標籤鍵值對。

g. 添加`Confidential`帶有值的鍵`True`。

h. 在 [資料庫權限] 區段中，選取 [資料庫權限] 和 [授與權限] 的說明。

i. 在 [資料表權限] 區段中，針對 [資料表權限] 和 [授與權限] 選取 [說明]、[選取] 和 [變更]。

j. 選擇 Grant (授予)。

5. 接下來，`lf-data-engineer`授與 LF 標籤運算式關聯的資源可授與權限。`Confidential=False`

a. 在功能窗格的 [權限] 下，選擇 [資料湖權限]。

b. 選擇 Grant (授予)。

c. 選取 IAM 使用者和角色。

d. 選擇角色`lf-data-engineer`。
e. 選擇 LF-標籤匹配的資源。
f. 選擇「新增 LF 標籤」。
g. 添加`Confidential`帶有值的鍵`False`。
h. 在 [資料庫權限] 區段中，選取 [資料庫權限] 和 [授與權限] 的說明。
i. 在 [資料表和資料行權限] 區段中，請勿選取任何項目。
j. 選擇 Grant (授予)。

6. 接下來，我們授`lf-data-engineer`予與 LF 標籤鍵值對和相關聯的資源授予可授予的權限。`Confidential=False Sensitive=True`

a. 在功能窗格的 [權限] 下，選擇 [資料權限]。
b. 選擇 Grant (授予)。
c. 選取 IAM 使用者和角色。
d. 選擇角色`lf-data-engineer`。
e. 在 LF 標籤或目錄資源部分下，選擇 LF 標籤匹配的資源。
f. 選擇「新增 LF 標籤」。
g. 添加`Confidential`帶有值的鍵`False`。
h. 選擇添加 LF-標籤鍵值對。
i. 添加`Sensitive`帶有值的鍵`True`。
j. 在 [資料庫權限] 區段中，選取 [資料庫權限] 和 [授與權限] 的說明。
k. 在 [資料表權限] 區段中，針對 [資料表權限] 和 [授與權限] 選取 [說明]、[選取] 和 [變更]。
l. 選擇 Grant (授予)。

步驟 3：建立 Lake Formation 資料庫

在此步驟中，您會建立兩個資料庫，並將 LF 標籤附加至資料庫和特定資料行以供測試之用。

建立資料庫和資料表以供資料庫層級存取

1. 首先，創建數據庫`tag_database`，表`source_data`，並附加適當的 LF 標籤。

a. 在「Lake Formation」主控台 (https://console.aws.amazon.com/lakeformation/) 的「資料目錄」下，選擇「資料庫」。
b. 選擇建立資料庫。

c. 針對名稱，輸入 tag_database。

d. 在位置中，輸入由 AWS CloudFormation 範本建立的 Amazon S3 位置(s3://lf-tagbased-demo-*Account-ID*/tag_database/)。

e. 取消選取僅對此資料庫中的新資料表使用 IAM 存取控制。

f. 選擇建立資料庫。

2. 接下來，在中創建一個新表tag_database。

a. 在 [資料庫] 頁面上，選取資料庫tag_database。

b. 選擇檢視表格，然後按一下建立表格。

c. 針對名稱，輸入 source_data。

d. 在 Database (資料庫) 中，選擇 tag_database 資料庫。

e. 對於「表格格式」，選擇「標準 AWS Glue 表」。

f. 對於「資料位於」，請選取「我的帳戶中的指定路徑」。

g. 在「包含路徑」中，輸入要由 AWS CloudFormation 範本tag_database建立的路徑(s3://lf-tagbased-demo*Account-ID*/tag_database/)。

h. 對於「資料格式」，選取「CSV」。

i. 在 [上傳結構定義] 下，輸入下列資料行結構的 JSON 陣列，以建立結構定義：

```
[
              {
                   "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. 選擇上傳。上傳結構描述後，資料表結構定義應如下列螢幕擷取畫面所示：

#	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. 選擇提交。

3. 接下來，在數據庫級別附加 LF 標籤。

a. 在「資料庫」頁面上，尋找並選取tag_database。

b. 在 [動作] 功能表上，選擇 [編輯 LF 標籤]。

c. 選擇「指派新 LF 標籤」。

d. 對於「已指定的金鑰」，請選擇您先前建立的 Confidential LF 標籤。

e. 對於「值」，請選擇True。

f. 選擇儲存。

這樣就完成了 LF 標籤分配給標籤數據庫。

創建用於列級訪問的數據庫和表

重複下列步驟以建立資料庫col_tag_database和資料表source_data_col_lvl，並在欄層級附加 LF 標籤。

1. 在「資料庫」頁面上，選擇建立資料庫。
2. 針對名稱，輸入 col_tag_database。
3. 在位置中，輸入由 AWS CloudFormation 範本建立的 Amazon S3 位置(s3://lf-tagbased-demo-*Account-ID*/col_tag_database/)。
4. 取消選取僅對此資料庫中的新資料表使用 IAM 存取控制。
5. 選擇建立資料庫。
6. 在 [資料庫] 頁面上，選取您的新資料庫(col_tag_database)。
7. 選擇檢視表格，然後按一下建立表格。
8. 針對名稱，輸入 source_data_col_lvl。
9. 對於「資料庫」，請選擇新的資料庫(col_tag_database)。
10. 對於「表格格式」，選擇「標準 AWS Glue 表」。
11. 對於「資料位於」，請選取「我的帳戶中的指定路徑」。
12. 輸入的 Amazon S3 路徑col_tag_database(s3://lf-tagbased-demo-*Account-ID*/col_tag_database/)。
13. 對於「資料格式」，選取CSV。
14. 在下Upload schema，輸入下列結構定義 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. 選擇 Upload。上傳模式後，表模式應該看起來像下面的屏幕截圖。

#	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. 選擇「送出」以完成表格的建立。
17. 現在，將 Sensitive=True LF 標籤關聯到列vendorid和。fare_amount

    a. 在「表格」頁面上，選取您建立的表格(source_data_col_lvl)。
    b. 在 [動作] 功能表上，選擇 [綱要]。
    c. 選取欄vendorid，然後選擇「編輯 LF 標籤」。
    d. 在「指定的金鑰」中，選擇「敏感
    e. 對於「值」，選擇「真」。
    f. 選擇儲存。
18. 接下來，將 Confidential=False LF 標籤關聯到。col_tag_database這是必需的，以lf-data-analyst便能夠在登錄col_tag_database時描述數據庫 Amazon Athena。

    a. 在「資料庫」頁面上，尋找並選取col_tag_database。
    b. 在 [動作] 功能表上，選擇 [編輯 LF 標籤]。
    c. 選擇「指派新 LF 標籤」。
    d. 在「指定的金鑰」中，選擇您先前建立的 Confidential LF 標籤。
    e. 對於「值」，請選擇False。
    f. 選擇儲存。

步驟 4：授予資料表權限

col_tag_database使用 LF 標籤Confidential和向資料分析師授與使用資料庫tag_database和資料表的權限。Sensitive

1. 請依照下列步驟，將權限授與lf-data-analyst使用者 LF 標籤 Confidential=True (資料庫:TAG_Database) 相關聯的物件，以便擁有Describe資料庫和資料表的權限。Select

    a. 登錄到 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/ 作為lf-data-engineer.
    b. 在 [權限] 下，選取 [資料湖權限]。
    c. 選擇 Grant (授予)。
    d. 在主體下，選取 IAM 使用者和角色。
    e. 對於 IAM 使用者和角色，請選擇lf-data-analyst。
    f. 在 LF 標籤或目錄資源下，選取「LF 標籤符合的資源」。

g. 選擇「新增 LF 標籤」。

h. 對於「金鑰」，請選擇Confidential。

i. 對於「值」，請選擇True。

j. 對於資料庫權限，請選取Describe。

k. 對於「表格」權限，請選擇「選取並說明」

l. 選擇 Grant (授予)。

2. 接下來，重複這些步驟，將權限授與資料分析師的 LF 標籤運算式。Confidential=False此 LF 標籤用於描述col_tag_database和表在lf-data-analyst從 Amazon Athena 登錄source_data_col_lvl時。

a. 登錄到 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/ 作為lf-data-engineer.

b. 在 [資料庫] 頁面上，選取資料庫col_tag_database。

c. 選擇「動作與授權」。

d. 在主體下，選取 IAM 使用者和角色。

e. 對於 IAM 使用者和角色，請選擇lf-data-analyst。

f. 選擇 LF-標籤匹配的資源。

g. 選擇「新增 LF 標籤」。

h. 對於「金鑰」，請選擇Confidential。

i. 對於價值觀選擇False。

j. 對於資料庫權限，請選取Describe。

k. 對於資料表權限，請勿選取任何項目。

l. 選擇 Grant (授予)。

3. 接下來，重複這些步驟，以授與與的 LF 標籤運算式的資料分析師權限。Confidential=False Sensitive=True此 LF 標籤用於描述col_tag_database和表source_data_col_lvl（列級）從 Amazon Athena 登錄時。lf-data-analyst

a. 登錄到 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/ 作為lf-data-engineer.

b. 在 [資料庫] 頁面上，選取資料庫col_tag_database。

c. 選擇「動作與授權」。

d. 在主體下，選取 IAM 使用者和角色。

e. 對於 IAM 使用者和角色，請選擇lf-data-analyst。

f. 選擇 LF-標籤匹配的資源。

g. 選擇「新增 LF 標籤」。

h. 對於「金鑰」，請選擇Confidential。

i. 對於價值觀選擇False。

j. 選擇「新增 LF 標籤」。

k. 對於「金鑰」，請選擇Sensitive。

l. 對於價值觀選擇True。

m. 對於資料庫權限，請選取Describe。

n. 對於資料表權限，請選取Select和Describe。

o. 選擇 Grant (授予)。

步驟 5：在 Amazon Athena 執行查詢以驗證許可

在此步驟中，請使用 Amazon Athena 對這兩個資料表執行SELECT查詢(source_data and source_data_col_lvl)。使用 Amazon S3 路徑做為查詢結果位置(s3://lf-tagbased-demo-*Account-ID*/athena-results/)。

1. 登錄到 Athena 控制台 https://console.aws.amazon.com/athena/ 作為lf-data-analyst.
2. 在 Athena 查詢編輯器tag_database中，選擇左側面板中的。
3. 選擇旁邊的其他選單選項圖示 (三個垂直點)，source_data然後選擇「預覽表格」。
4. 選擇 Run query (執行查詢)。

   查詢應該花費幾分鐘的時間來運行。查詢會顯示輸出中的所有資料行，因為 LF 標籤是在資料庫層級相關聯，而且資料source_data表會自動LF-tag從資料庫繼承。tag_database
5. 使用col_tag_database和執行另一個查詢source_data_col_lvl。

   第二個查詢會傳回標記為Non-Confidential和的兩個資料行Sensitive。
6. 您也可以檢查以查看您沒有原則授與的欄上，以 Lake Formation 標籤為基礎的存取原則行為。從表格中選取未標記的欄時source_data_col_lvl，Athena 會傳回錯誤訊息。例如，您可以執行下列查詢來選擇未標記的資料欄geolocationid：

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

步驟 6：清理 AWS 資源

若要避免不必要的費用 AWS 帳戶，您可以刪除您在本教學課程中使用的 AWS 資源。

1. 登錄到 Lake Formation 控制台，`lf-data-engineer`並刪除數據庫`tag_database`和`col_tag_database`.
2. 接下來，登錄為`lf-data-steward`並清理所有 LF 標籤權限，數據權限和數據位置權限上面授予的被授予`lf-data-engineer`和。`lf-data-analyst`.
3. 使用您用來部署 AWS CloudFormation 堆疊的 IAM 登入資料，以帳戶擁有者身分登入 Amazon S3 主控台。
4. 刪除下列值區：
   - lf-tagbased-demo-accesslogs-*#####*
   - lf-tagbased-demo-*#####*
5. 在 https://console.aws.amazon.com/cloudformation 登入 AWS CloudFormation 主控台，然後刪除您建立的堆疊。等待堆疊狀態變更為`DELETE_COMPLETE`。

使用資料列層級存取控制來保護資料湖

AWS Lake Formation 資料列層級權限可讓您根據資料合規性和控管原則，提供資料表中特定資料列的存取權。如果您有儲存數十億筆記錄的大型資料表，您需要一種方法來讓不同的使用者和團隊只存取允許他們查看的資料。列層級存取控制是保護資料的一種簡單且高效能的方式，同時讓使用者能夠存取執行工作所需的資料。Lake Formation 透過識別哪些主體存取哪些資料、時間和透過哪些服務存取哪些資料，以提供集中式稽核與合規性報告

在本教學課程中，您將學習列層級存取控制在 Lake Formation 中的運作方式，以及如何設定它們。

本教程包括用於快速設置所需資源的 AWS CloudFormation 模板。您可以查看和自定義它以滿足您的需求。

主題

- 步驟 4：使用資料篩選器進行查詢
- 步驟 5：清理 AWS 資源

目標對象

本教學課程適用於資料管理員、資料工程師和資料分析師。下表列出資料擁有者和資料用戶的角色和職責。

角色	描述
IAM 管理員	可建立使用者和角色以及 Amazon Simple Storage Service (Amazon S3) 貯體的使用者。具有受AdministratorAccess AWS 管策略。
資料湖管理員	負責設定資料湖、建立資料篩選器以及授與權限給資料分析師的使用者。
資料分析	可對資料湖執行查詢的使用者。居住在不同國家/地區的資料分析師（針對我們的使用案例，美國和日本）只能針對位於自己國家/地區的客戶分析產品評論，並且基於法規遵循原因，應該無法查看位於其他國家/地區的客戶資料。

必要條件

在開始本教學課程之前，您必須擁有 AWS 帳戶 可用來以具有正確權限的系統管理使用者身分登入。如需詳細資訊，請參閱 完成初始 AWS 設定工作。

本教學課程假設您熟悉 IAM。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

變更 Lake Formation 設定

> **Important**
>
> 在啟動 AWS CloudFormation 範本之前，請按照下列步驟停用「僅對 Lake Formation 中的新資料庫/表格使用 IAM 存取控制」選項：

1. 在美國東部 (維吉尼亞北部) 區域或美國西部 (奧勒岡) 區域的 https://console.aws.amazon.com/lakeformation/ 登入 Lake Formation 主控台。
2. 在「資料目錄」下選擇「設定」。
3. 取消選取僅對新資料庫使用 IAM 存取控制，並針對新資料庫中的新資料表僅使用 IAM 存取控制。
4. 選擇儲存。

步驟 1：佈建資源

本自學課程包括用於快速設置的 AWS CloudFormation 樣板。您可以查看和自定義它以滿足您的需求。該 AWS CloudFormation 模板生成以下資源：

- 下列項目的使用者和政策
  - DataLakeAdmin
  - DataAnalyst美國
  - DataAnalyst太平紳
- Lake Formation 數據湖的設置和權限
- Lambda 函數 (適用於 Lambda 支援的 AWS CloudFormation 自訂資源)，用於將範例資料檔案從公有 Amazon S3 儲存貯體複製到 Amazon S3 儲存貯體
- 作為我們的資料湖使用的 Amazon S3 儲存貯體
- 資 AWS Glue Data Catalog 料庫、資料表和分割區

建立您的資源

請依照下列步驟使用 AWS CloudFormation 範本建立資源。

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/cloudformation 登入 AWS CloudFormation 主控台。
2. 選擇「啟動堆疊」。
3. 在「建立堆疊」畫面中選擇「下一步」
4. 輸入堆疊名稱。
5. 在DatalakeAdminUserName和中 DatalakeAdminUserPassword，輸入資料湖管理員使用者的 IAM 使用者名稱和密碼。
6. 對於DataAnalystUsUserName和 DataAnalystUsUserPassword，輸入負責美國市場的資料分析師使用者所需的使用者名稱和密碼。

7. 對於DataAnalystJpUserName和 DataAnalystJpUserPassword，輸入負責日本市場的資料分析師使用者所需的使用者名稱和密碼。
8. 在中 DataLakeBucketName，輸入資料儲存貯體的名稱。
9. 對於 DatabaseName，並TableName保留為預設值。
10. 選擇下一步
11. 在下一頁上，選擇 [下一步]。
12. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
13. 選擇建立。

    堆疊建立可能需要一分鐘的時間才能完成。

步驟 2：不含資料篩選器的查詢

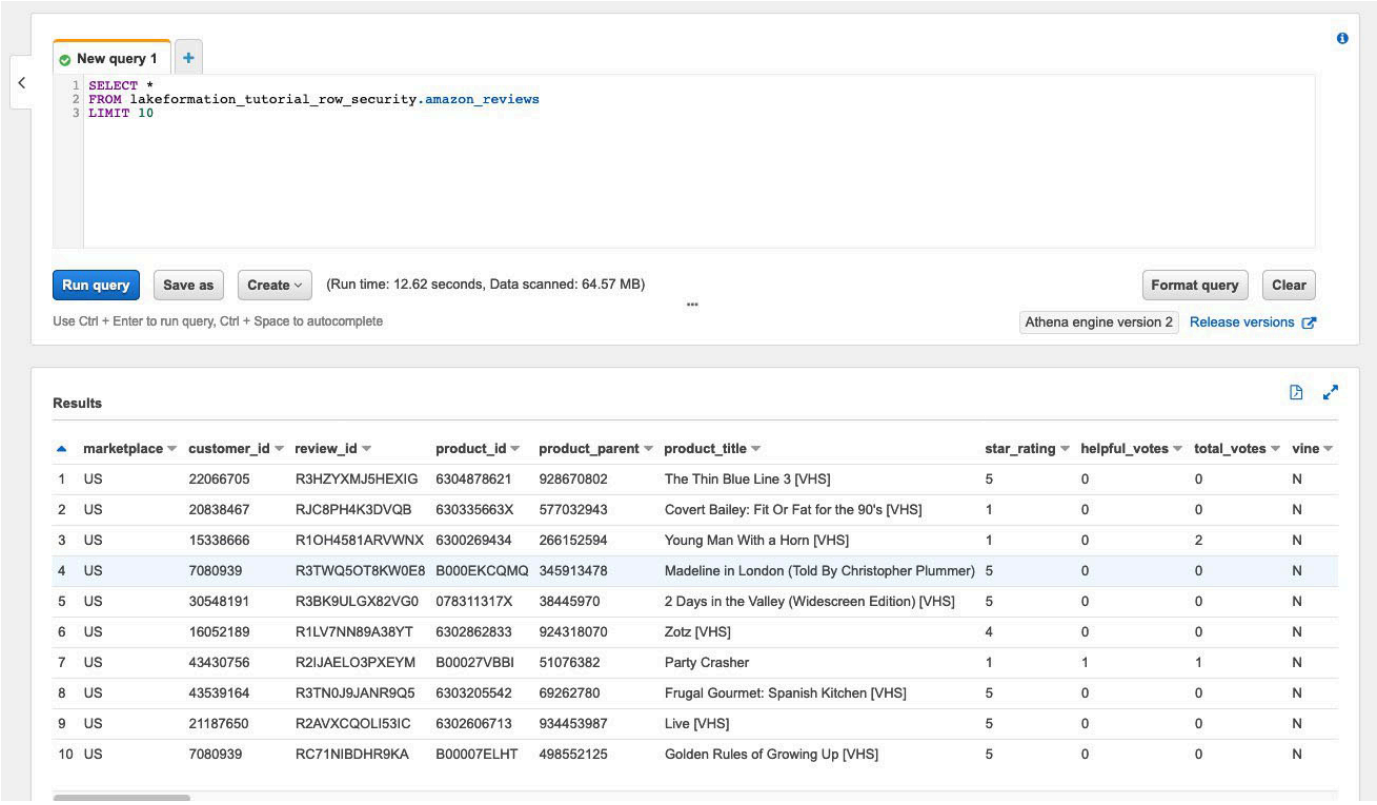
設定環境之後，您可以查詢產品評論表格。首先查詢沒有列級訪問控制的表，以確保您可以看到數據。如果您是第一次在 Amazon Athena 執行查詢，則需要設定查詢結果位置。

查詢不含資料列層級存取控制的資料表

1. 以DatalakeAdmin使用者身分登入Athena主控台 https://console.aws.amazon.com/athena/，然後執行下列查詢：

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

   下面的屏幕截圖顯示了查詢結果。此表格只有一個分割區product_category=Video，因此每筆記錄都是影片產品的評論註解。

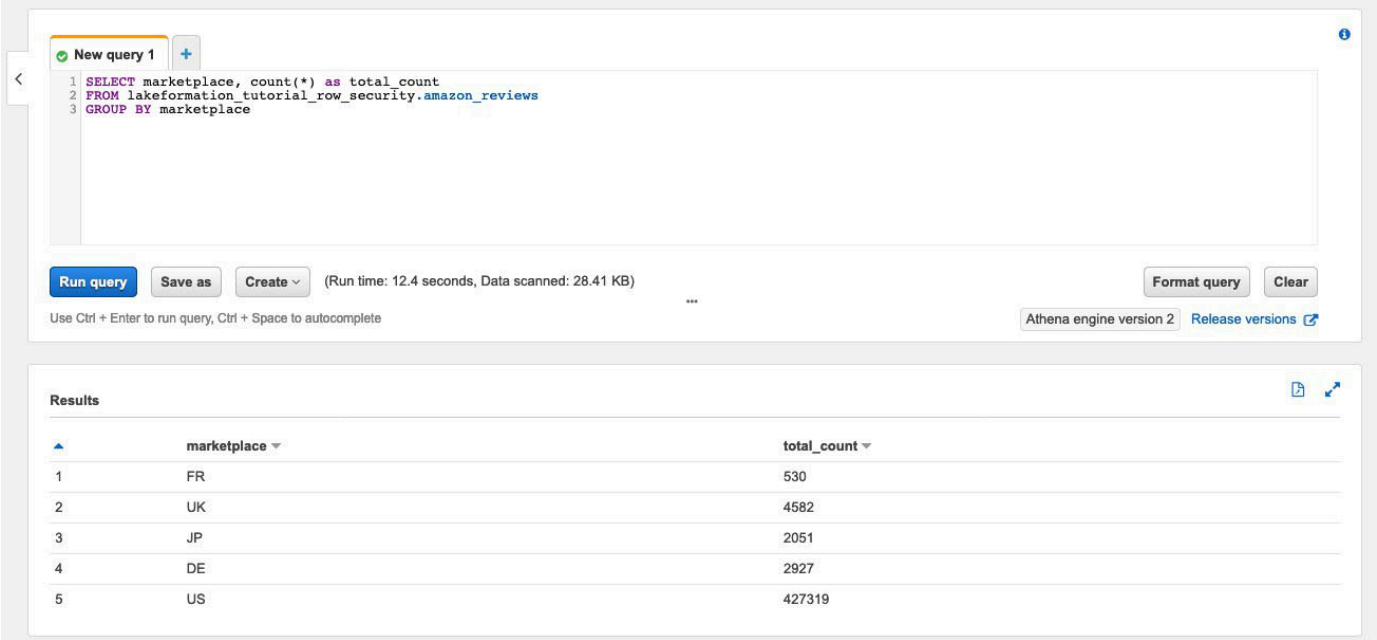


	marketplace	customer_id	review_id	product_id	product_parent	product_title	star_rating	helpful_votes	total_votes	vine
1	US	22066705	R3HZYXMJ5HEXIG	6304878621	928670802	The Thin Blue Line 3 [VHS]	5	0	0	N
2	US	20838467	RJC8PH4K3DVQB	630335663X	577032943	Covert Bailey: Fit Or Fat for the 90's [VHS]	1	0	0	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. 接下來，運行聚合查詢以檢索每個記錄的總數marketplace。

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

下面的屏幕截圖顯示了查詢結果。該marketplace列有五個不同的值。在後續步驟中，您將使用欄設定以marketplace資料列為基礎的篩選器。



步驟 3：設定資料篩選器並授予權限

本教程使用兩個數據分析師：一個負責美國市場，另一個負責日本市場。每位分析師都使用 Athena 僅針對其特定市場分析客戶評論。建立兩個不同的資料篩選器，一個是負責美國市場的分析師，另一個用於負責日本市場的分析師。然後，授予分析師各自的權限。

建立資料篩選器並授予權限

1. 創建一個過濾器來限制對USmarketplace數據的訪問。

   a. 以DatalakeAdmin使用者身分登入美國東部 (維吉尼亞北部) 區域 https://console.aws.amazon.com/lakeformation/ 的 Lake Formation 主控台。

   b. 選擇數據過濾器。

   c. 選擇 [建立新篩選器]。

   d. 對於「資料篩選器名稱」，輸入amazon_reviews_US。

   e. 在「目標資料庫」中，選擇資料庫lakeformation_tutorial_row_security。

   f. 針對「目標」表格，選擇表格amazon_reviews。

   g. 對於列級訪問，保留為默認值。

   h. 對於列篩選表示式，請輸入marketplace='US'。

   i. 選擇 Create filter (建立篩選條件)。

2. 建立篩選條件以限制對日文marketplace資料的存取。

a. 在 [資料篩選] 頁面上，選擇 [建立新篩選器]。

b. 對於「資料篩選器名稱」，輸入amazon_reviews_JP。

c. 在「目標資料庫」中，選擇資料庫lakeformation_tutorial_row_security。

d. 針對「目標」表格，選擇table amazon_reviews。

e. 對於列級訪問，保留為默認值。

f. 對於列篩選表示式，請輸入marketplace='JP'。

g. 選擇 Create filter (建立篩選條件)。

3. 接下來，使用這些資料篩選器將權限授予資料分析師。請依照下列步驟將權限授予美國資料分析師 (DataAnalystUS) :

a. 在 [權限] 下，選擇 [資料湖權限]。

b. 在 [資料權限] 下，選擇 [授權]

c. 對於主體，請選擇 IAM 使用者和角色，然後選取角DataAnalystUS色。

d. 對於 LF 標籤或目錄資源，請選擇具名資料目錄資源。

e. 針對 Database (資料庫)，輸入 lakeformation_tutorial_row_security。

f. 對於表格-選擇性，請選擇。amazon_reviews

g. 對於數據過濾器-可選的選擇amazon_reviews_US。

h. 針對資料篩選權限，選取選取。

i. 選擇 Grant (授予)。

4. 請依照下列步驟將權限授予日本資料分析師 (DataAnalystJP) :

a. 在 [權限] 下，選擇 [資料湖權限]。

b. 在 [資料權限] 下，選擇 [授權]

c. 對於主體，請選擇 IAM 使用者和角色，然後選取角DataAnalystJP色。

d. 對於 LF 標籤或目錄資源，請選擇具名資料目錄資源。

e. 針對 Database (資料庫)，輸入 lakeformation_tutorial_row_security。

f. 對於表格-選擇性，請選擇。amazon_reviews

g. 對於數據過濾器-可選的選擇amazon_reviews_JP。

h. 針對資料篩選權限，選取選取。

i. 選擇 Grant (授予)。

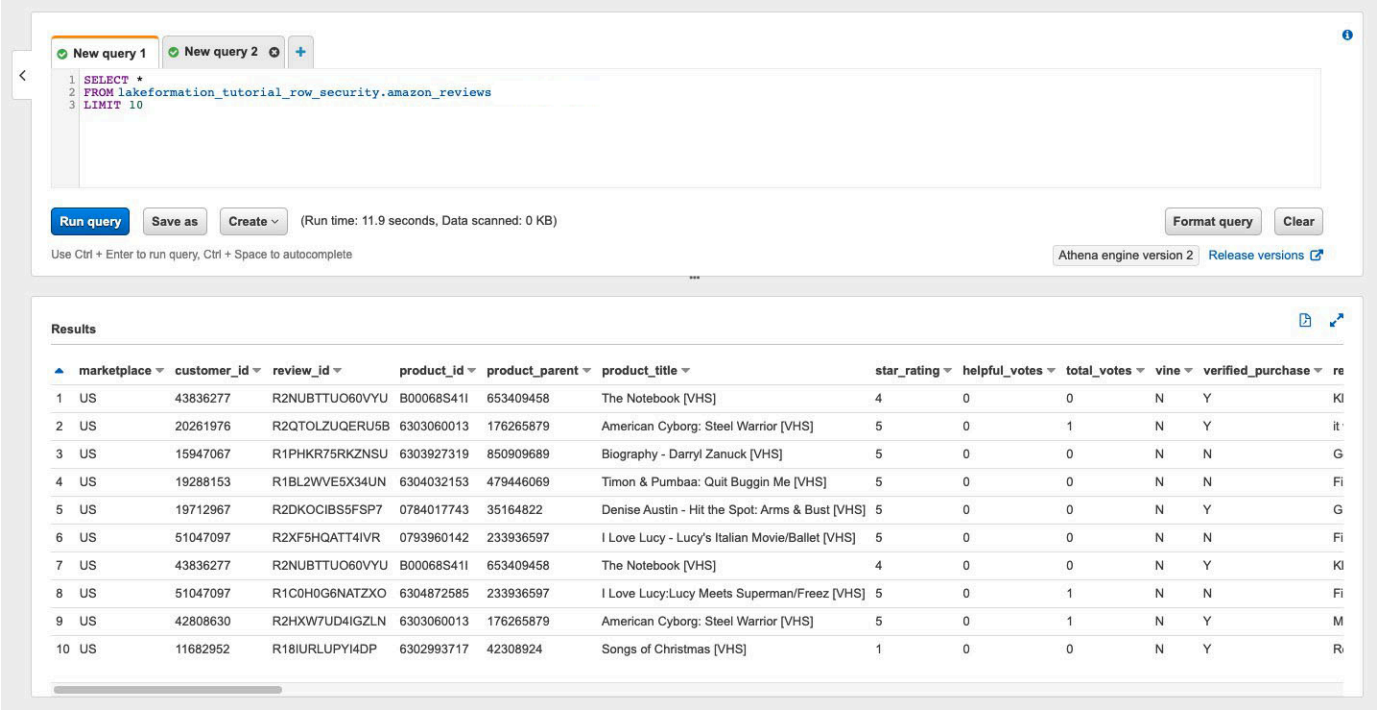
步驟 4：使用資料篩選器進行查詢

將資料篩選器附加到產品評論表後，執行一些查詢，並查看 Lake Formation 如何強制執行權限。

1. 以DataAnalystUS使用者身分登入 Athena 主控台，網址為 https://console.aws.amazon.com/athena/。
2. 執行下列查詢以擷取一些記錄，這些記錄會根據我們定義的資料列層級權限進行篩選：

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

下面的屏幕截圖顯示了查詢結果。



	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. 同樣地，執行查詢來計算每個市集的記錄總數。

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

查詢結果只會在結果marketplaceUS中顯示。這是因為只允許用戶查看marketplace列值等於的行US。

4. 切換到用DataAnalystJP戶並運行相同的查詢。

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

查詢結果僅顯示屬於的記錄JPmarketplace。

5. 執行查詢以計算每筆記錄的總數marketplace。

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

查詢結果只會顯示屬於的資料列JPmarketplace。

步驟 5：清理 AWS 資源

清除資源

若要避免不必要的費用 AWS 帳戶，您可以刪除您在本教學課程中使用的 AWS 資源。

- 刪除雲形成堆棧。

使用 Lake Formation 標籤式存取控制和具名資源共用資料湖

本教學課程示範如 AWS Lake Formation 何設定為與多家公司、組織或業務單位安全地共用儲存在資料湖中的資料，而不必複製整個資料庫。使用 Lake Formation 跨帳戶存取控制，有兩個選項可與其他 AWS 帳戶 人共用您的資料庫和資料表：

- 基於 Lake Formation 標籤的訪問控制（推薦）

  以 Lake Formation 標籤為基礎的存取控制是一種授權策略，可根據屬性定義權限。在 Lake Formation，這些屬性被稱為 LF- 標籤。如需更多詳細資訊，請參閱 使用以 Lake Formation 標籤為基礎的存取控制來管理資料湖。
- Lake Formation 命名資源

  名為資源方法的 Lake Formation 是一種定義資源權限的授權策略。資源包括資料庫、資料表和資料行。資料湖管理員可以指派和撤銷 Lake Formation 資源的權限。如需更多詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。

如果資料湖管理員偏好明確授與個別資源的權限，我們建議您使用具名資源。當您使用具名資源方法將資料目錄資源的 Lake Formation 權限授與外部帳號時，Lake Formation 會使用 AWS Resource Access Manager (AWS RAM) 來共用資源。

主題

目標對象

本教學課程適用於資料管理員、資料工程師和資料分析師。在 Lake Formation 中共用資料目錄表格 AWS Glue 和管理權限時，生產帳戶內的資料管理員會根據其支援的功能擁有功能所有權，並且可以授予各種消費者、外部組織和帳戶的存取權。下表列出了本教學課程中使用的角色：

角色	描述
DataLakeAdminProducer	資料湖管理員 IAM 使用者具有下列存取權： • 對資料目錄中所有資源的完整讀取、寫入和更新存取權 • 授予資源權限的能力 • 可以建立共用資料表的資源連結 • 可以將 LF 標籤附加至資源，此資源可根據資料管理員建立的任何原則，提供對主參與者的存取

角色	描述
DataLakeAdminConsumer	資料湖管理員 IAM 使用者具有下列存取權： • 對資料目錄中所有資源的完整讀取、寫入和更新存取權 • 授予資源權限的能力 • 可以建立共用資料表的資源連結 • 可以將 LF 標籤附加至資源，此資源可根據資料管理員建立的任何原則，提供對主參與者的存取
DataAnalyst	DataAnalyst 使用者具有下列存取權： • 對基於 Lake Formation 標籤的訪問策略或使用命名資源方法共享的資源進行細粒度訪問

在生產者帳戶中設定 Lake Formation 資料目錄設定

在開始本教學課程之前，您必須擁有 AWS 帳戶 可用來以具有正確權限的系統管理使用者身分登入。如需詳細資訊，請參閱 完成初始 AWS 設定工作。

本教學課程假設您熟悉 IAM。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

在生產者帳戶中設定 Lake Formation 資料目錄設定

> **Note**
>
> 在本教學課程中，具有來源資料表的帳戶稱為生產者帳戶，而需要存取來源資料表的帳戶稱為消費者帳戶。

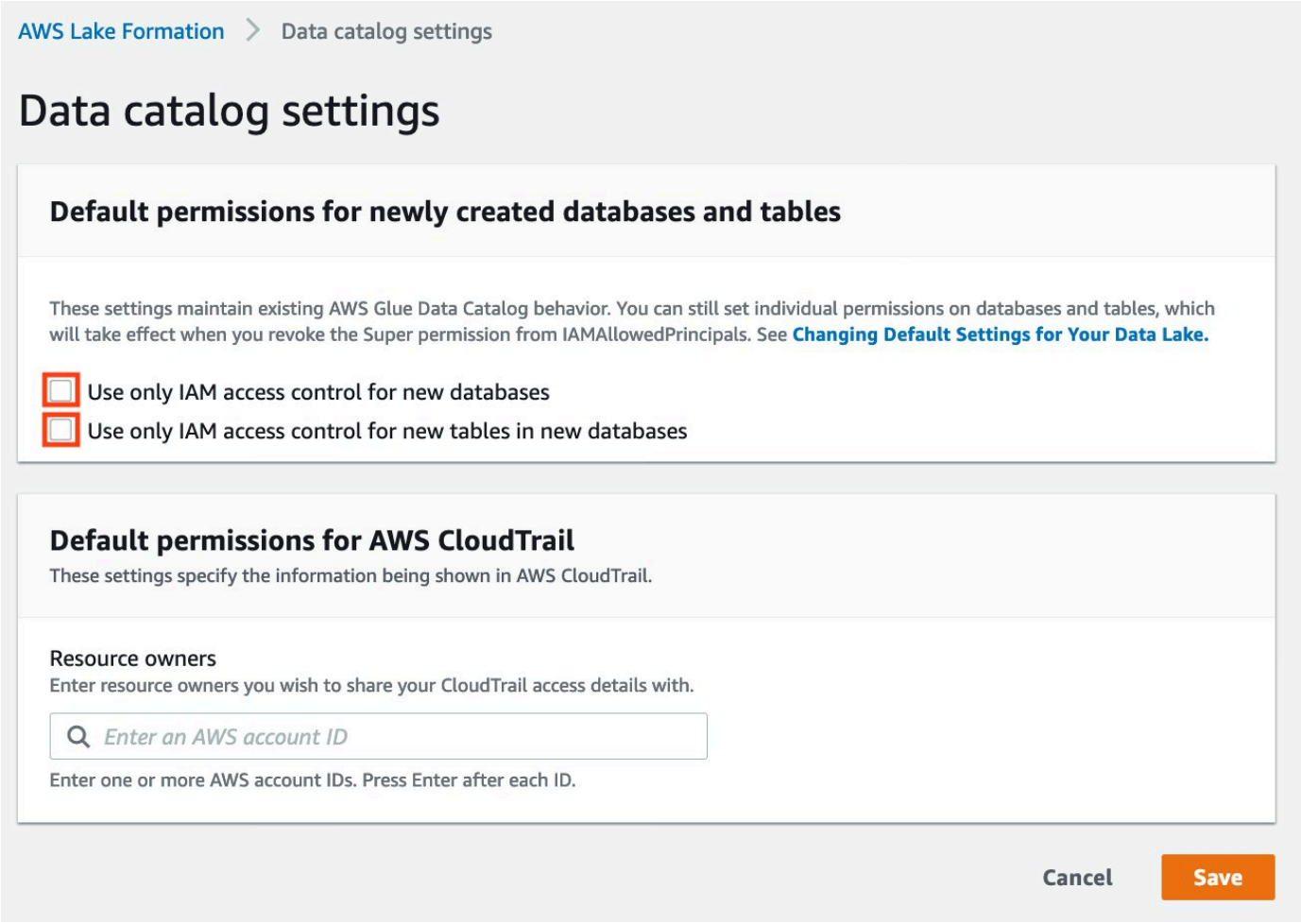
Lake Formation 提供了自己的許可管理模型。為了維持與 IAM 權限模型的回溯相容性，預設情況下會將`Super`權限授與給群組`IAMAllowedPrincipals`上的所有現有 AWS Glue Data Catalog 資源。此外，新資料目錄資源也會啟用「僅使用 IAM 存取控制」設定。本教學課程使用 Lake Formation 許可的精細存取控制，並使用 IAM 政策進行粗略的存取控制。如需詳細資訊，請參閱 細粒度存取控制的方法。因此，在您使用 AWS CloudFormation 範本進行快速設定之前，您需要在生產者帳戶中變更 Lake Formation 資料目錄設定。

> **Important**
>
> 此設定會影響所有新建立的資料庫和資料表，因此我們強烈建議您在非生產帳戶或新帳戶中完成此教學課程。此外，如果您使用的是共用帳戶 (例如貴公司的開發帳戶)，請確定該帳戶不會影響其他資源。如果您想要保留預設的安全性設定，則在與其他帳戶共用資源時，必須完成額外的步驟，在此步驟中撤銷資料庫或表格`IAMAllowedPrincipals`上的預設「超級」權限。我們在本教程後面討論的細節。

若要在生產者帳戶中設定 Lake Formation 資料目錄設定，請完成以下步驟：

1. 以管理員 AWS Management Console 使用者身分或具有 Lake Formation `PutDataLakeSettings` API 權限的使用者身分登入使用生產者帳戶。
2. 在「Lake Formation」主控台的導覽窗格的「資料目錄」下，選擇「設定」。
3. 取消選取僅對新資料庫使用 IAM 存取控制，並針對新資料庫中的新資料表僅使用 IAM 存取控制

   選擇儲存。



此外，您可以`IAMAllowedPrincipals`在 [系統管理角色和工作] [資料庫建立者] 下移除`CREATE_DATABASE`權限。只有這樣，您才能管理誰可以透過 Lake Formation 權限建立新資料庫。

步驟 1：使用 AWS CloudFormation 範本佈建資源

製作者帳號的 CloudFormation 範本會產生下列資源：

- 作為資料湖使用的 Amazon S3 儲存貯體。
- Lambda 函數 (適用於支援 Lambda 的 AWS CloudFormation 自訂資源)。我們使用該函數將範例資料檔案從公有 Amazon S3 儲存貯體複製到您的 Amazon S3 儲存貯體。
- IAM 使用者和政策： DataLakeAdminProducer.
- 適當的 Lake Formation 設置和權限包括：
  - 在生產者帳戶中定義 Lake Formation 資料湖管理員

  - 將 Amazon S3 儲存貯體註冊為 Lake Formation 資料湖位置 (生產者帳戶)
- 資 AWS Glue Data Catalog 料庫、資料表和磁碟分割。由於共用資源有兩個選項 AWS 帳戶，因此此範本會建立兩組獨立的資料庫和資料表。

消費者帳戶的 AWS CloudFormation 範本會產生下列資源：

- IAM 使用者和政策：
  - DataLakeAdminConsumer
  - DataAnalyst
- 資 AWS Glue Data Catalog 料庫。此資料庫用於建立共用資源的資源連結。

在生產者帳戶中建立您的資源

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/cloudformation 登入 AWS CloudFormation 主控台。
2. 選擇啟動堆疊。
3. 選擇下一步。
4. 在堆疊名稱中，輸入堆疊名稱，例如`stack-producer`。
5. 在「使用者組態」區段中，輸入和的使用者名稱`ProducerDatalakeAdminUserName`和密碼`ProducerDatalakeAdminUserPassword`。
6. 在中 DataLakeBucketName，輸入資料湖值區的名稱。此名稱必須是全域唯一的。
7. 對於DatabaseName和 TableName，保留預設值。
8. 選擇下一步。
9. 在下一頁上，選擇 [下一步]。
10. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
11. 選擇建立。

    堆疊建立最多可能需要一分鐘的時間。

在消費者帳戶中建立資源

1. 在美國東部 (維吉尼亞北部) 區域的 https://console.aws.amazon.com/cloudformation 登入 AWS CloudFormation 主控台。
2. 選擇啟動堆疊。

3. 選擇下一步。
4. 在堆疊名稱中，輸入堆疊名稱，例如stack-consumer。
5. 在「使用者組態」區段中，輸入和的使用者名稱ConsumerDatalakeAdminUserName和密碼ConsumerDatalakeAdminUserPassword。
6. 針對DataAnalystUserName和DataAnalystUserPassword，輸入資料分析師 IAM 使用者所需的使用者名稱和密碼。
7. 在中 DataLakeBucketName，輸入資料湖值區的名稱。此名稱必須是全域唯一的。
8. 對於 DatabaseName，保留預設值。
9. 對於AthenaQueryResultS3BucketName，輸入存放亞馬遜雅典娜查詢結果的 Amazon S3 儲存貯體的名稱。如果您沒有儲存貯體，請建立一個 Amazon S3 儲存貯體。
10. 選擇下一步。
11. 在下一頁上，選擇 [下一步]。
12. 檢閱最後一頁上的詳細資訊，然後選取 [我確認 AWS CloudFormation 可能會建立 IAM 資源]。
13. 選擇建立。

    堆疊建立最多可能需要一分鐘的時間。

> **Note**
>
> 完成教學課程後，請刪除中的堆疊 AWS CloudFormation 以避免產生費用。確認已成功刪除堆疊事件狀態中的資源。

步驟 2：Lake Formation 跨帳戶共享的先決條件

在與 Lake Formation 共用資源之前，基於標籤的存取控制方法和具名資源方法都有先決條件。

完整的標籤式存取控制跨帳戶資料共用先決條件

- 如需跨帳戶資料共用需求的詳細資訊，請參閱跨帳戶資料共用一章中的必要條件章節。

  若要與跨帳戶版本設定的第 3 版或更新版本共用 Data Catalog 資源，授與者需要在您帳戶的 AWS 受管政策AWSLakeFormationCrossAccountManager中定義 IAM 許可。

  如果您使用的是第 1 版或第 2 版的交叉帳戶版本設定，則必須先將下列權限物件新增至生產者帳號中的資料目錄資源策略，才能使用以標籤為基礎的存取控制方法授與跨帳戶存取JSON權限。這

會在成立時授予取用者帳戶存取資料目錄glue:EvaluatedByLakeFormationTags的權限。此外，對於您使用 Lake Formation 權限標籤向消費者帳戶授予權限的資源，此情況也成為真。您要授與權限的每 AWS 帳戶 個項目都需要此原則。

下列策略必須位於Statement元素內。我們在下一節討論完整的 IAM 政策。

```
{
    "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
        }
    }
}
```

完成具名資源方式跨帳戶共用先決條件

1. 如果您的帳號中沒有資料目錄資源政策，則 Lake Formation 跨帳戶會授予您照常進行的操作。但是，如果資料目錄資源策略存在，您必須在其中新增下列陳述式，以允許跨帳戶授權在使用具名資源方法建立時成功。如果您計劃只使用具名的資源方法，或僅使用以標籤為基礎的存取控制方法，則可以略過此步驟。在本教程中，我們評估這兩種方法，我們需要添加以下策略。

   下列策略必須位於Statement元素內。我們在下一節討論完整的 IAM 政策。

```
{
          "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. 接下來，使用 AWS Command Line Interface (AWS CLI) 新增 AWS Glue Data Catalog 資源策略。

   如果您同時使用以標籤為基礎的存取控制方法和具名的資源方法來授與跨帳戶權限，則在新增前述原則時必須將EnableHybrid引數設定為「true」。由於控制台目前不支持此選項，因此您必須使用 `glue:PutResourcePolicy` API 和 AWS CLI.

   首先，建立政策文件 (例如 policy.json)，並新增前兩個政策。以 AWS 帳戶 接收授權*consumer-account-id*的*## ID*、*##*取代為資料類別目錄的區域 (包含您要授與權限的資料庫和表格)，以及使用生產者 *AWS ## ID ##*戶識別碼。

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

輸入以下 AWS CLI 命令。以正確*glue-resource-policy*的值取代 (例如檔案://政策y.json)。

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

如需詳細資訊，請參閱put-resource-policy。

步驟 3：使用以標籤為基礎的存取控制方法，實作跨帳戶共用

在本節中，我們將引導您完成下列高階步驟：

1. 定義 LF 標籤。
2. 將 LF 標籤指派給目標資源。
3. 將 LF 標籤權限授予消費者帳戶。
4. 將資料權限授與消費者帳戶。
5. 或者，撤銷資料庫、資料表和資料行的權限。IAMAllowedPrincipals
6. 建立共用資料表的資源連結。
7. 創建一個 LF 標籤並將其分配給目標數據庫。
8. 將 LF 標籤資料權限授予消費者帳戶。

定義 LF 標籤

> **Note**
>
> 如果您已登入製作者帳戶，請先登出再完成以下步驟。

1. 以資料湖管理員身分登入製作者帳戶，網址為 https://console.aws.amazon.com/lakeformation/。使用您在 AWS CloudFormation 堆疊建立期間指定的生產者帳號、IAM 使用者名稱 (預設值為`DatalakeAdminProducer`) 和密碼。
2. 在「Lake Formation」主控台 (https://console.aws.amazon.com/lakeformation/) 的導覽窗格的「權限」下，並在「系統管理角色和工作」下，選擇「LF 標籤」。
3. 選擇「新增 LF 標籤」。

將 LF 標籤分配給目標資源

將 LF 標籤分配給目標資源，並將數據權限授予另一個帳戶

身為資料湖管理員，您可以將標籤附加至資源。如果您打算使用單獨的角色，則可能必須授予描述並將權限附加到單獨的角色。

1. 在導覽窗格的 [資料目錄] 下，選取 [資料庫]。
2. 選取目標資料庫，(`lakeformation_tutorial_cross_account_database_tbac`)然後在 [動作] 功能表上選擇 [編輯 LF 標籤]。

   在本教學課程中，您將 LF 標籤指派給資料庫，但也可以將 LF 標籤指派給資料表和資料行。
3. 選擇「指派新 LF 標籤」。
4. 添加鍵`Confidentiality`和值`public`。
5. 選擇儲存。

向消費者帳戶授予 LF 標籤權限

仍在生產者帳戶中，授予消費者帳戶訪問 LF 標籤的權限。

1. 在瀏覽窗格的 [權限]、[系統管理角色和工作] 下的 [LF 標籤權限] 下，選擇 [授與]。
2. 針對主參與者，選擇外部帳戶。
3. 輸入目標 AWS 帳戶 ID。

AWS 帳戶 在同一個組織內會自動顯示。否則，您必須手動輸入 AWS 帳戶 ID。在撰寫本文時，以 Lake Formation 標籤為基礎的存取控制不支援授予組織或組織單位的權限。

4. 對於 LF 標籤，請選擇要與消費者帳戶共用的 LF 標籤的金鑰和值 (金鑰**Confidentiality**和值)。 public
5. 對於權限，選取 LF 標籤權限的說明。

   LF 標籤權限是授予消費者帳戶的權限。可授與的權限是取用者帳戶可以授與其他主體的權限。
6. 選擇 Grant (授予)。

   此時，消費者資料湖系統管理員應該能夠在 [權限]、[系統管理角色和工作 LF-tag] 底下，找到透過使用者帳戶 L ake Formation 主控台共用的原則標籤。

將資料權限授予消費者帳戶

現在，我們將透過指定 LF-tag 運算式，並授與取用者帳戶對符合運算式的任何資料表或資料庫的存取權，來提供對消費者帳戶的資料存取權。

1. 在功能窗格的 [權限] 下的 [資料湖權限] 下，選擇 [授與]。
2. 對於主參與者，請選擇外部帳戶，然後輸入目標 AWS 帳戶 ID。
3. 對於 LF 標籤或目錄資源，請選擇要與消費者帳戶共用的 LF 標籤的索引鍵和值 (金鑰**Confidentiality**和值)。 public
4. 對於「權限」，在「由 LF 標籤匹配的資源（推薦）」下選擇「添加 LF 標籤」。
5. 選取要與消費者帳戶共用之標籤的金鑰和值 (金鑰Confidentiality和值public)。
6. 對於資料庫權限，請選取 [資料庫權限] 下的 [說明] 以授與資料庫層級的存取權限
7. 消費者資料湖系統管理員應該能夠在 Lake Formation 主控台 https://console.aws.amazon.com/lakeformation/ 的「權限」、「系統管理角色和工作」、「LF- tag」下，找到透過消費者帳戶共用的原則標籤。
8. 選取 [可授與權限] 下的 [說明]，讓消費者帳戶可以將資料庫層級權限授與其使用者。
9. 對於資料表和資料行權限，請選取表格權限下的選取並說明。
10. 選取 [可授與權限] 下的 [選取並說明]。
11. 選擇 Grant (授予)。

撤銷資料庫、**IAMAllowedPrincipals**資料表和資料行的權限 (選擇性)。

在本自學課程開始時，您已變更「Lake Formation 資料目錄」設定。如果您跳過該零件，則需要執行此步驟。如果您變更了 Lake Formation 資料目錄設定，則可以略過此步驟。

在此步驟中，我們需要撤銷數據庫或表格IAMAllowedPrincipals上的默認 S uper 權限。如需詳細資訊，請參閱 步驟 4：將資料存放區切換至 Lake Formation 型權限模型。

在撤銷的許可之前IAMAllowedPrincipals，請確保您已透過 Lake Formation 授與必要許可的現有 IAM 主體。這包括三個步驟：

1. 透過「Lake Formation」GetDataAccess 動作 (使用 IAM 政策)，將 IAM 權限新增至目標 IAM 使用者或角色。
2. 使用 Lake Formation 資料許可授與目標 IAM 使用者或角色 (更改、選取等)。
3. 然後，撤銷的權限IAMAllowedPrincipals。否則，撤銷的許可後IAMAllowedPrincipals，現有的 IAM 主體可能無法再存取目標資料庫或資料目錄。

   如果您想要套用 IAMAllowedPrincipals Lake Formation 權限模型 (而非 IAM 政策模型)，以管理單一帳戶內或使用 Lake Formation 權限模型的多個帳戶之間的使用者存取，則需要撤銷的超級權限。您不必撤銷要保留傳統 IAM 政策模型的IAMAllowedPrincipals其他表格的許可。

   此時，消費者帳戶資料湖管理員應該能夠在 Lake Formation 主控台 https://console.aws.amazon.com/lakeformation/ 的「資料目錄」資料庫下找到要透過使用者帳戶共用的資料庫和資料表。如果沒有，請確認下列項目是否已正確設定：

   1. 正確的原則標記和值會指派給目標資料庫和表格。
   2. 正確的標籤權限和數據權限被分配給消費者帳戶。
   3. 撤銷資料庫或資料表IAMAllowedPrincipals的預設超級權限。

建立共用資料表的資源連結

在帳號之間共用資源，且共用資源未放入用戶帳戶的「資料目錄」時。為了使它們可用，並使用 Athena 等服務查詢共享表的基礎數據，我們需要創建一個指向共享表的資源鏈接。資源連結是指向本機或共用資料庫或表格的連結的資料目錄物件。如需詳細資訊，請參閱 建立資源連結。透過建立資源連結，您可以：

- 為符合「資料目錄」資源命名原則的資料庫或表格指定不同的名稱。
- 使用 Athena 和 Redshift 頻譜等服務來查詢共用的資料庫或資料表。

若要建立資源連結，請完成以下步驟：

1. 如果您已登入消費者帳戶，請登出。
2. 以消費者帳戶資料湖管理員身分登入。使用您在 AWS CloudFormation 堆疊建立期間指定的消費者帳戶 ID、IAM 使用者名稱 (預設 DatalakeAdminConsumer) 和密碼。
3. 在「Lake Formation」主控台 (https://console.aws.amazon.com/lakeformation/) 的導覽窗格的「資料目錄」下，選擇共用資料庫`lakeformation_tutorial_cross_account_database_tbac`。

   如果您沒有看到資料庫，請重新瀏覽上述步驟，以查看是否所有項目都已正確設定。
4. 選擇「檢視表格」。
5. 選擇共用資料表`amazon_reviews_table_tbac`。
6. 在 [動作] 功能表上，選擇 [建立資源連結]。
7. 對於資源連結名稱，請輸入名稱 (在此自學課程中，請參閱`amazon_reviews_table_tbac_resource_link`)。
8. 在 [資料庫] 下，選取建立資源連結的資料庫 (針對此貼文，建立資料庫的 AWS CloudFormation n 個堆疊`lakeformation_tutorial_cross_account_database_consumer`)。
9. 選擇建立。

   資源連結會顯示在「資料目錄」的「表格」下。

創建一個 LF 標籤並將其分配給目標數據庫

Lake Formation 標籤與資源位於相同的資料目錄中。這表示在授與取用者帳戶中資源連結的存取權時，無法使用在生產者帳號中建立的標籤。在使用者帳戶中共用資源連結時，您需要在使用者帳戶中建立一組個別的 LF 標籤，才能使用以 LF 標籤為基礎的存取控制。

1. 定義消費者帳戶中的 LF 標籤。在本教學課程中，我們使用鍵`Division`和值`salesmarketing`、和`analyst`。
2. 將 LF-tag 鍵`Division`和值指派`analyst`給建立資源連結的資料庫`lakeformation_tutorial_cross_account_database_consumer`。

向消費者授予 LF 標籤數據權限

作為最後一步，將 LF 標籤數據許可授予消費者。

1. 在功能窗格的 [權限] 下的 [資料湖權限] 下，選擇 [授與]。
2. 對於主體，請選擇 IAM 使用者和角色，然後選擇使用者`DataAnalyst`。
3. 對於 LF 標籤或目錄資源，請選擇「符合 LF 標籤的資源」(建議選項)。
4. 選擇關鍵部門和價值分析師。
5. 對於資料庫權限，請選取資料庫權限下的描述。
6. 對於資料表和資料行權限，請選取表格權限下的選取並說明。
7. 選擇 Grant (授予)。
8. 對 LF 標籤鍵所在`DataAnalystConfidentiality`且值為的使用者重複這些步驟。`public`

   此時，消費者帳戶中的資料分析師使用者應該能夠找到資料庫和資源連結，並透過 Athena 主控台 https://console.aws.amazon.com/athena/ 查詢共用資料表。如果沒有，請確認下列項目是否已正確設定：

   - 即會針對共用表格建立資源連結
   - 您授予用戶訪問生產者帳戶共享的 LF 標籤
   - 您授與使用者存取與資源連結建立資源連結的資源連結和資料庫相關聯的 LF 標籤
   - 檢查您是否將正確的 LF 標籤指定給資源連結，以及建立資源連結的資料庫

第 4 步：實現指定的資源方法

若要使用指定的資源方法，我們會引導您完成下列高階步驟：

1. 您也可以選擇撤銷資料庫、資料表和資料行的權限。`IAMAllowedPrincipals`
2. 將資料權限授與消費者帳戶。
3. 接受來源的資源共用 AWS Resource Access Manager。
4. 建立共用資料表的資源連結。
5. 將共用資料表的資料權限授與取用者。
6. 將資源連結的資料權限授與取用者。

撤銷資料庫、**`IAMAllowedPrincipals`**資料表和資料行的權限 (選擇性)

- 在本自學課程開始時，我們變更了 Lake Formation 資料目錄設定。如果您跳過該零件，則需要執行此步驟。如需指示，請參閱上一節中的選用步驟。

將資料權限授予消費者帳戶

1.
> **Note**
>
> 如果您以其他使用者身分登入製作者帳戶，請先登出。

使用生產者帳戶資料湖管理員使用 AWS 帳戶 ID、IAM 使用者名稱 (預設值為`DatalakeAdminProducer`) 和在 AWS CloudFormation 堆疊建立期間指定的密碼，登入位於 https://console.aws.amazon.com/lakeformation/ 的 Lake Formation 主控台。

2. 在 [權限] 頁面的 [資料湖權限] 下，選擇 [授與]。
3. 在主參與者下，選擇外部帳戶，然後輸入一或多個 AWS 帳戶 ID 或 AWS 組織 ID。如需更多資訊，請參閱：AWS Organizations。

   生產者帳號所屬且 AWS 帳戶 位於相同組織內的組織會自動出現。否則，請手動輸入帳號 ID 或組織 ID。
4. 對於 LF 標籤或目錄資源，請選擇。`Named data catalog resources`
5. 在「資料庫」下，選擇資料庫`lakeformation_tutorial_cross_account_database_named_resource`。
6. 選擇「新增 LF 標籤」。
7. 在表格下，選擇所有表格。
8. 對於表列權限，請選擇表權限下的選擇和描述。
9. 選取「可授與權限」下的「選取並說明」。
10. 或者，對於資料權限，如果需要資料行層級權限管理，請選擇「簡單欄式存取」。
11. 選擇 Grant (授予)。

如果您尚未撤銷的權限`IAMAllowedPrincipals`，則會收到「授予權限失敗」錯誤。此時，您應該在「權限」，「數據」權限下看到通過 AWS RAM 與消費者帳戶共享的目標表。

接受資源共用來源 AWS RAM

> **Note**
>
> 此步驟僅適用 AWS 帳戶於以組織為基礎的共用，而不需要進行以組織為基礎的共用。

1. 使用消費者帳戶資料湖管理員使用 IAM 使用者名稱 (預設值為 DatalakeAdminConsumer) 和在 AWS CloudFormation 堆疊建立期間指定的密碼登入 AWS 主控台。 https://console.aws.amazon.com/connect/
2. 在 AWS RAM 主控台的導覽窗格的 [與我共用] 下的 [資源共用] 下，選擇共用的 Lake Formation 資源。「狀態」 應為「待處理」。
3. 選擇「動作與授權」。
4. 確認資源詳細資訊，然後選擇 [接受資源共用]。

   此時，消費者帳戶資料湖管理員應該能夠在「資料目錄」、「資料庫」下的 Lake Formation 主控台 (https://console.aws.amazon.com/lakeformation/) 中找到共用資源。

建立共用資料表的資源連結

- 遵循步驟 3：使用以標籤為基礎的存取控制方法，實作跨帳戶共用步驟 6 中的指示，為共用資料表建立資源連結。命名資源連結`amazon_reviews_table_named_resource_resource_link`。在資料庫中建立資源連結`lakeformation_tutorial_cross_account_database_consumer`。

將共用資料表的資料權限授與取用者

若要將共用資料表的資料權限授與用戶，請完成下列步驟：

1. 在湖泊格式化主控台 (https://console.aws.amazon.com/lakeformation/) 的 [權限] 下的 [資料湖權限] 下，選擇 [授與]。
2. 對於主體，請選擇 IAM 使用者和角色，然後選擇使用者`DataAnalyst`。
3. 對於 LF 標籤或目錄資源，請選擇具名資料目錄資源。
4. 在「資料庫」下，選擇資料庫`lakeformation_tutorial_cross_account_database_named_resource`。如果您在下拉式清單中沒有看到資料庫，請選擇 [載入更多]。
5. 在「表格」下，選擇表格`amazon_reviews_table_named_resource`。
6. 對於資料表和資料行權限，請選取表格權限下的選取並說明。
7. 選擇 Grant (授予)。

將資源連結的資料權限授與取用者

除了授與資料湖使用者存取共用資料表的權限之外，您還需要授與資料湖使用者存取資源連結的權限。

1. 在 Lake Formation 控制台 (https://console.aws.amazon.com/lakeformation/) 的「權限」下的「資料湖權限」下，選擇「授予」。
2. 對於主體，請選擇 IAM 使用者和角色，然後選擇使用者`DataAnalyst`。
3. 對於 LF 標籤或目錄資源，請選擇具名資料目錄資源。
4. 在「資料庫」下，選擇資料庫`lakeformation_tutorial_cross_account_database_consumer`。如果您在下拉式清單中沒有看到資料庫，請選擇 [載入更多]。
5. 在「表格」下，選擇表格`amazon_reviews_table_named_resource_resource_link`。
6. 對於資源連結權限，請選取資源連結權限下的描述。
7. 選擇 Grant (授予)。

   此時，消費者帳戶中的資料分析師使用者應該能夠找到資料庫和資源連結，並透過 Athena 主控台查詢共用資料表。

   如果沒有，請確認下列項目是否已正確設定：

   - 即會針對共用表格建立資源連結
   - 您授與使用者存取製作者帳戶共用之資料表的權限
   - 您已將建立資源連結的資源連結和資料庫的存取權授與使用者

步驟 5：清理 AWS 資源

若要避免不必要的費用 AWS 帳戶，您可以刪除您在本教學課程中使用的 AWS 資源。

1. 使用生產者帳戶登錄到 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/ 並刪除或更改以下內容：

   - AWS Resource Access Manager 資源共享
   - Lake Formation 標籤
   - AWS CloudFormation 堆疊
   - Lake Formation 設置
   - AWS Glue Data Catalog
2. 使用消費者帳戶登錄到 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/ 並刪除或更改以下內容：

   - Lake Formation 標籤

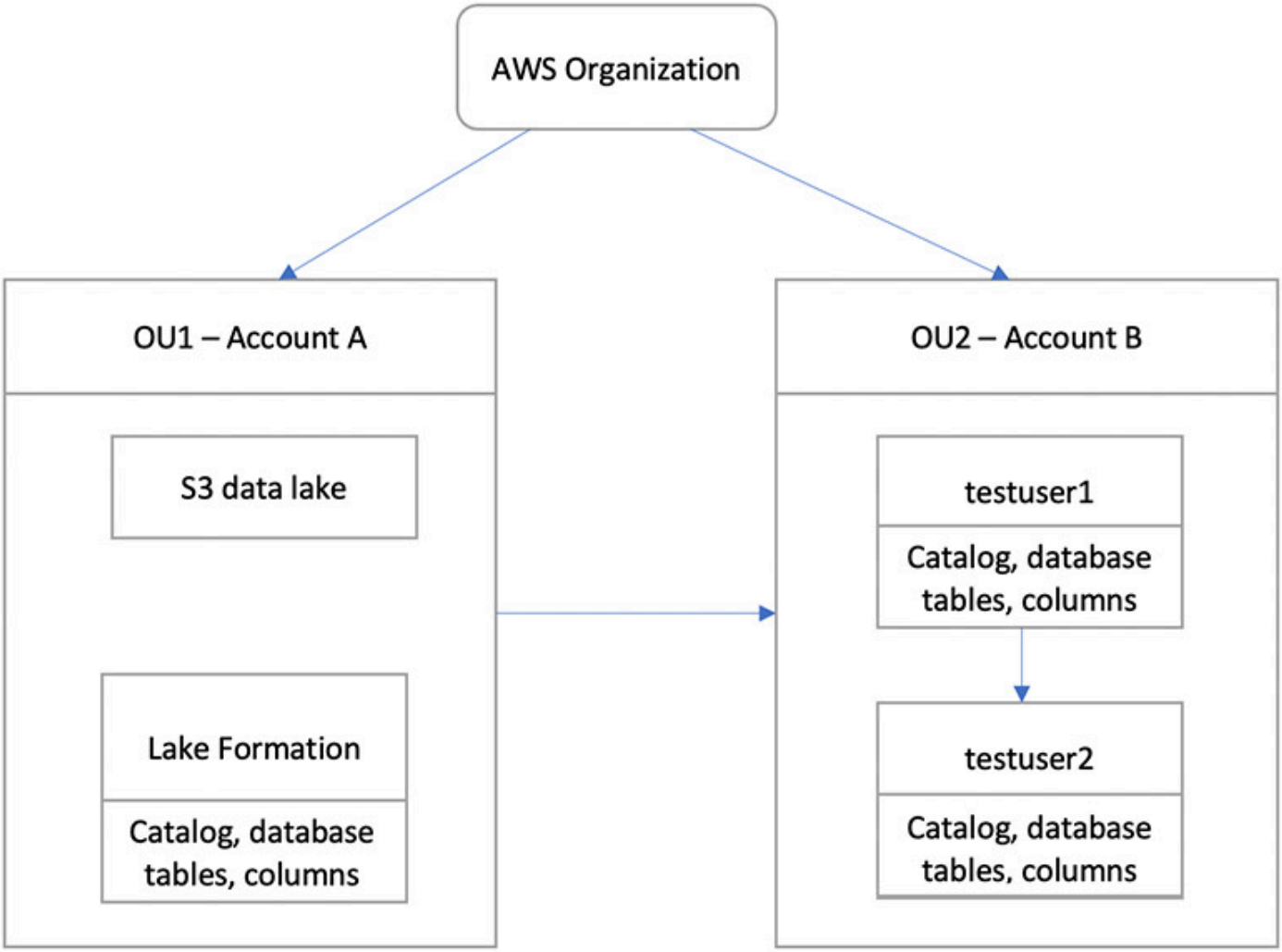
- AWS CloudFormation 堆疊

使用 Lake Formation 精細存取控制共用資料湖

本教學課程提供 AWS 帳戶 有 step-by-step 關如何在管理多個項目時，使用 Lake Formation 快速輕鬆地共用資料集的指示 AWS Organizations。您可以定義精細的權限來控制對敏感資料的存取。

下列程序也會顯示帳戶 A 的資料湖管理員如何為帳戶 B 提供精細的存取權，以及帳戶 B 中的使用者如何擔任資料管理員，為其帳戶中的其他使用者授與對共用資料表的細微存取權。每個帳戶內的資料管理員都可以獨立將存取權委派給自己的使用者，讓每個團隊或企業單位 (LOB) 擁有自主權。

使用案例假設您正在使用 AWS Organizations 來管理 AWS 帳戶. 單一組織單位 (OU1) 中帳戶 A 的使用者會授與 OU2 中帳戶 B 的使用者存取權。當您不使用「Organizations」時，您可以使用相同的方法，例如只有少數帳戶。下圖說明資料湖中資料集的精細存取控制。帳戶 A 中可使用資料湖。帳戶 A 的資料湖管理員可提供帳戶 B 的精細存取權。此圖表也顯示帳戶 B 的使用者提供帳戶 A 資料湖表的欄層級存取權給帳戶 B 中的其他使用者。



主題

目標對象

本教學課程適用於資料管理員、資料工程師和資料分析師。下表列出了本教學課程中使用的角色：

角色	描述
IAM 管理員	具有受 AWS 管理策略的使用者：AdministratorAccess 。
資料湖管理員	具有受 AWS 管理策略的使用者：AWSLakeFormationDataAdmin 附加至角色。
資料分析	具有受 AWS 管理策略的使用者：已AmazonAthenaFullAccess 附加。

必要條件

在開始本教學課程之前，您必須擁有 AWS 帳戶 可用來以具有正確權限的系統管理使用者身分登入。如需詳細資訊，請參閱 完成初始 AWS 設定工作。

本教學課程假設您熟悉 IAM。如需 IAM 的相關資訊，請參閱 IAM 使用者指南。

本教學課程需要下列資源：

- 兩個組織單位：
  - OU1 — 包含帳戶 A
  - OU2 — 包含帳戶 B
- 帳戶 A 中的 Amazon S3 資料湖位置 (儲存貯體)。
- 帳戶 A 中的資料湖管理員使用者。您可以使用湖泊形成主控台 (https://console.aws.amazon.com/lakeformation/) 或PutDataLakeSettings操作湖泊形成 API 來建立資料湖管理員。
- 帳戶 A 中設定的 Lake Formation，並在帳戶 A 中向湖泊形成註冊的 Amazon S3 資料湖位置。
- 帳戶 B 中的兩個使用者具有下列 IAM 受管政策：
  - 測試用戶 1 — 附加了 AWS 受管理的策略AWSLakeFormationDataAdmin。
  - 測試用戶 2 — 已附加受 AWS 管理策略AmazonAthenaFullAccess。
- 帳戶 B 的 Lake Formation 資料庫中的資料庫 testdb。

步驟 1：提供對另一個帳戶的精細訪問權限

瞭解帳戶 A 的資料湖管理員如何為帳戶 B 提供精細的存取權。

授予對另一個帳戶的細微訪問權限

1. 以資料湖管理員身分登入 AWS Management Console 帳戶 A 中的 https://console.aws.amazon.com/connect/。
2. 開啟「Lake Formation」主控台 (https://console.aws.amazon.com/lakeformation/)，然後選擇「開始使用」。
3. 在導覽窗格中，選擇 [資料庫]。
4. 選擇 Create database (建立資料庫)。
5. 在資料庫詳細資訊區段中，選取資料庫。
6. 在「名稱」中，輸入名稱 (在本自學課程中，我們使用`sampledb01`)。
7. 確定未選取「僅對此資料庫中的新資料表使用 IAM 存取控制」。不選擇此選項可以讓我們控制來自 Lake Formation 的訪問。
8. 選擇建立資料庫。
9. 在 [資料庫] 頁面上，選擇您的資料庫`sampledb01`。
10. 在 [動作] 功能表上，選擇 [授權]。
11. 在 [授與權限] 區段中，選取 [外部帳戶]。
12. 針對 AWS 帳戶 ID 或 AWS 組織 ID，請在 OU2 中輸入帳戶 B 的帳戶 ID。
13. 在表格中，選擇您希望帳戶 B 可存取的表格 (針對此貼文，我們使用表格`acc_a_area`)。或者，您可以授予對表中列的訪問權限，這是我們在這篇文章中做的。
14. 在 [包含欄] 中，選擇您希望帳戶 B 擁有存取權的欄 (對於此貼文，我們授予類型、名稱和識別碼的權限)。
15. 對於「欄」，請選擇「包含欄」。
16. 針對資料表權限，選取選取。
17. 針對可授與的權限，選取選取。需要可授予的權限，因此帳戶 B 中的管理員使用者可以將權限授與帳戶 B 中的其他使用者。
18. 選擇 Grant (授予)。
19. 在導覽窗格中，選擇 Tables (資料表)。
20. 您可以在 [AWS 帳戶 和具有存取權的 AWS 組織] 區段中看到一個使用中的連線。

建立資源連結

像 Amazon Athena 這樣的整合式服務無法直接跨帳戶存取資料庫或表格。因此，您必須建立資源連結，以便 Athena 可以存取您帳戶中的資源連結，連至其他帳戶中的資料庫和表格。建立資料表 (acc_a_area) 的資源連結，以便帳戶 B 使用者可以向 Athena 查詢其資料。

1. 請以帳戶 B 的身分登入 AWS 主控台 https://console.aws.amazon.com/connect/ testuser1。
2. 在「Lake Formation」主控台 (https://console.aws.amazon.com/lakeformation/) 的導覽窗格中，選擇「表格」。您應該會看到帳戶 A 提供存取權的表格。
3. 選擇 acc_a_area 資料表。
4. 在 [動作] 功能表上，選擇 [建立資源連結]。
5. 對於資源連結名稱，請輸入名稱 (在此自學課程中，請參閱acc_a_area_rl)。
6. 對於「資料庫」，請選擇您的資料庫 (testdb)。
7. 選擇建立。
8. 在導覽窗格中，選擇 Tables (資料表)。
9. 選擇 acc_b_area_rl 資料表。
10. 在 [動作] 功能表上，選擇 [檢視資料]。

    系統會將您重新導向至 Athena 主控台，您應該會在其中看到資料庫和資料表。

    您現在可以在資料表上執行查詢，以查看從帳戶 B 提供存取權給 testuser1 的資料行值。

步驟 2：為同一帳戶中的用戶提供精細訪問權限

本節說明帳戶 B (testuser1) 中的使用者 (擔任資料管理員) 如何針對共用資料表中的資料行名稱，提供對相同帳戶 (testuser2) 中的另一位使用者的精細存取權。aac_b_area_rl

向同一帳戶中的用戶授予細粒度訪問權限

1. 請以帳戶 B 的身分登入 AWS 主控台 https://console.aws.amazon.com/connect/ testuser1。
2. 在「Lake Formation」主控台的導覽窗格中，選擇「表格」。

   您可以透過資源連結授與表格的權限。若要這麼做，請在 [表格] 頁面上選取資源連結acc_b_area_rl，然後在 [動作] 功能表上選擇 [授與目標]。
3. 在 [授與權限] 區段中，選取 [我的帳戶]。
4. 對於 IAM 使用者和角色，請選擇使用者testuser2。

5. 在「欄」中，選擇欄名稱。
6. 針對資料表權限，選取選取。
7. 選擇 Grant (授予)。

   建立資源連結時，只有您可以檢視和存取它。若要允許您帳號中的其他使用者存取資源連結，您需要授與資源連結本身的權限。您需要授予描述或刪除權限。在 [表格] 頁面上，再次選取您的表格，然後在 [動作] 功能表上選擇 [授與]。
8. 在 [授與權限] 區段中，選取 [我的帳戶]。
9. 對於 IAM 使用者和角色，請選取使用者`testuser2`。
10. 對於資源連結權限，請選取描述。
11. 選擇 Grant (授予)。
12. 使用帳戶 B 的身分登入 AWS 主控台`testuser2`。

    在 Athena 控制台（https://console.aws.amazon.com/athena/）上，您應該會看到數據庫和表格`acc_b_area_rl`。您現在可以在資料表上執行查詢，以查看`testuser2`有權存取的資料行值。

入職 Lake Formation 權限

AWS Lake Formation 使用 AWS Glue Data Catalog 以資料庫和表格的形式存放 Amazon S3 資料的中繼資料。資料表儲存基礎資料的相關資訊，包括結構描述資訊、分割區資訊和資料位置。數據庫是表的集合。資料目錄也包含資源連結，這些連結是外部帳戶中共用資料庫和表格的連結，可用於跨帳戶存取資料湖中的資料。每個 AWS 帳戶每個 AWS 區域都有一個資料目錄。

Lake Formation 提供關聯式資料庫管理系統 (RDBMS) 許可模型，可授與或撤銷資料目錄中資料庫、表格和欄的存取權，以及 Amazon S3 中的基礎資料。

在瞭解 Lake Formation 權限模型的詳細資訊之前，先檢閱下列背景資訊會很有幫助：

- 由 Lake Form 管理的資料湖位於 Amazon Simple Storage Service (Amazon S3) 的指定位置。
- Lake Formation 會維護一個資料目錄，其中包含有關要匯入資料湖的來源資料的中繼資料，例如日誌和關聯式資料庫中的資料，以及 Amazon S3 中資料湖中資料的相關資料。中繼資料會組織為資料庫和資料表。中繼資料表包含結構描述、位置、分割區及其他關於它們所代表之資料的資訊。元數據庫是表的集合。
- 「Lake Formation 資料目錄」與使用的「資料目錄」相同AWS Glue。您可以使用AWS Glue編目器建立「資料目錄」表格，也可以使用AWS Glue擷取、轉換和載入 (ETL) 工作將基礎資料填入資料湖中。
- 「資料目錄」中的資料庫和表格稱為「資料目錄」資源。資料目錄中的表稱為中繼資料表，以區分資料來源中的表格或 Amazon S3 中的表格資料。中繼資料表在 Amazon S3 或資料來源中指向的資料稱為基礎資料。
- 主體是指使用者或角色、Amazon QuickSight 使用者或群組、透過 SAML 提供者向 Lake Formation 進行驗證的使用者或群組，或針對跨帳戶存取控制、 AWS 帳戶 ID、組織 ID 或組織單位 ID。
- AWS Glue檢索器會建立中繼資料表，但您也可以使用 Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 手動建立中繼資料表格。建立中繼資料表時，您必須指定位置。當您建立資料庫時，該位置是選擇性的。表格位置可以是 Amazon S3 位置或資料來源位置，例如 Amazon Relational Database Service 服務 (Amazon RDS) 資料庫。資料庫位置一律是 Amazon S3 位置。
- 與 Lake Formation 整合的服務 (例如 Amazon Athena 和 Amazon Redshift) 可以存取資料目錄以取得中繼資料並檢查執行查詢的授權。如需整合式服務的完整清單，請參閱AWS 與 Lake Formation 的服務集成。

主題

Lake Formation 許可權概述

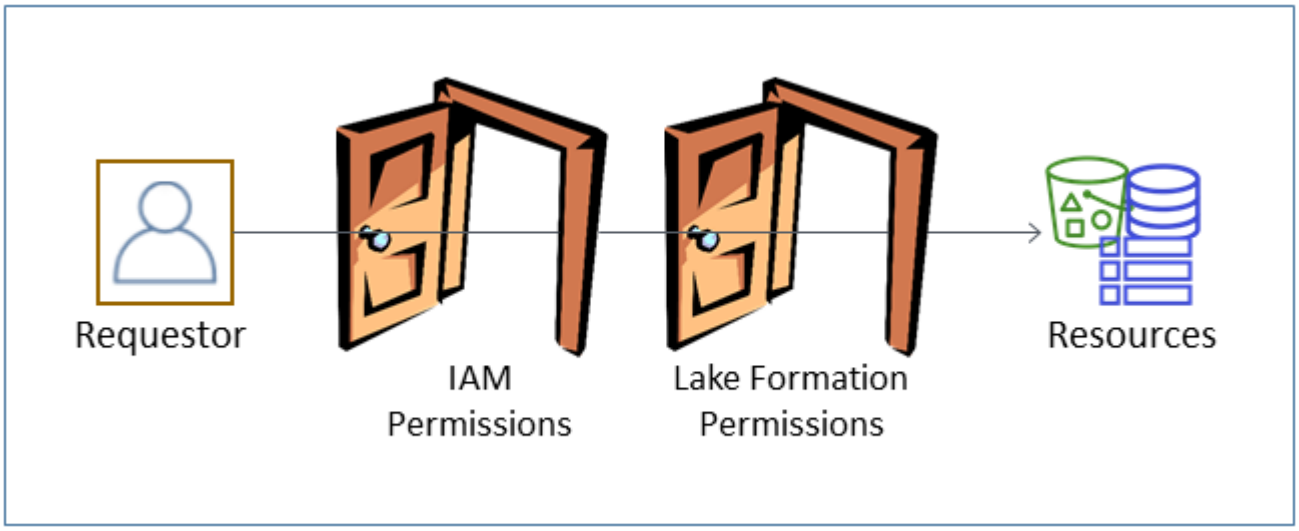
中有兩種主要的權限類型 AWS Lake Formation：

- 中繼資料存取 — 資料目錄資源的權限 (資料目錄權限)。

  這些權限可讓主參與者建立、讀取、更新及刪除「資料目錄」中的中繼資料資料庫和表格。
- 基礎資料存取 — Amazon Simple Storage Service (Amazon S3) 中位置的許可 (資料存取許可和資料位置許可)。
  - 資料湖許可可讓主體讀取和寫入資料至基礎 Amazon S3 位置 — 資料目錄資源所指向的資料。
  - 資料位置許可可讓主體建立和更改指向特定 Amazon S3 位置的中繼資料庫和表格。

對於這兩個區域，Lake Formation 使用 Lake Formation 許可和 AWS Identity and Access Management （ IAM ）許可的組合。IAM 許可模型包含 IAM 政策。Lake Formation 權限模型被實現為 DBMS 風格的授權/撤銷命令，例如。`Grant SELECT on` *`tableName`* `to` *`userName`*

當主體提出存取資料目錄資源或基礎資料的請求時，要求成功，必須通過 IAM 和 Lake Formation 的權限檢查。



Lake Formation 許可控制對資料目錄資源、Amazon S3 位置和這些位置基礎資料的存取。IAM 許可控制對 Lake Formation 以及 AWS Glue API 和資源的存取。因此，雖然您可能具有在資料型錄 (CREATE_TABLE) 中建立中繼資料表的 Lake Formation 權限的 Lake Formation 權限，但是如果您沒有 `glue:CreateTable` API 的 IAM 權限，您的作業就會失敗。（為什麼要獲`glue:`得許可？ 因為 Lake Formation 使用AWS Glue數據目錄。）

> **Note**
>
> Lake Formation 許可權僅適用於其被授予的地區。

AWS Lake Formation 需要授權每個主參與者 (使用者或角色) 對 Lake Formation 執行動作 — 管理的資源。主參與者會由資料湖管理員或其他主參與者授與必要的授權，並具有授與 Lake Formation 權限的權限。

當您將 Lake Formation 權限授與主體時，您可以選擇性地授與將該權限傳遞給另一位主體的能力。

您可以使用湖泊形成控制台的 Lake Formation API、 AWS Command Line Interface （AWS CLI）或「資料」權限和「資料位置」頁面來授予和撤銷 Lake Formation 權限。

細粒度存取控制的方法

透過資料湖，目標是對資料進行精細的存取控制。在 Lake Formation 中，這意味著對資料目錄資源和 Amazon S3 位置進行精細的存取控制。您可以使用下列其中一種方法來實現精細的存取控制。

方法	Lake Formation 權限	IAM 許可	說明
方法 1	開啟	細粒度	這是向後相容性的預設方法AWS Glue。 • Open 表示會將特殊權限Super授與給群組IAMAllowedPrincipals ，群組會自動建立，並包含任何 IAM 政策允許存取您資料目錄資源的 IAM 使用者和角色，而且該Super權限可讓主體在授與該群組的資料庫或表格上執行每個受支援的 Lake Formation 作業。IAMAllowedPrincipals 這有效地導致對資料目錄資源和 Amazon S3 位置的存取僅由 IAM 政策控制。如需詳細資訊，請參閱變更資料湖的預設設定 及 將AWS Glue資料權限升級至 AWS Lake Formation 模型。 • 精細程度表示 IAM 政策可控制對資料目錄資源和個別 Amazon S3 儲存貯體的所有存取。 在 Lake Formation 主控台上，此方法會顯示為僅限 IAM 存取控制。
方法 2	細粒度	粗粒	這是建議使用的方法。 • 細微的存取意味著將有限的 Lake Formation 許可授予資料目錄資源、Amazon S3 位置以及這些位置中的基礎資料的個別主體。 • 粗粒度意味著對個別操作和 Amazon S3 位置的存取權有更廣泛的許可。例如，粗略的 IAM 政策可能包含"glue:*"或"glue:Create*" 而非保留 Lake Formation 權限"glue:Cre

方法	Lake Formation 權限	IAM 許可	說明
			ateTables" ，以控制主體是否可以建立目錄物件。這也意味著授予主體對他們執行工作所需的 API 的訪問權限，但鎖定其他 API 和資源。例如，您可以建立 IAM 政策，讓主體建立 Data Catalog 資源並建立和執行工作流程，但不允許建立 AWS Glue連線或使用者定義函數。請參閱本節稍後的範例。

Important

請注意以下事項：

- 根據預設，Lake Formation 會啟用「僅使用 IAM 存取控制」設定，以便與現有「AWS Glue 資料目錄」行為相容。我們建議您在轉換為使用 Lake Formation 權限後停用這些設定。如需詳細資訊，請參閱 變更資料湖的預設設定。
- 資料湖管理員和資料庫建立者具有您必須瞭解的隱含 Lake Formation 權限。如需詳細資訊，請參閱 隱含 Lake Formation 權限。

元數據訪問控制

對於資料目錄資源的存取控制，以下討論假設使用 Lake Formation 許可進行精細的存取控制，以及使用 IAM 政策進行粗略的存取控制。

有兩種不同的方法可授與資料目錄資源的 Lake Formation 權限：

- 具名資源存取控制 — 使用此方法，您可以透過指定資料庫或表格名稱來授與特定資料庫或表格的權限。補助金有這種形式：

  將權限授與資源上的主參與者 [使用授與選項]。

  使用授與選項，您可以允許受權者將權限授與其他主參與者。
- 以標籤為基礎的存取控制 — 使用此方法，您可以將一或多個 LF 標籤指派給資料目錄資料庫、資料表和欄，並將一或多個 LF 標籤的權限授與主體。每個 LF 標籤是一個鍵值對，例

如。department=sales具有 LF 標籤且符合資料目錄資源上 LF 標籤的主參與者可以存取該資源。對於具有大量資料庫和表格的資料湖，建議使用此方法。它在中詳細說明基於 Lake Formation 標籤的訪問控制。

主參與者對資源所擁有的權限是這兩種方法所授與之權限的聯集。

下表摘要說明「資料目錄」資源的可用「Lake Formation」權限。欄標題指出授與權限的資源。

目錄	資料庫	資料表
CREATE_DATABASE	CREATE_TABLE	ALTER
	ALTER	DROP
	DROP	DESCRIBE
	DESCRIBE	SELECT*
		INSERT*
		DELETE*

例如，授與資料庫的CREATE_TABLE權限。這表示主體可以在該資料庫中建立資料表。

帶有星號 (*) 的權限已授與資料目錄資源，但它們適用於基礎資料。例如，中繼資料表的DROP權限可讓您從「資料目錄」刪除表格。不過，在相同資料表上授與的DELETE權限可讓您使用 SQL DELETE 陳述式，在 Amazon S3 中刪除資料表的基礎資料。透過這些權限，您也可以在 Lake Formation 主控台上檢視資料表，並使用 AWS Glue API 擷取有關資料表的資訊。因此、SELECTINSERT、和都DELETE是資料目錄權限和資料存取權限。

授與SELECT資料表時，您可以新增包含或排除一或多個資料行的篩選器。這允許對中繼資料表資料行進行精細的存取控制，從而限制整合式服務的使用者在執行查詢時可以看到的資料行。僅使用 IAM 政策時，無法使用此功能。

還有一個名為的特殊權限Super。此Super權限可讓主體在授與該作業的資料庫或表格上執行每個受支援的 Lake Formation 作業。此權限可與其他 Lake Formation 權限共存。例如，您可以在中繼資料表INSERT上授Super與SELECT、和。主參與者可以在表格上執行所有支援的動作，而當您撤銷時Super，SELECT和INSERT權限仍會保留。

如需每個權限的詳細資訊，請參閱Lake Formation 權限參考。

> **Important**
>
> 若要查看由其他使用者建立的「資料目錄」表格，您必須至少獲得一個表格的 Lake Formation 權限。如果您在資料表上獲得至少一個權限，您也可以看到資料表包含的資料庫。

您可以使用 Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 授與或撤銷資料目錄權限。以下是授與使用者在`retail`資料庫中建立資料表之`datalake_user1`權限的 AWS CLI 命令範例。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

以下是粗粒度存取控制 IAM 政策的範例，該政策可透過 Lake Formation 權限補充精細的存取控制。它允許在任何元數據庫或表上的所有操作。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "glue:*Database*",
                "glue:*Table*",
                "glue:*Partition*"
            ],
            "Resource": "*"
        }
    ]
}
```

下一個例子也是粗粒度，但更嚴格。它允許對指定帳戶和區域中「資料目錄」中的所有中繼資料資料庫和表格執行唯讀作業。

```
{
    "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:*"
        }
    ]
}
```

將這些原則與下列原則進行比較，這些原則會實作以 IAM 為基礎的精細存取控制。它只會授與指定帳戶和區域中客戶關係管理 (CRM) 中繼資料資料庫中資料表子集的權限。

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

如需粗略存取控制原則的更多範例，請參閱。Lake Formation 角色和 IAM 許可參考

基礎資料存取控制

當整合式 AWS 服務請求存取由存取控制之 Amazon S3 位置中的資料時 AWS Lake Formation，Lake Formation 會提供臨時登入資料來存取資料。

若要讓湖泊形成控制對 Amazon S3 位置基礎資料的存取，請向 Lake Formation 註冊該位置。

註冊 Amazon S3 位置後，您可以開始授予以下 Lake Formation 許可：

- 資料存取權限 (SELECTINSERT、和DELETE)指向該位置的「資料目錄」表格。
- 該位置的資料位置權限。

Lake Formation 資料位置許可控制建立指向特定 Amazon S3 位置的資料目錄資源的能力。資料位置權限為資料湖中的位置提供額外的安全層。當您授與CREATE_TABLE或ALTER權限給主參與者時，您也會授與資料位置權限，以限制主參與者可以建立或變更中繼資料表格的位置。

Amazon S3 位置是儲存貯體下的儲存貯體或前置詞，但不是個別的 Amazon S3 物件。

您可以使用 Lake Formation 主控台、API 或 AWS CLI. 授予書的一般形式如下：

```
grant DATA_LOCATION_ACCESS to principal on S3 location [with grant option]
```

如果您包含with grant option，受權者可以將權限授與其他主參與者。

回想一下，Lake Formation 許可始終與 AWS Identity and Access Management （IAM）許可結合使用，以實現精細的訪問控制。對於基礎 Amazon S3 資料的讀取/寫入許可，IAM 許可的授與方式如下：

註冊位置時，您可以指定 IAM 角色，以授與該位置的讀取/寫入許可。在向集成 AWS 服務提供臨時憑據時，Lake Formation 承擔該角色。一般角色可能會附加下列原則，其中註冊位置是值區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 提供了一個服務鏈接的角色，您可以在註冊過程中使用它來自動創建這樣的策略。如需詳細資訊，請參閱 使用服務連結角色進行 Lake Formation。

因此，註冊 Amazon S3 位置會授予該位置上所需的 s3: IAM 許可，其中許可由用於註冊該位置的角色指定。

> **Important**
>
> 避免註冊已啟用請求者付費的 Amazon S3 儲存貯體。對於在 Lake Formation 註冊的值區，用於註冊值區的角色一律會被視為請求者。如果值區是由其他 AWS 帳戶存取，如果該角色與值區擁有者屬於相同的帳戶，則值區擁有者會收取資料存取費用。

對於基礎資料的讀取/寫入存取權，除了 Lake Formation 許可外，主體還需要以下 IAM 權限：

`lakeformation:GetDataAccess`

有了此許可，Lake Formation 就會授與要求存取資料所需的臨時憑證。

> **Note**
>
> Amazon Athena 要求用戶`lakeformation:GetDataAccess`獲得許可。其他整合式服務需要其基礎執行角色才能擁有`lakeformation:GetDataAccess`權限。

此權限包含在中的建議策略中Lake Formation 角色和 IAM 許可參考。

總而言之，為了使 Lake Form 校長能夠使用由 Lake Formation 權限控制的訪問權限讀取和寫入基礎數據：

- 向 Lake Formation 註冊包含資料的 Amazon S3 位置。
- 建立指向基礎資料位置之「資料目錄」表格的主參與者必須具有資料位置權限。
- 讀取和寫入基礎資料的主參與者必須對指向基礎資料位置的「資料目錄」表格具有 Lake Formation 資料存取權限。
- 在 Lake Formation 註冊基礎資料位置時，讀取和寫入基礎資料的主體必須具有 `lakeformation:GetDataAccess` IAM 權限。

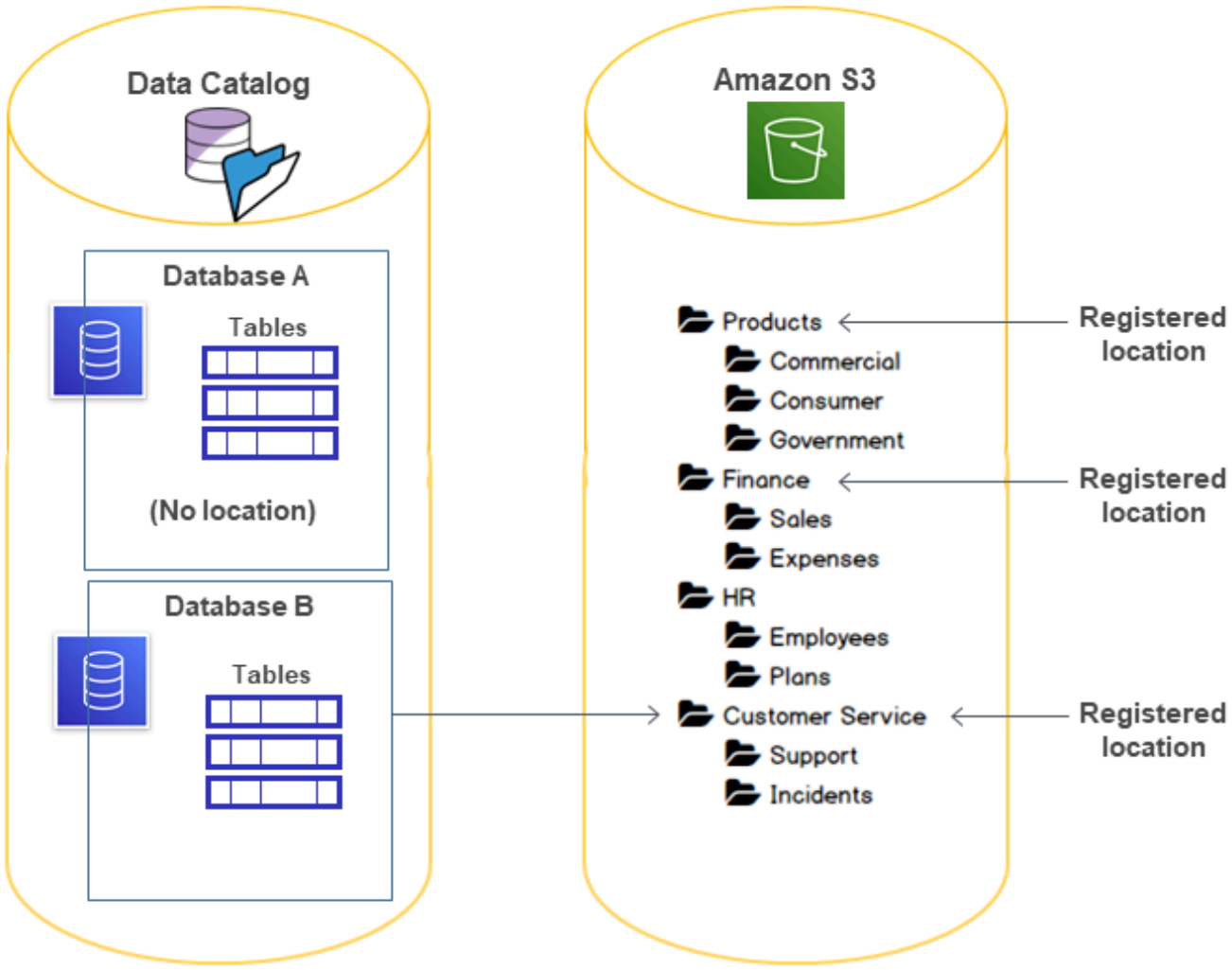
> **Note**
>
> 如果您可以透過 IAM 或 Amazon S3 政策存取這些位置，則 Lake Formation 許可模型不會阻止透過 Amazon S3 API 或主控台存取 Amazon S3 位置。您可以將 IAM 政策附加到主體以封鎖此存取權。

有關數據位置權限的更多

資料位置權限可控制資料目錄資料庫和表格上建立和更新作業的結果。規則如下：

- 主體必須在 Amazon S3 位置具有明確或隱含的資料位置許可，才能建立或更新指定該位置的資料庫或表格。
- 使用控制台、API 或授予明確權限`DATA_LOCATION_ACCESS` AWS CLI。
- 當資料庫具有指向已註冊位置的 location 屬性、主體具有資料庫的權限，且主體嘗試在該位置或子位置建立資料表時，就會授與隱含權限。`CREATE_TABLE`
- 如果主參與者被授與某個位置的資料位置權限，則主參與者對所有子位置都具有資料位置權限。
- 主體不需要資料位置權限，即可對基礎資料執行讀取/寫入作業。擁有`SELECT`或`INSERT`資料存取權限就足夠了。資料位置權限僅適用於建立指向該位置的資料目錄資源。

考慮下圖中顯示的場景。



在這張圖中：

- Amazon S3 桶 `ProductsFinance`，並`Customer Service`在 Lake Formation 註冊。
- `Database A`沒有位置屬性，且`Database B`具有指向`Customer Service`值區的位置屬性。
- 用戶`CREATE_TABLE`在兩個數據庫上都`datalake_user`有。
- 使用`datalake_user`者僅獲得`Products`儲存貯體的資料位置權限。

以下是使用者`datalake_user`嘗試在特定資料庫中的特定位置建立目錄表格時的結果。

datalake_user嘗試建立資料表的位置

資料庫和位置	成功或失敗	原因
資料庫 A 在 Finance/Sales	失敗	沒有資料位置權限
資料庫 A 在 Products	成功	具有資料位置權限
資料庫 A 在 HR/Plans	成功	地點尚未註冊
資料庫 B 在 Customer Service/Incidents	成功	資料庫的位置屬性位於 Customer Service

如需詳細資訊，請參閱下列內容：

- 將 Amazon S3 位置新增至您的資料湖
- Lake Formation 權限參考
- Lake Formation 角色和 IAM 許可參考

Lake Formation 角色和 IAM 許可參考

本節列出了一些建議的 Lake Formation 角色及其建議的 AWS Identity and Access Management（IAM）許可。如需 Lake Formation 權限的相關資訊，請參閱the section called “Lake Formation 權限參考”。

AWS Lake Formation 人物

下表列出建議的 AWS Lake Formation 人物角色。

Lake Formation 角色

人物	描述
IAM 管理員 (超級使用者)	(必要) 可建立 IAM 使用者和角色的使用者。具有受AdministratorAccess AWS 管策略。擁有所有 Lake Formation 資源的所有權限。可以新增資料湖管理員。如果未指定資料湖管理員，則無法授與 Lake Formation 權限。

人物	描述
資料湖管理員	(必要) 可註冊 Amazon S3 位置、存取資料目錄、建立資料庫、建立和執行工作流程、授與 Lake Formation 權限給其他使用者，以及檢視 AWS CloudTrail 日誌的使用者。IAM 許可比 IAM 管理員少，但足以管理資料湖。無法新增其他資料湖管理員。
唯讀管理員	(選擇性) 可檢視主參與者、資料目錄資源、權限和 AWS CloudTrail 記錄檔的使用者，而無需進行更新的權限。
數據工程師	(選擇性) 可建立資料庫、建立和執行編目器和工作流程，以及授與檢索器和工作流程所建立之「資料目錄」表格 Lake Formation 權限的使用者。我們建議您讓所有的資料工程師資料庫建立者。如需詳細資訊，請參閱 建立資料庫。
資料分析	(選擇性) 可以使用，例如，對資料湖執行查詢的使用者 Amazon Athena。只有足夠的權限來執行查詢。
工作流角色	(必要) 代表使用者執行工作流程的角色。您可以在從藍圖建立工作流程時指定此角色。

AWS 對 Lake Formation 的管理政策

您可以使用 AWS 受管政策和內嵌政策，授予使用 AWS Lake Formation 所需的 AWS Identity and Access Management (IAM) 許可。以下 AWS 管理政策適用於 Lake Formation。

AWS 受管理的策略：AWSLakeFormationDataAdmin

AWSLakeFormationDataAdmin原則會授予管理存取權限以 AWS Lake Formation 及相關服務 (例 AWS Glue 如管理資料湖)。

您可以附加`AWSLakeFormationDataAdmin`至您的使用者、群組和角色。

許可權詳細

- `CloudTrail`— 允許主參與者檢視記 AWS CloudTrail 錄檔。若要檢閱資料湖設定中的任何錯誤，則必須執行此動作。

- Glue— 可讓主參與者檢視、建立及更新「資料目錄」中的中繼資料表格和資料庫。這包括以Get、、、ListCreate、Update和開頭的 API 作業Search。Delete這是管理資料湖表格的中繼資料所必需的。
- IAM— 允許主體擷取附加至角色的 IAM 使用者、角色和政策的相關資訊。資料管理員必須檢閱和列出 IAM 使用者和角色，才能授予 Lake Formation 權限。
- Lake Formation— 授予資料湖管理員需要 Lake Formation 權限才能管理資料湖。
- S3— 允許主體擷取 Amazon S3 儲存貯體及其位置的相關資訊，以便設定資料湖的資料位置。

```
"Statement": [
        {
            "Sid": "AWSLakeFormationDataAdminAllow",
            "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:ListWorkflows",
                "glue:BatchGetWorkflows",
                "glue:DeleteWorkflow",
                "glue:GetWorkflowRuns",
                "glue:StartWorkflowRun",
                "glue:GetWorkflow",
                "s3:ListBucket",
                "s3:GetBucketLocation",
                "s3:ListAllMyBuckets",
                "s3:GetBucketAcl",
                "iam:ListUsers",
```

```
                "iam:ListRoles",
                "iam:GetRole",
                "iam:GetRolePolicy"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AWSLakeFormationDataAdminDeny",
            "Effect": "Deny",
            "Action": [
                "lakeformation:PutDataLakeSettings"
            ],
                "Resource": "*"
        }
    ]
}
```

> **Note**
>
> 此原AWSLakeFormationDataAdmin則不會授與資料湖管理員所需的所有權限。需要其他權限才能建立和執行工作流程，以及使用服務連結角色註冊位置AWSServiceRoleForLakeFormationDataAccess。如需詳細資訊，請參閱 建立資料湖管理員 及 使用服務連結角色進行 Lake Formation。

AWS 受管理的策略：AWSLakeFormationCrossAccountManager

AWSLakeFormationCrossAccountManager政策通過 Lake Formation 提供跨帳戶訪問 AWS Glue 資源的權限，並授予對其他必要服務的讀取訪問權限，例如 AWS Organizations 和 AWS RAM。

您可以附加AWSLakeFormationCrossAccountManager至您的使用者、群組和角色。

許可權詳細

此政策包含以下許可。

- Glue— 允許主參與者設定或刪除資料目錄資源原則以進行存取控制。
- Organizations— 允許主參與者擷取組織的帳戶與組織單位 (OU) 資訊。
- ram:CreateResourceShare— 允許主參與者建立資源共用。
- ram:UpdateResourceShare允許主參與者修改指定資源共用的某些內容。

- ram:DeleteResourceShare— 允許主參與者刪除指定的資源共用。
- ram:AssociateResourceShare— 允許主參與者將指定的主參與者清單和資源清單新增至資源共用。
- ram:DisassociateResourceShare— 允許主參與者從參與指定的資源共用中移除指定的主參與者或資源。
- ram:GetResourceShares— 允許主參與者擷取有關您擁有或與您共用之資源共用的詳細資訊。
- ram:RequestedResourceType— 允許主參與者擷取資源類型 (資料庫、表格或目錄)。
- AssociateResourceSharePermission— 允許主參與者新增或取代資源共用中包含之資源類型的 AWS RAM 權限。您可以只有一個與資源共用中的每個資源類型相關聯的權限。

```
{
    "Version": "2012-10-17",
    "Statement": [{
            "Sid": "AllowCreateResourceShare",
            "Effect": "Allow",
            "Action": [
                "ram:CreateResourceShare"
            ],
            "Resource": "*",
            "Condition": {
                "StringLikeIfExists": {
                    "ram:RequestedResourceType": [
                        "glue:Table",
                        "glue:Database",
                        "glue:Catalog"
                    ]
                }
            }
        },
        {
            "Sid": "AllowManageResourceShare",
            "Effect": "Allow",
            "Action": [
                "ram:UpdateResourceShare",
                "ram:DeleteResourceShare",
                "ram:AssociateResourceShare",
                "ram:DisassociateResourceShare",
                "ram:GetResourceShares"
            ],
            "Resource": "*",
```

```
            "Condition": {
                "StringLike": {
                    "ram:ResourceShareName": [
                        "LakeFormation*"
                    ]
                }
            }
        },
        {
            "Sid": "AllowManageResourceSharePermissions",
            "Effect": "Allow",
            "Action": [
                "ram:AssociateResourceSharePermission"
            ],
            "Resource": "*",
            "Condition": {
                "StringLike": {
                    "ram:PermissionArn": [
                        "arn:aws:ram::aws:permission/AWSRAMLFEnabled*"
                    ]
                }
            }
        },
        {
            "Sid": "AllowXAcctManagerPermissions",
            "Effect": "Allow",
            "Action": [
                "glue:PutResourcePolicy",
                "glue:DeleteResourcePolicy",
                "organizations:DescribeOrganization",
                "organizations:DescribeAccount",
                "ram:Get*",
                "ram:List*"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AllowOrganizationsPermissions",
            "Effect": "Allow",
            "Action": [
                "organizations:ListRoots",
                "organizations:ListAccountsForParent",
                "organizations:ListOrganizationalUnitsForParent"
            ],
```

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

AWS 受管理的策略：AWSGlueConsoleFullAccess

AWSGlueConsoleFullAccess當附加原則的身分識別使用時，策略會授與 AWS Glue 資源的完整存取權 AWS Management Console。如果您依照此政策中指定的資源命名慣例，使用者就能擁有完整的主控台功能。此原則通常會附加至 AWS Glue 主控台的使用者。

此外AWS Glue，Lake Formation 擔任服務角色，AWSGlueServiceRole允許存取相關服務，包括 Amazon 彈性運算雲 (Amazon EC2)、亞馬遜簡單儲存服務 (Amazon S3) 和亞馬遜 CloudWatch。

AWS managed policy:LakeFormationDataAccessServiceRolePolicy

此原則會附加至名為的服務連結角色，ServiceRoleForLakeFormationDataAccess該角色可讓服務根據您的要求對資源執行動作。您無法將此政策附加到 IAM 身分。

此政策允許 Lake Formation 整合 AWS 服務 (例如 Amazon Athena 或 Amazon Redshift) 使用服務連結角色來探索 Amazon S3 資源。

若要取得更多資訊，請參閱使用服務連結角色進行 Lake Formation。

許可權詳細

此原則包含下列權限。

- s3:ListAllMyBuckets— 傳回要求已驗證寄件者所擁有之所有值區的清單。

```
{
 "Version": "2012-10-17",
 "Statement": [
  {
   "Sid": "LakeFormationDataAccessServiceRolePolicy",
   "Effect": "Allow",
   "Action": [
    "s3:ListAllMyBuckets"
   ],
   "Resource": [
    "arn:aws:s3:::*"
```

```
    ]
   }
  ]
}
```

Lake Formation AWS 管理政策的更新

檢視自此服務開始追蹤這些變更以來，對 Lake Formation AWS 管理政策的更新相關詳細資訊。

變更	描述	日期
Lake Formation 更新`AWSLakeFormationCrossAccountManager` 政策。	Lake Formation 通過在AWSLakeFormationCrossAccountManager政策聲明中添加 Sid 元素來增強了政策。	2024 年三月
Lake Formation 更新`AWSLakeFormationDataAdmin` 政策。	Lake Formation 透過將 Sid 元素新增至AWSLakeFormationDataAdmin政策陳述式並移除多餘動作，藉此增強了政策。	2024 年三月
Lake Formation 更新`LakeFormationDataAccessServiceRolePolicy` 政策。	Lake Formation 通過在LakeFormationDataAccessServiceRolePolicy策略聲明中添加 Sid 元素來增強了策略。	2024年2月份
Lake Formation 更新`AWSLakeFormationCrossAccountManager` 政策。	Lake Formation 通過添加新的權限來在混合訪問模式下啟用跨帳戶數據共享來增強了AWSLakeFormationCrossAccountManager政策。	2023 年十月
Lake Formation 更新`AWSLakeFormationCr`	Lake Formation 增強了AWSLakeFormationCrossAccountManager策略，以便在首次共用資源時，每個收件者帳號	2022 年 5 月 6 日

變更	描述	日期
ossAccoun tManager 政策。	僅建立一個資源共用。之後使用相同帳號共用的所有資源都會附加至相同的資源共用。	
Lake Formation 開始跟踪變化。	Lake Formation 開始跟踪其 AWS 管理政策的變化。	2022 年 5 月 6 日

人物角色建議的權限

以下是每個角色的建議權限。不包括 IAM 管理員，因為該使用者擁有所有資源的所有權限。

主題

資料湖管理員權限

> **Important**
>
> 在下列策略中，請以<account-id>有效的 AWS 帳號取代，並取代為<workflow_role>具有執行工作流程權限的角色名稱 (如中所定義) 工作流角色權限。

政策類型	政策
AWS 受管理政策	• AWSLakeFormationDataAdmin • LakeFormationDataAccessServiceRolePo licy (服務連結角色政策) • AWSGlueConsoleFullAccess (選用) • CloudWatchLogsReadOnlyAccess (選用)

政策類型	政策
	• AWSLakeFormationCrossAccountManager (選用) • AmazonAthenaFullAccess (選用) 如需選擇性 AWS 受管理原則的相關資訊，請參閱the section called “建立資料湖管理員”。

內聯政策（用於創建 Lake Formation 服務鏈接角色）	{
	"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"
	}
	]
	}

政策類型	政策

(選擇性) 內嵌原則 (工作流程角色的密碼原則)。只有在資料湖管理員建立並執行工作流程時，才需要這樣做。	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *<account-id>* :role/*<workflow_role>* "
	]
	}
	]
	}

（可選）內嵌政策（如果您的帳戶正在授予或接收跨帳戶 Lake Formation 權限）。此原則適用於接受或拒絕 AWS RAM 資源共用邀請，以及允許授與跨帳號權限給組織。`ram:EnableSharingWithAwsOrganization` 只有 AWS Organizations 管理帳戶中的資料湖管理員才需要。	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"ram:AcceptResourceShareInvitation",
	"ram:RejectResourceShareInvitation",
	"ec2:DescribeAvailabilityZones",
	"ram:EnableSharingWithAwsOrganization"
	],
	"Resource": "*"
	}
	]
	}

唯讀管理員權限

Policy type (政策類型)	政策
內嵌原則 (基本)	{ "Version":"2012-10-17", "Statement":[{ "Effect":"Allow", "Action":["lakeformation:GetEffectivePermissio nsForPath", "lakeformation:ListPermissions", "lakeformation:ListDataCellsFilter", "lakeformation:GetDataCellsFilter", "lakeformation:SearchDatabasesByLFTa gs", "lakeformation:SearchTablesByLFTags", "lakeformation:GetLFTag", "lakeformation:ListLFTags", "lakeformation:GetResourceLFTags", "lakeformation:ListLakeFormationOpti ns", "cloudtrail:DescribeTrails", "cloudtrail:LookupEvents", "glue:GetDatabase", "glue:GetDatabases", "glue:GetConnections", "glue:SearchTables", "glue:GetTable", "glue:GetTableVersions", "glue:GetPartitions", "glue:GetTables", "glue:GetWorkflow", "glue:ListWorkflows", "glue:BatchGetWorkflows", "glue:GetWorkflowRuns", "glue:GetWorkflow", "s3:ListBucket", "s3:GetBucketLocation", "s3:ListAllMyBuckets", "s3:GetBucketAcl", "iam:ListUsers",

Policy type (政策類型)	政策

	"iam:ListRoles",
	"iam:GetRole",
	"iam:GetRolePolicy"
	],
	"Resource":"*"
	},
	{
	"Effect":"Deny",
	"Action":[
	"lakeformation:PutDataLakeSettings"
	],
	"Resource":"*"
	}
	]
	}

資料工程師權限

> **Important**
>
> 在下列策略中，請以<account-id>有效的 AWS 帳號取代，並取代為<workflow_role>工作流程角色的名稱。

政策類型	政策
AWS 受管理政策	`AWSGlueConsoleFullAccess`
內嵌原則 (基本)	{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:GetDataAccess", "lakeformation:GrantPermissions", "lakeformation:RevokePermissions", "lakeformation:BatchGrantPermissions",

政策類型	政策
	"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": "*" }] }

政策類型	政策

內嵌政策 (適用於受控資料表上的作業，包括交易中的作業)	{
	"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": "*"
	}
	]
	}

政策類型	政策

內嵌原則 (適用於使用以 Lake Formation 標記為基礎的存取控制 (LF-TBAC) 方法的中繼資料存取控制)	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": [
	"lakeformation:AddLFTagsToResource",
	"lakeformation:RemoveLFTagsFromResou
	rce",
	"lakeformation:GetResourceLFTags",
	"lakeformation:ListLFTags",
	"lakeformation:GetLFTag",
	"lakeformation:SearchTablesByLFTags",
	"lakeformation:SearchDatabasesByLFTags"
	],
	"Resource": "*"
	}
	]
	}

內嵌原則 (工作流程角色的密碼原則)	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *<account-id>* :role/*<workflow
	_role>* "
	]
	}
	]
	}

資料分析師權限

政策類型	政策
AWS 受管理政策	AmazonAthenaFullAccess

內嵌原則 (基本)	{
	"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": "*"
	}
	]
	}

(選擇性) 內嵌政策 (適用於受控資料表上的作業，包括交易內的作業)	{ "Version": "2012-10-17", "Statement": [{ "Effect": "Allow", "Action": ["lakeformation:StartTransaction", "lakeformation:CommitTransaction", "lakeformation:CancelTransaction", "lakeformation:ExtendTransaction", "lakeformation:DescribeTransaction",

政策類型	政策

	<pre>"lakeformation:ListTransactions",
	"lakeformation:GetTableObjects",
	"lakeformation:UpdateTableObjects",
	"lakeformation:DeleteObjectsOnCancel"
	],
	"Resource": "*"
	}
	]
	}</pre>

工作流角色權限

此角色具有執行工作流程所需的權限。建立工作流程時，您可以指定具有這些權限的角色。

> **Important**
>
> 在下列政策中，請以<region>有效的 AWS 區域識別碼 (例如us-east-1) 取代<account-id>為有效的 AWS 帳戶號碼、<workflow_role>工作流程角色的名稱，以及 *AWS CloudTrail* <your-s3-cloudtrail-bucket>日誌的 Amazon S3 路徑。

政策類型	政策
AWS 受管理政策	AWSGlueServiceRole

內嵌政策 (資料存取)	<pre>{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "Lakeformation",
	"Effect": "Allow",
	"Action": [
	"lakeformation:GetDataAccess",
	"lakeformation:GrantPermissions"
	],
	"Resource": "*"
	}
	]

政策類型	政策
	} }

內嵌原則 (工作流程角色的密碼原則)	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Sid": "PassRolePermissions",
	"Effect": "Allow",
	"Action": [
	"iam:PassRole"
	],
	"Resource": [
	"arn:aws:iam:: *<account-id>* :role/*<workflow_role>* "
	]
	}
	]
	}

內嵌政策 (用於擷取資料湖外的資料，例如 AWS CloudTrail 記錄)	{
	"Version": "2012-10-17",
	"Statement": [
	{
	"Effect": "Allow",
	"Action": ["s3:GetObject", "s3:ListBucket"],
	"Resource": ["arn:aws:s3::: *<your-s3-cloudtrail-bucket>* /*"]
	}
	]
	}

變更資料湖的預設設定

若要維持與的回溯相容性AWS Glue， AWS Lake Formation 請使用下列初始安全性設定：

- 此Super權限會授與所有現有資AWS Glue料目錄資源IAMAllowedPrincipals上的群組。
- 新資料目錄資源會啟用「僅使用 IAM 存取控制」設定。

這些設定有效地導致資料目錄資源和 Amazon S3 位置的存取僅由 AWS Identity and Access Management (IAM) 政策控制。個別 Lake Formation 權限未生效。

該`IAMAllowedPrincipals`群組包括您的 IAM 政策允許存取資料目錄資源的所有 IAM 使用者和角色。此`Super`權限可讓主體在授與該作業的資料庫或表格上執行每個受支援的 Lake Formation 作業。

若要變更安全性設定，使資料目錄資源 (資料庫和表格) 的存取由 Lake Formation 權限管理，請執行下列動作：

1. 變更新資源的預設安全性設定。如需說明，請參閱變更預設權限模型或使用混合存取模式。
2. 變更現有資料目錄資源的設定。如需說明，請參閱將AWS Glue資料權限升級至 AWS Lake Formation 模型。

使用 Lake Formation **`PutDataLakeSettings`** API 操作更改默認安全設置

您也可以使用 Lake Formation PutDataLakeSettingsAPI 操作來更改默認安全設置。此動作會將選用的目錄 ID 和DataLakeSettings結構當做引數。

若要在新資料庫和資料表上強制執行 Lake Formation 的中繼資料和基礎資料存取控制，請依下列方式撰寫`DataLakeSettings`結構程式

> **Note**
>
> <AccountID>以有效的 AWS 帳戶 ID 和有<Username>效的 IAM 使用者名稱取代。您可以將多個使用者指定為資料湖管理員。

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ],
        "CreateDatabaseDefaultPermissions": [],
        "CreateTableDefaultPermissions": []
    }
}
```

您也可以編寫如下結構的代碼。省略CreateDatabaseDefaultPermissions或CreateTableDefaultPermissions參數等同於傳遞空白清單。

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ]
    }
}
```

此動作可有效地撤銷IAMAllowedPrincipals群組對新資料庫和資料表的所有 Lake Formation 權限。建立資料庫時，您可以覆寫此設定。

若要僅由 IAM 在新資料庫和資料表上強制執行中繼資料和基礎資料存取控制，請依下列方式編寫DataLakeSettings結構的程式碼

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

這會將 Super Lake Formation 的權限授予新資料庫和資料表上的`IAMAllowedPrincipals`群組。建立資料庫時，您可以覆寫此設定。

> **Note**
>
> 在先前的`DataLakeSettings`結構中，的唯一允許的值`DataLakePrincipalIdentifier`是`IAM_ALLOWED_PRINCIPALS`，且唯一允許的值`Permissions`為`ALL`。

隱含 Lake Formation 權限

AWS Lake Formation 將下列隱含權限授與資料湖管理員、資料庫建立者和表格建立者。

資料湖管理員

- 可`Describe`存取資料目錄中的所有資源，但從另一個帳戶直接共用至不同主體的資源除外。系統管理員無法撤銷此存取權。
- 擁有資料湖中任何位置的資料位置權限。
- 可以授與或撤銷對資料目錄中任何資源的存取權給任何主體 (包括自己)。系統管理員無法撤銷此存取權。
- 可以在「資料目錄」中建立資料庫。
- 可以將建立資料庫的權限授與其他使用者。

> **Note**
>
> 資料湖管理員只有在具有 IAM 許可的情況下才能註冊 Amazon S3 位置。本指南中建議的資料湖管理員原則會授與這些權限。此外，資料湖管理員沒有隱含的權限，可以刪除其他人建立的資料庫或變更/刪除表格。但是，他們可以授予自己這樣做的權限。

如需資料湖管理員的詳細資訊，請參閱建立資料湖管理員。

數據庫創建

- 擁有他們所建立之資料庫的所有資料庫權限、擁有在資料庫中建立之資料表的權限，以及可授與相同 AWS 帳戶中其他主體在資料庫中建立資料表的權限。也具有AWSLakeFormationCrossAccountManager AWS 受管理策略的資料庫建立者可以將資料庫權限授與其他 AWS 帳戶或組織。

  資料湖管理員可以使用 Lake Formation 主控台或 API 來指定資料庫建立者。

  > **Note**
  >
  > 資料庫建立者對於其他人在資料庫中建立的資料表沒有隱含的權限。

如需詳細資訊，請參閱 建立資料庫。

表格建立者

- 擁有他們所建立之資料表的所有權限。
- 可以將這些資料表所建立之所有資料表的權限授與相同 AWS 帳戶中的主參與者。
- 如果帳戶或組織具有AWSLakeFormationCrossAccountManager AWS 受管理的策略，則可以將其建立的所有表格的權限授與給其他 AWS 帳戶或組織。
- 可以檢視包含它們所建立之資料表的資料庫。

Lake Formation 權限參考

若要執行 AWS Lake Formation 作業，主體需要 Lake Formation 許可和 AWS Identity and Access Management (IAM) 許可。您通常會使用粗略的存取控制政策來授予 IAM 許可，如中所述。the section called “Lake Formation 許可權概述 ”您可以使用主控台、API 或 AWS Command Line Interface (AWS CLI) 授予 Lake Formation 權限。

要了解如何授予或撤銷 Lake Formation 權限，請參閱the section called “授與和撤銷資料目錄權限”和the section called “授與資料位置權限”。

> **Note**
>
> 本節中的範例顯示如何將權限授與相同 AWS 帳戶中的主參與者。如需跨帳戶補助的範例，請參閱the section called “跨帳戶資料共用”。

每種資源類型的 Lake Formation 型權限

以下是可用於每種資源類型的有效 Lake Formation 權限：

資源	權限
Database	ALL (Super)
	ALTER
	CREATE_TABLE
	DESCRIBE
	DROP
Table	ALL (Super)
	ALTER
	DELETE
	DESCRIBE
	DROP
	INSERT
	SELECT
View	ALL (Super)
	SELECT
	DESCRIBE
	DROP
Data Catalog	CREATE_DATABASE
Amazon S3 location	DATA_LOCATION_ACCESS

資源	權限	
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	
	DELETE	
	DROP	
	INSERT	
	SELECT	
Resource link - Database or Table	DESCRIBE	
	DROP	

資源	權限
Table with data filters	DESCRIBE
	DROP
	SELECT
Table with column filter	SELECT

主題

- Lake Formation 授予和撤銷 AWS CLI 命令
- Lake Formation 權限

Lake Formation 授予和撤銷 AWS CLI 命令

本節中的每個權限描述都包含使用 AWS CLI 命令授與權限的範例。以下是湖的形成grant-permissions和revoke-permissions AWS CLI 命令的概要。

```
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>]
```

如需這些命令的詳細說明，請參閱命令參考中的授與權限和撤銷權限。AWS CLI 本節提供有關此--principal選項的其他資訊。

--principal選項的值為下列其中一項：

- (IAM) 使用者或角色的 Amazon 資源名稱 AWS Identity and Access Management (ARN)
- ARN 適用於透過 SAML 提供者進行驗證的使用者或群組，例如 Microsoft 作用中目錄同盟服務 (AD FS)
- 適用於 Amazon QuickSight 用戶或組的 ARN
- 針對跨帳戶權限、 AWS 帳戶 ID、組織 ID 或組織單位 ID

以下是所有--principal類型的語法和範例。

主體是 IAM 使用者

語法:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/<user-name>
```

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1
```

主體是 IAM 角色

語法:

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:role/<role-name>
```

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:role/workflowrole
```

主體是透過 SAML 提供者驗證的使用者

語法：

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:user/<user-name>
```

範例：

```
--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
```

主體是透過 SAML 提供者驗證的群組

語法：

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:saml-
provider/<SAMLproviderName>:group/<group-name>
```

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
idp1:group/data-scientists
```

```
--principal DataLakePrincipalIdentifier=arn:aws:iam::111122223333:saml-provider/
AthenaLakeFormationOkta:group/my-group
```

校長是 Amazon QuickSight 企業版用戶

語法：

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-
id>:user/<namespace>/<user-name>
```

> **Note**
>
> 對於<namespace>，您必須指定default。

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-
east-1:111122223333:user/default/bi_user1
```

校長是 Amazon QuickSight 企業版集團

語法：

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:<region>:<account-
id>:group/<namespace>/<group-name>
```

> **Note**
>
> 對於<namespace>，您必須指定default。

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:quicksight:us-
east-1:111122223333:group/default/data_scientists
```

本金是一個 AWS 帳戶

語法:

```
--principal DataLakePrincipalIdentifier=<account-id>
```

範例：

```
--principal DataLakePrincipalIdentifier=111122223333
```

主參與者是一個組織

語法:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-
id>:organization/<organization-id>
```

範例：

```
--principal
 DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:organization/o-
abcdefghijkl
```

校長是一個組織單位

語法:

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::<account-
id>:ou/<organization-id>/<organizational-unit-id>
```

範例：

```
--principal DataLakePrincipalIdentifier=arn:aws:organizations::111122223333:ou/o-
abcdefghijkl/ou-ab00-cdefghij
```

主體是 IAM 身分中心身分識別使用者或群組

範例：使用者

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserID>
```

範例：群組：

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::group/<GroupID>
```

主體是 IAM 群組-**`IAMAllowedPrincipals`**

Lake Formation 會將資料目錄中所有資料庫和資料表的Super權限設定為預設呼叫IAMAllowedPrincipals的群組。如果此群組權限存在於 AWS Glue資料庫或資料表上，您帳戶中的所有主體都可以透過的 IAM 主體政策存取資源。當您開始使用 Lake Formation 許可以保護先前受 IAM 政策保護的資料目錄資源時，它會提供回溯相容性 AWS Glue。

當您使用 Lake Formation 管理「資料目錄」資源的權限時，您需要先撤銷資源的IAMAllowedPrincipals權限，或選擇主參與者和資源使用混合式存取模式，Lake Formation 權限才能運作。

範例：

```
--principal DataLakePrincipalIdentifier=IAM_Allowed_Principals
```

主體是 IAM 群組-**ALLIAMPrincipals**

當您授與資料目錄資源上的ALLIAMPrincipals群組權限時，帳戶中的每個主體都可以使用 Lake Formation 許可和 IAM 權限存取資料目錄資源。

範例：

```
--principal DataLakePrincipalIdentifier=123456789012:IAMPrincipals
```

Lake Formation 權限

本節包含您可以授與主參與者的可用「Lake Formation」權限。

ALTER

權限	授與此資源	受贈人也需要
ALTER	DATABASE	glue:UpdateDatabase
ALTER	TABLE	glue:UpdateTable
ALTER	LF-Tag	lakeformation:UpdateLFTag

具有此權限的主體可以變更「資料目錄」中資料庫或表格的中繼資料。對於表格，您可以變更資料欄結構定義並新增資料欄參數。您無法變更中繼資料表所指向之基礎資料中的資料行。

如果要變更的屬性是已註冊的 Amazon Simple Storage Service (Amazon S3) 位置，則主體必須具有新位置的資料位置許可。

Example

下列範例會將ALTER權限授與 AWS 帳戶 1111-2222-3333 retail 中資料庫datalake_user1上的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "ALTER" --resource '{ "Database": {"Name":"retail"}}'
```

Example

下列範例會授與ALTER資料庫中資料表datalake_user1上inventory的使用者retail。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

CREATE_DATABASE

權限	授與此資源	受贈人也需要
CREATE_DATABASE	Data Catalog	glue:CreateDatabase

具有此權限的主參與者可以在「資料目錄」中建立中繼資料資料庫或資源連結。主參與者也可以在資料庫中建立資料表。

Example

下列範例會授與CREATE_DATABASE AWS 帳號 1111-2222-datalake_user1 3333 中的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

當主體在「資料目錄」中建立資料庫時，不會授與任何基礎資料的權限。已授與下列其他中繼資料權限 (以及將這些權限授與其他使用者的能力)：

- CREATE_TABLE在資料庫中
- ALTER 資料庫
- DROP 資料庫

建立資料庫時，主體可以選擇性地指定 Amazon S3 位置。視主體是否具有資料位置權限而定，在所有情況下，CREATE_DATABASE權限可能不足以建立資料庫。牢記以下三種情況很重要。

建立資料庫使用案例	需要的許可
位置屬性未指定。	CREATE_DATABASE 就足夠了。
指定了位置屬性，並且該位置不由 Lake Formation 管理（未註冊）。	CREATE_DATABASE 就足夠了。
位置屬性被指定，並且該位置由 Lake Formation（已註冊）管理。	CREATE_DATABASE 需要加上指定位置的資料位置權限。

CREATE_TABLE

權限	授與此資源	受贈人也需要
CREATE_TABLE	DATABASE	glue:CreateTable

具有此權限的主參與者可以在指定資料庫內的「資料目錄」中建立中繼資料表格或資源連結。

Example

下列範例會授與使用datalake_user1者在 AWS 帳戶 1111-2222-3333 中建立資retail料庫中資料表的權限。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

當主參與者在「資料目錄」中建立資料表時，該資料表上的所有 Lake Formation 權限都會授與主參與者，並能夠將這些權限授與其他人。

跨帳戶補助

如果資料庫擁有者帳戶授CREATE_TABLE與收件者帳戶，且收件者帳戶中的使用者在擁有者帳戶的資料庫中成功建立資料表，則適用下列規則：

- 收件者帳戶中的使用者和資料湖管理員對資料表具有所有 Lake Formation 權限。他們可以將表格的權限授與其帳戶中的其他主參與者。他們無法將權限授與擁有者帳戶或任何其他帳戶中的主體。
- 擁有者帳戶中的資料湖管理員可以將表格的權限授與其帳戶中的其他主體。

資料位置權限

當您嘗試建立指向 Amazon S3 位置的表格時，視您是否具有資料位置CREATE_TABLE許可而定，權限可能不足以建立表格。牢記以下三種情況很重要。

建立資料表使用案例	需要的許可
指定的地點不由 Lake Formation 管理（未註冊）。	CREATE_TABLE 就足夠了。
指定的位置由 Lake Formation 管理 (已註冊)，且包含的資料庫沒有位置屬性，或具有不是表格位置 Amazon S3 前綴的位置屬性。	CREATE_TABLE 需要加上指定位置的資料位置權限。
指定的位置由 Lake Formation 管理 (已註冊)，且包含的資料庫具有指向已註冊位置的位置屬性，並且是表格位置的 Amazon S3 前置詞。	CREATE_TABLE 就足夠了。

DATA_LOCATION_ACCESS

權限	授與此資源	受贈人也需要
DATA_LOCATION_ACCESS	Amazon S3 位置	(該位置上的 Amazon S3 許可，必須由用於註冊位置的角色指定。)

這是唯一的資料位置權限。具有此權限的主體可以建立指向指定 Amazon S3 位置的中繼資料庫或表格。必須註冊地點。對某個位置具有資料位置權限的主參與者也具有子位置的位置權限。

Example

下列範例會s3://products/retail將資料位置權限授與 AWS 帳戶 1111-2222-3333 datalake_user1 中的使用者。

```
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不需要查詢或更新基礎資料。此權限僅適用於建立資料目錄資源。

如需資料位置權限的詳細資訊，請參閱Underlying data access control。

DELETE

權限	授與此資源	受贈人也需要
DELETE	TABLE	如果位置已註冊，則不需要額外的 IAM 許可。)

具有此權限的主體可以刪除表格指定之 Amazon S3 位置的基礎資料。主體也可以在 Lake Formation 主控台上檢視表格，並使用 AWS Glue API 擷取資料表的相關資訊。

Example

下列範例會將DELETE權限授與 AWS 帳戶 1111-2222-3333 資料庫inventoryretail中資料表datalake_user1上的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DELETE" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

此權限僅適用於 Amazon S3 中的資料，而不適用於其他資料存放區中的資料，例如 Amazon Relational Database Service (Amazon RDS)。

DESCRIBE

權限	授與此資源	受贈人也需要
DESCRIBE	表格資源連結	glue:GetTable
	資料庫資源連結	glue:GetDatabase

權限	授與此資源	受贈人也需要
DESCRIBE	DATABASE	glue:GetDatabase
DESCRIBE	TABLE	glue:GetTable
DESCRIBE	LF-Tag	glue:GetTable glue:GetDatabase lakeformation:GetResourceLFTags lakeformation:ListLFTags lakeformation:GetLFTag lakeformation:SearchTablesByLFTags lakeformation:SearchDatabasesByLFTags

具有此權限的主參與者可以檢視指定的資料庫、表格或資源連結。不會以隱含方式授與其他資料目錄權限，也不會以隱含方式授與資料存取權限。資料庫和資料表會顯示在整合式服務的查詢編輯器中，但除非授與其他 Lake Formation 權限 (例如SELECT)，否則無法對其進行任何查詢。

例如，擁有資料庫DESCRIBE的使用者可以看到資料庫和所有資料庫中繼資料 (描述、位置等)。但是，用戶無法找出數據庫包含哪些表，也無法刪除，更改或創建數據庫中的表。同樣地，擁有資料表DESCRIBE的使用者可以看到資料表和資料表中繼資料 (說明、結構描述、位置等)，但無法捨棄、變更或執行對資料表的查詢。

以下是一些額外的規則DESCRIBE：

- 如果使用者對資料庫、資料表或資源連結具有其他 Lake Formation 權限，DESCRIBE則會隱含授與。
- 如果用戶只有SELECT一個表（部分SELECT）列的子集，則用戶只能看到這些列。

- 您無法授DESCRIBE予在表格上選取部分功能的使用者。相反地，您無法為授與的表格指定欄包含清單或排除清單。DESCRIBE

Example

下列範例會將DESCRIBE權限授與 AWS 帳號 1111-2222-3333 之資料庫inventory-link中資料表資源連結datalake_user1上retail的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory-link"}}'
```

DROP

權限	授與此資源	受贈人也需要
DROP	DATABASE	glue:DeleteDatabase
DROP	TABLE	glue:DeleteTable
DROP	LF-Tag	lakeformation:DeleteLFTag
DROP	資料庫資源連結 表格資源連結	glue:DeleteDatabase glue:DeleteTable

具有此權限的主參與者可以卸除「資料目錄」中的資料庫、表格或資源連結。您無法將資料庫上的 DROP 授與外部帳戶或組織。

> **Warning**
>
> 刪除數據庫丟棄數據庫中的所有表。

Example

下列範例會將DROP權限授與 AWS 帳戶 1111-2222-3333 retail 中資料庫datalake_user1上的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Database": {"Name":"retail"}}'
```

Example

下列範例會授與DROP資料庫中資料表datalake_user1上inventory的使用者retail。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

Example

下列範例會授與DROP資料庫中資源連結datalake_user1上inventory-link的使用者retail。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "DROP" --resource '{ "Table": {"DatabaseName":"retail", "Name":"inventory-
link"}}'
```

INSERT

權限	授與此資源	受贈人也需要
INSERT	TABLE	如果位置已註冊，則不需要額外的 IAM 許可。）

具有此權限的主體可以在表格指定的 Amazon S3 位置插入、更新和讀取基礎資料。主體也可以在 Lake Formation 主控台中檢視表格，並使用 AWS Glue API 擷取資料表的相關資訊。

Example

下列範例會將INSERT權限授與 AWS 帳戶 1111-2222-3333 資料庫inventoryretail中資料表datalake_user1上的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "INSERT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

此權限僅適用於 Amazon S3 中的資料，而不適用於其他資料存放區 (例如 Amazon RDS) 中的資料。

SELECT

權限	授與此資源	受贈人也需要
SELECT	• TABLE	如果位置已註冊，則不需要額外的 IAM 許可。)

具有此權限的主體可以檢視資料目錄中的資料表，並且可以在資料表指定的位置查詢 Amazon S3 中的基礎資料。主體可以在 Lake Formation 主控台中檢視表格，並使用 AWS Glue API 擷取表格的相關資訊。如果在授與此權限時套用資料行篩選，主體只能檢視內含資料行的中繼資料，而且只能從內含的資料行查詢資料。

> **Note**
>
> 整合式分析服務有責任在處理查詢時套用欄篩選。

Example

下列範例會將SELECT權限授與 AWS 帳戶 1111-2222-3333 資料庫inventoryretail中資料表datalake_user1上的使用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "SELECT" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

此權限僅適用於 Amazon S3 中的資料，而不適用於其他資料存放區 (例如 Amazon RDS) 中的資料。

您可以使用選擇性包含清單或排除清單來篩選 (限制存取) 特定欄。包含列表指定可以訪問的列。排除清單會指定無法存取的資料欄。如果沒有包含或排除清單，則可存取所有表格欄。

結果只會glue:GetTable傳回呼叫者有權檢視的資料行。亞馬遜雅典娜和亞馬 Amazon Redshift 等集成服務榮譽列包含和排除列表。

Example

下列範例會使SELECT用包含清單授與資料表datalake_user1上的inventory使用者。

```
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

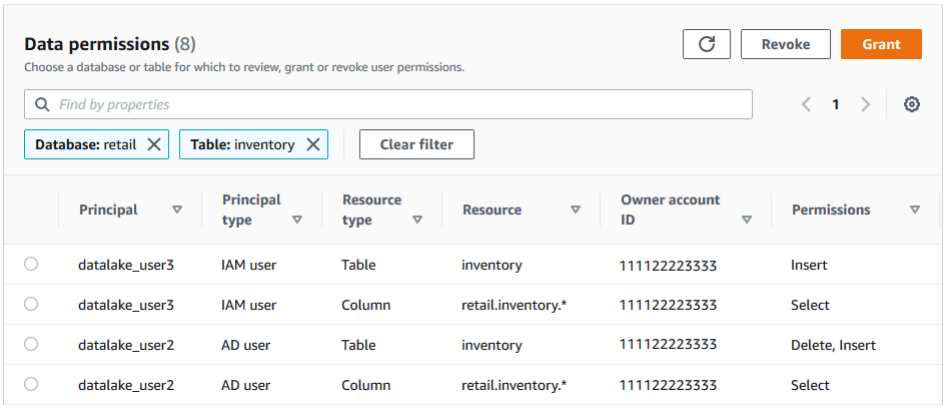
下一個範例會使用排除清單授與資SELECTinventory料表。

```
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"]}}}'
```

下列限制適用於SELECT權限：

- 授與時SELECT，如果套用欄篩選，則無法包含授與選項。
- 您無法限制屬於分割區索引鍵之資料行的存取控制。
- 具有資料表中資料行子集SELECT權限的主參與者無法授與該資料表的ALTERDROPDELETE、、或INSERT權限。同樣地，對資料表具有ALTERDROPDELETE、或INSERT權限的主參與者，也無法授與資料行篩選的SELECT權限。

權限一律會以個別列的形式顯示在 Lake Formation 主控台的 **SELECT** [資料] 權限頁面上。下列影像顯示SELECT授datalake_user2與使用者和datalake_user3inventory表格中所有資料欄的資料行。



Super

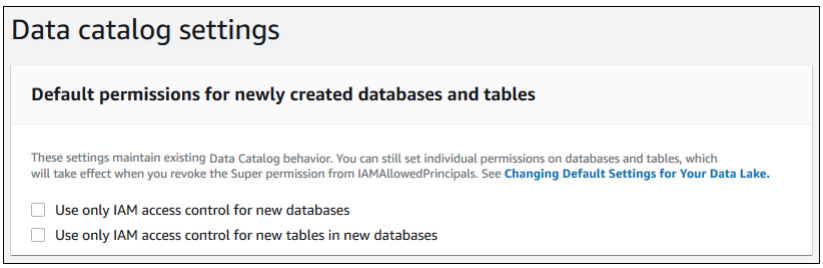
權限	對此資源授與	受贈人也需要
Super	DATABASE	glue:*Database*
Super	TABLE	glue:*Table*, glue:*Partition*

此權限可讓主體在資料庫或資料表上執行所有支援的 Lake Formation 作業。您無法將資料庫授與Super外部帳戶。

此權限可與其他 Lake Formation 權限共存。例如，您可以授與SuperSELECT、和中繼資料表的INSERT權限。主參與者接著就可以在表格上執行所有支援的作業。撤銷時Super，SELECT和INSERT權限會保留，主參與者只能執行選取和插入作業。

您可以Super將其授與群組，而不是授與個別主參與者IAMAllowedPrincipals。系統會自動建立IAMAllowedPrincipals群組，其中包含 IAM 政策允許存取資料目錄資源的所有 IAM 使用者和角色。當Super授與資料目錄資源時，IAMAllowedPrincipals對資源的存取權限僅由 IAM 政策進行有效控制。

您可以利用 Lake Formation 主控台IAMAllowedPrincipals的 「設定」 頁面上的選項，來擁有自動授與新目錄資源的Super權限。



- 若Super要IAMAllowedPrincipals授予所有新資料庫，請選取僅對新資料庫使用 IAM 存取控制。
- 若Super要授予新IAMAllowedPrincipals資料庫中所有新資料表的權限，請針對新資料庫中的新資料表選取僅使用 IAM 存取控制。

> **Note**
>
> 此選項會預設選取 [建立資料庫] 對話方塊中的 [僅對此資料庫中的新資料表使用 IAM 存取控制] 核取方塊。它沒有什麼比這更多。它是 [建立資料庫] 對話方塊中的核取方塊，可啟用授與Super給IAMAllowedPrincipals。

這些 「設定」 頁面選項預設為啟用狀態。如需詳細資訊，請參閱下列內容：

- the section called “變更資料湖的預設設定”
- the section called “將AWS Glue資料權限升級至 Lake Formation 型模型”

ASSOCIATE

權限	授與此資源	受贈人也需要
ASSOCIATE	LF-Tag	glue:GetDatabase glue:GetTable lakeformation:AddLFTagsToResource" lakeformation:RemoveLFTagsFromResource"

權限	授與此資源	受贈人也需要
		lakeformation:GetResourceLFTags lakeformation:ListLFTags lakeformation:GetLFTag lakeformation:SearchTablesByLFTags lakeformation:SearchDatabasesByLFTags

對 LF 標籤具有此權限的主參與者可以將 LF 標籤指派給資料目錄資源。授予ASSOCIATE隱含授予。DESCRIBE

Example

這個例子授予用戶datalake_user1對 LF 標籤的ASSOCIATE權限與密鑰。module它授予檢視和指派該機碼所有值的權限，如星號 (*) 所示。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

整合 IAM 身分識別中心

您可以透過連線至身分識別提供者 (IdPs) AWS IAM Identity Center，並跨 AWS 分析服務集中管理使用者和群組的存取。您可以將身分識別提供者 (例如 Okta、Ping 和 Microsoft Entra ID (舊稱為 Azure Active Directory) 與 IAM 身分識別中心整合，讓組織中的使用者使用單一登入體驗來存取資料。IAM 身分中心也支援連接其他第三方身分識別提供者。

如需詳細資訊，請參閱 AWS IAM Identity Center 使用指南中的支援身分識別提供者。

您可以在 IAM Identity Center 中設定 AWS Lake Formation 為已啟用的應用程式，資料湖管理員可以將精細的許可授與授權的使用者和 AWS Glue Data Catalog 資源群組。

您組織的使用者可以使用您組織的身分提供者登入任何已啟用 Identity Center 的應用程式，並查詢套用 Lake Formation 權限的資料集。透過此整合，您可以管理 AWS 服務的存取，而無需建立多個 IAM 角色。

> **Note**
>
> 受信任的身分傳播可讓使用者現有的使用者和群組成員資格存取所有 AWS 分析服務中的資料。透過受信任的身分傳播，使用者可以登入應用程式，而應用程式可以在要求中傳遞使用者的身分識別，以存取 AWS 服務中的資料。您不需要執行任何服務特定的身分識別提供者設定或 IAM 角色設定。如需詳細資訊，請參閱 AWS IAM Identity Center 使用指南中的跨應用程式的受信任身分傳播。

如需限制的詳細資訊，請參閱IAM 身分識別中心整合限制。

主題

必要條件

以下是整合 IAM 身分中心與 Lake Formation 的先決條件。

1. 啟用 IAM 身分中心 — 啟用 IAM 身分中心是支援身分驗證和身分傳播的先決條件。
2. 選擇您的身分來源 — 啟用 IAM 身分中心後，您必須擁有識別提供者來管理使用者和群組。您可以使用內建的身分識別中心目錄做為身分識別來源，也可以使用外部 IdP，例如 Microsoft Entra ID 或 Okta。

   如需詳細資訊，請參閱 AWS IAM Identity Center 使用指南中的管理您的身分識別來源和 Connect 至外部身分識別提供者。

3. 建立 IAM 角色 — 建立 IAM 身分中心連線的角色需要在 Lake Formation 和 IAM 身分中心中建立和修改應用程式組態的許可，如下列內嵌政策所示。

   您需要根據 IAM 最佳做法新增許可。下列程序會詳細說明特定權限。如需詳細資訊，請參閱開始使用 IAM 身分中心。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:CreateLakeFormationIdentityCenterConfiguration",
                "sso:CreateApplication",
                "sso:PutApplicationAssignmentConfiguration",
                "sso:PutApplicationAuthenticationMethod",
                "sso:PutApplicationGrant",
                "sso:PutApplicationAccessScope",
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

   如果您要與外部 AWS 帳戶 或組織共用資料目錄資源，則必須具有 AWS Resource Access Manager (AWS RAM) 權限才能建立資源共用。如需共用資源所需權限的詳細資訊，請參閱跨帳戶資料共用先決條件。

下列內嵌政策包含檢視、更新和刪除 Lake Formation 與 IAM 身分中心整合屬性所需的特定權限。

- 使用下列內嵌政策允許 IAM 角色檢視與 IAM 身分中心的 Lake Formation 整合。

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

```
                "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication"
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- 使用下列內嵌政策允許 IAM 角色更新與 IAM 身分中心的 Lake Formation 整合。此策略也包含與外部帳號共用資源所需的選用權限。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:UpdateLakeFormationIdentityCenterConfiguration",
                "lakeformation:DescribeLakeFormationIdentityCenterConfiguration",
                "sso:DescribeApplication",
                "sso:UpdateApplication",
            ],
            "Resource": [
                "*"
            ]
        }
    ]
}
```

- 使用下列內嵌政策允許 IAM 角色刪除與 IAM 身分中心的 Lake Formation 整合。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:DeleteLakeFormationIdentityCenterConfiguration",
                "sso:DeleteApplication",
            ],
```

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

- 如需授與或撤銷 IAM 身分中心使用者和群組資料湖許可所需的 IAM 許可，請參閱授予或撤銷 Lake Formation 許可所需的 IAM 許可。

權限說明

- `lakeformation:CreateLakeFormationIdentityCenterConfiguration`— 建立 Lake Formation IdC 配置。
- `lakeformation:DescribeLakeFormationIdentityCenterConfiguration`— 描述現有的 IdC 組態。
- `lakeformation:DeleteLakeFormationIdentityCenterConfiguration`— 能夠刪除現有的 Lake Formation IdC 配置。
- `lakeformation:UpdateLakeFormationIdentityCenterConfiguration`-用於更改現有的 Lake Formation 配置。
- `sso:CreateApplication`— 用於建立 IAM Identity Center 應用程式。
- `sso:DeleteApplication`— 用於刪除 IAM Identity Center 應用程式。
- `sso:UpdateApplication`— 用於更新 IAM Identity Center 應用程式。
- `sso:PutApplicationGrant`— 用於變更受信任的字符發行者資訊。
- `sso:PutApplicationAuthenticationMethod`— 授予 Lake Formation 身份驗證訪問權限
- `sso:GetApplicationGrant`— 用於列出受信任的字符發行者資訊。
- `sso:DeleteApplicationGrant` – 刪除信任權杖發行者資訊。
- `sso:PutApplicationAccessScope`— 新增或更新應用程式 IAM 身分中心存取範圍的授權目標清單。
- `sso:PutApplicationAssignmentConfiguration`— 用於設定使用者如何取得應用程式的存取權。

將 Lake Formation 與 IAM 身份中心連接

您必須先完 Lake Formation 下列步驟，才能使用 IAM 身分中心管理身分以授與資料目錄資源的存取權限。您可以使用 Lake Formation 主控台或建立 IAM 身分中心整合 AWS CLI。

AWS Management Console

連接 Lake Formation 與 IAM 身份中心

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在左側導覽窗格中，選取 IAM 身分中心整合。

Create IAM Identity Center Integration

Enable IAM Identify Center and then create Lake Formation - IAM Identity Center integration to manage identities from IAM Identity Center (external IdPs like Azure AD or Okta Universal Directory). Learn more

▼ **How it works**

Enable IAM Identity Center
Enable IAM Identity Center for your account or organization and select an identity provider.

Create Lake Formation integration
Integrate Lake Formation with IAM Identity Center to permit Lake Formation to access users from your selected identity provider.

Grant permissions
Grant permissions to users on Data Catalog databases and tables using fine-grained Lake Formation permissions.

Connect Lake Formation to IAM Identity Center

Connect to organization instance of IAM Identity Center

Manage access to Lake Formation by assigning users and groups from the Identity Center directory for your organization. Learn more

Recommended

Connect to account instance of IAM Identity Center

Manage access to Lake Formation by assigning existing or creating dedicated users and groups from your Identity Center directory. Learn more

instance of IAM Identity Center
Manage access to Lake Formation by assigning users and groups from your Identity Center directory.
arn:aws:sso:::instance/ssoins-6987513bf5410c2f

Add AWS account and organization IDs

Add AWS accounts and organizations whose users need access to Lake Formation managed resources.

AWS Accounts and AWS organizations
Enter one or more AWS account IDs and AWS organization IDs. Press Enter after each ID.

Choose AWS account, AWS organization ID

▶ **Lake Formation application integration - *optional***
Add applications that can access S3 data locations registered with Lake Formation on behalf of the user.

After this step, you can't edit the connection. You can edit AWS accounts, organizations, and applications. If you want to modify the connection, delete it and create a new connection.

3. (選擇性) 輸入一或多個有效 AWS 帳戶 ID、組織 ID 及/或組織單位 ID，以允許外部帳戶存取資料目錄資源。當 IAM 身分中心使用者或群組嘗試存取 Lake Formation 受管資料目錄資源時，Lake Formation 會擔任 IAM 角色來授權中繼資料存取權。如果 IAM 角色屬於沒有 AWS Glue 資源政策和資 AWS RAM 源共用的外部帳戶，則 IAM 身分中心使用者和群組即使擁有 Lake Formation 許可，也無法存取該資源。

   Lake Formation 使用 AWS Resource Access Manager （ AWS RAM ）服務與外部帳戶和組織共享資源。 AWS RAM 傳送邀請給受權者帳號以接受或拒絕資源共用。

   如需詳細資訊，請參閱 接受來自的資源共用邀請 AWS RAM。

   > **Note**
   >
   > Lake Formation 允許來自外部帳戶的 IAM 角色，代表 IAM Identity Center 使用者和群組用於存取資料目錄資源的承運人角色，但只能授與擁有帳戶內的資料目錄資源的許可。如果您嘗試將權限授與外部帳戶中資料目錄資源的 IAM Identity Center 使用者和群組，Lake Formation 會擲回下列錯誤-「主體不支援跨帳戶授與」。

4. (選擇性) 在「建立湖泊形成」整合畫面上，指定第三方應用程式的 ARN，這些應用程式可以存取向 Lake Formation 註冊的 Amazon S3 位置中的資料。Lake Formation 會根據有效許可，以 AWS STS 令牌形式將臨時登入資料劃分到已註冊的 Amazon S3 位置，以便授權的應用程式可以代表使用者存取資料。
5. 選取提交。

   Lake Formation 管理員完成步驟並建立整合後，IAM 身分中心屬性會顯示在 Lake Formation 主控台中。完成這些任務使 Lake Formation 成為啟用 IAM 身份中心的應用程序。主控台內的屬性包括整合狀態。整合狀態Success會顯示完成時間。此狀態指出 IAM 身分中心組態是否已完成。

AWS CLI

- 下面的示例演示了如何創建 Lake Formation 與 IAM 身份中心集成。您也可以指定應用程式的 `Status` (ENABLED,DISABLED)。

```
aws lakeformation create-lake-formation-identity-center-configuration \
    --catalog-id <123456789012> \
    --instance-arn <arn:aws:sso:::instance/ssoins-112111f12ca1122p> \
    --share-recipients '[{"DataLakePrincipalIdentifier": "<123456789012>"},
                         {"DataLakePrincipalIdentifier": "<555555555555>"}]' \
```

```
    --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"],
 "Status": "ENABLED"}'
```

- 下列範例說明如何檢視與 IAM 身分中心的 Lake Formation 整合。

```
aws lakeformation describe-lake-formation-identity-center-configuration
 --catalog-id <123456789012>
```

更新 IAM 身分中心整合

建立連線後，您可以為 IAM 身分中心整合新增第三方應用程式，以便與 Lake Formation 整合，並代表使用者存取 Amazon S3 資料。您也可以從 IAM 身分中心整合移除現有的應用程式。您可以使用 Lake Formation 控制台添加或刪除應用程序 AWS CLI，並使用UpdateLakeFormationIdentityCenterConfiguration操作。

> **Note**
>
> 建立 IAM 身分中心整合後，您無法更新執行個體ARN。

AWS Management Console

更新與 Lake Formation 的現有 IAM 身份中心連接

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在左側導覽窗格中，選取 IAM 身分中心整合。
3. 在 IAM 身分中心整合頁面上選取新增。
4. 輸入一或多個有效 AWS 帳戶 ID、組織 ID 和/或組織單位 ID，以允許外部帳戶存取資料目錄資源。
5. 在「新增應用程式」畫面上，輸入您要與 Lake Formation 整合的第三方應用程式的應用程式 ID。
6. 選取新增。

AWS CLI

您可以執行下列 AWS CLI 命令，為 IAM 身分中心整合新增或移除第三方應用程式。將外部篩選狀態設定為時ENABLED，可讓 IAM 身分中心為第三方應用程式提供身分識別管理，以存取由 Lake Formation 管理的資料。您也可以透過設定應用程式狀態來啟用或停用 IAM 身分中心整合。

```
aws lakeformation update-lake-formation-identity-center-configuration \
 --external-filtering '{"AuthorizedTargets": ["<app arn1>", "<app arn2>"], "Status":
 "ENABLED"}'\
 --share-recipients '[{"DataLakePrincipalIdentifier": "<444455556666>"}
                     {"DataLakePrincipalIdentifier": "<777788889999>"}]' \
 --application-status ENABLED
```

刪除與 IAM 身分中心的 Lake Formation 連接

如果您想要刪除現有的 IAM 身分中心整合，可以使用 Lake Formation 主控台或DeleteLakeFormationIdentityCenterConfiguration作業來執行此作業。 AWS CLI

AWS Management Console

若要刪除與 Lake Formation 的現有 IAM 身分中心連線

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在左側導覽窗格中，選取 IAM 身分中心整合。
3. 在 IAM 身分中心整合頁面上選取刪除。
4. 在「確認整合」畫面上，確認動作，然後選取「刪除」。

AWS CLI

您可以執行下列 AWS CLI 命令來刪除 IAM 身分中心整合。

```
aws lakeformation delete-lake-formation-identity-center-configuration \
    --catalog-id <123456789012>
```

授與權限給使用者和群組

您的資料湖管理員可以將資料目錄資源 (資料庫、表格和檢視) 上的權限授與 IAM Identity Center 使用者和群組，以便輕鬆存取資料。若要授與或撤銷資料湖許可，授與者需要下列 IAM 身分中心動作的許可。

- DescribeUser
- DescribeGroup
- DescribeInstance

您可以使用 Lake Formation 控制台、API 或 AWS CLI.

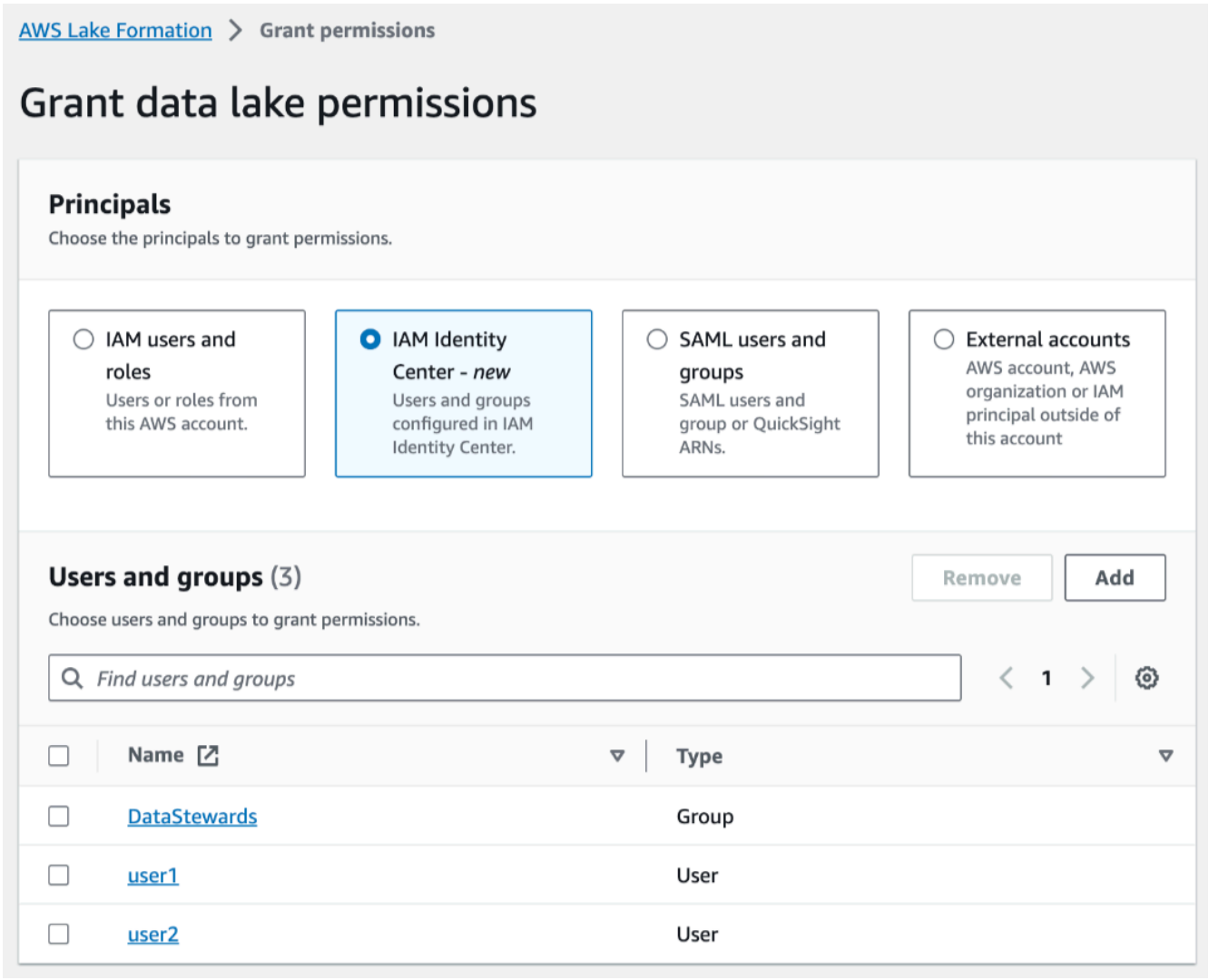
如需授與權限的詳細資訊，請參閱the section called “授與和撤銷資料目錄權限”。

> **Note**
>
> 您只能授與帳戶中資源的權限。若要將權限重疊顯示給與您共用之資源的使用者和群組，您必須使用 AWS RAM 資源共用。

AWS Management Console

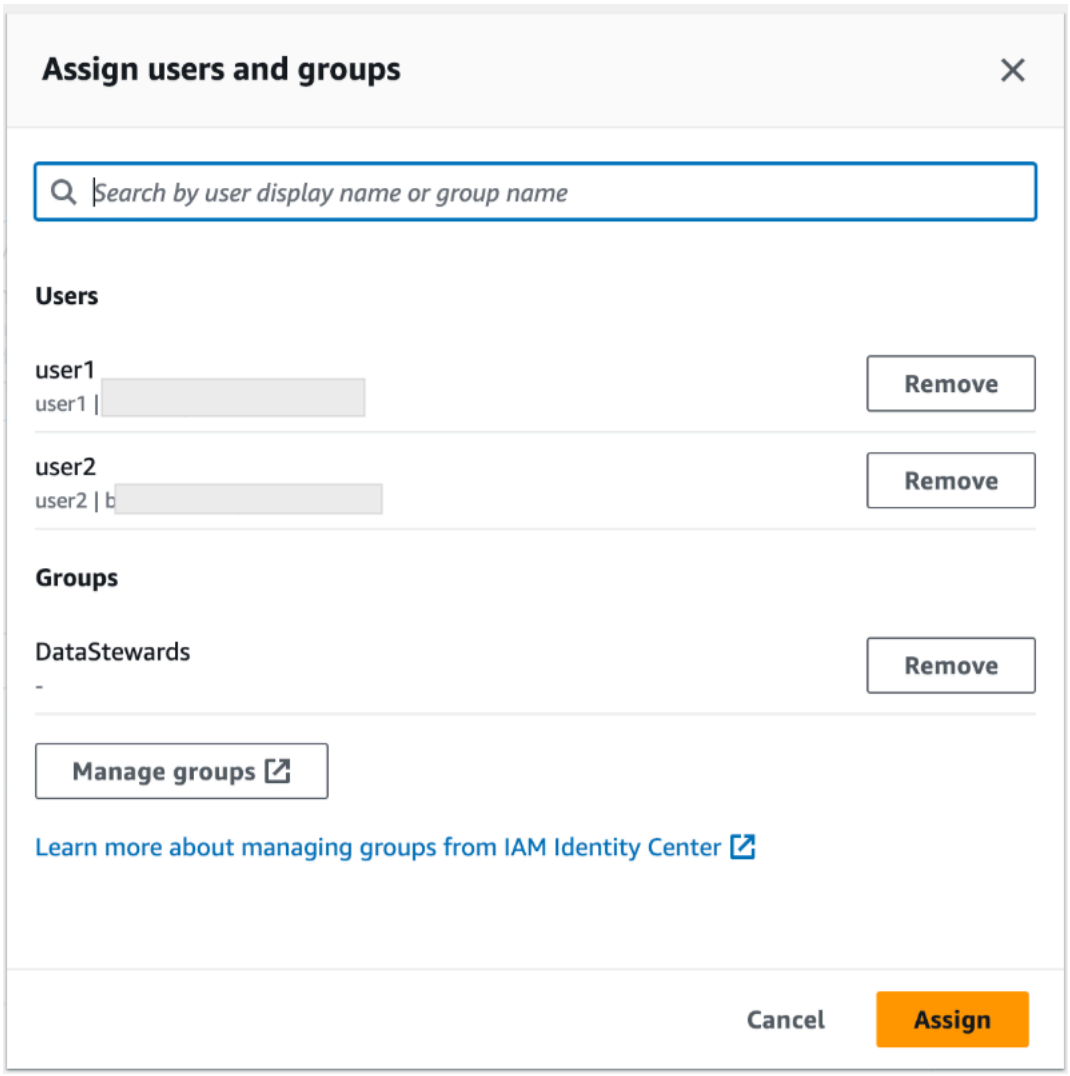
授與使用者和群組的權限

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 選取湖泊形成主控台中「權限」下的「資料湖權限」。
3. 選取授權。
4. 在 [授與資料湖權限] 頁面上，選擇 [SSM 使用者和群組]。
5. 選取 [新增] 以選擇要授與權限的使用者和群組。



6. 在「指派使用者和群組」畫面上，選擇要授與權限的使用者和/或群組。

   選取「指派」。



7. 接下來，選擇授予權限的方法。

   如需使用具名資源方法授與權限的指示，請參閱使用具名資源方法授與資料湖權限。

   如需使用 LF 標籤授予權限的指示，請參閱。使用 LF-TBAC 方法授與資料湖權限

8. 選擇您要授與權限的資料目錄資源。
9. 選擇要授與的資料目錄權限。
10. 選取授權。

AWS CLI

下列範例顯示如何授與資料表的 IAM 身分中心使用者SELECT權限。

```
aws lakeformation grant-permissions \
```

```
--principal DataLakePrincipalIdentifier=arn:aws:identitystore:::user/<UserId> \
--permissions "SELECT" \
--resource '{ "Table": { "DatabaseName": "retail", "TableWildcard": {} } }'
```

若要UserId從 IAM 身分中心擷取，請參閱 IAM 身分中心 API 參考中的GetUserId操作。

將 Amazon S3 位置新增至您的資料湖

若要將 Amazon Simple Storage Service (Amazon S3) 位置新增為資料湖中的儲存，請在中註冊該位置 AWS Lake Formation。然後，您可以使用 Lake Formation 權限，對指向此位置的 AWS Glue Data Catalog 物件以及位置中的基礎資料進行精細的存取控制。

Lake Formation 還允許以混合式存取模式註冊資料位置，並提供您選擇性地啟用資料目錄中資料庫和表格的 Lake Formation 權限的彈性。使用混合式存取模式，您將擁有一個增量路徑，可讓您為一組特定使用者設定 Lake Formation 權限，而不會中斷其他現有使用者或工作負載的權限原則。

如需設定混合存取模式的詳細資訊，請參閱 混合存取模式

當您註冊位置時，Amazon S3 路徑和該路徑下的所有資料夾都會註冊。

例如，假設您有如下所示的 Amazon S3 路徑組織：

`/mybucket/accounting/sales/`

如果您註冊S3://mybucket/accounting，該sales文件夾也被註冊並在 Lake Formation 管理下。

如需註冊位置的詳細資訊，請參閱Underlying data access control。

> **Note**
>
> 建議使用 Lake Formation m 權限用於結構化資料 (以包含列和欄的表格排列)。如果您的資料包含物件型非結構化資料，請考慮使用 Amazon S3 的 IAM 許可來管理資料存取。

主題

- 用於註冊地點的角色需求
- 註冊 Amazon S3 位置
- 註冊加密的 Amazon S3 位置

- 在另一個 AWS 帳戶中註冊 Amazon S3 位置
- 跨 AWS 帳戶註冊加密的 Amazon S3 位置
- 註銷 Amazon S3 位置

用於註冊地點的角色需求

註冊亞馬遜簡單儲存服務 AWS Identity and Access Management (Amazon S3) 位置時，必須指定 (IAM) 角色。 AWS Lake Formation 存取該位置中的資料時，會假設該角色。

您可以使用下列其中一種角色類型來註冊位置：

- Lake Formation 服務相關的角色。此角色會授與位置的必要權限。使用此角色是註冊位置的最簡單方法。如需詳細資訊，請參閱 使用服務連結角色進行 Lake Formation。
- 使用者定義的角色。當您需要授與超過服務連結角色所提供的權限時，請使用使用者定義的角色。

  在下列情況下，您必須使用使用者定義的角色：

  - 在另一個帳戶中註冊位置時。

    如需詳細資訊，請參閱 the section called “在另一個 AWS 帳戶中註冊 Amazon S3 位置” 及 the section called “跨 AWS 帳戶註冊加密的 Amazon S3 位置”。
  - 如果您使用 AWS 受管 CMK (aws/s3) 來加密 Amazon S3 位置。

    如需詳細資訊，請參閱 註冊加密的 Amazon S3 位置。
  - 如果您打算使用 Amazon EMR 訪問該位置。

    如果您已使用服務連結角色註冊位置，並且想要開始透過 Amazon EMR 存取該位置，則必須取消註冊該位置，然後使用使用者定義的角色重新註冊該位置。如需詳細資訊，請參閱 the section called “註銷 Amazon S3 位置”。

使用服務連結角色進行 Lake Formation

AWS Lake Formation 使用 AWS Identity and Access Management (IAM) 服務連結角色。服務連結角色是一種獨特的 IAM 角色類型，直接連結至 Lake Formation。服務連結角色由 Lake Formation 預先定義，並包含服務代表您呼叫其他服 AWS 務所需的所有權限。

服務鏈接的角色使設置 Lake Formation 更容易，因為您不必創建角色並手動添加必要的權限。Lake Formation m 定義了其服務鏈接角色的權限，除非另有定義，否則只有 Lake Formation 可以擔任其角色。定義的許可包括信任政策和許可政策，並且該許可政策不能連接到任何其他 IAM 實體。

此服務連結角色會信任下列服務擔任該角色：

- lakeformation.amazonaws.com

當您在帳戶 A 中使用服務連結角色註冊帳戶 B 擁有的 Amazon S3 位置時，帳戶 B 中的 Amazon S3 儲存貯體政策 (以資源為基礎的政策) 必須授予帳戶 A 中服務連結角色的存取權限。

> **Note**
>
> 服務控制原則 (SCP) 不會影響服務連結角色。
> 如需詳細資訊，請參閱AWS Organizations 使用指南中的服務控制原則 (SCP)。

Lake Formation 的服務連結角色權限

Lake Formation 使用名為AWSServiceRoleForLakeFormationDataAccess的服務鏈接角色。此角色提供一組 Amazon Simple Storage Service (Amazon S3) 許可，可讓 Lake Formation 整合服務 (例如 Amazon Athena) 存取已註冊的位置。註冊資料湖位置時，必須提供在該位置具有所需 Amazon S3 讀取/寫入許可的角色。您可以使用此服務連結角色，而不是建立具有所需 Amazon S3 許可的角色。

第一次將服務連結角色命名為註冊路徑的角色時，會代表您建立服務連結角色和新的 IAM 政策。Lake Formation 將路徑添加到內聯策略，並將其附加到服務鏈接的角色。當您使用服務連結角色註冊後續路徑時，Lake Formation 會將路徑新增至現有策略。

以資料湖管理員身分登入時，請註冊資料湖位置。然後，在 IAM 主控台中搜尋角色AWSServiceRoleForLakeFormationDataAccess並檢視其附加政策。

例如，在您註冊位置之後s3://my-kinesis-test/logs，Lake Formation 會建立下列內嵌政策並將其附加至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"
            ]
        }
    ]
}
```

為 Lake Formation 成建立與服務相關的角色

您不需要手動建立一個服務連結角色。當您在 AWS Management Console、或 AWS API 中向湖泊形成註冊 Amazon S3 位置時 AWS CLI，Lake Formation 會為您建立服務連結角色。

> **Important**
>
> 此服務連結角色可以顯示在您的帳戶，如果您於其他服務中完成一項動作時，可以使用支援此角色的功能。若要進一步了解，請參閱我的 IAM 帳戶中出現的新角色。

若您刪除此服務連結角色，之後需要再次建立，您可以在帳戶中使用相同程序重新建立角色。當您向 Lake Formation 成註冊 Amazon S3 位置時，Lake Form 會再次為您建立服務連結角色。

您也可以使用 IAM 主控台建立具有 Lake Formation 使用案例的服務連結角色。在 AWS CLI 或 AWS API 中，使用`lakeformation.amazonaws.com`服務名稱建立服務連結角色。如需詳細資訊，請參閱 IAM 使用者指南中的建立服務連結角色。如果您刪除此服務連結角色，您可以使用此相同的程序以再次建立該角色。

編輯 Lake Formation 的服務連結角色

Lake Formation 不允許您編輯AWSServiceRoleForLakeFormationDataAccess服務鏈接的角色。因為有各種實體可能會參考服務連結角色，所以您無法在建立角色之後變更角色名稱。然而，您可使用 IAM 來編輯角色描述。如需更多資訊，請參閱 IAM 使用者指南中的編輯服務連結角色。

刪除 Lake Formation 的服務連結角色

若您不再使用需要服務連結角色的功能或服務，我們建議您刪除該角色。如此一來，您就沒有未主動監控或維護的未使用實體。然而，在手動刪除服務連結角色之前，您必須先清除資源。

> **Note**
>
> 如果當您嘗試刪除資源時，Lake Formation 服務正在使用此角色，則刪除可能會失敗。若此情況發生，請等待數分鐘後並再次嘗試操作。

刪除 Lake Formation 所使用的 Lake Formation 資源

- 如果您已使用服務連結角色向 Lake Formation 註冊 Amazon S3 位置，則在刪除服務連結角色之前，您需要取消註冊該位置，然後使用自訂角色重新註冊該位置。

使用 IAM 手動刪除服務連結角色

使用 IAM 主控台或 AWS API 刪除AWSServiceRoleForLakeFormationDataAccess服務連結角色。 AWS CLI如需詳細資訊，請參閱《IAM 使用者指南》中的刪除服務連結角色。

以下是使用者定義角色的需求：

- 建立新角色時，請在 IAM 主控台的 [建立角色] 頁面上選擇 [AWS 服務]，然後在 [選擇使用案例] 下選擇 [Lake Formation]。

  如果您使用不同的路徑建立角色，請確定角色與之間的信任關係`lakeformation.amazonaws.com`。如需詳細資訊，請參閱修改角色信任原則 (主控台)。
- 角色必須與下列實體具有信任關係：
  - `glue.amazonaws.com`
  - `lakeformation.amazonaws.com`

  如需詳細資訊，請參閱修改角色信任原則 (主控台)。

- 該角色必須具有內嵌政策，以授與該位置的 Amazon S3 讀取/寫入許可。以下是典型政策。

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

- 將下列信任政策新增至 IAM 角色，以允許 Lake Formation 服務擔任該角色，並將臨時信任歸納至整合式分析引擎。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "DataCatalogViewDefinerAssumeRole1",
            "Effect": "Allow",
            "Principal": {
               "Service": [
                    "glue.amazonaws.com",
                    "lakeformation.amazonaws.com"
                ]
            },
```

```
            "Action": [
                "sts:AssumeRole"
            ]
        }
    ]
}
```

- 註冊該位置的資料湖管理員必須具有角色的`iam:PassRole`權限。

  以下是授與此權限的內嵌政策。以<account-id>有效的 AWS 帳號取代，並取代<role-name>為角色的名稱。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "PassRolePermissions",
            "Effect": "Allow",
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::<account-id>:role/<role-name>"
            ]
        }
    ]
}
```

- 要允許 Lake Formation 在日誌中添加 CloudWatch 日誌並發布指標，請添加以下內聯策略。

> **Note**
>
> 寫入記 CloudWatch 錄會產生費用。

```
{
    "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:*"
            ]
        }
    ]
}
```

註冊 Amazon S3 位置

註冊亞馬遜簡單儲存服務 AWS Identity and Access Management (Amazon S3) 位置時，必須指定 (IAM) 角色。Lake Formation 會在將臨時登入資料授予存取該位置資料的整合式 AWS 服務時擔任該角色。

> **Important**
>
> 避免註冊已啟用請求者付費的 Amazon S3 儲存貯體。對於在 Lake Formation 註冊的值區，用於註冊值區的角色一律會被視為請求者。如果值區是由其他 AWS 帳戶存取，如果該角色與值區擁有者屬於相同的帳戶，則值區擁有者會收取資料存取費用。

您可以使用 AWS Lake Formation 主控台、Lake Formation API 或 AWS Command Line Interface (AWS CLI) 註冊 Amazon S3 位置。

開始之前

檢閱用來註冊位置之角色的需求。

若要註冊位置 (主控台)

> **Important**
>
> 下列程序假設 Amazon S3 位置與資料目錄位於相同的 AWS 帳戶中，且該位置中的資料未加密。本章中的其他章節涵蓋跨帳戶註冊和加密位置的註冊。

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入，或以具有 `lakeformation:RegisterResource` IAM 權限的使用者身分登入。
2. 在導覽窗格的 [系統管理] 下，選取 [資料湖位置]。
3. 選擇 [註冊位置]，然後選擇 [瀏覽] 以選取 Amazon Simple Storage Service (Amazon S3) 路徑。
4. (選用，但強烈建議使用) 選取檢閱位置許可以檢視所選 Amazon S3 位置中所有現有資源及其許可的清單。

   註冊選取的位置可能會導致您的 Lake Formation 使用者取得該位置已存取資料的存取權。檢視此清單可協助您確保現有資料保持安全。
5. 對於 IAM 角色，請選擇符合中the section called “用於註冊地點的角色需求”要求的`AWSServiceRoleForLakeFormationDataAccess`服務連結角色 (預設) 或自訂 IAM 角色。

   只有在使用自訂 IAM 角色註冊時，才能更新已註冊位置或其他詳細資料。若要編輯使用服務連結角色註冊的位置，您應該取消註冊該位置並重新註冊。
6. 選擇啟用資料目錄聯合選項，以允許 Lake Formation 擔任角色，並向整合式 AWS 服務提供暫時證明資料，以存取聯合資料庫下的表格。如果某個位置已向 Lake Formation 註冊，且您想要在聯合資料庫下的表格使用相同的位置，則需要使用「啟用資料目錄聯合」選項註冊相同的位置。
7. 選擇混合存取模式，預設不啟用 Lake Formation 權限。當您以混合式存取模式註冊 Amazon S3 位置時，可以透過選擇該位置下資料庫和表格的主體來啟用 Lake Formation 許可。

   如需設定混合存取模式的詳細資訊，請參閱混合存取模式。
8. 選取 [註冊地點]。

若要註冊位置 (AWS CLI)

1. 註冊一個新的地點與 Lake Formation

   此範例使用服務連結角色來註冊位置。您可以使用`--role-arn`參數來提供自己的角色。

取代<s3-path><s3-access-role>為有效的 Amazon S3 路徑、帳戶號碼為有效 AWS 帳戶，以及具有註冊資料位置許可的 IAM 角色。

> **Note**
>
> 如果已註冊位置是使用服務連結角色進行註冊，則無法編輯該位置的屬性。

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --use-service-linked-role
```

下列範例會使用自訂角色來註冊位置。

```
aws lakeformation register-resource \
 --resource-arn arn:aws:s3:::<s3-path> \
 --role-arn arn:aws:iam::<123456789012>:role/<s3-access-role>
```

2. 更新登錄於 Lake Formation 的地點

只有使用自訂 IAM 角色註冊的已註冊位置時，您才可以編輯該位置。對於以服務連結角色註冊的位置，您應該取消註冊該位置並重新註冊。如需詳細資訊，請參閱 the section called “註銷 Amazon S3 位置”。

```
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. 在混合存取模式下向同盟註冊資料位置

```
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
```

如需詳細資訊，請參閱 RegisterResourceAPI 作業。

> **Note**
>
> 註冊 Amazon S3 位置後，任何指向該位置 (或其任何子位置) 的 AWS Glue 表格都會傳回`GetTable`呼叫`true`中的`IsRegisteredWithLakeFormation`參數值。資料目錄 API 作業 (例如`GetTables`和`SearchTables`不會更新`IsRegisteredWithLakeFormation`參數的值)，並傳回預設值 (即 false) 存在已知限制。建議您使用 `GetTable` API 來檢視`IsRegisteredWithLakeFormation`參數的正確值。

註冊加密的 Amazon S3 位置

Lake Formation 與 AWS Key Management Service(AWS KMS) 整合，可讓您更輕鬆地設定其他整合服務，以加密和解密 Amazon Simple Storage Service (Amazon S3) 位置中的資料。

客戶管理 AWS KMS keys 和 AWS 受管金鑰 支持。目前，只有 Athena 才支援用戶端加密/解密。

您必須在註冊 Amazon S3 位置時指定 AWS Identity and Access Management (IAM) 角色。對於加密的 Amazon S3 位置，角色必須具有使用加密和解密資料的權限 AWS KMS key，或者 KMS 金鑰政策必須將金鑰的許可授與角色。

> **Important**
>
> 避免註冊已啟用請求者付費的 Amazon S3 儲存貯體。對於在 Lake Formation 註冊的值區，用於註冊值區的角色一律會被視為請求者。如果值區是由其他 AWS 帳戶存取，如果該角色與值區擁有者屬於相同的帳戶，則值區擁有者會收取資料存取費用。

註冊位置的最簡單方法是使用 Lake Formation 服務鏈接的角色。此角色會授與位置所需的讀取/寫入權限。您也可以使用自訂角色來註冊位置，前提是該位置符合中的要求the section called “用於註冊地點的角色需求”。

> **Important**
>
> 如果您使用 AWS 受管金鑰 加密 Amazon S3 位置，則無法使用 Lake Formation 服務連結角色。您必須使用自訂角色，並將金鑰的 IAM 許可新增至角色。本節稍後將提供詳細資訊。

下列程序說明如何註冊使用客戶受管金鑰或加密的 Amazon S3 位置 AWS 受管金鑰。

- 註冊使用客戶管理金鑰加密的位置
- 註冊一個加密的位置 AWS 受管金鑰

開始之前

檢閱用來註冊位置之角色的需求。

註冊使用客戶受管金鑰加密的 Amazon S3 位置

> **Note**
>
> 如果 KMS 金鑰或 Amazon S3 位置與資料目錄不在相同的 AWS 帳戶中，請the section called “跨 AWS 帳戶註冊加密的 Amazon S3 位置”改為遵循中的指示。

1. 在 https://console.aws.amazon.com/kms 開啟 AWS KMS 主控台，然後以 AWS Identity and Access Management (IAM) 管理使用者或可修改用於加密位置之 KMS 金鑰金鑰金鑰政策的使用者身分登入。
2. 在瀏覽窗格中，選擇 [客戶受管金鑰]，然後選擇所需 KMS 金鑰的名稱。

3. 在 KMS 金鑰詳細資料頁面上，選擇金鑰原則索引標籤，然後執行下列其中一項動作，以 KMS 金鑰使用者身分新增您的自訂角色或 Lake Formation 服務連結角色：

   - 如果顯示預設檢視 (包含 [金鑰管理員]、[金鑰刪除]、[金鑰使用者] 和 [其他 AWS 帳戶] 區段) — 在 [關鍵使用者] 區段下，新增您的自訂角色或 Lake Formation 服務連結角色`AWSServiceRoleForLakeFormationDataAccess`。
   - 如果顯示金鑰原則 (JSON) — 編輯原則，將您的自訂角色或 Lake Formation 服務連結角色新增`AWSServiceRoleForLakeFormationDataAccess`至「允許使用金鑰」物件，如下列範例所示。

   > **Note**
   >
   > 如果該物件遺失，請使用範例中顯示的權限來新增該物件。此範例使用服務連結角色。

```
        ...
        {
            "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. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入，或以具有 `lakeformation:RegisterResource` IAM 權限的使用者身分登入。
5. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。
6. 選擇 [註冊位置]，然後選擇 [瀏覽] 以選取 Amazon Simple Storage Service (Amazon S3) 路徑。
7. (選用，但強烈建議使用) 選擇檢閱位置許可以檢視所選 Amazon S3 位置中所有現有資源及其許可的清單。

   註冊選取的位置可能會導致您的 Lake Formation 使用者取得該位置已存取資料的存取權。檢視此清單可協助您確保現有資料保持安全。
8. 對於 IAM 角色，請選擇`AWSServiceRoleForLakeFormationDataAccess`服務連結角色 (預設) 或符合the section called “用於註冊地點的角色需求”.
9. 選擇註冊地點。

如需服務連結角色的詳細資訊，請參閱Lake Formation 的服務連結角色權限。

若要註冊使用加密的 Amazon S3 位置 AWS 受管金鑰

> **Important**
>
> 如果 Amazon S3 位置與資料目錄不在相同的 AWS 帳戶中，請the section called “跨 AWS 帳戶註冊加密的 Amazon S3 位置”改為遵循中的指示。

1. 建立 IAM 角色以用來註冊位置。請確定其符合中列出的需求the section called “用於註冊地點的角色需求”。
2. 將下列內嵌原則新增至角色。它會授與角色金鑰的權限。該`Resource`規格必須指定的 Amazon 資源名稱 (ARN). AWS 受管金鑰您可以從 AWS KMS 控制台獲取 ARN。若要取得正確的 ARN，請確定您使用與用來加密位置的 AWS 帳戶和 [區域] 相同的 AWS 受管金鑰 帳戶和 [區域] 登入 AWS KMS 主控台。

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

```
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:GenerateDataKey*",
        "kms:DescribeKey"
      ],
      "Resource": "<AWS #### ARN>"
    }
  ]
}
```

3. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入，或以具有 `lakeformation:RegisterResource` IAM 權限的使用者身分登入。
4. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。
5. 選擇 [註冊位置]，然後選擇 [瀏覽] 以選取 Amazon S3 路徑。
6. (選用，但強烈建議使用) 選擇檢閱位置許可以檢視所選 Amazon S3 位置中所有現有資源及其許可的清單。

   註冊選取的位置可能會導致您的 Lake Formation 使用者取得該位置已存取資料的存取權。檢視此清單可協助您確保現有資料保持安全。
7. 對於 IAM 角色，請選擇您在步驟 1 中建立的角色。
8. 選擇註冊地點。

在另一個 AWS 帳戶中註冊 Amazon S3 位置

AWS Lake Formation 讓您能夠跨 AWS 帳戶註冊亞馬遜簡單儲存服務 (Amazon S3) 位置。例如，如果帳戶 A 中，帳戶 A 中的使用者可以在帳戶 B 中註冊 Amazon S3 儲存貯體。 AWS Glue Data Catalog

使用帳戶 A 中的 AWS Identity and Access Management (IAM) 角色在 AWS 帳戶 B 中 AWS 註冊 Amazon S3 儲存貯體需要下列許可：

- 帳戶 A 中的角色必須授與帳戶 B 中值區的權限。
- 帳戶 B 中的值區政策必須授與帳戶 A 中角色的存取權限。

> **Important**
>
> 避免註冊已啟用請求者付費的 Amazon S3 儲存貯體。對於在 Lake Formation 註冊的值區，用於註冊值區的角色一律會被視為請求者。如果值區是由其他 AWS 帳戶存取，如果該角色與值區擁有者屬於相同的帳戶，則值區擁有者會收取資料存取費用。
> 您無法使用 Lake Formation 服務連結角色在其他帳戶中註冊位置。您必須改用使用者定義的角色。角色必須符合中的需求the section called “用於註冊地點的角色需求”。如需服務連結角色的詳細資訊，請參閱Lake Formation 的服務連結角色權限。

開始之前

檢閱用來註冊位置之角色的需求。

在其他 AWS 帳戶中註冊地點

> **Note**
>
> 如果位置已加密，請改為遵循中的指the section called “跨 AWS 帳戶註冊加密的 Amazon S3 位置”示。

下列程序假設帳戶 1111-2222-3333 中包含資料目錄的主體想要註冊帳戶 1234-5678-9012 的 Amazon S3 儲存貯`awsexamplebucket1`體。

1. 在帳戶 1111-2222-3333 中，請登入 AWS Management Console 並開啟身分與存取權管理主控台，網址為。https://console.aws.amazon.com/iam/
2. 建立新角色或檢視符合中需求的現有角色the section called “用於註冊地點的角色需求”。確保該角色在授予 Amazon S3 許可`awsexamplebucket1`。
3. 前往 https://console.aws.amazon.com/s3/ 開啟的 Amazon Simple Storage Service (Amazon S3) 主控台。請使用帳號登入。
4. 在「值區名稱」清單中，選擇值區名稱、`awsexamplebucket1`。
5. 選擇許可。
6. 在 [權限] 頁面上，選擇 [值區政策]。
7. 在值區政策編輯器中，貼上下列原則。<role-name>以您的角色名稱取代。

```
{
```

```
    "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. 選擇儲存。
9. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分或具有足夠權限的使用者身分登入帳戶 1111-2222-3333。
10. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。
11. 在 [資料湖位置] 頁面上，選擇 [註冊位置]。
12. 在 [註冊位置] 頁面上，對於 Amazon S3 路徑，輸入儲存貯體名稱`s3://awsexamplebucket1`。

> **Note**
>
> 您必須輸入值區名稱，因為當您選擇「瀏覽」時，跨帳戶時段不會出現在清單中。

13. 對於 IAM 角色，請選擇您的角色。
14. 選擇註冊地點。

跨 AWS 帳戶註冊加密的 Amazon S3 位置

AWS Lake Formation 與 AWS Key Management Service(AWS KMS) 整合，讓您能夠更輕鬆地設定其他整合服務，以加密和解密 Amazon Simple Storage Service (Amazon S3) 位置中的資料。

客戶管理的金鑰和 AWS 受管金鑰 均受支援。不支援用戶端加密/解密。

> **Important**
>
> 避免註冊已啟用請求者付費的 Amazon S3 儲存貯體。對於在 Lake Formation 註冊的值區，用於註冊值區的角色一律會被視為請求者。如果值區是由其他 AWS 帳戶存取，如果該角色與值區擁有者屬於相同的帳戶，則值區擁有者會收取資料存取費用。

本節說明如何在下列情況下註冊 Amazon S3 位置：

- Amazon S3 位置中的資料會使用中建立的 KMS 金鑰加密 AWS KMS。
- Amazon S3 位置與 AWS AWS Glue Data Catalog.
- KMS 金鑰是否與資料目錄位於相同的 AWS 帳戶中。

使用帳戶 A 中的 AWS Identity and Access Management (IAM) 角色在 AWS 帳戶 B 中註冊 AWS KMS — 加密的 Amazon S3 儲存貯體需要下列許可： AWS

- 帳戶 A 中的角色必須授與帳戶 B 中值區的權限。
- 帳戶 B 中的值區政策必須授與帳戶 A 中角色的存取權限。
- 如果 KMS 金鑰位於帳戶 B 中，則金鑰原則必須授與帳戶 A 中角色的存取權，而帳戶 A 中的角色必須授與 KMS 金鑰的權限。

在下列程序中，您會在包含資料目錄的 AWS 帳戶中建立角色 (先前討論中的帳戶 A)。然後，您可以使用此角色來註冊位置。在存取 Amazon S3 中的基礎資料時，Lake Formation 會擔任此角色。假定的角色具有 KMS 金鑰的必要權限。因此，您不需要將 KMS 金鑰的權限授與主體透過 ETL 工作存取基礎資料，或使用整合式服務 (例如. Amazon Athena

> **Important**
>
> 您無法使用 Lake Formation 服務連結角色在其他帳戶中註冊位置。您必須改用使用者定義的角色。角色必須符合中的需求the section called “用於註冊地點的角色需求”。如需服務連結角色的詳細資訊，請參閱Lake Formation 的服務連結角色權限。

開始之前

檢閱用來註冊位置之角色的需求。

若要跨 AWS 帳戶註冊加密的 Amazon S3 位置

1. 在與資料目錄相同的 AWS 帳戶中，登入 AWS Management Console 並在開啟 IAM 主控台https://console.aws.amazon.com/iam/。
2. 建立新角色或檢視符合中需求的現有角色the section called “用於註冊地點的角色需求”。確保角色包含授予該位置 Amazon S3 許可的政策。
3. 如果 KMS 金鑰與資料目錄不在相同的帳戶中，請將內嵌政策新增至角色，以授與 KMS 金鑰的必要權限。政策範例如下。以 <cmk-region>KMS 金鑰的區域和帳戶號碼取代和 *< cmk-account-id >*。<key-id>以金鑰識別碼取代。

```
{
    "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. 在 Amazon S3 主控台上，新增儲存貯體政策，將所需的 Amazon S3 許可授予該角色。以下為儲存貯體政策的範例。將 *< catalog-account-id >* 取代為資料目錄的 AWS 帳戶號碼<role-name>、角色名稱<bucket-name>以及值區的名稱。

```
{
    "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. 在中 AWS KMS，以 KMS 金鑰使用者身分新增角色。

   a. 請在以下位置開啟 AWS KMS 主控台。 https://console.aws.amazon.com/kms 然後，以系統管理員使用者或使用者身分登入，該使用者可以修改用於加密位置之 KMS 金鑰的金鑰原則。

   b. 在功能窗格中，選擇 [客戶受管金鑰]，然後選擇 KMS 金鑰的名稱。

   c. 在 KMS 金鑰詳細資料頁面的 [金鑰原則] 索引標籤下，如果金鑰原則的 JSON 檢視未顯示，請選擇 [切換到原則檢視]。

   d. 在金鑰政策區段中，選擇編輯，然後將角色的 Amazon 資源名稱 (ARN) 新增至`Allow use of the key`物件，如以下範例所示。

> **Note**
>
> 如果該物件遺失，請使用範例中顯示的權限來新增該物件。

```
...
{
    "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": "*"
},
...
```

如需詳細資訊，請參閱AWS Key Management Service 開發人員指南中的允許其他帳戶中的使用者使用 KMS 金鑰。

6. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入資料目錄 AWS 帳戶。
7. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。
8. 選擇註冊地點。
9. 在 [註冊位置] 頁面上，對於 Amazon S3 路徑，輸入位置路徑為**s3://*<bucket>*/*<prefix>***。取代<bucket>為值區的名稱，<prefix>以及位置路徑的其餘部分。

> **Note**
>
> 您必須輸入路徑，因為當您選擇「瀏覽」時，跨帳戶值區不會出現在清單中。

10. 對於 IAM 角色，請從步驟 2 中選擇角色。
11. 選擇註冊地點。

註銷 Amazon S3 位置

如果您不想再由 Lake Formation 管理，可以取消註冊 Amazon 簡單儲存服務 (Amazon S3) 位置。取消註冊位置不會影響在該位置授予的 Lake Formation ms 資料位置權限。您可以重新註冊已取消註冊的位置，而資料位置權限仍然有效。您可以使用不同的角色來重新註冊位置。

取消註冊位置 (主控台)

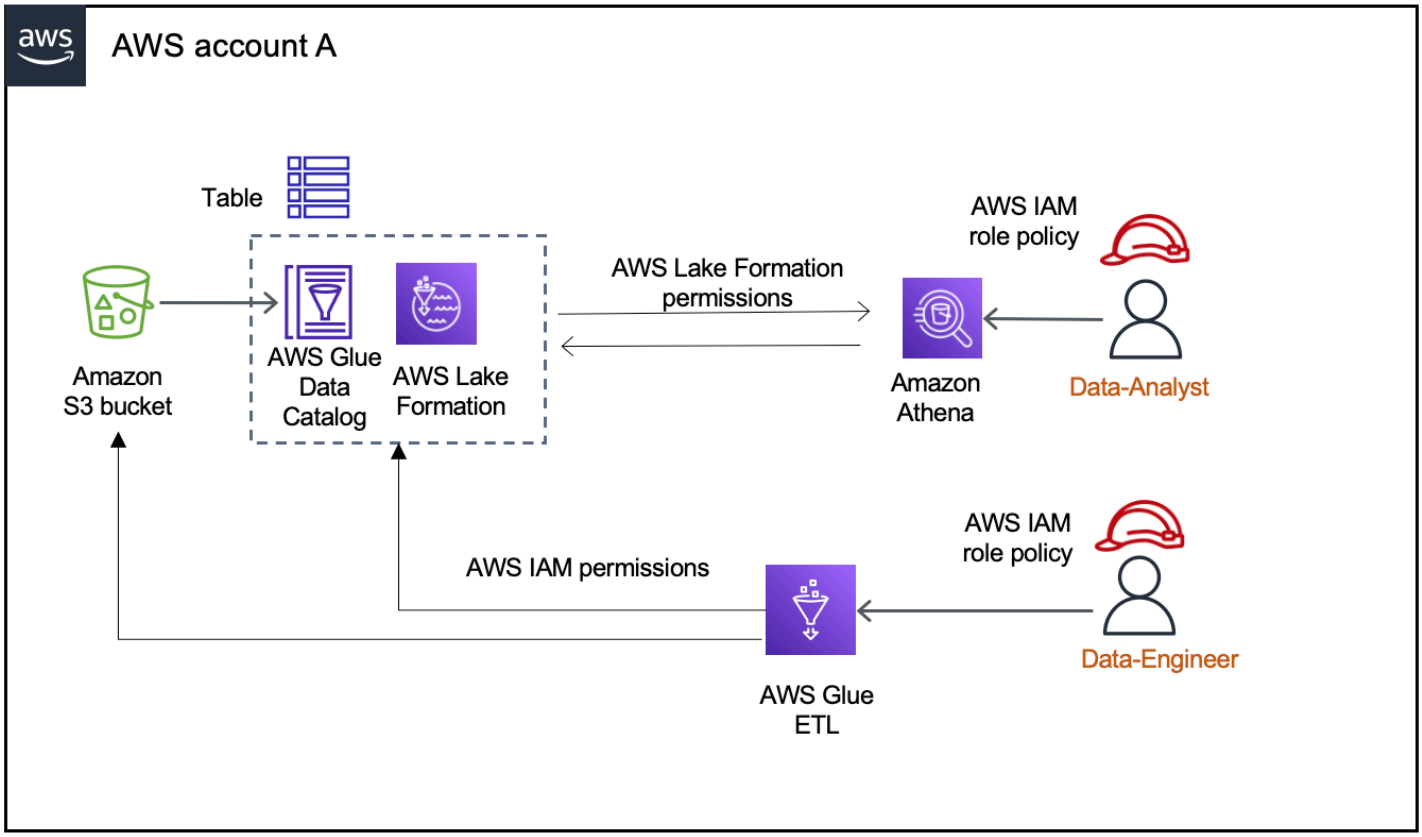
1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入，或以具有 `lakeformation:RegisterResource` IAM 權限的使用者身分登入。
2. 在導覽窗格的 [系統管理] 下，選擇 [資料湖位置]。
3. 選取位置，然後在 [動作] 功能表上選擇 [移除]。
4. 出現確認提示時，請選擇「移除」。

混合存取模式

AWS Lake Formation 混合存取模式支援兩種權限路徑到相同的 AWS Glue Data Catalog 資料庫和資料表。
在第一個途徑中，Lake Formation 可讓您選取特定主體，並透過選擇加入，授予他們 Lake Formation 存取資料庫和表格的權限。第二個途徑允許所有其他主體透過 Amazon S3 的預設 IAM 主體政策和 AWS Glue 動作存取這些資源。

向 Lake Formation 註冊 Amazon S3 位置時，您可以選擇對此位置的所有資源強制執行 Lake Formation 許可，或使用混合存取模式。依預設`CREATE_TABLE`，`CREATE_PARTITION`混合式存取模式僅強制執行`UPDATE_TABLE`權限。當 Amazon S3 位置處於混合模式時，您可以透過選擇該位置下資料庫和表格的主體來啟用 Lake Formation 許可。

因此，混合式存取模式提供了靈活性，可為特定使用者集選擇性地為資料目錄中的資料庫和表格啟用 Lake Formation，而不會中斷其他現有使用者或工作負載的存取。



如需注意事項和限制，請參閱 混合式存取模式考量與限制 。

術語和定義

以下是根據您設定存取權限的方式，資料目錄資源的定義：

Lake Formation 資源

在 Lake Formation 成註冊的資源。使用者需要 Lake Formation m 權限才能存取資源。

AWS Glue 資源

未在 Lake Formation 註冊的資源。使用者只需要 IAM 許可即可存取資源，因為它具有`IAMAllowedPrincipals`群組權限。不強制執行 Lake Formation 許可權。

如需`IAMAllowedPrincipals`群組權限的詳細資訊，請參閱元數據權限。

混合式資源

以混合式存取模式註冊的資源。根據存取資源的使用者，資源會在作為 Lake Formation 資源或資源之間動態切換。 AWS Glue

常見的混合存取模式使用案例

您可以使用混合存取模式，在單一帳戶和跨帳戶資料共用案例中提供存取權：

單一帳戶案例

- 將 AWS Glue 資源轉換為混合資源 — 在此案例中，您目前並未使用 Lake Formation，但想要針對資料目錄資料庫和表格採用 Lake Formation 權限。當您以混合式存取模式註冊 Amazon S3 位置時，可以將 Lake Formation 權限授與選擇指向該位置的特定資料庫和表格的使用者。
- 將 Lake Formation 資源轉換為混合資源 — 目前，您正在使用 Lake Formation 許可來控制資料目錄資料庫的存取，但想要使用 Amazon S3 的 IAM 許可提供新主體的存取權， AWS Glue 而且不會中斷現有的 Lake Formation 許可。

  當您將資料位置註冊更新為混合式存取模式時，新主體可以使用 IAM 許可政策存取指向 Amazon S3 位置的資料目錄資料庫，而不會中斷現有使用者的 Lake Formation 許可。

  在更新資料位置註冊以啟用混合式存取模式之前，您必須先選擇使用 Lake Formation 權限存取資源的主參與者。
  這是為了防止目前工作流程的潛在中斷。
  您也需要將資料庫中資料表的`Super`權限授與`IAMAllowedPrincipal`群組。

跨帳戶資料共用案例

- 使用混合存取模式共用 AWS Glue 資源 — 在此案例中，生產者帳戶在資料庫中具有目前與使用 Amazon S3 的 IAM 許可政策和 AWS Glue 動作與消費者帳戶共用的表格。資料庫的資料位置未在 Lake Formation 註冊。

  在以混合訪問模式註冊數據位置之前，您需要將 Cross 帳戶版本設置更新為版本 4。第 4 版提供`IAMAllowedPrincipal`群組對資源的 AWS RAM 權限時，跨帳戶共用所需的新`Super`權限原則。對於具有`IAMAllowedPrincipal`群組權限的資源，您可以將 Lake Formation 權限授予外部帳戶，並選擇他們使用 Lake Formation 權限。收件者帳戶中的資料湖管理員可以將 Lake Formation 權限授與帳戶中的主體，並選擇他們加入以強制執行 Lake Formation 權限。
- 使用混合存取模式共用 Lake Formation 資源 — 目前，生產者帳戶在資料庫中有資料表，這些表格與執行 Lake Formation 權限的消費者帳戶共用。資料庫的資料位置已在 Lake Formation 註冊。

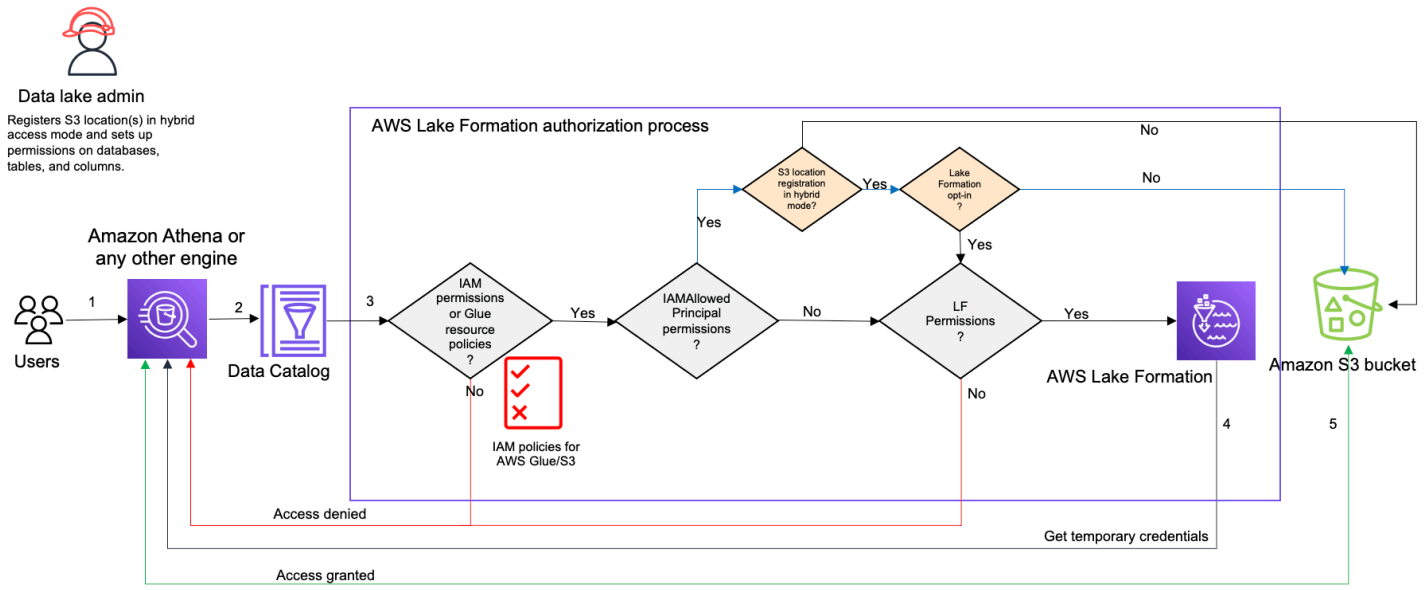
  在這種情況下，您可以將 Amazon S3 位置註冊更新為混合存取模式，並使用 Amazon S3 儲存貯體政策和資料目錄資源政策將來自 Amazon S3 的資料和資料目錄資源政策中的中繼資料共用給消費者帳戶中的主體。在更新 Amazon S3 位置註冊之前，您需要重新授予現有的 Lake Formation m 許可

並選擇加入主體。此外，您需要將資料庫中資料表的Super權限授與IAMAllowedPrincipals群組。

主題

混合存取模式的運作方式

下圖顯示當您查詢資料目錄資源時，Lake Formation 授權如何在混合式存取模式下運作。



在存取資料湖中的資料之前，資料湖管理員或具有管理權限的使用者會設定個別的「資料目錄」表格使用者原則，以允許或拒絕存取「資料目錄」中的表格。然後，具有執行RegisterResource操作許可的主體會以混合存取模式向 Lake Formation 註冊表格的 Amazon S3 位置。系統管理員會將 Lake Formation 權限授與資料目錄資料庫和資料表上的特定使用者，並選擇他們在混合式存取模式下對這些資料庫和表格使用 Lake Formation 權限。

1. 提交查詢-主體使用 Amazon 雅典娜， AWS Glue亞馬遜 EMR 或 Amazon Redshift Spectrum 等整合服務提交查詢或 ETL 腳本。

2. 請求資料-整合式分析引擎會識別要求的資料表，並將中繼資料要求傳送至資料目錄 (GetTable、GetDatabase)。
3. 檢查權限-「資料目錄」會使用 Lake Formation 驗證查詢主體的存取權限。
   a. 如果資料表沒有附加的IAMAllowedPrincipals群組權限，則會強制執行 Lake Formation 權限。
   b. 如果主體已選擇在混合式存取模式中使用 Lake Formation 權限，且資料表具有附加的IAMAllowedPrincipals群組權限，則會強制執行 Lake Formation 權限。查詢引擎應用它從 Lake Formation 接收的過濾器，並將數據返回給用戶。
   c. 如果資料表位置未向 Lake Formation 註冊，且主體尚未選擇在混合式存取模式中使用 Lake Formation 權限，則「資料目錄」會檢查資料表是否有附加IAMAllowedPrincipals群組權限。如果此權限存在於資料表上，則帳戶中的所有主參與者都會取得Super或表格的All權限。
4. 取得認證 —「資料目錄」會檢查並讓引擎知道表格位置是否已在 Lake Formation 中註冊。如果基礎資料已向 Lake Formation 註冊，則分析引擎會要求 Lake Formation 提供臨時登入資料，以存取 Amazon S3 儲存貯體中的資料。
5. 取得資料 — 如果主體獲得存取資料表資料的授權，則 Lake Formation 會提供對整合式分析引擎的暫時存取權。分析引擎會使用暫時存取從 Amazon S3 擷取資料，並執行必要的篩選，例如欄、列或儲存格篩選。當引擎完成執行作業時，會將結果傳回給使用者。此程序稱為憑證販售程序。欲了解更多信息， 請參閱與 Lake Formation 面整合。
6. 如果表格的資料位置未向 Lake Formation 註冊，則會直接向 Amazon S3 進行分析引擎的第二個呼叫。系統會針對資料存取評估相關的 Amazon S3 儲存貯體政策和 IAM 使用者政策。每當您使用 IAM 政策時，請務必遵循 IAM 最佳實務。如需詳細資訊，請參閱 IAM 使用者指南中的 IAM 中的安全性最佳實務。

設定混合式存取模式-常見案例

與 Lake Formation 權限一樣，您通常會有兩種情況，在這兩種情況下，您可以使用混合式存取模式來管理資料存取：提供對一個主參與者的存取權， AWS 帳戶 並提供對外部 AWS 帳戶 或主體的存取權。

本節提供在下列案例中設定混合式存取模式的指示：

在一個混合存取模式中管理權限 AWS 帳戶

- 將 AWS Glue 資源轉換為混合資源 — 您目前正在使用 Amazon S3 的 IAM 許可，為帳戶中的所有主體提供資料庫中表格的存取權， AWS Glue 但想要採用 Lake Formation 以增量方式管理許可。

- 將 Lake Formation 資源轉換為混合資源 — 您目前正在使用 Lake Formation 來管理您帳戶中所有主參與者之資料庫中表格的存取權，但只想針對特定主體使用 Lake Formation。您想要在相同的資料庫和資料表上使用 AWS Glue 和 Amazon S3 的 IAM 許可，提供對新主體的存取權。

以混合式存取模式管理權 AWS 帳戶限

- 使用混合存取模式共用 AWS Glue 資源— 您目前並未使用 Lake Formation 來管理資料表的權限，但想要套用 Lake Formation 權限，以便為另一個帳戶中的主體提供存取權限。
- 使用混合接入模式共享 Lake Formation 資源— 您正在使用 Lake Formation 來管理資料表的存取權，但想要透過在相同資料庫和資料表上使用 AWS Glue 和 Amazon S3 的 IAM 許可，為另一個帳戶中的主體提供存取權。

設定混合存取模式 — 高階步驟

1. 透過選取混合存取模式，向 Lake Formation 註冊 Amazon S3 資料位置。
2. 主參與者必須具有資料湖位置的DATA_LOCATION權限，才能建立指向該位置的資料目錄表格或資料庫。
3. 將跨帳戶版本設定設定為版本 4。
4. 向特定 IAM 使用者或資料庫和資料表上的角色授予精細許可。同時，請確定Super對資料庫上的IAMAllowedPrincipals群組以及資料庫中所有或選取的資料表設定或All權限。
5. 選擇加入主參與者和資源。帳戶中的其他主體可以使用適用於和 Amazon S3 動作的 IAM 許可政策繼續存取資料庫 AWS Glue 和表格。
6. 選擇性地為選擇加入使用 Lake Formation 許可的主體清除 Amazon S3 的 IAM 許可政策。

設定混合式存取模式的先決條件

以下是設定混合式存取模式的先決條件：

> **Note**
>
> 我們建議 Lake Formation 管理員以混合存取模式註冊 Amazon S3 位置，並選擇加入主體和資源。

1. 授與資料位置權限 (DATA_LOCATION_ACCESS) 以建立指向 Amazon S3 位置的資料目錄資源。資料位置許可控制建立指向特定 Amazon S3 位置的資料目錄資料庫和表格的能力。

2. 若要在混合式存取模式下與其他帳戶共用資料目錄資源 (不從資源移除IAMAllowedPrincipals群組權限)，您需要將「跨帳戶」版本設定更新為第 4 版。若要使用 Lake Formation 主控台更新版本，請在「資料目錄」設定頁面的「跨帳戶版本設定」 下選擇「版本 4」。

   您也可以使用put-data-lake-settings AWS CLI 指令將CROSS_ACCOUNT_VERSION參數設定為版本 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. 若要在混合式存取模式下授予跨帳戶許可，授與者必須具有 AWS Glue 和 AWS RAM 服務的必要 IAM 許可。受 AWS 管理的原則會AWSLakeFormationCrossAccountManager授與必要的權限。
   為了在混合式存取模式下啟用跨帳戶資料共用，我們新增了兩個新的 IAM 許可來更新AWSLakeFormationCrossAccountManager受管政策：

   - 公羊：ListResourceSharePermissions
   - 公羊：AssociateResourceSharePermission

> **Note**
>
> 如果您未針對授與者角色使用 AWS 受管理的原則，請將上述原則新增至您的自訂原則。

將 AWS Glue 資源轉換為混合資源

請遵循以下步驟以混合存取模式註冊 Amazon S3 位置，並在不中斷現有資料目錄使用者資料存取的情況下啟動新的 Lake Formation 使用者。

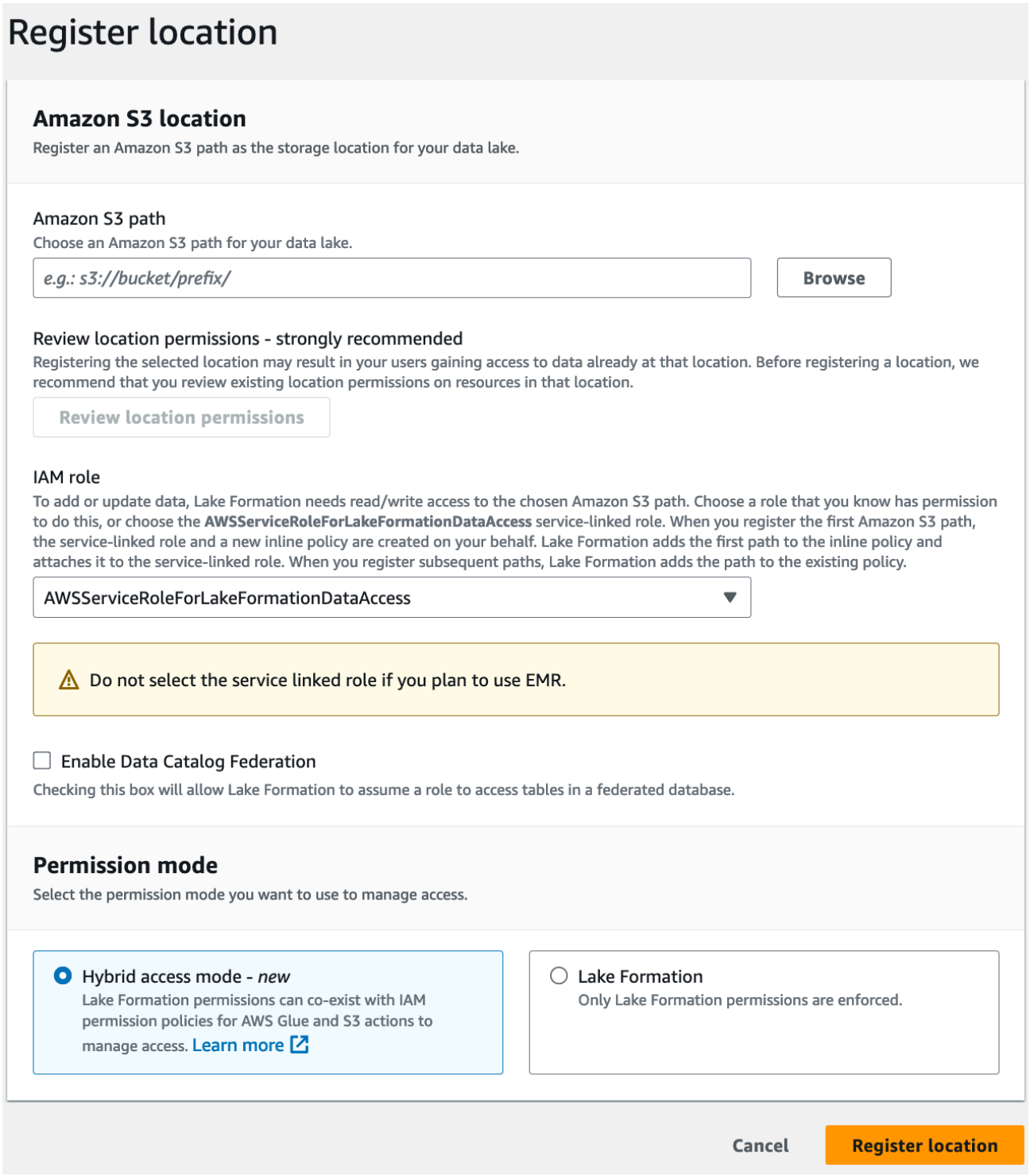
案例說明-資料位置未向 Lake Formation 註冊，使用者對資料目錄資料庫和表格的存取權由 Amazon S3 的 IAM 許可政策和 AWS Glue 動作決定。
依預設，`IAMAllowedPrincipals`群組具有資料庫中所有資料表的`Super`權限。

為未在 Lake Formation 註冊的資料位置啟用混合存取模式

1. 註冊啟用混合存取模式的 Amazon S3 位置。

   Console

   1. 以資料湖管理員身分登入 Lake Formation 主控台。
   2. 在導覽窗格中，選擇 [管理] 下的 [資料湖位置]。
   3. 選擇註冊地點。



4. 在 [註冊位置] 視窗中，選擇您要向 Lake Formation 註冊的 Amazon S3 路徑。
5. 對於 IAM 角色，請選擇`AWSServiceRoleForLakeFormationDataAccess`服務連結角色 (預設) 或自訂 IAM 符合中需求的角色[用於註冊地點的角色需求](#)。
6. 選擇混合式存取模式，將精細的 Lake Formation 存取控制原則套用至選擇加入主參與者，以及指向已註冊位置的資料目錄資料庫和表格。

選擇 Lake Formation 以允許 Lake Formation 授權對註冊位置的訪問請求。

7. 選擇註冊地點。

AWS CLI

以下是與 Lake Formation 註冊數據位置的例子：真/ HybridAccessEnabled假。該HybridAccessEnabled參數的默認值是假的。將 Amazon S3 路徑、角色名稱和 AWS 帳戶識別碼取代為有效值。

```
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. 授予權限並選擇加入主體，以便在混合式存取模式下對資源使用 Lake Formation 權限

在您選擇以混合式存取模式加入主參與Super者和資源之前，請確認在具有在混合式存取模式下向 Lake Formation 註冊位置的資料庫和表格上是否存在IAMAllowedPrincipals群組的授與或All權限。

> **Note**
>
> 您無法在資料庫All tables內授與IAMAllowedPrincipals群組權限。您需要從下拉菜單中單獨選擇每個表，並授予權限。此外，在資料庫中建立新表格時，您可以選擇使用「資料目錄設定」**Use only IAM access control for new tables in new databases** 中的。當您在資料庫中建立新資料表時，此選項會自動將Super權限授與IAMAllowedPrincipals群組。

Console

1. 在 Lake Formation 主控台的「資料目錄」下，選擇「資料庫」或「表格」。
2. 從清單中選取資料庫或表格，然後從「動作」功能表選擇「授權」。
3. 選擇主參與者，使用具名的資源方法或 LF 標籤來授與資料庫、資料表和資料行的權限。

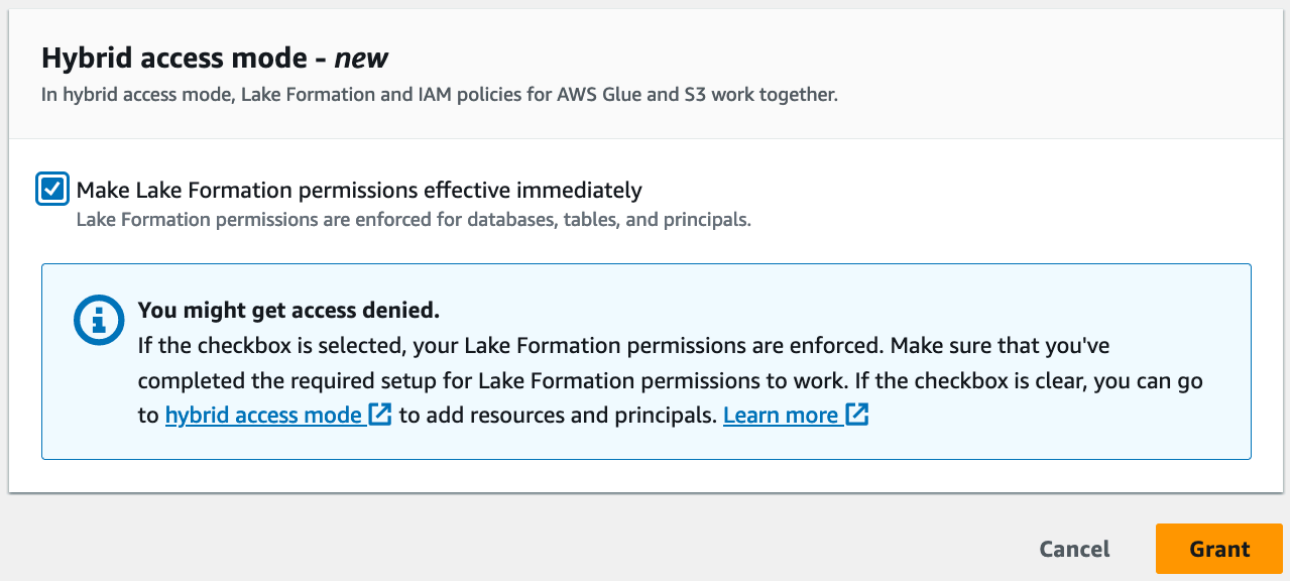
或者，選擇資料湖權限，從清單中選取要授與權限的主參與者，然後選擇授與。

如需授與資料權限的詳細資訊，請參閱授與和撤銷資料目錄資源的權限。

> **Note**
>
> 如果您要授與主體「建立」資料表權限，您也必須將資料位置權限 (DATA_LOCATION_ACCESS) 授與主參與者。更新資料表不需要此權限。如需詳細資訊，請參閱 授與資料位置權限。

4. 當您使用具名資源方法授與權限時，可在 [授與資料] 權限頁面的下方區段中找到選擇加入主參與者和資源的選項。

   選擇「讓 Lake Formation」權限立即生效，以啟用主體和資源的「Lake Formation」權限。



5. 選擇 Grant (授予)。

   當您在指向資料位置的資料表 A 上選擇加入主體 A 時，如果資料位置是以混合模式註冊，則主體 A 可以使用 Lake Formation 權限存取此資料表的位置。

AWS CLI

以下是選擇在混合存取模式下的主體和資料表的範例。將角色名稱、 AWS 帳戶 ID、資料庫名稱和資料表名稱取代為有效值。

```
aws lakeformation create-lake-formation-opt-in --cli-input-json file://file path
```

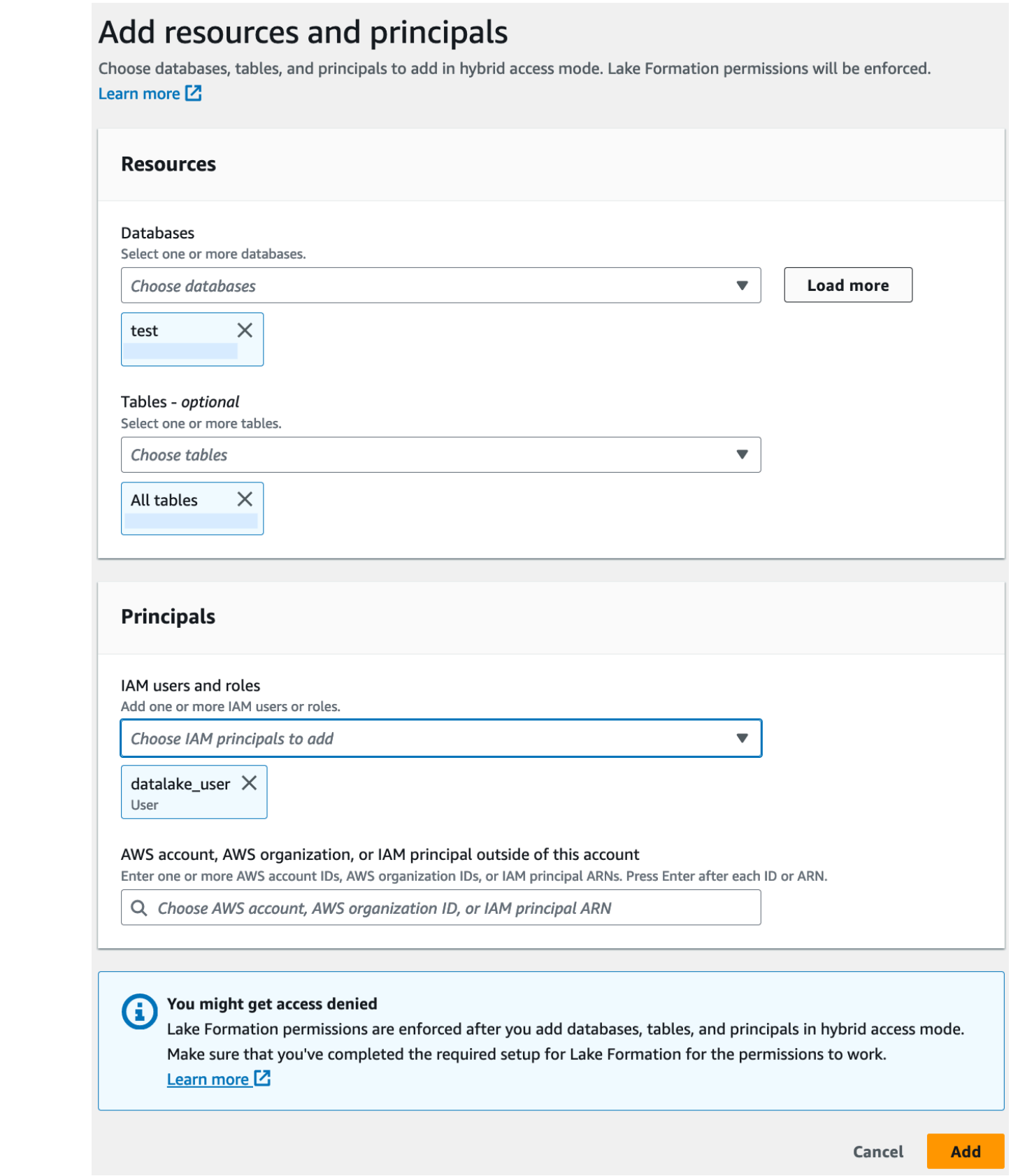
```
json:
  {
        "Principal": {
            "DataLakePrincipalIdentifier":
 "arn:aws:iam::<123456789012>:role/<hybrid-access-role>"
        },
        "Resource": {
            "Table": {
                "CatalogId": "<123456789012>",
                "DatabaseName": "<hybrid_test>",
                "Name": "<hybrid_test_table>"
            }
        }
    }
```

a. (Optional) 如果您選擇 LF-tags 來授與權限，您可以選擇加入主體，以便在單獨的步驟中使用 Lake Formation 權限。您可以從左側導覽列的 [權限] 下選擇 [混合存取模式] 來執行此操作。

b. 在 [混合式存取模式] 頁面的下方區段中，選擇 [新增]，將資源和主體新增至混合式存取模式。

c. 在 [新增資源和主體] 頁面上，選擇以混合存取模式註冊的資料庫和表格。選擇要選擇加入的主體，以便在混合式存取模式中使用 Lake Formation 權限。

   您可以在數據庫`All tables`下選擇授予訪問權限。


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

將 Lake Formation 資源轉換為混合資源

如果您目前對資料目錄資料庫和資料表使用 Lake Formation 權限，您可以編輯位置註冊屬性以啟用混合式存取模式。這可讓您使用 Amazon S3 的 IAM 許可政策和 AWS Glue 動作，為新的主體提供相同資源的存取權，而不會中斷現有的 Lake Formation 許可。

案例說明-下列步驟假設您已向 Lake Formation 註冊了資料位置，而且您已針對指向該位置的資料庫、資料表或資料行上的主體設定權限。如果位置已使用服務連結角色註冊，則無法更新位置參數並啟用混合存取模式。根據預設，`IAMAllowedPrincipals`群組對資料庫及其所有資料表具有「超級」權限。

> **Important**
>
> 如果不選擇存取此位置中資料的主體，請勿將位置註冊更新為混合式存取模式。

為在 Lake Formation 註冊的資料位置啟用混合存取模式

1.

   > **Warning**
   >
   > 我們不建議將 Lake Formation 管理的資料位置轉換為混合式存取模式，以避免中斷其他現有使用者或工作負載的權限原則。

   選擇擁有 Lake Formation 權限的現有主體。

   1. 列出並檢閱您授與資料庫和資料表上主參與者的權限。如需詳細資訊，請參閱 查看 Lake Formation 中的數據庫和表權限。
   2. 從左側導覽列選擇 [權限] 下的 [混合存取模式]，然後選擇 [新增]。
   3. 在 [新增主體和資源] 頁面上，從您要在混合存取模式下使用的 Amazon S3 資料位置選擇資料庫和表格。選擇已擁有 Lake Formation 權限的主體。
   4. 選擇 [新增] 以選擇加入主體，以便在混合式存取模式中使用 Lake Formation 權限。
2. 選擇混合存取模式選項，以更新 Amazon S3 儲存貯體/前置詞登錄。

   Console

   1. 以資料湖管理員身分登入 Lake Formation 主控台。
   2. 在導覽窗格中的 [註冊並擷取] 下，選擇 [資料湖位置]。
   3. 選取位置，然後在 [動作] 功能表上選擇 [編輯]。

4. 選擇混合存取模式。
5. 選擇儲存。
6. 在「資料目錄」下，選取資料庫或表格，並授與Super或All權限給稱為的虛擬群組IAMAllowedPrincipals。
7. 確認當您更新位置註冊屬性時，您現有的 Lake Formation m 使用者存取權限不會中斷。以 Lake Formation 主體身分登入 Athena 主控台，並在指向更新位置的資料表上執行範例查詢。

   同樣地，驗證使用 IAM 許可政策存取資料庫和資料表的使用 AWS Glue 者存取權。

AWS CLI

以下是與 Lake Formation 註冊數據位置的例子：真/ HybridAccessEnabled假。該HybridAccessEnabled參數的默認值是假的。將 Amazon S3 路徑、角色名稱和 AWS 帳戶識別碼取代為有效值。

```
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
}
```

使用混合存取模式共用 AWS Glue 資源

與另一個 AWS 帳戶 執行 Lake Formation 許可的其他人 AWS 帳戶 或主體共用資料，而不會中斷現有的「資料目錄」使用者基於 IAM 的存取。

案例說明-生產者帳戶具有資料目錄資料庫，該資料庫具有使用 Amazon S3 和 AWS Glue 動作的 IAM 主要政策進行存取控制。資料庫的資料位置未在 Lake Formation 註冊。根據預設，IAMAllowedPrincipals群組具有資料庫及其所有資料表的Super權限。

在混合式存取模式下授予跨帳戶 Lake Formation 權限

1. 生產者帳戶設定

   1. 使用具有 `lakeformation:PutDataLakeSettings` IAM 許可的角色登入 Lake Formation 主控台。
   2. 移至資料目錄設定，然後選擇`Version 4`跨帳戶版本設定。

      如果您目前使用的是版本 1 或 2，請參更新跨帳戶資料共用版本設定閱更新至版本 3 的指示。

      從版本 3 升級到 4 時，不需要變更權限原則。
   3. 註冊您計劃以混合存取模式共用的資料庫或表格的 Amazon S3 位置。
   4. 確認您在上述步驟中以混合式存取模式註冊資料位置的資料庫和資料表上存在`IAMAllowedPrincipals`群組的`Super`權限。
   5. 將 Lake Formation 權限授與組 AWS 織、組織單位 (OU)，或直接向其他帳戶中的 IAM 主體授與權限。
   6. 如果您要將權限直接授與 IAM 主體，請從消費者帳戶中選擇加入主體，以在混合式存取模式下強制執行 Lake Formation 許可權，方法是啟用「讓 Lake Formation」權限立即生效。

      如果您將跨帳戶權限授予另一個 AWS 帳戶，則當您選擇加入該帳戶時，僅對該帳戶的管理員強制執行 Lake Formation 權限。收件者帳戶資料湖管理員需要重疊顯示權限並選擇帳戶中的主體，以針對處於混合存取模式的共用資源強制執行 Lake Formation 權限。

      如果您選擇 [LF-標籤比對資源] 選項來授與跨帳戶權限，則必須先完成授與權限步驟。您可以在 Lake Formation 主控台左側導覽列的「權限」下選擇「混合式存取模式」，以個別步驟選擇將主體和資源加入混合式存取模式。然後選擇 [新增] 以新增您要強制執行 Lake Formation 權限的資源和主體。

2. 消費者帳戶設定

   1. 以資料湖管理員身分登入湖泊形成主控台 https://console.aws.amazon.com/lakeformation/。
   2. 轉到 https://console.aws.amazon.com/ram，然後接受資源共享邀請。 AWS RAM 主控台中的 [與我共用] 索引標籤會顯示與您帳戶共用的資料庫和資料表。
   3. 在 Lake Formation 中建立資源連結至共用資料庫和/或資料表。
   4. 將資源連結和`Grant on target`權限 (在原始共用資源上) 授`Describe`與您 (消費者) 帳戶中的 IAM 主體。

5. 將與您共用的資料庫或資料表上的 Lake Formation 權限授與您帳戶中的主體。透過啟用「使湖泊形成」權限立即生效選項，選擇加入主體和資源，以在混合式存取模式下強制執行 Lake Formation 權限。
6. 通過運行示例 Athena 查詢來測試主體的 Lake Formation 權限。使用 Amazon S3 的 IAM 主要政策和 AWS Glue 動作來測試 AWS Glue 使用者的現有存取權。

   (選擇性) 針對設定為使用 Lake Formation 許可的主體移除資料存取和 IAM 主體政策的 Amazon S3 儲存貯體政策，以 AWS Glue 及針對設定為使用湖泊形成許可的主體移除 Amazon S3 資料存取。

使用混合接入模式共享 Lake Formation 資源

允許外部帳戶中的新資料目錄使用者使用基於 IAM 的政策存取資料目錄資料庫和表格，而不會中斷現有的 Lake Formation 跨帳戶共用權限。

案例說明-生產者帳戶具有 Lake Formation 管理的資料庫和表格，這些資料庫與帳戶層級或 IAM 主要層級的外部 (消費者) 帳戶共用。資料庫的資料位置已在 Lake Formation 註冊。群`IAMAllowedPrincipals`組沒有資料庫及其資料表的`Super`權限。

透過基於 IAM 的政策授予跨帳戶存取權給新的資料目錄使用者，而不會中斷現有的 Lake Formation 權限

1. 生產者帳戶設定

   1. 使用角色登錄到 Lake Formation 控制台`lakeformation:PutDataLakeSettings`。
   2. 在資料目錄設定下，選擇`Version 4`跨帳戶版本設定。

      如果您目前使用的是版本 1 或 2，請參更新跨帳戶資料共用版本設定閱更新至版本 3 的指示。

      從版本 3 升級到 4 不需要變更權限原則。
   3. 列出您授與資料庫和資料表主體的權限。如需詳細資訊，請參閱 查看 Lake Formation 中的數據庫和表權限。
   4. 通過選擇主體和資源來重新授予現有的 Lake Formation 跨帳戶權限。

> **Note**
>
> 在將資料位置註冊更新為混合式存取模式以授予跨帳戶權限之前，您必須重新授予每個帳戶至少一個跨帳戶資料共用。此步驟對於更新附加至 AWS RAM 資源共用的 AWS RAM 受管理權限是必要的。
> 2023 年 7 月，Lake Formation m 更新了用於共享數據庫和表的 AWS RAM 託管權限：
>
> - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueAllTablesReadWriteForDatabase`(資料庫層級共用原則)
> - `arn:aws:ram::aws:permission/AWSRAMLFEnabledGlueTableReadWrite`(表格層級共用政策)
>
> 2023 年 7 月之前授予的跨帳戶權限沒有這些更新 AWS RAM 的權限。
> 如果您已將跨帳戶權限直接授與主體，則需要個別將這些權限重新授與主體。如果您略過此步驟，存取共用資源的主參與者可能會收到非法的組合錯誤。

5. 轉到 https://console.aws.amazon.com/ram。
6. 主 AWS RAM 控台中的 [由我共用] 索引標籤會顯示您與外部帳戶或主體共用的資料庫和資料表名稱。

   確保附加到共享資源的權限具有正確的 ARN。
7. 確認 AWS RAM 共用中的資源處於`Associated`狀態。如果狀態顯示為`Associating`，請等到它們進入狀`Associated`態。如果狀態變成`Failed`，請停止並聯繫 Lake Formation 服務團隊。
8. 從左側導覽列選擇 [權限] 下的 [混合存取模式]，然後選擇 [新增]。
9. 「新增主參與者與資源」頁面會顯示具有存取權的資料庫和/或表格以及主參與者。您可以透過新增或移除主參與者和資源來進行必要的更新。
10. 針對您要變更為混合式存取模式的資料庫和表格，選擇具有 Lake Formation 權限的主體。選擇資料庫和表格。
11. 選擇新增以選擇加入主體，以在混合式存取模式中強制執行 Lake Formation 權限。
12. 將`Super`權限授與資料庫和選取的資料表`IAMAllowedPrincipals`上的虛擬群組。
13. 將 Amazon S3 位置 Lake Formation 註冊編輯為混合式存取模式。
14. 使用適用於 Amazon S3 AWS Glue 動作的 IAM 許可政策，為外部 (消費 AWS Glue 者) 帳戶中的使用者授予許可。

2. 消費者帳戶設定

   1. 以資料湖管理員身分登入湖泊形成主控台 https://console.aws.amazon.com/lakeformation/。
   2. 轉到 https://console.aws.amazon.com/ram 並接受資源共享邀請。 AWS RAM 頁面中的 [與我共用的資源] 索引標籤會顯示與您帳戶共用的資料庫和表格名稱。

      針對 AWS RAM 共用，請確定附加的權限具有共用 AWS RAM 邀請的正確 ARN。檢查 AWS RAM 共用中的資源是否處於`Associated`狀態。如果狀態顯示為`Associating`，請等到它們進入狀`Associated`態。如果狀態變成`Failed`，請停止並聯繫 Lake Formation 服務團隊。
   3. 在 Lake Formation 中建立資源連結至共用資料庫和/或資料表。
   4. 將資源連結和`Grant on target`權限 (在原始共用資源上) 授`Describe`與您 (消費者) 帳戶中的 IAM 主體。
   5. 接下來，在共用資料庫或資料表上，為您帳戶中的主體設定 Lake Formation 權限。

      在左側導覽列的 [權限] 下，選擇 [混合存取模式]。
   6. 在「混合式存取模式」頁面下方區段中選擇「新增」，即可從生產者帳戶中選擇加入主參與者以及與您共用的資料庫或表格。
   7. 針對 Amazon S3 AWS Glue 動作使用 IAM 許可政策，為帳戶中的使用 AWS Glue 者授予許可。
   8. 使用 Athena 在資料表上執行個別的範例查詢，測試使用者的 Lake Formation m AWS Glue 權限和權限

      (選擇性) 針對處於混合存取模式的主體清除 Amazon S3 的 IAM 許可政策。

從混合式存取模式移除主體和資源

請遵循下列步驟，從混合式存取模式移除資料庫、資料表和主體。

Console

1. 在 https://console.aws.amazon.com/lakeformation/ 上登錄到 Lake Formation 控制台。
2. 在 [權限] 下選擇 [混合存取模式]。
3. 在 [混合存取模式] 頁面上，選取資料庫或表格名稱旁邊的核取方塊，然後選擇`Remove`。
4. 警告訊息會提示您確認動作。選擇移除。

   Lake Formation 不再對這些資源強制執行許可，並且將使用 IAM 和 AWS Glue 許可來控制對此資源的存取。如果使用者沒有適當的 IAM 許可，這可能會導致使用者無法再存取此資源。

AWS CLI

下列範例顯示如何從混合式存取模式移除資源。

```
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>"
          }
    }
}
```

以混合式存取模式檢視主參與者與資源

請遵循下列步驟，以混合式存取模式檢視資料庫、資料表和主體。

Console

1. 在 https://console.aws.amazon.com/lakeformation/ 上登錄到 Lake Formation 控制台。
2. 在 [權限] 下選擇 [混合存取模式]。
3. 「混合式存取模式」頁面會顯示目前處於混合式存取模式的資源和主體。

AWS CLI

下列範例顯示如何列出處於混合式存取模式的所有選擇加入主體和資源。

```
aws lakeformation list-lake-formation-opt-ins
```

下列範例顯示如何列出特定主要資源配對的選擇加入。

```
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>"
          }
    }
}
```

其他資源

在下列部落格文章中，我們將逐步引導您以混合式存取模式為所選使用者啟動 Lake Formation 許可，而其他使用者已透過 IAM 和 Amazon S3 許可存取該資料庫。我們將審查在一個帳戶內和兩個 AWS 帳戶之間設置混合訪問模式的說明。

- 介紹混合式存取模式， AWS Glue Data Catalog 以便使用 Lake Formation 和 IAM 和 Amazon S3 政策來保護存取安全。

建立資料目錄表格和資料庫

AWS Lake Formation 使用資AWS Glue料目錄來儲存有關資料湖、資料來源、轉換和目標的中繼資料。有關資料來源和目標的中繼資料採用資料庫和表格的形式。資料表儲存基礎資料的相關資訊，包括結構描述資訊、分割區資訊和資料位置。數據庫是表的集合。資料目錄也包含資源連結，這些連結是外部帳戶中共用資料庫和表格的連結，可用於跨帳戶存取資料湖中的資料。

每個 AWS 帳戶每個 AWS 區域都有一個資料目錄。

主題

- 建立資料庫
- 建立資料表

- 使用檢視

建立資料庫

「資料目錄」中的中繼資料表儲存在資料庫中。您可以根據需要創建任意數量的數據庫，並且可以對每個數據庫授予不同的 Lake Formation 權限。

數據庫可以有一個可選的位置屬性。此位置通常位於已向 Lake Formation 註冊的 Amazon 簡易儲存服務 (Amazon S3) 位置內。當您指定位置時，主體不需要資料位置權限，即可建立指向資料庫位置內位置的資料目錄表格。如需詳細資訊，請參閱 Underlying data access control。

若要使用 Lake Formation 主控台建立資料庫，您必須以資料湖管理員或資料庫建立者的身分登入。資料庫建立者是已獲得 Lake Formation CREATE_DATABASE 權限的主體。您可以在 Lake Formation 主控台的 [管理角色和作業] 頁面上查看資料庫建立者清單。若要檢視此清單，您必須擁有 `lakeformation:ListPermissions` IAM 權限，並以資料湖管理員或資料庫建立者身分登入，並具有CREATE_DATABASE權限的授與選項。

若要建立資料庫

1. 在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以資料湖管理員或資料庫建立者身分登入。
2. 在導覽窗格的 [資料目錄] 下，選擇 [資料庫]。
3. 選擇建立資料庫。
4. 在「建立資料庫」對話方塊中，輸入資料庫名稱、選擇性位置和可選描述。
5. 選擇性地針對此資料庫中的新資料表選取「僅使用 IAM 存取控制」。

   如需此選項的詳細資訊，請參閱the section called “變更資料湖的預設設定”。
6. 選擇建立資料庫。

建立資料表

AWS Lake Formation 中繼資料表包含資料湖中資料的相關資訊，包括結構描述資訊、分割區資訊和資料位置。這些表格儲存在「AWS Glue資料目錄」中。您可以使用它們存取資料湖中的基礎資料，並使用 Lake Formation 權限管理該資料。表格儲存在「資料目錄」的資料庫中。

有數種方法可以建立「資料目錄」表格：

- 在AWS Glue中執行爬蟲程式。請參閱AWS Glue 開發人員指南中的定義檢索器。
- 建立並執行工作流程。請參閱the section called “使用工作流程匯入資”。
- 使用 Lake Formation 控制台，AWS GlueAPI 或 AWS Command Line Interface （AWS CLI）手動創建表格。
- 使用建立資料表 Amazon Athena。
- 建立外部帳號中表格的資源連結。請參閱the section called “建立資源連結”。

創建阿帕奇冰山表

AWS Lake Formation 支援建立使用 Apache 實木複合地板資料格式的 Apache 冰山資料表， AWS Glue Data Catalog 且資料駐留在 Amazon S3 中。「資料目錄」中的表格是表示資料倉庫中資料的中繼資料定義。默認情況下，Lake Formation 創建冰山 v2 表。有關 v1 和 v2 資料表之間的區別，請參閱 Apache Iceberg 文件中的格式版本變更。

Apache Iceberg 是開放式的資料表格式，專用於非常大型的分析資料集。Iceberg 允許輕鬆更改模式，也稱為模式演進，這意味著用戶可以在不中斷基礎數據的情況下從數據表中添加，重命名或刪除列。Iceberg 還為數據版本控制提供支持，允許用戶跟踪數據超時的更改。這會啟用時間行程功能，讓使用者能夠存取和查詢資料的歷史版本，並分析更新和刪除之間的資料變更。

您可以使用 Lake Formation 控制台或 AWS Glue API 中的`CreateTable`操作在「資料目錄」中建立冰山表。如需詳細資訊，請參閱CreateTable 動作 (Python: 建立表格)。

在資料目錄中建立 Iceberg 表格時，必須在 Amazon S3 中指定表格式和中繼資料檔案路徑，才能執行讀取和寫入。

當您向 Amazon S3 資料位置註冊時，您可以使用 Lake Formation 使用精細的存取控制許可來保護您的 Iceberg 資料表。 AWS Lake Formation對於 Amazon S3 中的來源資料和未向 Lake Formation 註冊的中繼資料，存取權由 Amazon S3 的 IAM 許可政策和 AWS Glue 動作決定。如需詳細資訊，請參閱管理 Lake Formation 權限。

> **Note**
>
> 資料目錄不支援建立磁碟分割和新增 Iceberg 資料表屬性。

主題

- 必要條件

- 創建一個冰山表

必要條件

若要在資料目錄中建立 Iceberg 表格，並設定 Lake Formation 資料存取權限，您需要完成下列需求：

1. 創建冰山表所需的權限，而無需在 Lake Formation 註冊的數據。

   除了在「資料目錄」中建立資料表所需的權限外，資料表建立者還需要下列權限：

   - `s3:PutObject`在資源陣列上:aw:s3::: {儲存 bucketName}
   - `s3:GetObject`在資源陣列上:aw:s3::: {儲存 bucketName}
   - `s3:DeleteObject`在資源陣列上:aw:s3::: {儲存 bucketName}
2. 使用 Lake Formation 註冊的數據創建冰山表所需的權限：

   若要使用 Lake Formation 管理和保護資料湖中的資料，請使用 Lake Formation 註冊具有表格資料的 Amazon S3 位置。這是為了使 Lake Formation 可以將憑據 AWS 分析服務（例如 Athena，Redshift 頻譜和 Amazon EMR）出售以訪問數據。如需註冊 Amazon S3 位置的詳細資訊，請參閱將 Amazon S3 位置新增至您的資料湖。

   讀取和寫入在 Lake Formation 註冊的基礎資料的主體需要下列權限：

   - `lakeformation:GetDataAccess`
   - `DATA_LOCATION_ACCESS`

     對某個位置具有資料位置權限的主參與者也具有所有子位置的位置權限。

     如需資料位置權限的詳細資訊，請參閱基礎資料存取控制。

若要啟用壓縮，服務需要假設具有更新資料目錄中資料表的權限的 IAM 角色。如需詳細資訊，請參閱資料表最佳化先決條件

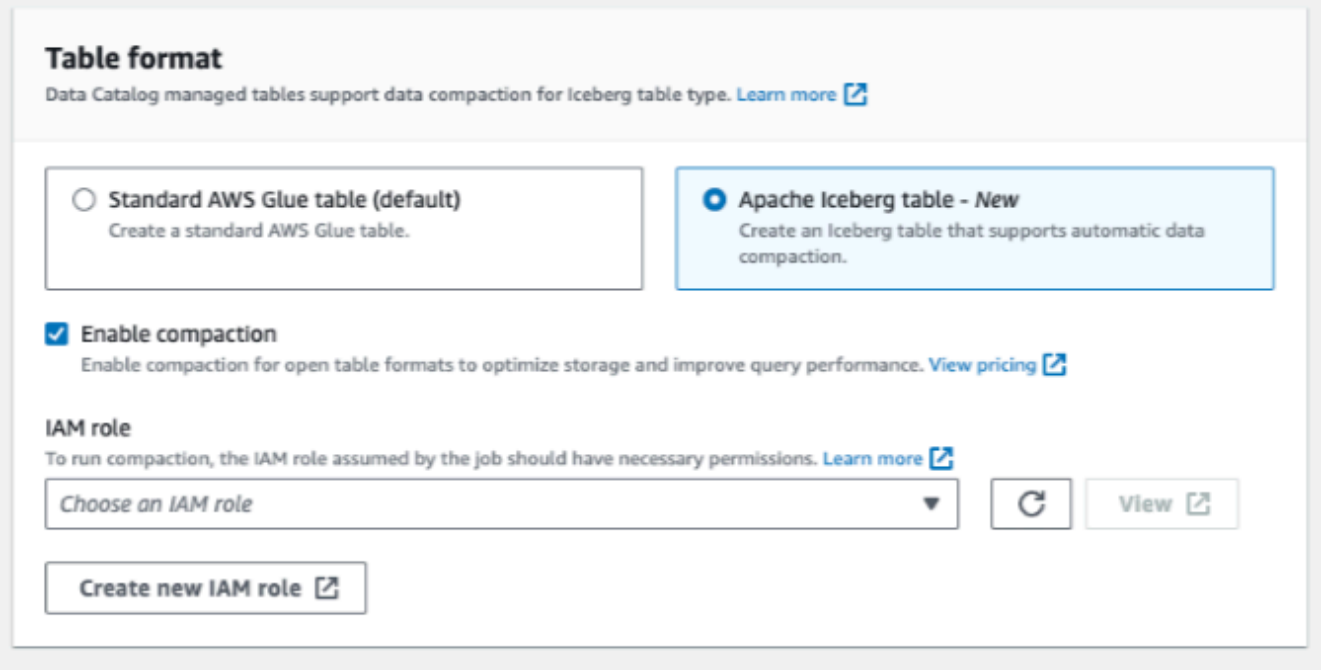
創建一個冰山表

您可以使用 Lake Formation 控制台創建冰山 v1 和 v2 表，或 AWS Command Line Interface 按照此頁面上的說明。您也可以使用 AWS Glue 控制台或 AWS Glue 編目程式. 如需詳細資訊，請參閱 AWS Glue 開發人員指南中的資料目錄和檢索器。

若要建立冰山表

Console

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在「資料目錄」下，選擇「表格」，然後使用「建立表格」按鈕指定下列屬性：
   - 表格名稱：輸入表格的名稱。如果您使用 Athena 存取表格，請參閱 Amazon Athena 使用者指南中的這些命名提示。
   - 資料庫：選擇現有的資料庫或建立新資料庫。
   - 說明：表格的說明。您可以撰寫說明，來協助您了解資料表的內容。
   - 表格格式：對於表格格式，請選擇 Apache 冰山。



   - 啟用壓縮：選擇「啟用壓縮」，將表格中的小型 Amazon S3 物件壓縮為更大的物件。
   - IAM 角色：若要執行壓縮，服務會代表您擔任 IAM 角色。您可以使用下拉式選單選擇 IAM 角色。請確認角色具有啟用壓縮的必要權限。

     若要進一步瞭解所需權限，請參閱 資料表最佳化先決條件 。
   - 位置：指定 Amazon S3 中存放中繼資料表之資料夾的路徑。Iceberg 需要資料目錄中的中繼資料檔案和位置，才能執行讀取和寫入。
   - 綱要：選擇新增資料欄來新增資料欄和資料欄的資料類型。您可以選擇建立空白資料表並稍後更新結構定義。數據目錄支持蜂巢數據類型。如需詳細資訊，請參閱 Hive 資料類型。

Iceberg 允許您在創建表後進化模式和分區。您可以使用 Athena 查詢來更新資料表結構描述和 Spark 查詢以更新分割區。

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/"
           }
       }'
```

最佳化處理 Iceberg 資料表

使用開放資料表格式 (例如，Apache Iceberg) 的 Amazon S3 資料湖會以 Amazon S3 物件形式儲存資料。如果資料湖資料表中具有數千個小型 Amazon S3 物件，會增加 Iceberg 資料表中的中繼資料額外負荷，並影響讀取效能。為了透過 Amazon EMR Amazon Athena 和 AWS Glue ETL 任務等 AWS 分析服務提供更好的讀取效能，請在資料目錄中為 Iceberg 表格 AWS Glue Data Catalog 提供受管壓縮 (將小型 Amazon S3 物件壓縮為較大物件的程序)。您可以使用 Lake Formation 主控 AWS Glue 台 AWS CLI、主控台或 AWS API 來啟用或停用資料目錄中個別 Iceberg 資料表的壓縮功能。

資料表最佳化工具會持續監視資料表分割區，並在檔案數量和檔案大小超過閾值時啟動壓縮程序。一個冰山表有資格壓縮，如果在寫入指定的文件大小。 target-file-size-bytes 屬性在 128MB 至 512 MB 範

圍內。在「資料目錄」中，如果資料表有五個以上的檔案 (每個檔案都小於寫入的 75%)，則壓縮程序便會開始。 target-file-size-bytes 財產。

例如，您在寫入時將檔案大小臨界值設定為 512MB 的資料表。 target-file-size-bytes 屬性（在 128MB 到 512MB 的規定範圍內），並且該表包含 10 個文件。如果 10 個文件中的 6 個小於 384MB（.75* 512），則數據目錄觸發壓縮。

Data Catalog 會在不干擾並行查詢的情況下執行壓縮程序。Data Catalog 僅支援 Parquet 格式資料表的資料壓縮。

如需支援的資料類型、壓縮格式和限制，請參閱受管理資料壓縮的支援格式和限制 。

主題

資料表最佳化先決條件

資料表最佳化工具會假設您在啟用資料表壓縮時指定的 AWS Identity and Access Management (IAM) 角色許可。IAM 角色必須具有讀取資料和更新 Data Catalog 中繼資料的權限。您可以建立 IAM 角色，並連接下列內嵌政策：

- 新增下列內嵌政策，針對未向 Lake Formation 註冊的資料的位置授予 Amazon S3 讀取/寫入權限。此原則也包含更新資料目錄中的表格，以及允許在記錄檔中 Amazon CloudWatch 新 AWS Glue 增記錄檔和發佈指標的權限。針對 Amazon S3 中未向 Lake Formation 註冊的來源資料，存取權將由 Amazon S3 和 AWS Glue 動作的 IAM 權限政策決定。

  在下列內嵌政策中，請將 `bucket-name` 取代為 Amazon S3 儲存貯體名稱，將 `aws-account-id` 和 `region` 取代為有效的 AWS 帳號和 Data Catalog 的區域，將 `database_name` 取代為資料庫的名稱，以及將 `table_name` 取代為資料表的名稱。

```
{
    "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:*"
        }
    ]
```

```
}
```

- 使用下列政策針對向 Lake Formation 註冊的資料啟用壓縮功能。

  如果壓縮角色沒有授與資料表的IAM_ALLOWED_PRINCIPALS群組權限，則該角色需要 Lake Formation ALTER INSERT 和資料表的DELETE權限。DESCRIBE

  如需使用 Lake Formation 註冊 Amazon S3 儲存貯體的詳細資訊，請參閱將 Amazon S3 位置新增至您的資料湖。

```
{
 "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:*"
```

```
        }
    ]
}
```

- (選用) 若要使用伺服器端加密加密之 Amazon S3 儲存貯體中的資料壓縮 Iceberg 資料表，則壓縮角色需要解密 Amazon S3 物件的權限，並產生新的資料金鑰以將物件寫入加密的儲存貯體。將下列原則新增至所需的 AWS KMS 金鑰。我們僅支援儲存貯體層級加密。

```
{
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::<aws-account-id>:role/<compaction-role-name>"
    },
    "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
    ],
    "Resource": "*"
}
```

- (選用) 針對向 Lake Formation 註冊的資料位置，用於註冊該位置的角色需要解密 Amazon S3 物件並產生新的資料金鑰以將物件寫入加密儲存貯體的權限。如需詳細資訊，請參閱 註冊加密的 Amazon S3 位置。
- (選擇性) 如果 AWS KMS 金鑰儲存在不同的 AWS 帳戶中，您必須包含壓縮角色的下列權限。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:GenerateDataKey"
            ],
            "Resource": ["arn:aws:kms:<REGION>:<KEY_OWNER_ACCOUNT_ID>:key/<KEY_ID>" ]
        }
    ]
}
```

- 您用來執行壓縮程序的角色必須具有該角色的 `iam:PassRole` 權限。

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

- 將下列信任政策新增至 AWS Glue 服務角色，以承擔 IAM 角色以執行壓縮程序。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "",
            "Effect": "Allow",
            "Principal": {
                "Service": "glue.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
 }
```

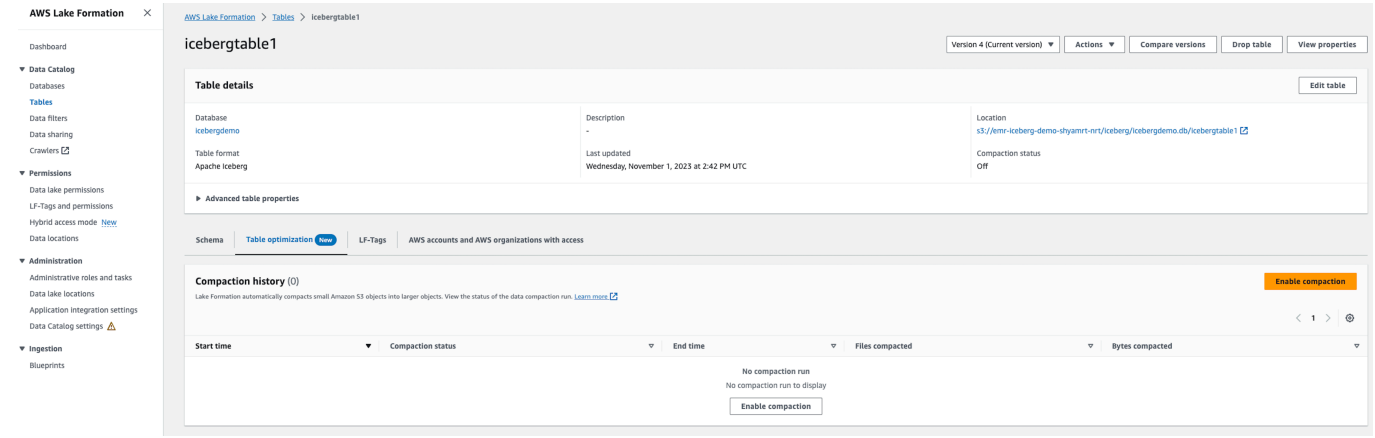
啟用壓縮功能

您可以使用 Lake Formation 主控 AWS Glue 台 AWS CLI、主控台或 AWS API，為資料目錄中的 Apache 冰山資料表啟用壓縮功能。針對新的資料表，您可以選擇 Apache Iceberg 作為資料表格式，並在您建立資料表時啟用壓縮功能。新資料表依預設會停用壓縮功能。

Console

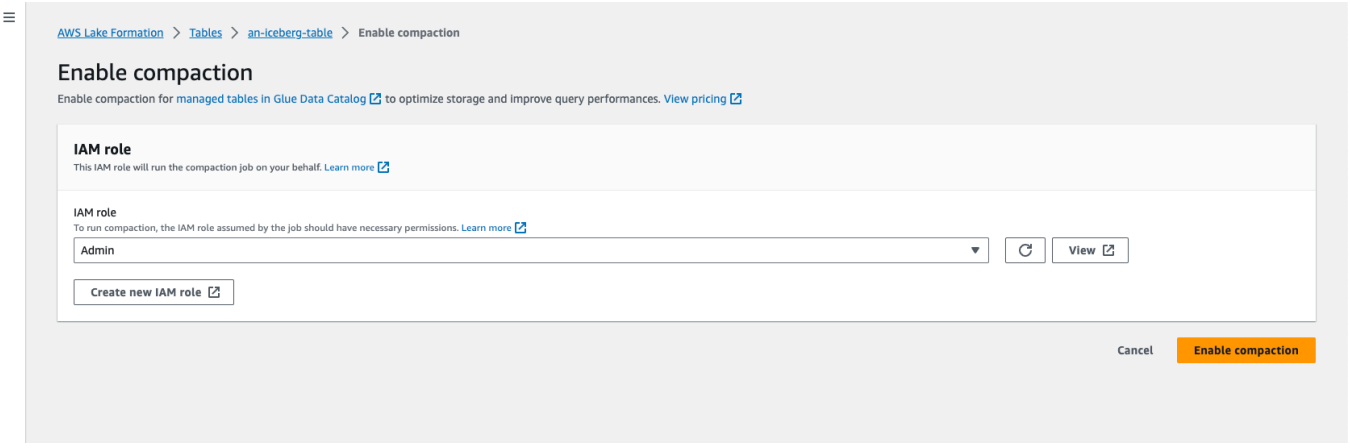
啟用壓縮功能

1. 在 https://console.aws.amazon.com/lakeformation/ 開啟 Lake Formation 主控台，並以資料湖管理員、表格建立者或已獲得資料表`glue:UpdateTable`和`lakeformation:GetDataAccess`權限的使用者身分登入。
2. 在導覽面板的 Data Catalog 下方，選擇資料表。
3. 在資料表頁面中，選擇您想要啟用壓縮功能之開放資料表格式的資料表，然後在動作功能表下選擇啟用壓縮。
4. 您也可以透過選取資料表並開啟資料表詳細資料頁面，來啟用壓縮功能。選擇頁面下半區段的資料表最佳化索引標籤，然後選擇啟用壓縮。



5. 接下來，從下拉式選單中選取現有的 IAM 角色，其權限會在 資料表最佳化先決條件 區段中顯示。

   當您選擇建立新 IAM 角色選項時，服務會建立具有執行壓縮程序之必要權限的自訂角色。



   請依照以下步驟更新現有 IAM 角色：

a. 若要更新 IAM 角色的權限政策，請在 IAM 主控台中，前往用於執行壓縮程序的 IAM 角色。
b. 在新增權限區段中，選擇建立政策。在新開啟的瀏覽器視窗中，建立要搭配您角色使用的新政策。
c. 在建立政策頁面上，選擇 JSON 標籤。將先決條件中顯示的 JSON 代碼複製到策略編輯器字段中。

AWS CLI

下列範例顯示如何啟用壓縮功能。將帳號 ID 取代為有效的 AWS 帳號 ID。將資料庫名稱和資料表名稱取代為實際的 Iceberg 資料表名稱和資料庫名稱。將其取`roleArn`代為 IAM 角色的 AWS 資源名稱 (ARN)，以及具有執行壓縮所需權限的 IAM 角色名稱。

```
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

呼叫 `CreateTableOptimizer` 操作以啟用資料表的壓縮功能。

啟用壓縮功能後，資料表最佳化索引標籤會顯示下列壓縮詳細資料 (大約 15-20 分鐘後)：

開始時間

在 Lake Formation 內開始壓實過程的時間。該值為以 UTC 時間為單位的時間戳記。

結束時間

壓實程序在資料目錄中結束的時間。該值為以 UTC 時間為單位的時間戳記。

Status

壓實執行的狀態。值會是 success 或 fail。

壓縮的檔案

壓縮的檔案總數。

字節壓縮

壓縮的字節總數。

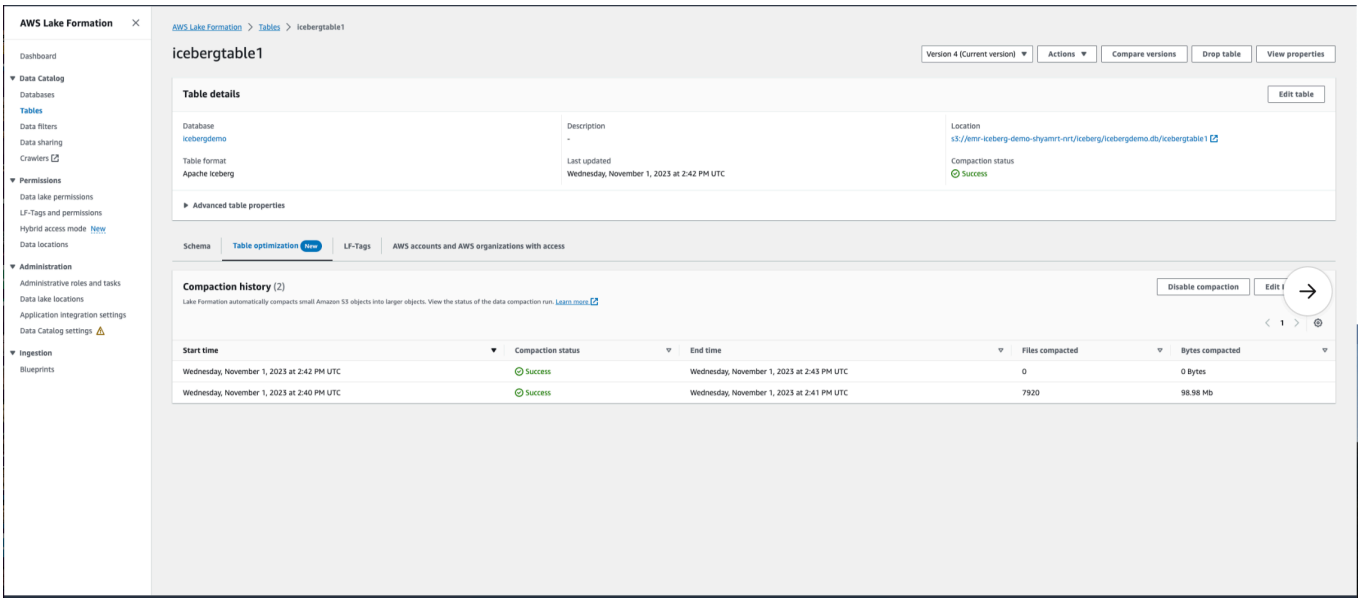
停用壓縮功能

您可以禁用自動壓縮使用 AWS Glue 控制台或特定的 Apache 冰山表。 AWS CLI

Console

1. 選擇 Data Catalog，然後選擇資料表。從資料表清單中，選擇您想要停用壓縮功能之開放資料表格式的資料表。
2. 您可以選擇 Iceberg 資料表，然後選擇動作下方的停用壓縮。

   您也可以透過選擇資料表詳細資料頁面下半區段的停用壓縮，來停用資料表的壓縮功能。



3. 在確認訊息中，選擇停用壓縮。您可以在稍後重新啟用壓縮功能。

   當您確認後，壓縮功能會停用，而資料表的壓縮狀態會變回 `Off`。

AWS CLI

在下列範例中，將帳戶 ID 取代為有效的 AWS 帳戶 ID。將資料庫名稱和資料表名稱取代為實際的 Iceberg 資料表名稱和資料庫名稱。將其取`roleArn`代為 IAM 角色的 AWS 資源名稱 (ARN)，以及具有執行壓縮所需權限的 IAM 角色的實際名稱。

```
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

調用 UpdateTableOptimizer 操作以禁用特定表的壓縮。

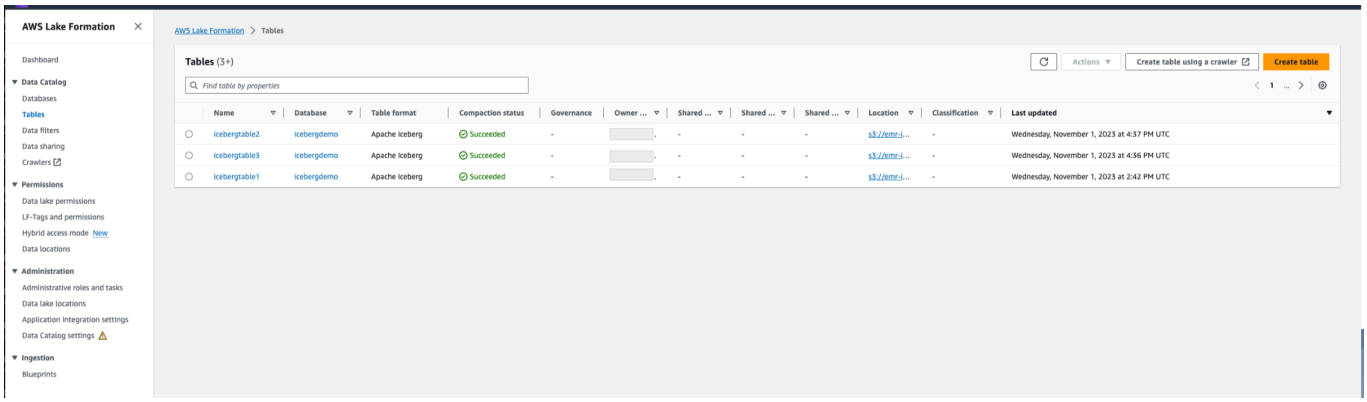
檢視壓縮詳細資料

您可以在 Lake Formation 控制台中查看 Apache 冰山的壓縮狀態 AWS CLI，或使用 AWS API 操作。

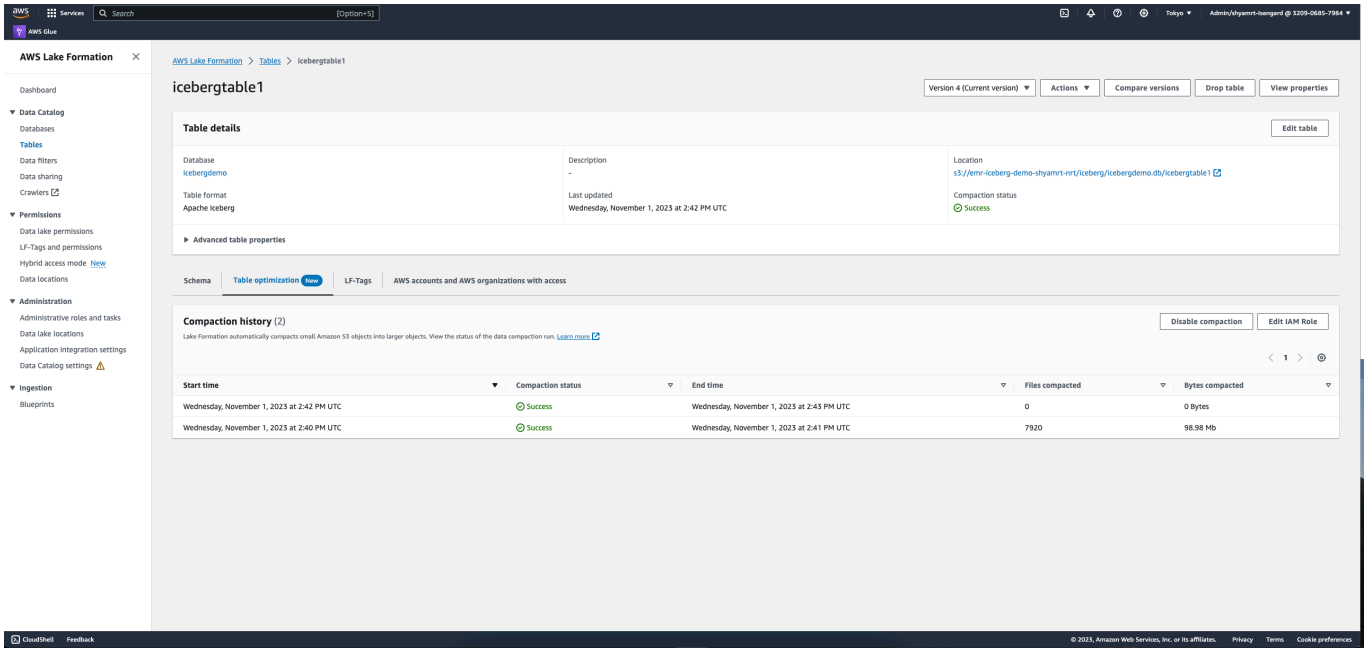
Console

若要檢視冰山表的壓縮狀態 (主控台)

- 您可以選擇「資料目錄」下的「表格」，在 Lake Formation 主控台上檢視冰山表的壓縮狀態。壓縮狀態欄位會顯示壓縮執行的狀態。您可以使用資料表偏好設定，來顯示資料表格式和壓縮狀態。



- 若要檢視特定表格的壓縮執行歷程記錄，請選擇 [下的表格] AWS Glue Data Catalog，然後選擇表格來檢視表格詳細資訊。資料表最佳化索引標籤會顯示資料表的壓縮歷史記錄。



AWS CLI

您可以使用檢視壓實詳細資訊 AWS CLI。

在下列範例中，請以有效 AWS 的帳戶 ID、資料庫名稱和資料表名稱取代為實際的 Iceberg 資料表名稱。

- 取得資料表的上次壓縮執行詳細資料

```
aws get-table-optimizer \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --type compaction
```

- 使用下列範例擷取特定資料表的最佳化工具歷史記錄。

```
aws list-table-optimizer-runs \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
  --table-name iceberg_table \
  --type compaction
```

- 下列範例顯示如何擷取多個最佳化工具的壓縮執行和組態詳細資料。您最多可以指定 20 個最佳化工具。

```
aws glue batch-get-table-optimizer \
--entries '[{"catalogId":"123456789012", "databaseName":"iceberg_db",
 "tableName":"iceberg_table", "type":"compaction"}]'
```

AWS API

- 使用 GetTableOptimizer 操作來擷取最佳化工具的上次執行詳細資料。
- 使用 ListTableOptimizerRuns 操作來擷取特定資料表中特定最佳化工具的歷史記錄。您可以在單一 API 呼叫中指定 20 個最佳化工具。
- 使用 BatchGetTableOptimizer 操作來擷取帳戶中多個最佳化工具的組態詳細資料。此操作不支援跨帳戶呼叫。

檢視 Amazon CloudWatch 量度

成功執行壓縮之後，服務會建立壓縮工作效能的 Amazon CloudWatch 指標。您可以轉到CloudWatch 指標並選擇度量標準，所有度量標準。您可以依特定命名空間 (例如 AWS Glue)、表格名稱或資料庫名稱來篩選測量結果。

如需詳細資訊，請參閱《Amazon CloudWatch 使用者指南》中的檢視可用指標。

- 壓縮的位元組數
- 壓縮的檔案數
- 配置給工作的 DPU 數目
- 任務持續時間 (小時)

刪除最佳化工具

您可以使用 AWS CLI 或 AWS API 操作刪除表的優化器和關聯的元數據。

執行下列 AWS CLI 命令以刪除表格的壓縮歷程記錄。

```
aws glue delete-table-optimizer \
  --catalog-id 123456789012 \
  --database-name iceberg_db \
```

```
  --table-name iceberg_table \
  --type compaction
```

使用 DeleteTableOptimizer 操作刪除資料表的最佳化工具。

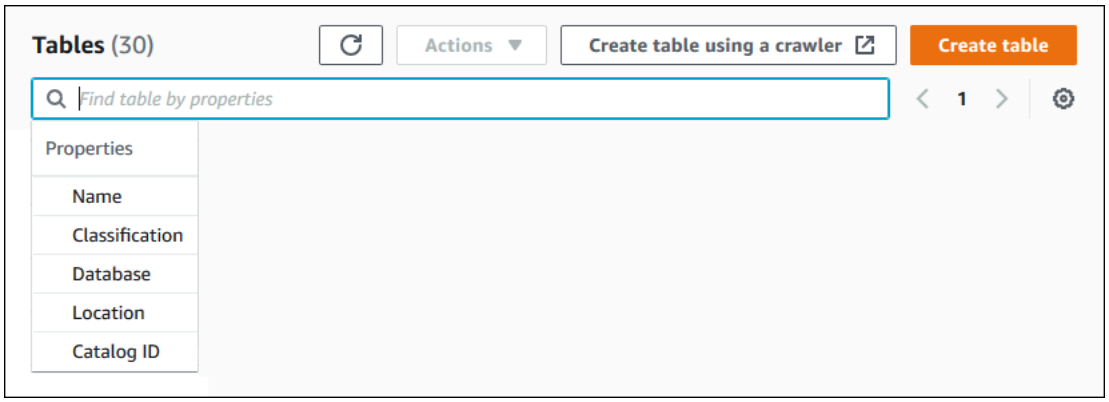
搜尋表格

您可以使用主 AWS Lake Formation 控台依名稱、位置、包含資料庫等搜尋「資料目錄」表格。搜尋結果只會顯示您擁有 Lake Formation 權限的表格。

若要搜尋表格 (主控台)

1. 登錄到 AWS Management Console 並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在導覽窗格中，選擇 Tables (資料表)。
3. 將游標置於頁面頂端的搜尋欄位中。此欄位具有預留位置文字「依屬性尋找表格」。

   這時系統顯示「屬性」菜單，其中顯示了要搜索的各種表格屬性。



4. 執行以下任意一項：

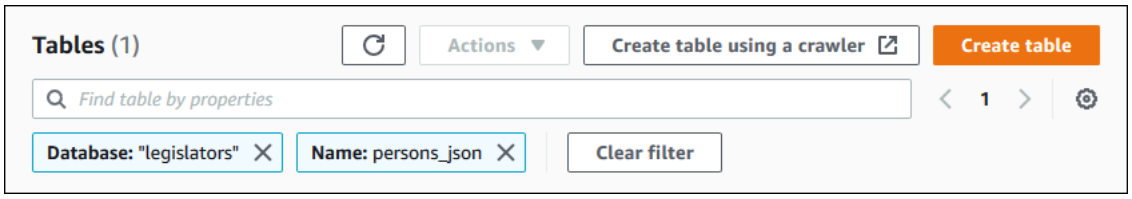
   - 按包含數據庫進行搜索。

     1. 從 [內容] 功能表選擇 [資料庫]，然後從顯示的 [資料庫] 功能表中選擇資料庫，或輸入資料庫名稱，然後按 Enter。

        列出您在資料庫中擁有權限的資料表。

     2. (選擇性) 若要將清單縮小為資料庫中的單一表格，請再次將游標置於搜尋欄位中，從「內容」功能表中選擇「名稱」，然後從顯示的「表格」功能表中選擇表格名稱，或輸入表格名稱並按 Enter。

此時會列出單一表格，而且資料庫名稱和表格名稱都會在搜尋欄位下顯示為並排。



若要調整濾鏡，請關閉其中一個拼貼，或選擇「清除濾鏡」。

- 按其他物業搜索。

  1. 從「內容」功能表中選擇搜尋屬性。

     若要依 AWS 帳戶 ID 進行搜尋，請從「內容」功能表選擇「目錄識別碼」，輸入有效的 AWS 帳戶識別碼 (例如，111122223333)，然後按 Enter 鍵。

     若要依位置搜尋，請從「內容」功能表選擇「位置」，然後從顯示的「位置」選單中選取位置。傳回所選位置 (例如 Amazon S3) 根位置中的所有表格。

跨 AWS 帳戶共用資料目錄表格和資料庫

您可以將資源的 Lake Formation 權限授與外部帳戶，以與外部 AWS 帳戶共用資料目錄資源 (資料庫和表格)。然後，使用者可以執行查詢和工作，以聯結和查詢多個帳戶的資料表。在某些限制下，當您與另一個帳號共用資料目錄資源時，該帳號中的主參與者可以在該資源上作業，就像資源位於其資料目錄中一樣。

您不會與外部 AWS 帳戶中的特定主參與者共用資源，而是與帳號或組織共用 AWS 資源。當您與 AWS 組織共用資源時，您將與該組織中所有層級的所有帳號共用資源。然後，每個外部帳戶中的資料湖管理員必須將共用資源的權限授與其帳戶中的主體。

如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享 及 授與和撤銷資料目錄資源的權限。

另請參閱:

- 存取和檢視共用資料目錄表格和資料庫
- 必要條件

使用檢視

這項功能目前在預覽版本中，並可能會有所變更。如需詳細資訊，請參閱 AWS 服務條款文件中的「測試版和預覽版」一節。

在中 AWS Glue Data Catalog，檢視是虛擬資料表，其中的內容是由參照一或多個資料表的查詢來定義。您可以使用適用於 Amazon 雅典娜、亞馬遜 Amazon Redshift 或亞馬遜 EMR 的 SQL 編輯器建立最多 10 個表格的檢視。視圖的基礎參考表可以屬於相同的數據庫或相同內的不同數據庫 AWS 帳戶。

SQL 是用於查詢表的編程語言，每個 AWS 分析引擎使用自己的 SQL 或 SQL 方言的變化。資料目錄支援使用不同的 SQL 方言建立檢視，只要每個方言參考相同的資料表、資料行和資料類型。透過定義可從多個引擎查詢的通用檢視結構描述和中繼資料物件，「資料目錄」檢視可讓您在整個資料湖中使用統一檢視。

當您管理資料目錄中的檢視時，您可以使用透過指 AWS Lake Formation 定的資源方法或使用 LF 標籤授與細微的權限，並在組織和 AWS 組織單位之間 AWS 帳戶共用這些權限。您也可以在各處共用資料目錄檢視 AWS 區域。這可讓使用者在 AWS 區域 不複製資料來源的情況下提供資料存取。

如需跨帳戶資料共用與跨區域資料存取的詳細資訊，請參閱：

- Lake Formation 的跨賬戶數據共享
- 跨區域存取表格

您可以使用「資料目錄」檢視來：

- 建立和管理單一檢視結構描述的權限。這可協助您避免在多個引擎中建立的重複檢視上存在不一致權限的風險。
- 在參考多個資料表的檢視上，將權限授與使用者，而不直接授與基礎參考資料表的權限。

有關限制，請參閱 資料目錄檢視考量和限制

主題

- 建立視圖的先決條件
- 建立檢視
- 授與資料目錄檢視的權限

建立視圖的先決條件

- 若要在資料目錄中建立檢視，您必須向 Lake Formation 註冊參考表的基礎 Amazon S3 資料位置。

  有關向 Lake Formation 註冊數據的詳細信息，請參閱將 Amazon S3 位置新增至您的資料湖。
- 檢視定義器必須是 IAM 角色。其他 IAM 身分無法建立資料目錄檢視。
- 定義檢視的 IAM 角色必須具有下列權限：
  - 完整的 Lake Formation SELECT 許可，並在所有參考表上Grantable選項。
  - 為 Lake Formation 和 AWS Glue 服務擔任該角色的信任政策。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "DataCatalogViewDefinerAssumeRole1",
            "Effect": "Allow",
            "Principal": {
               "Service": [
                    "glue.amazonaws.com",
                    "lakeformation.amazonaws.com"
                ]
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

  - IAM：PassRole 許可 AWS Glue 和 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"
                      ]
                }
            }
        }
    ]
}
```

- AWS Glue 和 Lake Formation 的權限。

```
{
    "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": "*"
        }
    ]
}
```

- 如果正在建立檢視的資料庫具有Super或授與該IAMAllowedPrincipals群組的ALL權限，則無法建立檢視表。若要撤銷資料庫上IAMAllowedPrincipals群組的Super權限，請參閱步驟 4：將資料存放區切換至 Lake Formation 型權限模型。

  如果您現有的資料湖設定不允許您為IAMAllowedPrincipals群組設定CreateTableDefaultPermissions空白，您可以建立新資料庫並使用下列結構編寫資料湖設定的程式碼。

```
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier":
 "arn:aws:iam::<AccountId>:user/<Username>"
            }
        ],
        CreateTableDefaultPermissions": [
            {
                "Principal": {
                    "DataLakePrincipalIdentifier": "IAM_ALLOWED_PRINCIPALS"
                },
                "Permissions": []
            }
        ]
 }
```

建立檢視

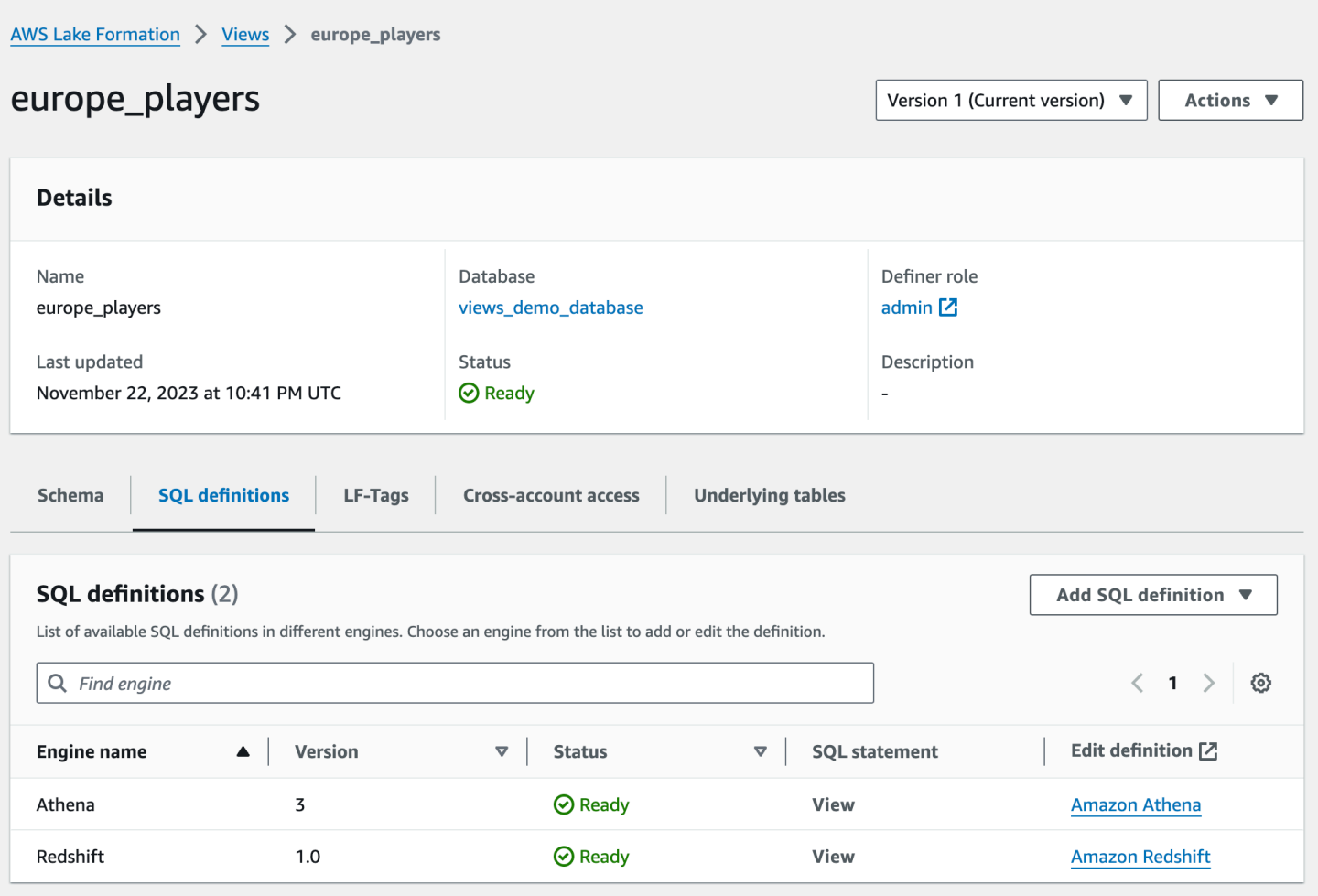
您可以使用 SQL 編輯器 Athena、亞馬 Amazon Redshift 或 Amazon EMR 在. AWS Glue Data Catalog

如需建立和管理資料目錄檢視之語法的詳細資訊，請參閱：

- 使用 Amazon Athena 用戶指南中的 AWS Glue Data Catalog 視圖。
- 在 Amazon Redshift 資料庫開發人員指南 AWS Glue Data Catalog中建立檢視。
- 使用 Amazon EMR 管理指南中的 AWS Glue Data Catalog 視圖。

建立「資料目錄」檢視之後，會顯示 Lake Formation 主控台中檢視的詳細資料。

1. 在「Lake Formation」主控台的「資料目錄」下選擇「檢視
2. 可用的檢視清單會顯示在「檢視」頁面上。
3. 從清單中選擇檢視，詳細資訊頁面就會顯示檢視的屬性。



結構描述

選擇一Column列，然後選取「編輯 LF 標籤」以更新標籤值或指定新的 LF 標籤。

SQL 定義

您可以看到可用的 SQL 定義清單。選取「新增 SQL 定義」，然後選擇查詢引擎以新增 SQL 定義。在Edit definition欄下選擇查詢引擎 (Athena 或 Amazon Redshift) 以更新 SQL 定義。

LF-標籤

選擇 「編輯 LF 標籤」以編輯標籤的值或指定新標籤。您可以使用 LF 標籤來授予檢視的權限。

跨帳戶存取權

您可以查看已共用資料目錄檢視的 AWS 帳戶組織和組織單位 (OU) 清單。

基礎資料表

用來建立檢視表的 SQL 定義中參照的基礎資料表會顯示在此索引標籤下。

授與資料目錄檢視的權限

建立檢視之後，您可以將檢視的資料湖權限授與組織和組織單位的主參與者。 AWS 帳戶如需授與權限的詳細資訊，請參閱 使用指定的資源方法授與檢視的權限 。

使用 Lake Formation 的工作流程匯入資料

使用 AWS Lake Formation，您可以使用工作流程匯入資料。工作流程定義了將資料匯入資料湖的資料來源和排程。它是用AWS Glue來協調載入和更新資料湖的處理序的編目器、作業和觸發程序的容器。

主題

- Lake Formation 的藍圖和工作流程
- 建立工作流程
- 執行工作流程

Lake Formation 的藍圖和工作流程

工作流程會封裝複雜的多工作擷取、轉換和載入 (ETL) 活動。工作流程會產生 AWS Glue 編目器、工作和觸發器，以協調資料的載入和更新。Lake Formation 會以單一實體的形式執行和追蹤工作流程。您可以將工作流程設定為依需求或按排程執行。

您在 Lake Formation 中建立的工作流程在AWS Glue主控台中會顯示為有向無環圖 (DAG)。每個 DAG 節點都是工作、爬行者程式或觸發程序。若要監控進度和疑難排解，您可以追蹤工作流程中每個節點的狀態。

完成 Lake Formation 工作流程後，執行工作流程的使用者將獲得工作流程所建立之「資料目錄」表格上的 Lake Formation SELECT 權限。

您也可以在中建立工作流程AWS Glue。但是，由於 Lake Formation 可讓您從藍圖建立工作流程，因此在 Lake Formation 中建立工作流程變得更簡單且更自動化。Lake Formation 提供了以下類型的藍圖：

- 資料庫快照集 — 從 JDBC 來源將所有表格中的資料載入或重新載入至資料湖。您可以根據排除模式從來源中排除某些資料。
- 增量資料庫 — 根據先前設定的書籤，僅將新資料從 JDBC 來源載入至資料湖。您可以在 JDBC 來源資料庫中指定要包含的個別表格。對於每個表格，您可以選擇書籤欄和書籤排序順序，以追蹤先前載入的資料。第一次對一組表格執行增量資料庫藍圖時，工作流程會從表格載入所有資料，並為下一次累加資料庫藍圖執行設定書籤。因此，您可以使用增量資料庫藍圖而不是資料庫快照藍圖來載入所有資料，前提是您將資料來源中的每個資料表指定為參數。
- 防護記錄檔 — 從記錄檔來源 (包括 AWS CloudTrail Elastic Load Balancing 記錄和應用程式負載平衡器記錄) 大量載入資料。

使用下表可協助決定是使用資料庫快照集還是增量資料庫藍圖。

在... 時使用資料庫快照集	使用增量資料庫的時機...
• 結構描述演進是靈活的。(欄會重新命名，先前的欄會被刪除，而新的欄會加入到它們的位置。) • 源和目標之間需要完全一致性。	• 架構演進是增量的。（只有連續添加列。） • 只會加入新列；先前的列不會更新。

> **Note**
>
> 使用者無法編輯由 Lake Formation 建立的藍圖和工作流程。

建立工作流程

開始之前，請確定您已將必要的資料權限和資料位置權限授與角色`LakeFormationWorkflowRole`。這樣，工作流程就可以在資料目錄中建立中繼資料表，並將資料寫入 Amazon S3 中的目標位置。如需詳細資訊，請參閱 (選擇性) 為工作流程建立 IAM 角色 及 Lake Formation 許可權概述 。

> **Note**
>
> Lake Formation 使用`GetTemplateInstance`、`GetTemplateInstances`、和`InstantiateTemplate`作業從藍圖建立工作流程。這些操作不可公開使用，並且僅在內部用於代表您創建資源。您會收到建立工作流程的 CloudTrail 事件。

若要從藍圖建立工作流程

1. 開啟主 AWS Lake Formation 控台，網址為 https://console.aws.amazon.com/lakeformation/。以資料湖管理員或具有資料工程師權限的使用者身分登入。如需詳細資訊，請參閱 Lake Formation 角色和 IAM 許可參考。
2. 在導覽窗格中，選擇 [藍圖]，然後選擇 [使用藍圖]。
3. 在 [使用藍圖] 頁面上，選擇動態磚以選取藍圖類型。
4. 在「匯入來源」下，指定資料來源。

   如果您是從 JDBC 來源匯入，請指定下列項目：

   - 資料庫連線 — 從清單中選擇連線。使用AWS Glue主控台建立其他連線。連線中的 JDBC 使用者名稱和密碼可決定工作流程可存取的資料庫物件。
   - 來源資料路徑 — <database><schema><table><database><table>根據資料庫產品，輸入//或/。Oracle 資料庫和 MySQL 不支援路徑中的結構描述。您可以用百分比 (%) 字元取代 *<schema>* 或 *<table>*。例如，對於具有系統 ID (SID) 的 Oracle 資料庫`orcl`，請輸入`orcl/%`以匯入連線中指名的使用者可存取的所有表格。

     > **Important**
     >
     > 此欄位區分大小寫。如果有任何元件的大小寫不相符，工作流程將會失敗。

     如果您指定了一個 MySQL 數據庫， AWS Glue ETL 默認情況下使用 Mysql5 JDBC 驅動程序，因此本地不支持 MySQL8。您可以編輯 ETL 工作命令檔，使用`customJdbcDriverS3Path`參數，如AWS Glue 開發人員指南中的 JDBC connectionType 值中所述，以使用支援 MySQL8 的不同 JDBC 驅動程式。

   如果您要從記錄檔匯入，請確定您為工作流程指定的角色 (「工作流程角色」) 具有存取資料來源所需的 IAM 許可。例如，若要匯入 AWS CloudTrail 日誌，使用者必須具有`cloudtrail:DescribeTrails`和`cloudtrail:LookupEvents`許可，才能在建立工作流程時查看 CloudTrail 日誌清單，而且工作流程角色必須具有 Amazon S3 中該 CloudTrail 位置的許可。
5. 執行以下任意一項：

   - 對於資料庫快照藍圖類型，選擇性地指定一或多個排除模式來識別要匯入的資料子集。這些排除模式是 Unix 風格`glob`的模式。它們會儲存為工作流程所建立之表格的屬性。

如需可用排除模式的詳細資訊，請參閱AWS Glue 開發人員指南中的包含和排除模式。

- 對於增量資料庫藍圖類型，指定下列欄位。為要匯入的每個表格新增一列。

  資料表名稱

  要匯入的表格。必須全部為小寫。

  書籤鍵

  以逗號分隔的定義書籤索引鍵的欄名稱清單。如果為空白，則使用主鍵來確定新的數據。每個欄的大小寫必須符合資料來源中定義的大小寫。

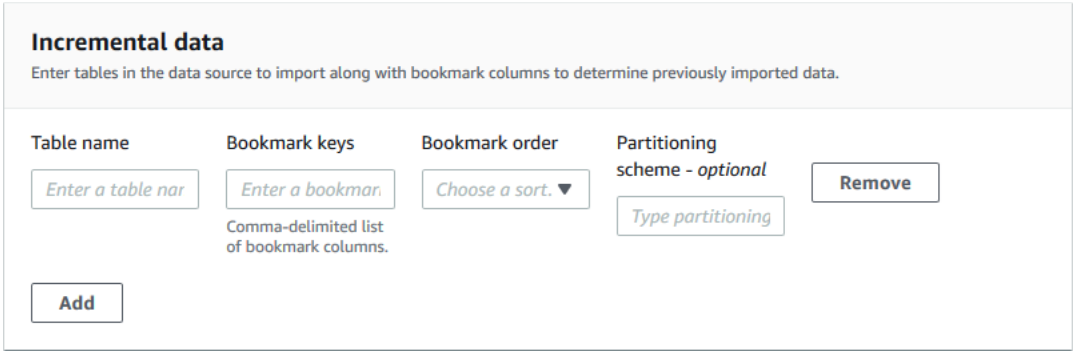
  > **Note**
  >
  > 只有在順序增加或減少（沒有間隙）時，主鍵才有資格作為默認書籤鍵。如果您想要使用主索引鍵作為書籤索引鍵，且它有間隙，您必須將主索引鍵資料行命名為書籤索引鍵。

  書籤順序

  當您選擇「升序」時，值大於書籤值的列會被識別為新列。當您選擇「遞減」時，值小於書籤值的列會識別為新列。

  分割結構

  (選擇性) 以斜線 (/) 分隔的分割索引鍵資料欄清單。例如： `year/month/day`.



如需詳細資訊，請參閱AWS Glue 開發人員指南中的使用 Job 書籤追蹤已處理的資料。

6. 在「匯入目標」下，指定目標資料庫、目標 Amazon S3 位置和資料格式。

   確保工作流程角色在資料庫和 Amazon S3 目標位置具有必要的 Lake Formation 許可。

> **Note**
>
> 目前，藍圖不支援在目標處加密資料。

7. 選擇匯入頻率。

   您可以使用「自訂」選項指定cron表示式。

8. 在匯入選項之下：

   a. 輸入工作流程名稱。

   b. 對於角色，請選擇您在中建立的角LakeFormationWorkflowRole色(選擇性) 為工作流程建立 IAM 角色。

   c. (可選) 指定資料表字首。字首會附加在工作流程建立的「資料目錄」表格名稱之前。

9. 選擇 [建立]，然後等待主控台回報工作流程已成功建立。

> **Tip**
>
> 您是否收到下列錯誤訊息？
> `User: arn:aws:iam::<account-id>:user/<username> is not authorized to perform: iam:PassRole on resource:arn:aws:iam::<account-id>:role/<rolename>...`
> 如果是這樣，請檢查您是否已<account-id>在所有策略中使用有效的 AWS 帳號替換。

> **另請參閱：**
>
> - Lake Formation 的藍圖和工作流程

執行工作流程

您可以使用 Lake Formation 主控台、主控AWS Glue台或AWS Glue命令列介面 (AWS CLI) 或 API 來執行工作流程。

執行工作流程 (Lake Formation 主控台)

1. 開啟主 AWS Lake Formation 控台，網址為 https://console.aws.amazon.com/lakeformation/。以資料湖管理員或具有資料工程師權限的使用者身分登入。如需詳細資訊，請參閱 Lake Formation 角色和 IAM 許可參考。
2. 在導覽窗格中，選擇 Blueprints (藍圖)。
3. 在 [藍圖] 頁面上，選取工作流程。然後在 [動作] 功能表上選擇 [開始]。
4. 工作流程執行時，請在「上次執行狀態」欄中檢視其進度。偶爾選擇刷新按鈕。

   狀態會從「執行中」、「探查」、「匯入」變為「已完成」。

   當工作流程完成時：

   - 資料目錄具有新的中繼資料表格。
   - 您的資料會擷取到資料湖中。

   如果工作流程失敗，請執行下列動作：

   a. 選取工作流程。選擇「動作」，然後選擇「檢視圖表」。

      工作流程會在AWS Glue主控台中開啟。

   b. 確認已選取工作流程，然後選擇 History (歷史記錄) 標籤。

   c. 在歷史記錄下，選擇最近的運行，然後選擇查看運行詳細信息。

   d. 在動態 (程式實際執行) 圖形中選取失敗的工作或爬行者程式，然後複查錯誤訊息。失敗的節點可能是紅色或黃色。

> 另請參閱：
>
> - Lake Formation 的藍圖和工作流程

管理 Lake Formation 權限

Lake Formation 為資料湖中的資料提供中央存取控制。您可以根據 Lake Formation 中的角色為使用者和應用程式定義安全原則型規則，並與 AWS Identity and Access Management 驗證這些使用者和角色的整合。規則定義完成後，Lake Formation 會對 Amazon Redshift 頻譜和亞馬 Amazon Athena 的使用者以表格和欄層級的粒度強制執行您的存取控制。

主題

授與資料位置權限

中的資料位置許可可 AWS Lake Formation 讓主體建立和更改指向指定已註冊 Amazon S3 位置的資料目錄資源。除了 Lake Formation 資料權限之外，資料位置權限還可以運作，以保護資料湖中的資訊。

Lake Formation 不會將 AWS Resource Access Manager (AWS RAM) 服務用於資料位置權限授與，因此您不需要接受資料位置權限的資源共用邀請。

您可以使用 Lake Formation 控制台、API 或 AWS Command Line Interface (AWS CLI) 來授予資料位置權限。

> **Note**
>
> 若要成功授予，您必須先向 Lake Formation 註冊資料位置。

另請參閱:

- Underlying data access control

主題

授予資料位置權限 (相同帳戶)

請依照下列步驟將資料位置權限授與您 AWS 帳戶中的主體。您可以使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ AWS CLI ）來授予權限。

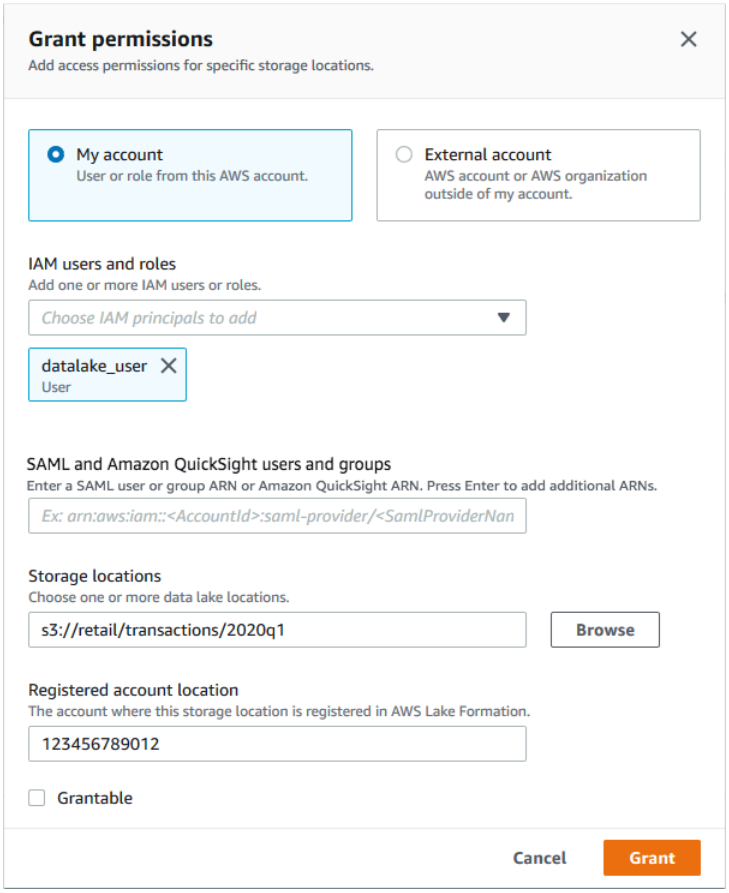
授與資料位置權限 (相同帳戶、主控台)

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員或擁有所需資料位置權限的主參與者身分登入。
2. 在功能窗格的 [權限] 下，選擇 [資料位置]。
3. 選擇 Grant (授予)。
4. 在 [授與權限] 對話方塊中，確定已選取 [我的帳戶] 磚。然後提供下列資訊：

   - 對於 IAM 使用者和角色，請選擇一或多個主體。
   - 對於 SAML 和 Amazon QuickSight 使用者和群組，請為使用者或透過 SAML 或 ARN 聯合的使用者或針對 Amazon 使用者或群組的 ARN 輸入一或多個 Amazon 資源名稱 (ARN)。QuickSight

     一次輸入一個 ARN，然後在每個 ARN 之後按 Enter。如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。
   - 對於儲存位置，請選擇瀏覽，然後選擇 Amazon Simple Storage Service (Amazon S3) 儲存位置。該地點必須在形成湖泊登記。再次選擇「瀏覽」以新增其他位置。您也可以鍵入位置，但請務必在位置之前使用`s3://`。

- 對於註冊帳戶地點，請輸入註冊地點的 AWS 帳戶 ID。這預設為您的帳戶 ID。在跨帳戶案例中，收件者帳戶中的資料湖管理員可以在將資料位置權限授與收件者帳戶中的其他主體時，在此處指定擁有者帳戶。
- (選擇性) 若要讓選取的主參與者授與所選位置的資料位置權限，請選取可授與。



5. 選擇 Grant (授予)。

若要授與資料位置權限 (相同帳戶， AWS CLI)

- 執行`grant-permissions`命令並授與主體，並`DATA_LOCATION_ACCESS`將 Amazon S3 路徑指定為資源。

  Example

  下列範例會將資料位置權限授與`s3://retail`使用者`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

下列範例會授與ALLIAMPrincipals群組的資料位置權限。s3://retail

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "111122223333"}}'
```

另請參閱:

- Lake Formation 權限參考

授與資料位置權限 (外部帳戶)

請依照下列步驟將資料位置權限授與外部 AWS 帳戶或組織。

您可以使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ AWS CLI ）來授予權限。

開始之前

確定已滿足所有跨帳戶存取先決條件。如需詳細資訊，請參閱 必要條件。

授與資料位置權限 (外部帳戶、主控台)

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員身分登入。
2. 在功能窗格的 [權限] 下，選擇 [資料位置]，然後選擇 [授與]。
3. 在 [授與權限] 對話方塊中，選擇 [外部帳戶] 動態磚。
4. 請提供下列資訊：

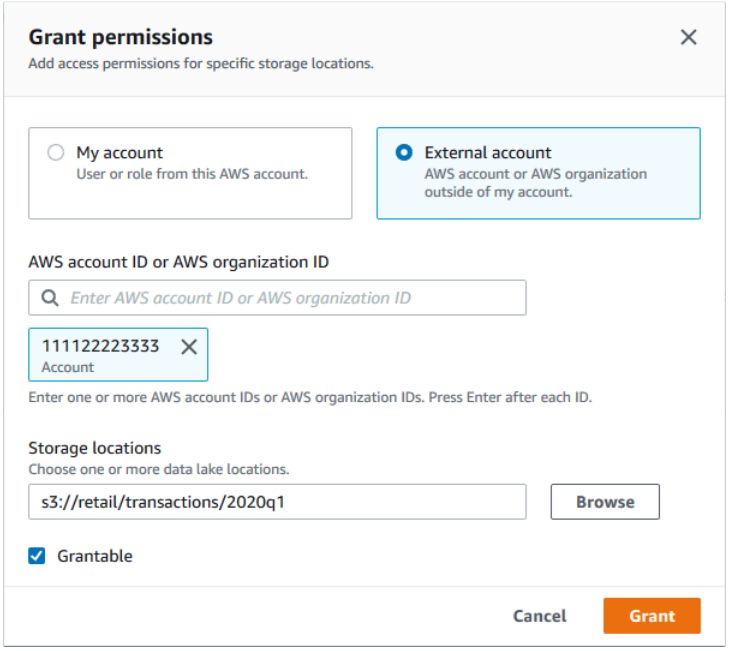
   - 針對AWS 帳號 ID 或 AWS 組織 ID，請輸入有效的 AWS 帳號、組織 ID 或組織單位 ID。

     在每個 ID 之後按 Enter 鍵。

組織 ID 由「o-」後跟 10 到 32 個小寫字母或數字組成。

組織單位 ID 由「ou-」後跟 4 到 32 個小寫字母或數字 (包含 OU 的根目錄識別碼) 組成。該字符串後跟第二個「-」（連字符）和 8 到 32 個其他小寫字母或數字。

- 在 [儲存位置] 下，選擇 [瀏覽]，然後選擇 Amazon Simple Storage Service (Amazon S3) 儲存位置。該地點必須在形成湖泊登記。



5. 選取「可授予」。
6. 選擇 Grant (授予)。

若要授與資料位置權限 (外部帳戶， AWS CLI)

- 若要授與外部 AWS 帳戶的權限，請輸入類似下列的命令。

```
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"}}'
```

此命令DATA_LOCATION_ACCESS使用授予選項授予帳戶 1111-2222-3333 Amazon S3 位置s3://retail/transactions/2020q1，該位置由帳戶 1234-5678-9012 擁有。

若要將權限授與組織，請輸入類似下列內容的命令。

```
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"}}'
```

此命令會向 Amazon S3 位置o-abcdefghijkl上的組織授予授予選項s3://retail/transactions/2020q1，該位置由帳戶 1234-5678-9012 擁有。DATA_LOCATION_ACCESS

若要將權限授與外部 AWS 帳戶中的主體，請輸入類似下列的命令。

```
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"}}'
```

此命令授DATA_LOCATION_ACCESS予 Amazon S3 位置的帳戶 1111-2222-3333 的主體s3://retail/transactions/2020q1，該位置由帳戶 1234-5678-9012 擁有。

Example

下列範例會授與外部帳戶中ALLIAMPrincipals群組的資料位置權限。s3://retail

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals --
permissions "DATA_LOCATION_ACCESS" --resource '{ "DataLocation":
 {"ResourceArn":"arn:aws:s3:::retail", "CatalogId": "123456789012"}}'
```

另請參閱:

- Lake Formation 權限參考

授予與您帳戶共用之資料位置的權限

資料目錄資源與您的 AWS 帳戶共用後，身為資料湖管理員，您可以將資源的權限授與帳戶中的其他主體。如果共用資料表授與權限，且表格指向已註冊的 Amazon S3 位置，則您還必須授與該位置的資料位置ALTER許可。同樣地，如果CREATE_TABLE或ALTER權限已授與共用資料庫，而資料庫具有指向已註冊位置的 location 屬性，您也必須授與該位置的資料位置權限。

若要將共用位置的資料位置權限授與帳戶中的主體，您的帳戶必須已獲得具有授與選項之位置的DATA_LOCATION_ACCESS權限。當您接著授與DATA_LOCATION_ACCESS帳戶中的其他主體時，您必須包含擁有者AWS 帳戶的資料目錄 ID (帳戶 ID)。所有者帳戶是註冊該地點的帳戶。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (授與AWS CLI 資料位置權限。

授予與您帳戶共用的資料位置的權限 (主控台)

- 請遵循 授予資料位置權限 (相同帳戶) 中的步驟。

  對於儲存位置，您必須輸入位置。對於註冊帳戶地點，請輸入所有者 AWS 帳戶的帳戶 ID。

授予與您帳戶共用之資料位置的權限 (AWS CLI)

- 輸入下列其中一個命令，將權限授與使用者或角色。

```
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>"}}'
```

授與和撤銷資料目錄資源的權限

您可以將資料湖權限授與中的主參與者，以 AWS Lake Formation 便主參與者可以建立和管理「資料目錄」資源，並可存取基礎資料。您可以授與資料庫、資料表和檢視的資料湖權限。當您授與資料表的權限時，您可以限制對特定資料表資料行或資料列的存取權，以進行更精細的存取控制。

您可以授與個別資料表和檢視表的權限，或是透過單一授與作業授與資料庫中所有資料表和檢視表的權限。如果您授與資料庫中所有資料表的權限，就會隱含地授與資料庫的DESCRIBE權限。然後資料庫會顯示在主控台的 [資料庫] 頁面上，並由 GetDatabases API 作業傳回。

您可以使用具名的資源方法或以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法來授與權限。

您可以將權限授與相同 AWS 帳戶 或外部帳戶或組織中的主參與者。當您授與外部帳戶或組織時，您將與這些帳戶或組織共用您擁有的資源。然後，這些帳戶或組織中的主參與者可以存取您擁有的資料目錄資源及基礎資料。

> **Note**
>
> 目前，LF-TBAC 方法支援將跨帳戶許可授與 IAM 主體 AWS 帳戶、組織和組織單位 (OU)。

當您授與外部帳戶或組織的權限時，您必須包含授與選項。只有外部帳戶中的資料湖管理員才能存取共用資源，直到管理員將共用資源的權限授與外部帳戶中的其他主體為止。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 授與資料目錄權限。

> **Note**
>
> 刪除資料目錄資源時，與該資源相關聯的所有權限都將變為無效。重新建立具有相同名稱的相同資源，將不會恢復 Lake Formation 權限。使用者必須再次設定新的權限。

> **另請參閱：**
>
> - 跨 AWS 帳戶共用資料目錄表格和資料庫
> - 元數據訪問控制
> - Lake Formation 權限參考

授予或撤銷 Lake Formation 許可所需的 IAM 許可

所有主體 (包括資料湖管理員) 都需要下列 AWS Identity and Access Management (IAM) 許可，才能使用 Lake Formation API 或撤銷「資 AWS Lake Formation 料目錄」權限或資料位置權限： AWS CLI

- `lakeformation:GrantPermissions`
- `lakeformation:BatchGrantPermissions`
- `lakeformation:RevokePermissions`
- `lakeformation:BatchRevokePermissions`
- `glue:GetTable`或`glue:GetDatabase`對於您正在使用指定資源方法授予權限的表或數據庫。

> **Note**
>
> 資料湖管理員具有隱含的湖泊形成權限，可授予和撤銷 Lake Formation 權限。但是他們仍然需要 Lake Formation 授予的 IAM 許可並撤銷 API 操作。
> 具有`AWSLakeFormationDataAdmin` AWS 受管政策的 IAM 角色無法新增新的資料湖管理員，因為此政策包含 Lake Formation API 作業的明確拒絕`PutDataLakeSetting`。

對於非資料湖管理員且想要使用 Lake Formation 主控台授與或撤銷權限的主體，建議使用下列 IAM 政策。

```
{
    "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": "*"
        }
    ]
}
```

此原則中的所有glue:和iam:權限都可在 AWS 受管理的原則中使用AWSGlueConsoleFullAccess。

若要使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 授予權限，主體需要其他 IAM 許可。如需詳細資訊，請參閱 基於 Lake Formation 標籤的訪問控制最佳實踐和考量 及 Lake Formation 角色和 IAM 許可參考。

跨帳戶 許可

想要使用具名資源方法授與跨帳戶 Lake Formation 權限的使用者，也必須具有AWSLakeFormationCrossAccountManager AWS 受管理策略中的權限。

資料湖管理員需要這些相同的權限來授與跨帳戶權限，再加上 AWS Resource Access Manager (AWS RAM) 權限才能啟用授與組織權限。如需詳細資訊，請參閱 資料湖管理員權限。

管理使用者

具有管理權限的主體 (例如，具有AdministratorAccess AWS 受管理的原則) 具有授與 Lake Formation 權限並建立資料湖管理員的權限。若要拒絕使用者或角色存取 Lake Formation 管理員作業，請在其原則中附加或新增管理員 API 作業的Deny陳述式。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "lakeformation:GetDataLakeSettings",
                "lakeformation:PutDataLakeSettings"
            ],
            "Effect": "Deny",
            "Resource": [
                "*"
            ]
        }
    ]
}
```

> **Important**
>
> 若要防止使用者透過擷取、轉換和載入 (ETL) 指令碼將自己新增為系統管理員，請確定拒絕所有非系統管理員使用者和角色存取這些 API 作業。`AWSLakeFormationDataAdmin` AWS 受管理的政策包含 Lake Formation API 作業的明確拒絕，可防止`PutDataLakeSetting`使用者新增新的資料湖管理員。

使用具名資源方法授與資料湖權限

您可以使用具名的資源方法授與特定「資料目錄」資料庫、資料表和檢視的 Lake Formation 權限。您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 來授予權限。

主題

使用指定的資源方法授與資料庫權限

下列步驟說明如何使用指定的資源方法來授與資料庫權限。

Console

使用 Lake Formation 主控台上的 [授與資料湖權限] 頁面。該頁面分為以下幾個部分：

- 主體 — 要授與權限的 IAM 使用者、角色、IAM 身分中心使用者和群組、SAML 使用者和群組、AWS 帳戶、組織或組織單位。
- LF 標籤或目錄資源 — 要授與權限的資料庫、表格、檢視或資源連結。
- 權限-要授予的 Lake Formation 許可權。

> **Note**
>
> 若要授與資料庫資源連結的權限，請參閱授與資源連結權限。

1. 開啟 [授與資料湖權限] 頁面。

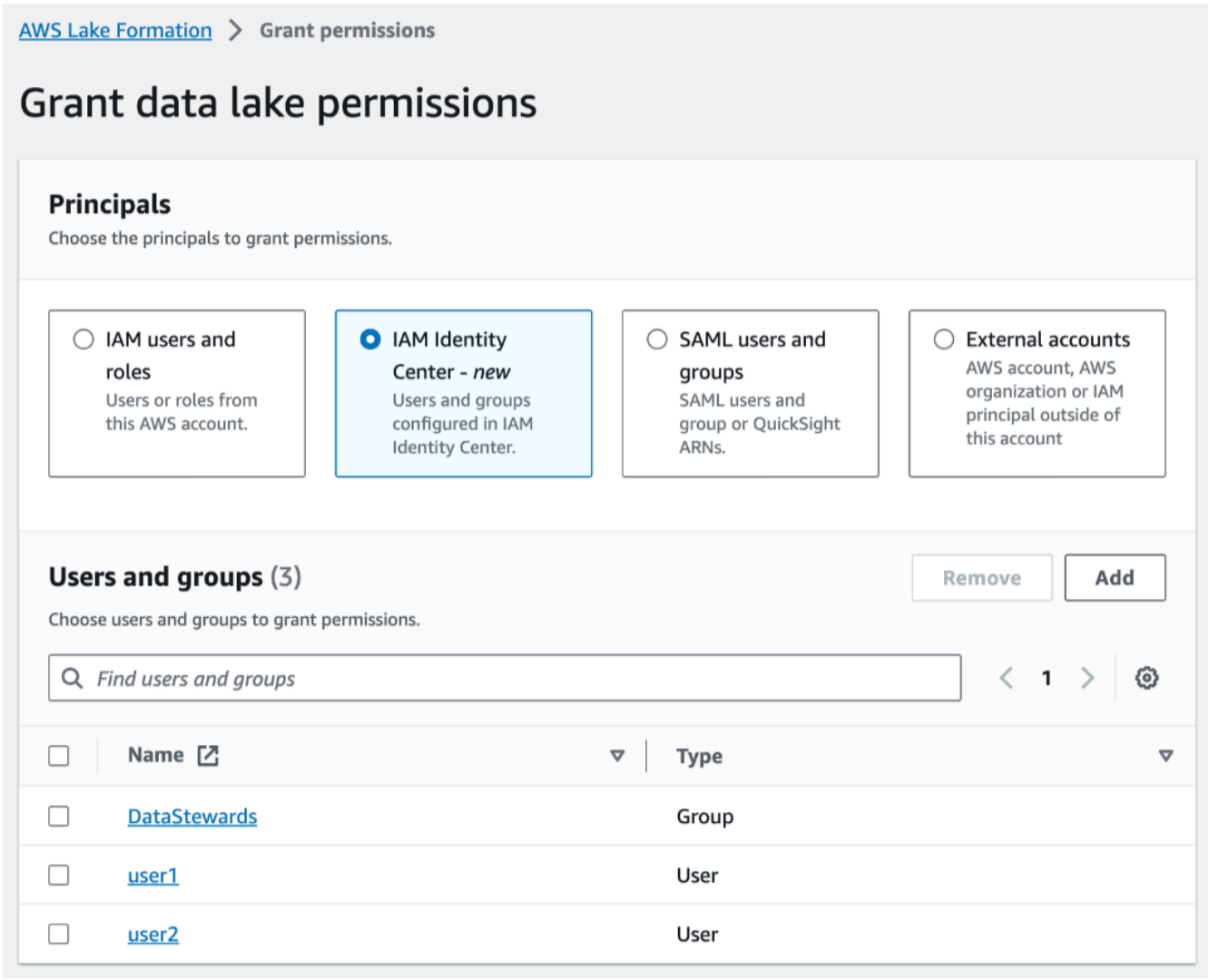
   在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以資料湖管理員、資料庫建立者或對資料庫具有可授予權限的 IAM 使用者身分登入。

   執行以下任意一項：

   - 在功能窗格的 [權限] 下，選擇 [資料湖權限]。然後選擇授予。
   - 在導覽窗格中，選擇 [資料目錄] 下的 [資料庫]。然後，在 [資料庫] 頁面上選擇資料庫，然後從 [動作] 功能表的 [權限] 下選擇 [授與]。

   > **Note**
   >
   > 您可以透過資源連結授與資料庫的權限。若要這樣做，請在「資料庫」頁面上選擇資源連結，然後在「動作」功能表上選擇「授與目標」。如需詳細資訊，請參閱 資源連結在 Lake Formation 中如何運作。

2. 接著，在 [主參與者] 區段中，選擇主參與者類型，然後指定要授與權限的主參與者。



IAM 使用者和角色

從 IAM 使用者和角色清單中選擇一或多個使用者或角色。

IAM Identity Center

從 [使用者和群組] 清單中選擇一或多個使用者或群組。選取「新增」以新增更多使用者或群組。

SAML 使用者和群組

對於 SAML 和 Amazon QuickSight 使用者和群組，請為透過 SAML 聯合的使用者或群組輸入一或多個 Amazon 資源名稱 (ARN)，或針對 Amazon 使用者或群組輸入 ARN。QuickSight 在每個 ARN 之後按 Enter 鍵。

如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。

> **Note**
>
> 僅支持 Amazon QuickSight 企業版與 Amazon QuickSight 的 Lake Formation 整合。

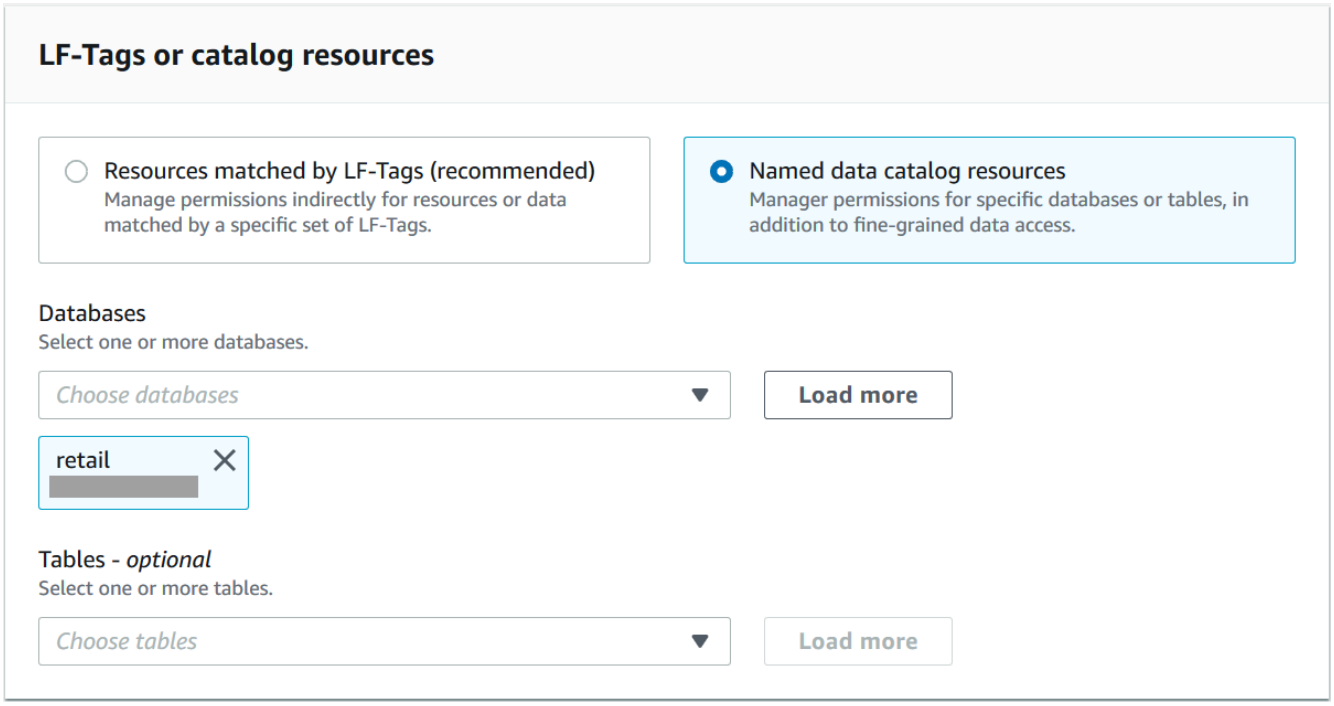
外部帳戶

針對AWS 帳戶、 AWS 組織或 IAM 主體，輸入 IAM 使用者或角色的一或多個有效 AWS 帳戶 ID、組織 ID、組織單位 ID 或 ARN。在每個 ID 之後按 Enter 鍵。

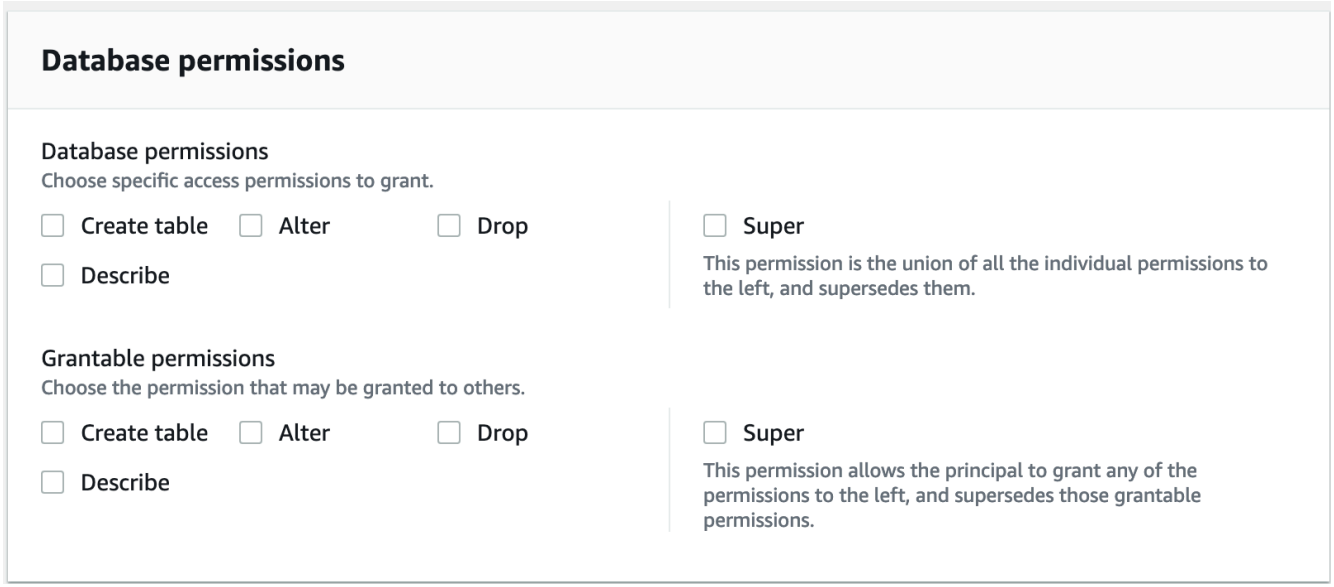
組織 ID 由「o-」後跟 10-32 個小寫字母或數字組成。

組織單位 ID 以「ou-」開頭，後面接著 4—32 個小寫字母或數字 (包含 OU 的根目錄識別碼)。該字符串後跟第二個「-」破折號和 8 到 32 個其他小寫字母或數字。

3. 在「LF 標籤」或「目錄資源」區段中，選擇「具名資料目錄資源」。



4. 從「資料庫」清單中選擇一或多個資料庫。您還可以選擇一個或多個表和/或數據過濾器。
5. 在「權限」區段中，選取權限和可授予的權限。在 [資料庫權限] 底下，選取一或多個要授與的權限。



> **Note**
>
> 授與`Create Table`或`Alter`對具有指向註冊位置之位置屬性的資料庫之後，請務必同時將該位置的資料位置權限授與主參與者。如需詳細資訊，請參閱 授與資料位置權限。

6. (選擇性) 在可授與權限下，選取授與收件者可以授與其帳戶中其他主體的權限。 AWS 當您從外部帳戶授與 IAM 主體許可時，不支援此選項。
7. 選擇 Grant (授予)。

AWS CLI

您可以使用指定的資源方法和 AWS Command Line Interface (AWS CLI) 授與資料庫權限。

若要授與資料庫權限，請使用 AWS CLI

- 執行`grant-permissions`命令，並根據授與的權限指定資料庫或資料目錄做為資源。

  在下列範例中，請以<account-id>有效的 AWS 帳戶 ID 取代。

  Example — 授予創建數據庫

  這個例子授`CREATE_DATABASE`予用戶`datalake_user1`。由於授與此權限的資源是「資料目錄」，因此指令會將空白`CatalogResource`結構指定為`resource`參數。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_DATABASE" --resource '{ "Catalog": {}}'
```

Example — 授予在指定數據庫中創建表

下一個例子retail將數據庫授予CREATE_TABLE用戶datalake_user1。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::<account-id>:user/datalake_user1 --
permissions "CREATE_TABLE" --resource '{ "Database": {"Name":"retail"}}'
```

Example — 授予外部 AWS 帳戶與授予選項

下一個範例會授CREATE_TABLE與資料庫上的授權選項retail給外部帳戶 1111-2222-3333。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333 --permissions "CREATE_TABLE"
 --permissions-with-grant-option "CREATE_TABLE" --resource '{ "Database":
 {"Name":"retail"}}'
```

Example — 授予組織

下一個範例會授ALTER與資料庫issues上的 grant 選項給組織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 -授予**ALLIAMPrincipals**同一帳戶

下一個範例會retail將資料庫的CREATE_TABLE權限授與相同帳戶中的所有主體。此選項可讓帳戶中的每個主體在資料庫中建立資料表，並建立表格資源連結，讓整合式查詢引擎存取共用的資料庫和資料表。當主參與者收到跨帳戶授權且沒有建立資源連結的權限時，此選項特別有用。在此案例中，資料湖管理員可以建立預留位置資料庫並將CREATE_TABLE權限授與ALLIAMPrincipal群組，讓帳戶中的每個 IAM 主體都能在預留位置資料庫中建立資源連結。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"temp","CatalogId":"111122223333"}}'
```

Example -授予**ALLIAMPrincipals**外部帳戶

下一個範例會retail將資料庫授CREATE_TABLE與外部帳戶中的所有主體。此選項可讓帳戶中的每個主體在資料庫中建立資料表。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333:IAMPrincipals
 --permissions "CREATE_TABLE"  --resource '{ "Database":
 {"Name":"retail","CatalogId":"123456789012"}}'
```

> **Note**
>
> 授與CREATE_TABLE或ALTER對具有指向註冊位置之位置屬性的資料庫之後，請務必同時將該位置的資料位置權限授與主參與者。如需詳細資訊，請參閱 授與資料位置權限。

> **另請參閱**
>
> - Lake Formation 權限參考
> - 授與與您帳戶共用的資料庫或資料表的權限
> - 存取和檢視共用資料目錄表格和資料庫

使用指定的資源方法授與資料表權限

您可以使用 Lake Formation 主控台，或 AWS CLI 授與「資料目錄」表格上的「Lake Formation」權限。您可以授與個別資料表的權限，或透過單一授與作業授與資料庫中所有資料表的權限。

如果您授與資料庫中所有資料表的權限，就會隱含地授與資料庫的DESCRIBE權限。然後資料庫會顯示在主控台的 [資料庫] 頁面上，並由 GetDatabases API 作業傳回。

當您選擇SELECT授與權限時，您可以選擇套用欄篩選、列篩選或儲存格篩選。

Console

下列步驟說明如何使用具名的資源方法和 Lake Formation 主控台上的 [授與資料湖權限] 頁面來授與資料表權限。該頁面分為以下幾個部分：

- 主參與者 — 要授與權限的使用者、角色、 AWS 帳號、組織或組織單位。
- LF 標籤或目錄資源 — 要授與權限的資料庫、表格或資源連結。
- 權限-要授予的 Lake Formation 許可權。

> **Note**
>
> 若要授與表格資源連結的權限，請參閱授與資源連結權限。

1. 開啟 [授與資料湖權限] 頁面。

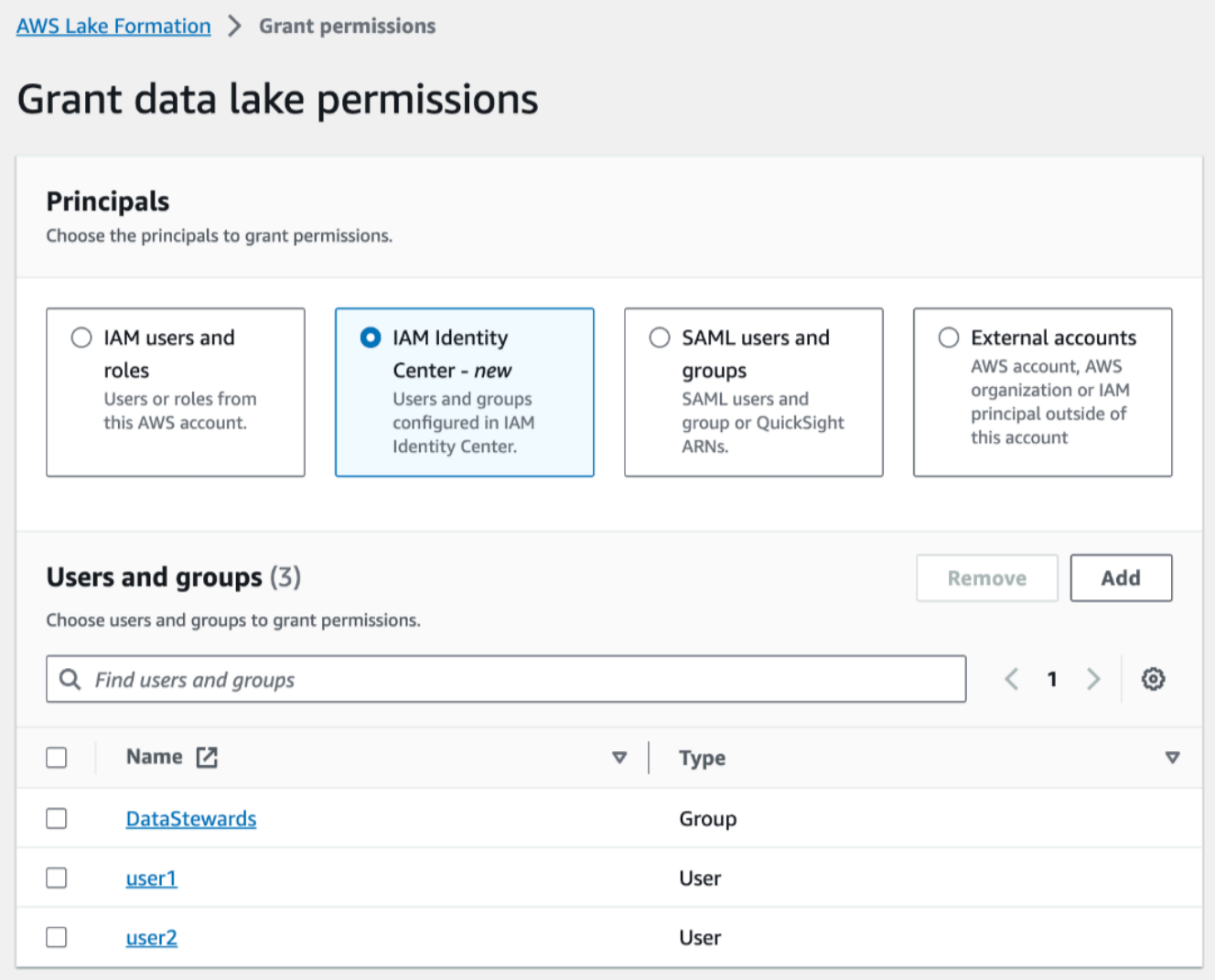
   在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以資料湖管理員、表格建立者或已獲授與具有授與選項之資料表權限的使用者身分登入。

   執行以下任意一項：

   - 在功能窗格中，選擇 [權限] 下的 [資料湖權限]。然後選擇授予。
   - 在導覽窗格中，選擇 Tables (資料表)。然後，在 [表格] 頁面上選擇一個表格，然後在 [動作] 功能表的 [權限] 下選擇 [授與]。

   > **Note**
   >
   > 您可以透過資源連結授與表格的權限。若要這樣做，請在「表格」頁面上選擇資源連結，然後在「動作」功能表上選擇「授與目標」。如需詳細資訊，請參閱 資源連結在 Lake Formation 中如何運作。

2. 接下來，在「主參與者」區段中，選擇主參與者類型並指定要授與權限的主參與者。



IAM 使用者和角色

從 IAM 使用者和角色清單中選擇一或多個使用者或角色。

IAM Identity Center

從 [使用者和群組] 清單中選擇一或多個使用者或群組。

SAML 使用者和群組

對於 SAML 和 Amazon QuickSight 使用者和群組，請為透過 SAML 聯合的使用者或群組輸入一或多個 Amazon 資源名稱 (ARN)，或針對 Amazon 使用者或群組輸入 ARN。QuickSight 在每個 ARN 之後按 Enter 鍵。

如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。

> **Note**
>
> 僅支持 Amazon QuickSight 企業版與 Amazon QuickSight 的 Lake Formation 整合。

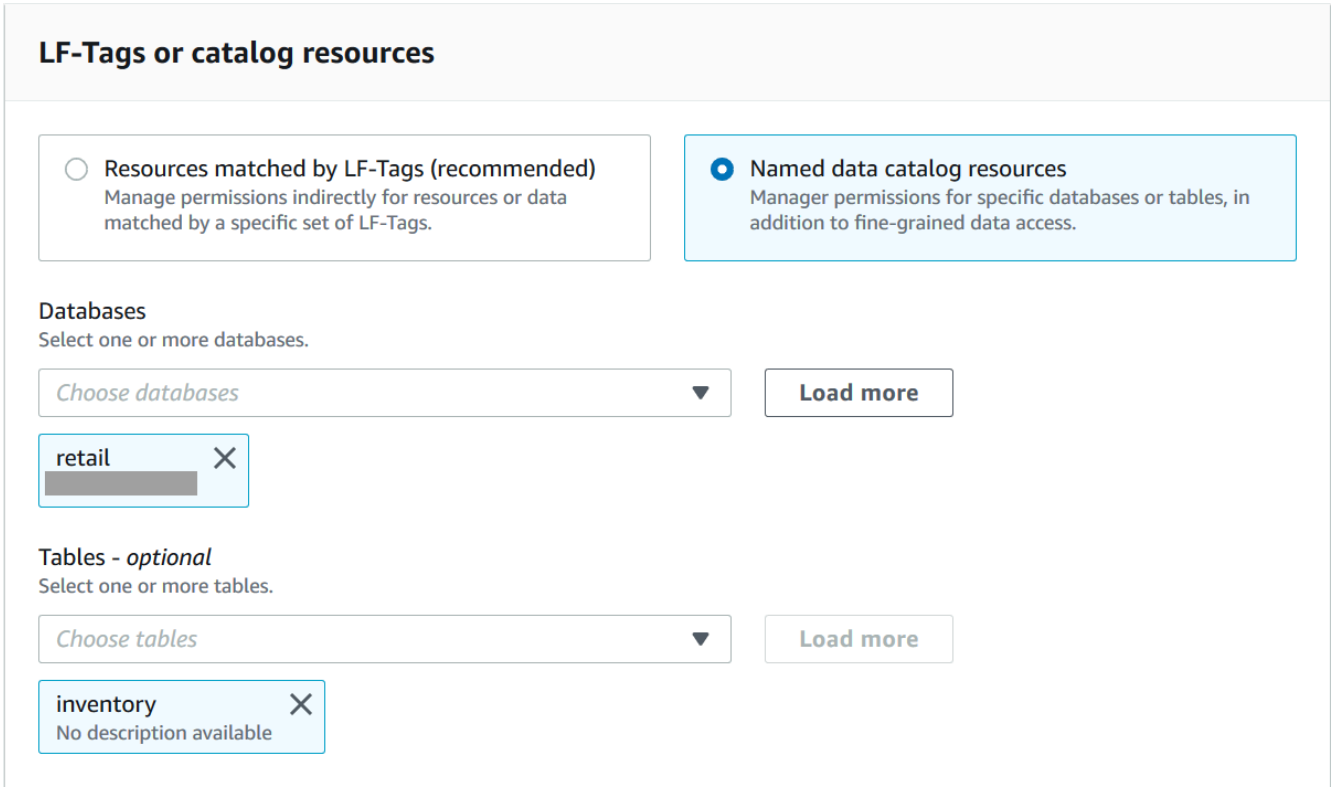
外部帳戶

針對AWS 帳戶 、 AWS 組織或 IAM 主體，輸入 IAM 使用者或角色的一或多個有效 AWS 帳戶 ID、組織 ID、組織單位 ID 或 ARN。在每個 ID 之後按 Enter 鍵。

組織 ID 由「o-」後跟 10-32 個小寫字母或數字組成。

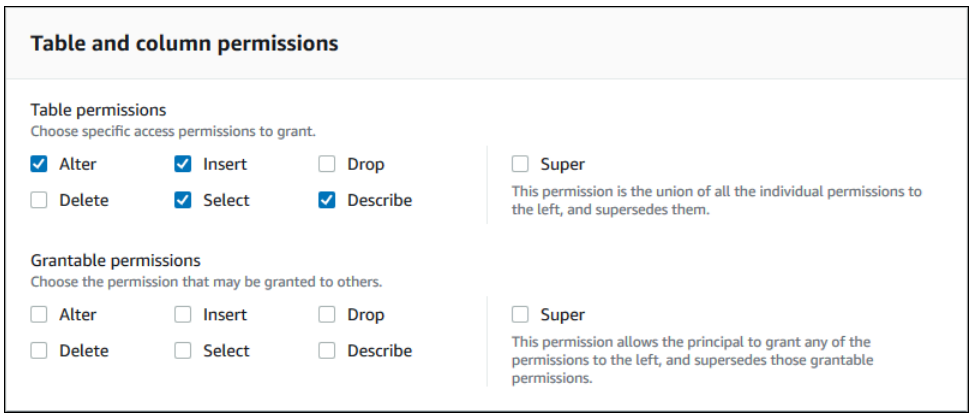
組織單位 ID 以「ou-」開頭，後面接著 4—32 個小寫字母或數字 (包含 OU 的根目錄識別碼)。該字符串後跟第二個「-」字符和 8 到 32 個其他小寫字母或數字。

3. 在 LF 標籤或目錄資源區段中，選擇資料庫。然後選擇一或多個表格，或選擇「所有表格」。

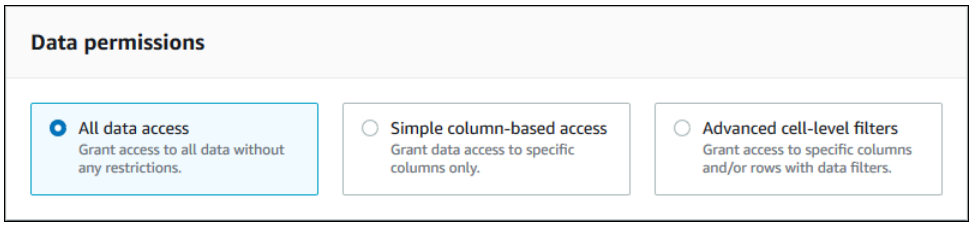


4. 指定沒有資料篩選的權限

在「權限」段落中，選取要授與的表格權限，並選擇性地選取可授與的權限。



如果您授與 [選取]，[資料] 權限區段會顯示在 [資料表和資料行權限] 區段下方，預設會選取 [所有資料存取] 選項。接受預設值。



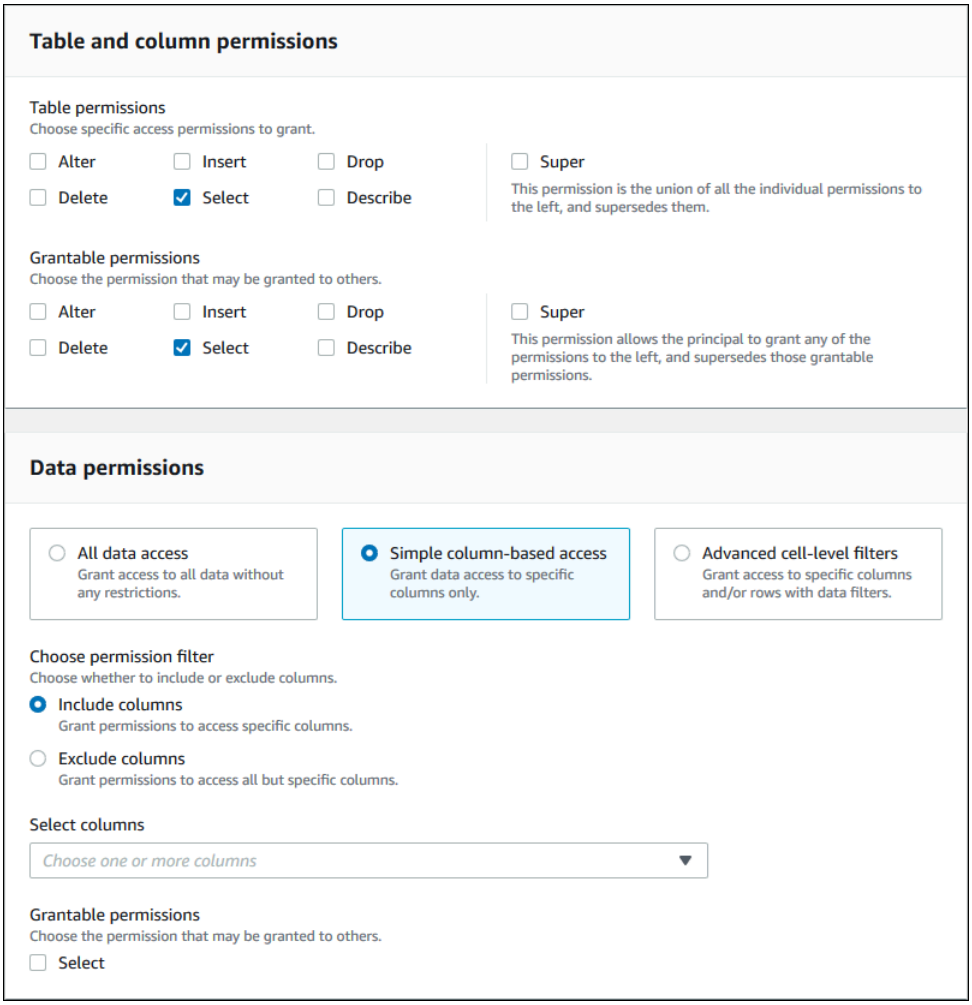
5. 選擇 Grant (授予)。
6. 使用資料篩選指定「選取」 權限

   選取 [選取] 權限。請勿選取任何其他權限。

   [資料權限] 區段會顯示在 [資料表和資料行權限] 區段下。
7. 執行以下任意一項：

   - 僅套用簡單欄篩選。

     1. 選擇「簡單基於列的訪問」。



2. 選擇是否要包含或排除資料欄，然後選擇要包含或排除的資料欄。

   授與外部 AWS 帳戶或組織權限時，僅支援包含清單。

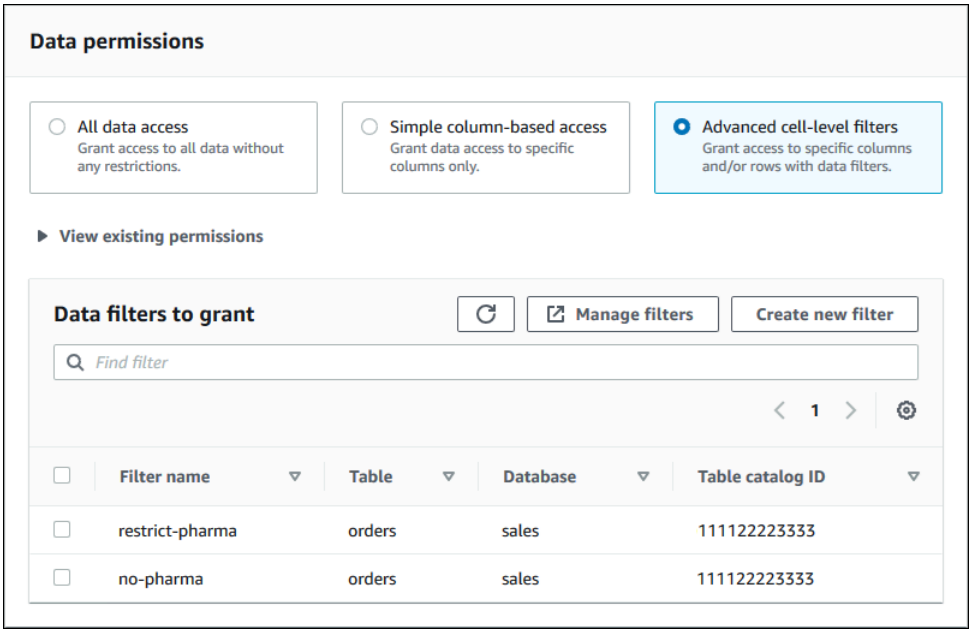
3. (選擇性) 在 [可授與的權限] 下，開啟 [選取] 權限的授與選項。

   如果您包含授與選項，授與收件者只能授與您授與他們的資料行的權限。

> **Note**
>
> 您也可以透過建立指定欄篩選並將所有列指定為資料列篩選的資料篩選來套用欄篩選。但是，這需要更多步驟。

- 套用欄、列或儲存格篩選。

  1. 選擇進階儲存格層級篩選器。



2. (選擇性) 展開「檢視現有權限」。
3. (選擇性) 選擇 [建立新篩選器]。
4. (選擇性) 若要檢視所列篩選器的詳細資訊，或建立新篩選或刪除現有篩選器，請選擇「管理篩選器」。

   [資料篩選] 頁面會在新的瀏覽器視窗中開啟。

   完成 [資料篩選] 頁面後，返回 [授與權限] 頁面，如有必要，請重新整理頁面以檢視您建立的任何新資料篩選器。
5. 選取要套用至授權的一或多個資料篩選器。

> **Note**
>
> 如果清單中沒有資料篩選，則表示沒有為選取的表格建立資料篩選。

8. 選擇 Grant (授予)。

AWS CLI

您可以使用指定的資源方法和 AWS Command Line Interface (AWS CLI) 授與資料表權限。

若要使用授與資料表權限 AWS CLI

- 運行`grant-permissions`命令，並指定一個表作為資源。

Example -授予單個表-沒有過濾

下列範例會授SELECT與使ALTERdatalake_user1用者 AWS 帳戶 1111-2222-3333 資料庫資料表上的使用者。inventory retail

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1 --
permissions "SELECT" "ALTER" --resource '{ "Table": {"DatabaseName":"retail",
 "Name":"inventory"}}'
```

> **Note**
>
> 如果您授與在ALTER已註冊位置中具有其基礎資料之表格的權限，請務必同時將該位置的資料位置權限授與主參與者。如需詳細資訊，請參閱 授與資料位置權限。

Example -使用「授予」選項對所有表進行授予-無過濾

下一個例子授予SELECT與在數據庫中的所有表的 grant 選項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 -授予簡單的列過濾

下一個範例會授與資SELECT料表中的資料行子集persons。它使用簡單的列過濾。

```
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 — 使用數據過濾器授予

此範例會授與SELECT資料orders表並套用資restrict-pharma料篩選器。

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

以下是文件的內容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"]
}
```

另請參閱

- Lake Formation 許可權概述
- Lake Formation 中的數據過濾和細胞級安全
- Lake Formation 角色和 IAM 許可參考
- 授與資源連結權限
- 存取和檢視共用資料目錄表格和資料庫

使用指定的資源方法授與檢視的權限

下列步驟說明如何使用具名資源方法和 [授與資料湖權限] 頁面來授與檢視的權限。該頁面分為以下幾個部分：

- 主體 — 要授與權限的 IAM 使用者、角色、IAM 身分中心使用者和群組 AWS 帳戶、組織或組織單位。
- LF 標籤或目錄資源 — 要授與權限的資料庫、表格、檢視或資源連結。
- 權限 — 要授與的資料湖權限。

開啟 [授與資料湖權限] 頁面

1. 在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以資料湖管理員、資料庫建立者或對資料庫具有可授予權限的 IAM 使用者身分登入。
2. 執行以下任意一項：

   - 在功能窗格的 [權限] 下，選擇 [資料湖權限]。然後選擇授予。
   - 在導覽窗格中，選擇 [資料目錄] 下的 [檢視]。然後，在 [檢視] 頁面上選擇檢視，然後從 [動作] 功能表的 [權限] 下選擇 [授與]。

> **Note**
>
> 您可以透過檢視的資源連結授與檢視的權限。若要這樣做，請在「檢視」頁面上選擇資源連結，然後在「動作」功能表上選擇「授與目標」。如需詳細資訊，請參閱 資源連結在 Lake Formation 中如何運作。

指定主參與者

在「主參與者」區段中，選擇主參與者類型，然後指定要授與權限的主參與者。

IAM 使用者和角色

從 IAM 使用者和角色清單中選擇一或多個使用者或角色。

IAM Identity Center

從 [使用者和群組] 清單中選擇一或多個使用者或群組。

SAML 使用者和群組

對於 SAML 和 Amazon QuickSight 使用者和群組，請為透過 SAML 聯合的使用者或群組輸入一或多個 Amazon 資源名稱 (ARN)，或針對 Amazon 使用者或群組輸入 ARN。 QuickSight 在每個 ARN 之後按 Enter 鍵。

如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。

> **Note**
>
> 僅支持 Amazon QuickSight 企業版與 Amazon QuickSight 的 Lake Formation 整合。

外部帳戶

針對AWS 帳戶、 AWS 組織或 IAM 主體，輸入 IAM 使用者或角色的一或多個有效 AWS 帳戶 ID、組織 ID、組織單位 ID 或 ARN。在每個 ID 之後按 Enter 鍵。

組織 ID 由「o-」後跟 10-32 個小寫字母或數字組成。

組織單位 ID 以「ou-」開頭，後面接著 4—32 個小寫字母或數字 (包含 OU 的根目錄識別碼)。該字符串後跟第二個「-」破折號和 8 到 32 個其他小寫字母或數字。

> 另請參閱
>
> - 存取和檢視共用資料目錄表格和資料庫

指定視圖

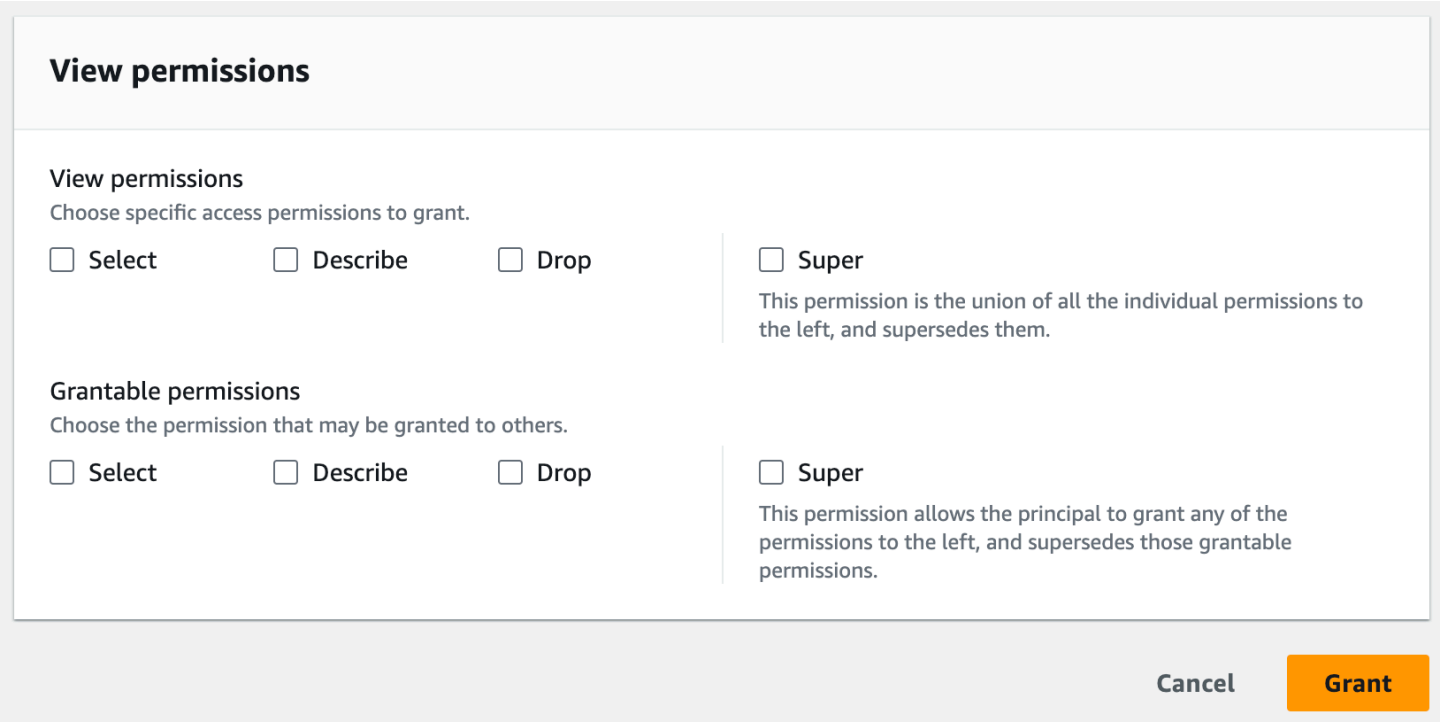
在 LF 標籤或目錄資源區段中，選擇要授與權限的一個或多個檢視。

1. 選擇具名資料目錄資源。
2. 從「檢視表」清單中選擇一或多個檢視。您也可以選擇一或多個「資料庫」、「表格」和/或「資料」篩選器。

   Grantng 資料湖權限在資料庫`All views`內，將導致受權者擁有資料庫中所有資料表和檢視的權限。

指定權限

在「權限」區段中，選取權限和可授予的權限。



1. 在「檢視權限」下，選取要授與的一或多個權限。
2. (選擇性) 在可授與權限下，選取授與收件者可以授與其中其他主體的權限。 AWS 帳戶當您從外部帳戶授與 IAM 主體許可時，不支援此選項。
3. 選擇 Grant (授予)。

另請參閱

- Lake Formation 權限參考
- 授與與您帳戶共用的資料庫或資料表的權限

基於 Lake Formation 標籤的訪問控制

基於 Lake Formation 標籤的訪問控制（LF-TBAC）是一種根據屬性定義權限的授權策略。在 Lake Formation，這些屬性被稱為 LF- 標籤。您可以將 LF 標籤附加至資料目錄資源，並授與使用這些 LF 標籤的這些資源上的 LLake Formation m 主體的權限。當主體的標籤值符合資源標籤值時，Lake Formation 允許對這些資源進行操作。LF-TBAC 在快速成長的環境中很有幫助，並有助於原則管理變得繁瑣的情況。

LF-TBAC 是當有大量資料目錄資源時，用來授予 Lake Formation 權限的建議方法。LF-TBAC 比指定的資源方法更具擴充性，而且需要較少的權限管理額外負荷。

> **Note**
>
> IAM 標籤與 LF 標籤不一樣。這些標籤不可互換。LF 標籤用於授予 Lake Formation 許可，IAM 標籤用於定義 IAM 政策。

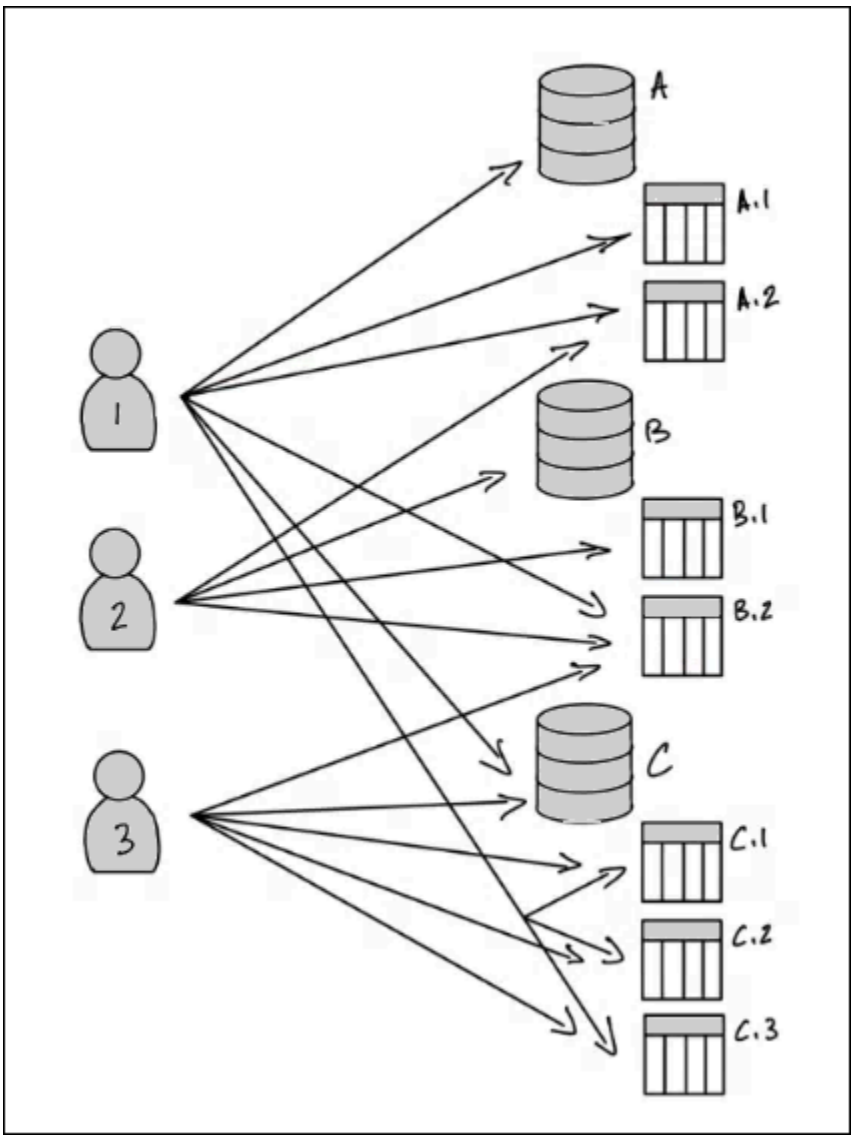
基於 Lake Formation 標籤的訪問控制如何工作

每個 LF 標籤都是一個鍵值對，例如或。`department=sales` `classification=restricted`一個鍵可以有多個定義的值，例如`department=sales,marketing,engineering,finance`。

若要使用 LF-TBAC 方法，資料湖管理員和資料工程師會執行下列工作。

任務	任務詳情
1. 定義 LF 標籤的性質和關係。	-
2. 在 Lake Formation 中創建 LF 標籤創建者。	添加 LF 標籤創建者
3. 在 Lake Formation 中創建 LF 標籤。	創建 LF-標籤
4. 將 LF 標籤指定給資料目錄資源。	將 LF 標籤指定給資料目錄資源
5. 授與權限給其他主體，以便將 LF 標籤指派給資源，選擇性地使用授與選項。	授予、撤銷和列出 LF 標籤值權限
6. 選擇性地使用授與選項將 LF 標籤運算式授與主參與者。	使用 LF-TBAC 方法授與資料湖權限
7. (建議) 確認主體可透過 LF-TBAC 方法存取正確的資源之後，請撤銷使用具名資源方法授與的權限。	-

考慮您必須將權限授與三個資料庫和七個資料表上的三個主體的情況。



若要透過使用指定的資源方法達到上述圖表中指示的權限，您必須進行 17 次授權，如下所示 (以偽程式碼)。

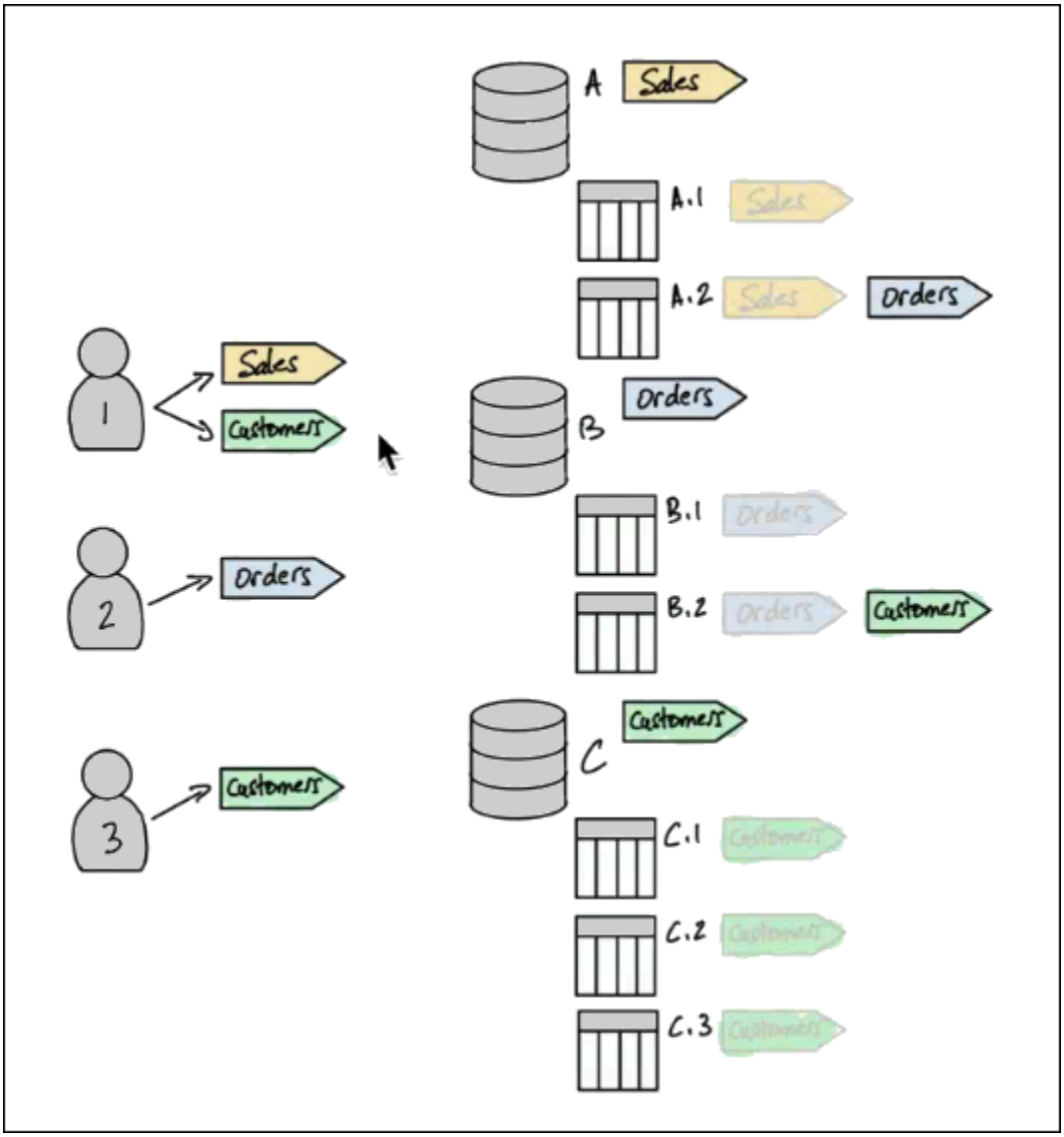
```
GRANT CREATE_TABLE ON Database A TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.1 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 1
GRANT SELECT, INSERT ON Table B.2 TO PRINCIPAL 1
...
GRANT SELECT, INSERT ON Table A.2 TO PRINCIPAL 2
GRANT CREATE_TABLE ON Database B TO PRINCIPAL 2
...
GRANT SELECT, INSERT ON Table C.3 TO PRINCIPAL 3
```

現在考慮如何使用 LF-TBAC 授予權限。下圖表指出您已將 LF 標籤指派給資料庫和資料表，並已將 LF 標籤的權限授與主體。

在此範例中，LF 標籤代表資料湖的區域，其中包含針對企業資源規劃 (ERP) 應用程式套件之不同模組的分析。您可以控制對各種模塊的分析數據的訪問。所有 LF 標籤都有鍵`module`和可能的值`SalesOrders`、和。`Customers`一個例子 LF 標籤看起來像這樣：

```
module=Sales
```

該圖僅顯示 LF 標籤值。



資料目錄資源和繼承的標籤指派

表從數據庫和列繼承 LF 標籤從表繼承 LF 標籤。繼承的值可以被覆蓋。在前面的圖表中，會繼承灰色的 LF 標籤。

由於繼承，資料湖管理員只需要對資源 (以虛擬程式碼為單位) 進行以下五個 LF 標籤指派。

```
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
```

標籤授予主體

將 LF 標籤指派給資料庫和資料表之後，資料湖管理員必須僅向主體授與四個 LF 標籤，如下所示 (以虛擬程式碼表示)。

```
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
```

現在，具有 LF 標籤的主體可以使用 `module=Sales` LF 標籤 (例如，資料庫 A) 存取資料目錄資源，具有 `module=Sales` LF 標籤的主體可以存取具有 `module=Customers` LF 標籤的資源，依`module=Customers`此類推。

前面的授予命令不完整。這是因為雖然它們透過 LF 標籤指示主體具有權限的資料目錄資源，但它們並不確切地指出主體對這些資源具有哪些 Lake Formation 權限 (例如 SELECTALTER)。因此，下列虛擬程式碼命令會更精確地表示如何透過 LF-tag 在資料目錄資源上授與 Lake Formation 權限。

```
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
```

將它放在一起-對資源產生的權限

指定給上圖中資料庫和資料表的 LF 標籤，以及授與圖表中主體的 LF 標籤，下表列出主參與者在資料庫和資料表上擁有的 Lake Formation 權限。

Principal	通過 LF 標籤授予的權限
主要項目 1	• CREATE_TABLE 在資料庫 A • SELECT, INSERT 在桌子 A.1 上 • SELECT, INSERT 在桌子 B.2 上 • CREATE_TABLE 在數據庫 C • SELECT, INSERT 在桌子 C.1 上 • SELECT，INSERT在桌子 C.2 上 • SELECT，INSERT在桌子 C.3 上
主要項目 2	• SELECT, INSERT 在桌子 A.2 上 • CREATE_TABLE 在資料庫 B 上 • SELECT, INSERT 在表 B.1 上 • SELECT, INSERT 在桌子 B.2 上
主要項目 3	• SELECT, INSERT 在桌子 B.2 上 • CREATE_TABLE 在數據庫 C • SELECT, INSERT 在桌子 C.1 上 • SELECT，INSERT在桌子 C.2 上 • SELECT，INSERT在桌子 C.3 上

底線

在這個簡單的範例中，使用五個指派作業和八個授與作業，資料湖管理員能夠指定 17 個權限。當有數十個資料庫和數百個資料表時，LF-TBAC 方法超過指定的資源方法的優點就會變得清楚。在需要授予每個資源的每個主體訪問權限的假設情況下，其中n(P)是主體的數量，n(R)是資源的數量：

- 使用指定的資源方法，所需的授權數量為 n(P) × n(R)。
- 使用 LF-TBAC 方法時，使用單一 LF 標籤，主參與者和資源指派的授權總數為 +。n(P) n(R)

另請參閱

- 管理 LF 標籤以進行中繼資料存取控制
- 使用 LF-TBAC 方法授與資料湖權限

主題

管理 LF 標籤以進行中繼資料存取控制

若要使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法來保護資料目錄資源 (資料庫、資料表和資料行)，請建立 LF 標籤、將它們指派給資源，並將 LF 標籤權限授與主體。

您必須先定義 LF 標籤，才能將 LF 標籤指派給資料目錄資源或授與權限給主體。只有資料湖管理員或具有 LF 標籤建立者權限的主參與者可以建立 LF 標籤。

LF 标签创建者

LF 標籤建立者是具有建立和管理 LF 標籤的權限的非系統管理員主體。資料湖管理員可以使用 Lake Formation 主控台或 CLI 新增 LF 標籤建立者。LF 標籤建立者具有隱含的 Lake Formation 權限，可更新和刪除 LF 標籤、將 LF 標籤指派給資源，以及授與 LF 標籤權限和 LF 標籤值權限給其他主體。

透過 LF 標籤建立者角色，資料湖管理員可以將標籤管理工作 (例如建立和更新標籤索引鍵和值) 委派給非管理員主參與者。資料湖管理員也可以授與 LF 標籤建立者可授與權`Create LF-Tag`限。然後，LF 標籤建立者可以將建立 LF 標籤的權限授與其他主體。

您可以在 LF 標籤上授予兩種類型的權限：

- LF 標籤權限-`Create LF-Tag` `Alter`、和。`Drop`建立、更新和刪除 LF 標籤需要這些權限。

  資料湖管理員和 LF 標籤建立者隱含地擁有他們建立的 LF 標籤的這些權限，並且可以明確地將這些權限授與主體，以管理資料湖中的標籤。
- LF-標記鍵值對權限-`Assign`、`Describe`和。`Grant with LF-Tag expressions`若要將 LF 標籤指派給資料目錄資料庫、資料表和欄，以及使用以 Lake Formation 標籤為基礎的存取控制將資源的權限授與主體，需要這些權限。LF 標籤創建者在創建 LF 標籤時隱式接收這些權限。

在收到Create LF-Tag權限並成功建立 LF 標籤之後，LF 標籤建立者可以將 LF 標籤指派給資源，並將 LF 標籤權限 (Create LF-Tag、、和) 授與其他非系統管理原則 AlterDrop，以管理資料湖中的標籤。您可以使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ ）AWS CLI來管理 LF 標籤。

> **Note**
>
> 資料湖管理員具有隱含的 Lake Formation 權限，可建立、更新和刪除 LF 標籤、將 LF 標籤指派給資源，以及授與 LF 標籤權限給主體。

如需最佳做法和考量事項，請參閱 基於 Lake Formation 標籤的訪問控制最佳實踐和考量

主題

- 添加 LF 標籤創建者
- 創建 LF-標籤
- 更新 LF-標籤
- 刪除 LF 標籤
- 列出 LF-標籤
- 將 LF 標籤指定給資料目錄資源
- 檢視指定給資源的 LF 標籤
- 檢視 LF 標籤指定給的資源
- LF 標籤的生命週期
- 基於 Lake Formation 標籤的訪問控制與基於 IAM 屬性的訪問控制的比較

> **另請參閱**
>
> - 授予、撤銷和列出 LF 標籤值權限
> - 使用 LF-TBAC 方法授與資料湖權限
> - 基於 Lake Formation 標籤的訪問控制

添加 LF 標籤創建者

依預設，資料湖管理員可以建立、更新和刪除 LF 標籤、將標籤指派給資料目錄資源，以及將標籤權限授與主體。如果您想要將標籤建立和管理作業委派給非管理員主參與者，資料湖管理員可以建立 LF 標籤建立者角色，並將 Lake Formation `Create LF-Tag` 權限授與角色。透過可`Create LF-Tag`授予的權限，LF 標籤建立者可以將標籤建立和維護任務委派給其他非管理主參與者。

> **Note**
>
> 跨帳戶權限授予只能包含`Describe`和`Associate`權限。您無法授與`Create LF-Tag`、`Drop`、`Alter`、和`Grant with LFTag expressions`權限給不同帳戶中的主體。

主題

- 建立 LF 標籤所需的 IAM 許可
- 新增 LF 標籤建立者

> **另請參閱**
>
> - 授予、撤銷和列出 LF 標籤值權限
> - 使用 LF-TBAC 方法授與資料湖權限
> - 基於 Lake Formation 標籤的訪問控制

建立 LF 標籤所需的 IAM 許可

您必須設定權限，以允許 Lake Formation 主體建立 LF 標籤。將下列陳述式新增至需要為 LF 標籤建立者之主參與者的權限原則。

> **Note**
>
> 雖然資料湖管理員具有建立、更新和刪除 LF 標籤、將 LF 標籤指派給資源，以及授與 LF 標籤給主體的隱含 LF-標籤，但資料湖管理員還需要下列 IAM 許可。

如需詳細資訊，請參閱 Lake Formation 角色和 IAM 許可參考。

```
{
"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"
     ]
 }
```

將 LF 標籤指派給資源並將 LF 標籤授與主參與者的主參與者必須具有相同的權限，但、和權限除外。`CreateLFTag` `UpdateLFTag` `DeleteLFTag`

新增 LF 標籤建立者

LF 標籤建立者可以使用 LF-TBAC 方法建立 LF 標籤、更新標籤索引鍵和值、刪除標籤、將標籤與資料目錄資源相關聯，以及將「資料目錄」資源的權限授與主體。LF 標籤建立者也可以將這些權限授與主參與者。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 來建立 LF 標籤建立者角色。

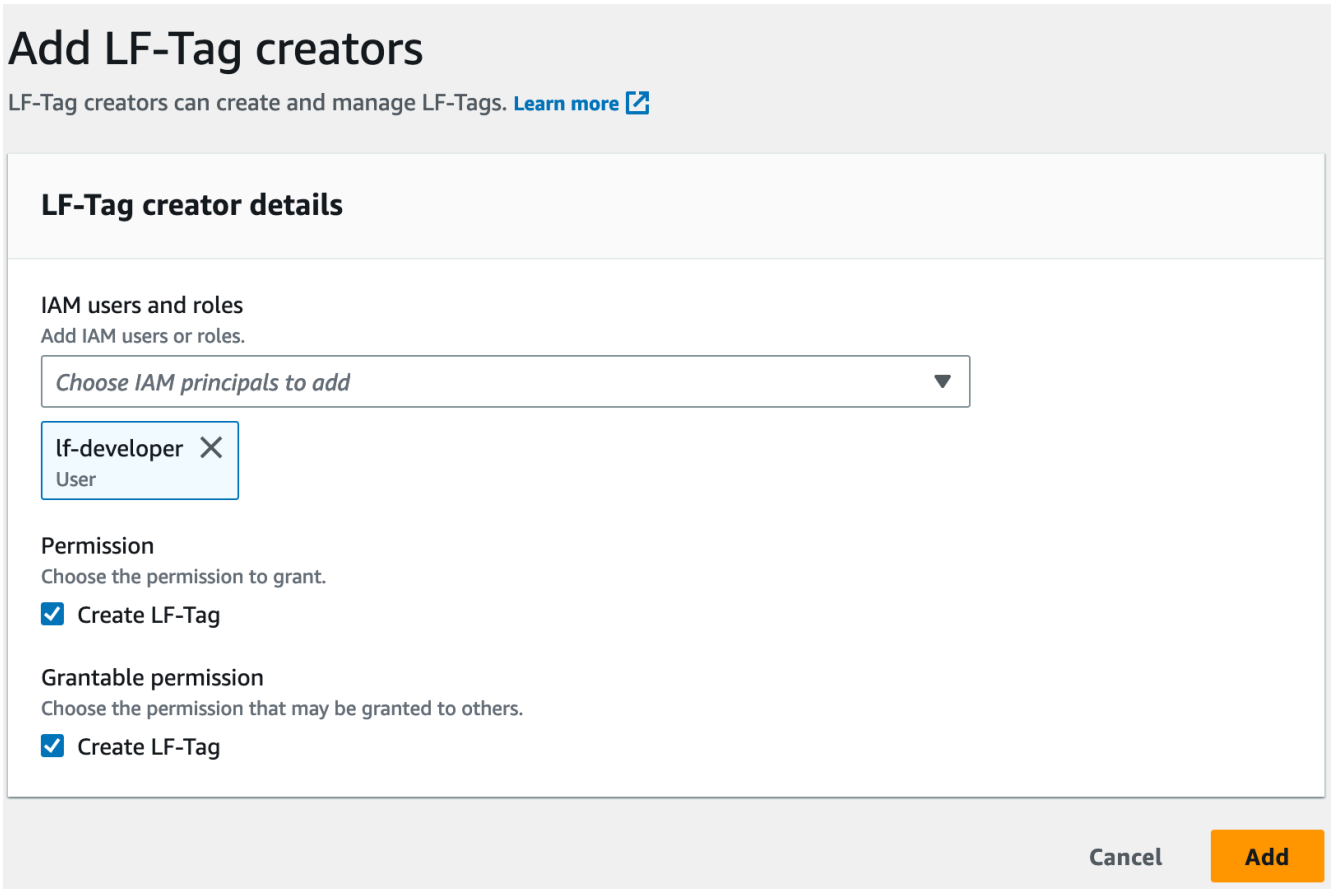
console

若要新增 LF 標籤建立者

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料管理員身分登入。
2. 在功能窗格的 [權限] 下，選擇 [LF 標籤和權限]。

   在 LF 標籤和權限頁面上，選擇 LF 標籤創建者部分，然後選擇添加 LF 標籤創建者。



3. 在 [新增 LF 標籤建立者] 頁面上，選擇具有建立 LF 標籤所需權限的 IAM 角色或使用者。
4. 「啟用Create LF-Tag權限」核取方塊。
5. (選擇性) 若要讓選取的主參與者授與Create LF-Tag權限給主參與者，請選擇可授與權Create LF-Tag限。
6. 選擇新增。

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"
    ]
}
```

以下是 LF 標籤建立者角色的可用權限：

權限	描述
Drop	對 LF 標籤具有此權限的主體可以從資料湖中刪除 LF 標籤。主體會取得 LF 標籤資源之所有標籤值的隱含Describe權限。
Alter	對 LF 標籤具有此權限的主體可以從 LF 標籤新增或移除標籤值。主體會取得 LF 標籤所有標籤值的隱含Alter權限。
Describe	在 LF 標籤上具有此權限的主體可以在將 LF 標籤指派給資源或授與 LF 標籤的權限時，檢視 LF 標籤及其值。您可以授予Describe所有鍵值或特定值。
Associate	對 LF 標籤具有此權限的主參與者可以將 LF 標籤指派給資料目錄資源。授予Associate 隱含授予。Describe
Grant with LF-Tag expression	對 LF 標籤具有此權限的主體可以使用 LF 標籤鍵和值授與資料目錄資源的權限。授予Grant with LF-Tag expression 隱含授予。Describe

這些權限是可授予的。已透過授與選項授與這些權限的主參與者可以將其授與其他主參與者。

創建 LF-標籤

所有 LF 標籤必須在 Lake Formation 中定義才能使用。LF 標籤由一個鍵和鍵的一個或多個可能的值。

在資料湖管理員為 LF 標籤建立者角色設定必要的 IAM 許可和 Lake Formation 權限之後，主體就可以建立 LF 標籤。LF 標籤創建者獲得隱含權限來更新或刪除 LF 標籤中的任何標籤值，並刪除 LF 標籤。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface ()AWS CLI來建立 LF 標籤。

Console

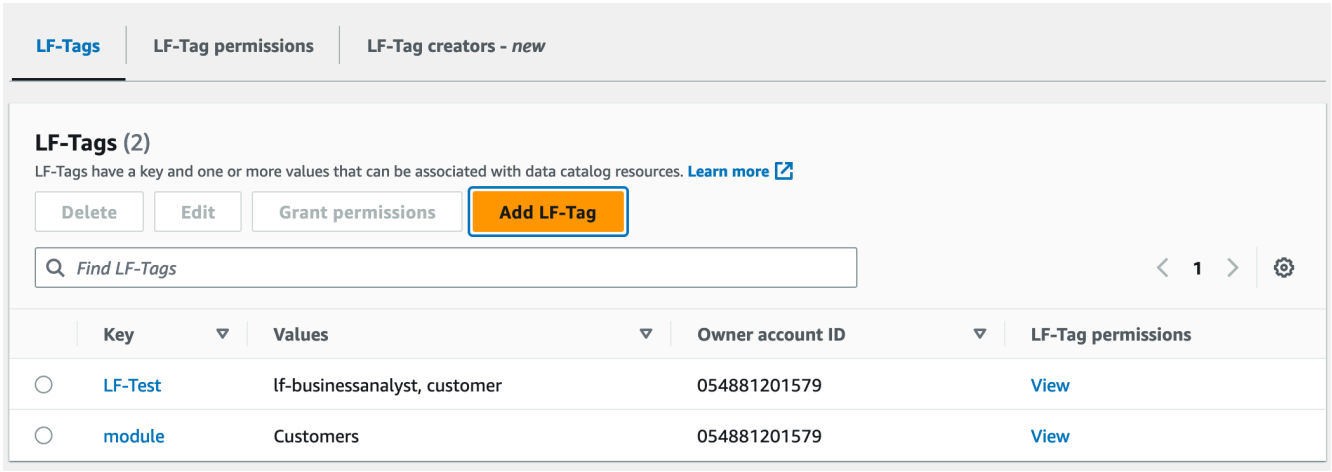
建立 LF 標籤的步驟

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以具有 LF 標籤建立者權限的主參與者身分登入，或以資料湖管理員身分登入。

2. 在功能窗格的 LF 標籤和權限下，選擇 LF 標籤。

   便會顯示「LF-標籤」頁面。



3. 選擇「新增 LF 標籤」。
4. 在「加入 LF 標籤」對話方塊中，輸入鍵和一個或多個值。

   每個鍵必須至少有一個值。若要輸入多個值，請輸入以逗號分隔的清單，然後按 Enter，或一次輸入一個值，然後在每個值之後選擇「新增」。允許的最大值數目為 1000。

5. 選擇 Add tag (新增標籤)。

AWS CLI

建立 LF 標籤的步驟

- 輸入`create-lf-tag`指令。

  下面的例子創建一個 LF 標籤與鍵`module`和值`Customers。Orders`

```
aws lakeformation create-lf-tag --tag-key module --tag-values Customers Orders
```

身為標籤建立者，主參與者會取得此 LF 標籤的Alter權限，並且可以從此 LF 標籤更新或移除任何標籤值。LF 標籤建立者主參與者也可以將Alter權限授與另一個主參與者，以更新及移除此 LF 標籤上的標籤值。

更新 LF-標籤

您可以透過新增或刪除允許的索引鍵值來更新您擁有Alter權限的 LF 標籤。您無法變更 LF 標籤金鑰。若要變更金鑰，請刪除 LF 標籤，然後使用所需金鑰新增一個標籤。除了Alter權限之外，您還需要 lakeformation:UpdateLFTag IAM 許可才能更新值。

刪除 LF 標籤值時，不會檢查任何資料目錄資源上是否存在該 LF 標籤值。如果刪除的 LF 標籤值與資源相關聯，則該資源將不再顯示該值，並且在該鍵值對上授予權限的任何主體都不再具有權限。

刪除 LF 標籤值之前，您可以選擇性地使用命remove-lf-tags-from-resource令，從具有要刪除值的資料目錄資源中移除 LF 標籤，然後使用您要保留的值重新標記資源。

只有資料湖管理員、LF 標籤建立者以及具有 LF 標籤Alter權限的主參與者可以更新 LF 標籤。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface ()AWS CLI來更新 LF 標籤。

Console

要更新 LF 標籤（控制台）

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、LF 標籤建立者或具有 LF 標籤Alter權限的主參與者身分登入。
2. 在功能窗格的 LF 標籤和權限下，選擇 LF 標籤。
3. 在 LF 標籤頁面上，選取 LF 標籤，然後選擇 [編輯]。
4. 在「編輯 LF 標籤」對話方塊中，加入或移除 LF 標籤值。

   若要加入多個值，請在「值」欄位中輸入以逗號分隔的清單並按 Enter，或一次輸入一個值，或在每個值之後選擇「加入」。
5. 選擇儲存。

AWS CLI

若要更新 LF 標籤 ()AWS CLI

- 輸入update-lf-tag指令。提供下列其中一個或兩個引數：
  - --tag-values-to-add
  - --tag-values-to-delete

Example

下列範例會以 LF-tag 鍵vice-president的值取代該值。vp level

```
aws lakeformation update-lf-tag --tag-key level --tag-values-to-add vice-president
--tag-values-to-delete vp
```

刪除 LF 標籤

您可以刪除不再使用的 LF 標籤。不會檢查資料目錄資源上是否存在 LF 標籤。如果已刪除的 LF 標籤與資源相關聯，則該資源將不再可見，且授與該 LF 標籤權限的任何主參與者都不再具有權限。

刪除 LF 標籤之前，您可以選擇性地使用remove-lf-tags-from-resource指令從所有資源中移除 LF 標籤。

只有資料湖管理員、LF 標籤建立者或具有 LF 標籤Drop權限的原則才能刪除 LF 標籤。除了Drop權限之外，主體還需要 lakeformation:DeleteLFTag IAM 許可才能刪除 LF 標籤。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface ()AWS CLI刪除 LF 標籤。

Console

若要刪除 LF 標籤 (控制台)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員身分登入。
2. 在功能窗格的 LF 標籤和權限下，選擇 LF 標籤。
3. 在 LF 標籤頁面上，選取 LF 標籤，然後選擇 [刪除]。

4. 在刪除標籤環境中？」對話方塊中，若要確認刪除，請在指定的欄位中輸入 LF-標籤鍵值，然後選擇「刪除」。

AWS CLI

若要刪除 LF 標籤 ()AWS CLI

- 輸入`delete-lf-tag`指令。提供要刪除的 LF 標籤的金鑰。

  Example

  下面的例子刪除帶有密鑰的 LF 標籤。`region`

```
aws lakeformation delete-lf-tag --tag-key region
```

列出 LF-標籤

您可以列出您具有`Describe`或`Associate`權限的 LF 標籤。與每個 LF 標籤鍵一起列出的值是您擁有權限的值。

LF-標籤創建者具有隱式權限，以查看他們創建的 LF 標籤。

資料湖系統管理員可以看到本機 AWS 帳戶中定義的所有 LF 標籤，以及已從外部帳戶授`Describe`與`Associate`權限給本機帳戶的所有 LF 標籤。資料湖管理員可以看到所有 LF 標籤的所有值。

您可以使用 AWS Lake Formation 控制台，API 或 AWS Command Line Interface （ ）AWS CLI列出 LF 標籤。

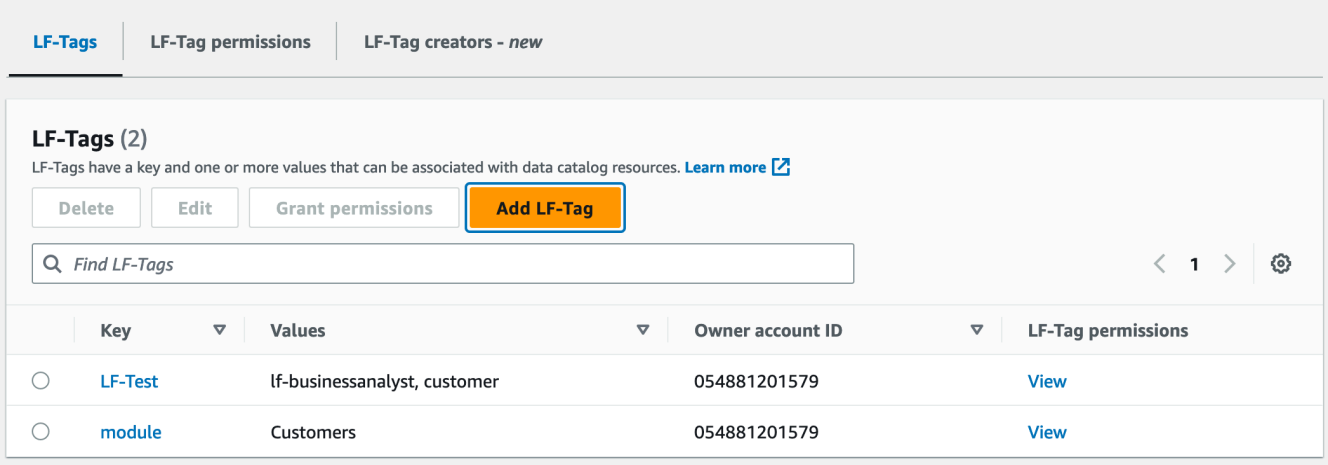
Console

列出 LF 標籤（控制台）

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以 LF 標籤建立者身分、以資料湖管理員身分登入，或以已獲授與 LF 標籤許可且具有 IAM 權限的主體身分登入。`lakeformation:ListLFTags`
2. 在功能窗格的 LF 標籤和權限下，選擇 LF 標籤。

   便會顯示「LF-標籤」頁面。



檢查所有者帳戶 ID 列，以確定從外部帳戶與您的帳戶共享的 LF 標籤。

AWS CLI

列出 LF-標籤（ ）AWS CLI

- 以資料湖管理員身分執行下列命令，或以已獲授與 LF 標籤權限且具有 `lakeformation:ListLFTags` IAM 權限的主體身分執行下列命令。

```
aws lakeformation list-lf-tags
```

輸出類似如下。

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "Orders",
```

```
                "Sales",
                "Customers"
            ]
        }
    ]
}
```

若要查看從外部帳戶授與的 LF 標籤，請包含命令選項。--resource-share-type ALL

```
aws lakeformation list-lf-tags --resource-share-type ALL
```

輸出類似如下。請注意索NextToken引鍵，這表示還有更多要列出的項目。

```
{
    "LFTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "level",
            "TagValues": [
                "director",
                "vp",
                "c-level"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
            "TagValues": [
                "Orders",
                "Sales",
                "Customers"
            ]
        }
    ],
    "NextToken": "eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ=="
}
```

重複此命令，然後新增--next-token引數以檢視從外部帳戶授與的任何剩餘本機 LF 標籤和 LF 標籤。來自外部帳戶的 LF 標籤始終位於單獨的頁面上。

```
aws lakeformation list-lf-tags --resource-share-type ALL
```

```
--next-token eyJleHBpcmF0aW...ZXh0Ijp0cnVlfQ==
```

```
{
    "LFTags": [
        {
            "CatalogId": "123456789012",
            "TagKey": "region",
            "TagValues": [
                "central",
                "south"
            ]
        }
    ]
}
```

API

您可以使用可供 Lake Formation 的 SDK 列出請求者有權查看的標籤。

```
import boto3

client = boto3.client('lakeformation')
...

response = client.list_lf_tags(
    CatalogId='string',
    ResourceShareType='ALL',
    MaxResults=50'
)
```

此命令會傳回具有下列結構的dict物件：

```
{
    'LFTags': [
        {
            'CatalogId': 'string',
            'TagKey': 'string',
            'TagValues': [
                'string',
            ]
        },
```

```
    ],
    'NextToken': 'string'
}
```

如需所需許可的詳細資訊，請參閱Lake Formation 角色和 IAM 許可參考。

將 LF 標籤指定給資料目錄資源

您可以將 LF 標籤指派給資料目錄資源 (資料庫、表格和欄)，以控制對這些資源的存取。只有被授與相符 LF 標籤的主參與者 (以及使用具名資源方法授與存取權的主參與者) 才能存取資源。

如果資料表從資料庫繼承 LF 標籤，或是資料行從資料表繼承 LF 標籤，您可以將新值指派給 LF-tag 索引鍵，以覆寫繼承的值。

您可以指派給資源的 LF 標籤數目上限為 50。

主題

- 管理指派給資源之標籤的需求
- 將 LF 標籤分配給表列
- 將 LF 標籤指定給資料目錄資源
- 更新資源的 LF 標籤
- 從資源中刪除 LF 標籤

管理指派給資源之標籤的需求

若要將 LF 標籤指定給資料目錄資源，您必須：

- 獲得 LF 標籤上的 Lake Formation ASSOCIATE 許可。
- 擁有 IAM `lakeformation:AddLFTagsToResource` 許可。
- 擁有膠水：Glue 資料庫的GetDatabase 權限。
- 身為資源擁有者 (建立者)、具有該GRANT選項的資源擁有 Super Lake Formation 權限，或具有下列GRANT選項的權限：
  - 對於同一 AWS 帳戶中的資料庫：DESCRIBECREATE_TABLE、ALTER、和 DROP
  - 對於外部帳戶中的資料庫：DESCRIBE、CREATE_TABLE和 ALTER
  - 對於表格 (和欄)：DESCRIBEALTER、DROP、INSERT、SELECT、和 DELETE

此外，LF 標籤及其指定目標的資源必須位於相同 AWS 的帳號中。

若要從資料目錄資源移除 LF 標籤，您必須符合這些要求，並且還具有 `lakeformation:RemoveLFTagsFromResource` IAM 權限。

將 LF 標籤分配給表列

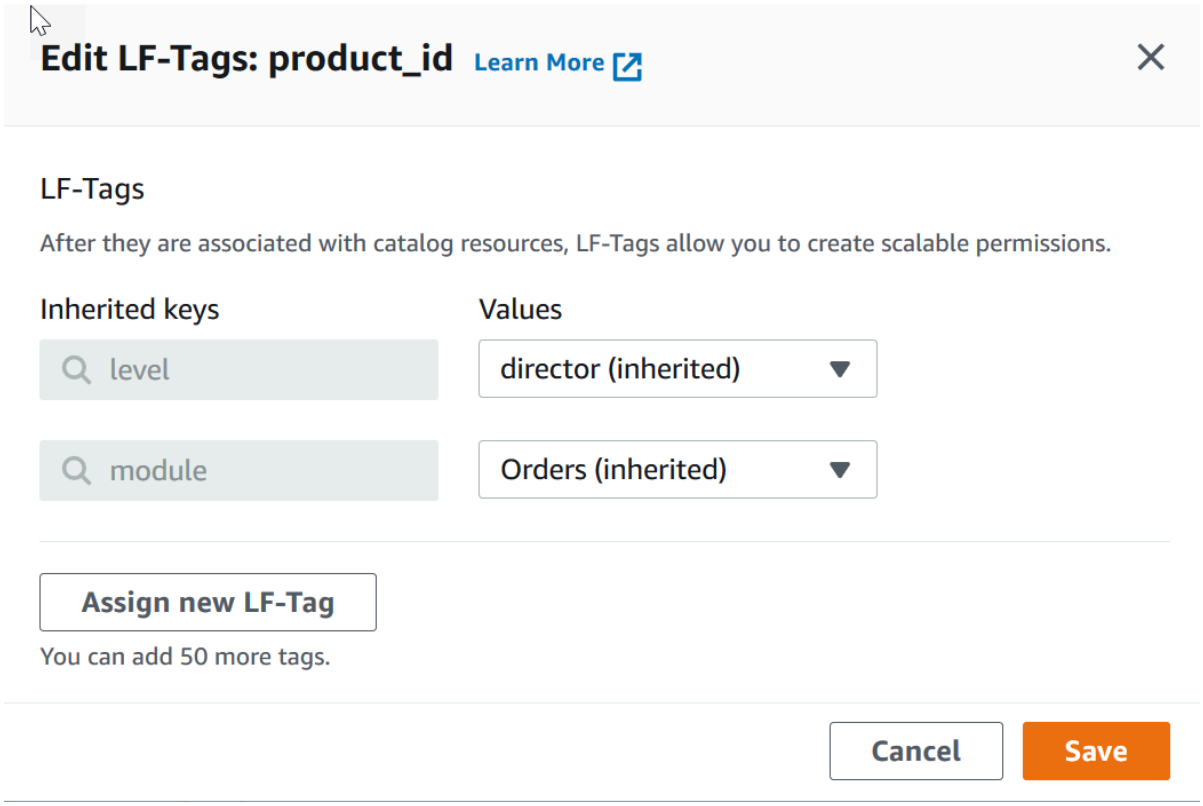
若要將 LF 標籤指派給資料表資料行 (主控台)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

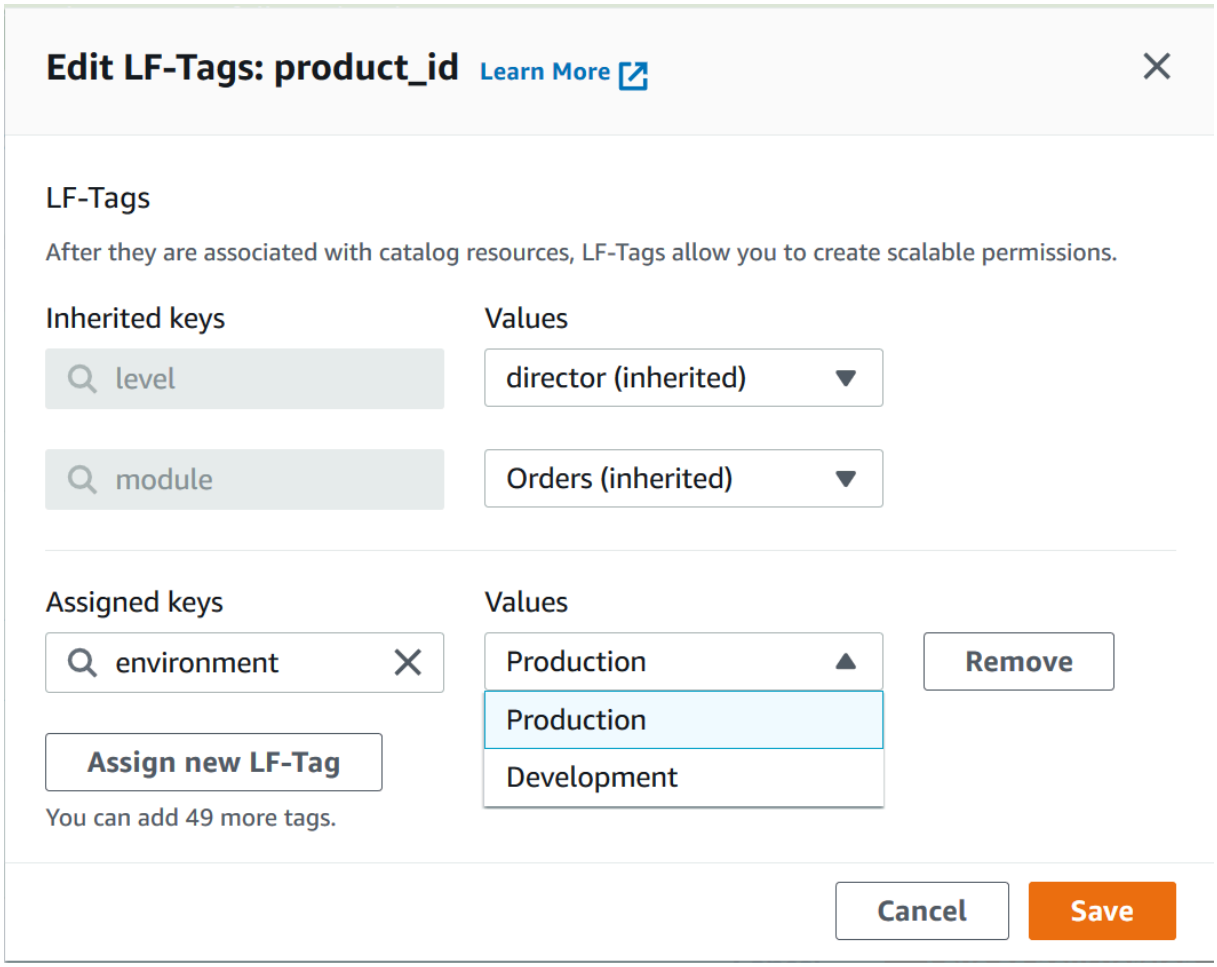
   以符合上述需求的使用者身分登入。
2. 在導覽窗格中，選擇 Tables (資料表)。
3. 選擇表格名稱 (而非表格名稱旁的選項按鈕)。
4. 在表格詳細資訊頁面的「綱要」段落中，選擇編輯綱要。
5. 在 [編輯結構描述] 頁面上，選取一或多個資料欄，然後選擇 [編輯標記]。

   > **Note**
   >
   > 如果您打算添加或刪除列並保存新版本，請先執行此操作。然後編輯 LF 標籤。

   「編輯 LF 標籤」對話方塊隨即出現，並顯示繼承自表格的任何 LF 標籤。



6. (選擇性) 對於「繼承的索引鍵」欄位旁的「值」清單，請選擇一個值來覆寫繼承的值。
7. (選擇性) 選擇「指派新 LF 標籤」。然後針對 [指派的索引鍵] 選擇一個機碼，並針對 [值] 選擇金鑰的值。



8. (選擇性) 再次選擇「指派新 LF 標籤」以新增另一個 LF 標籤。
9. 選擇儲存。

將 LF 標籤指定給資料目錄資源

Console

將 LF 標籤指定給資料目錄資料庫或表格的步驟

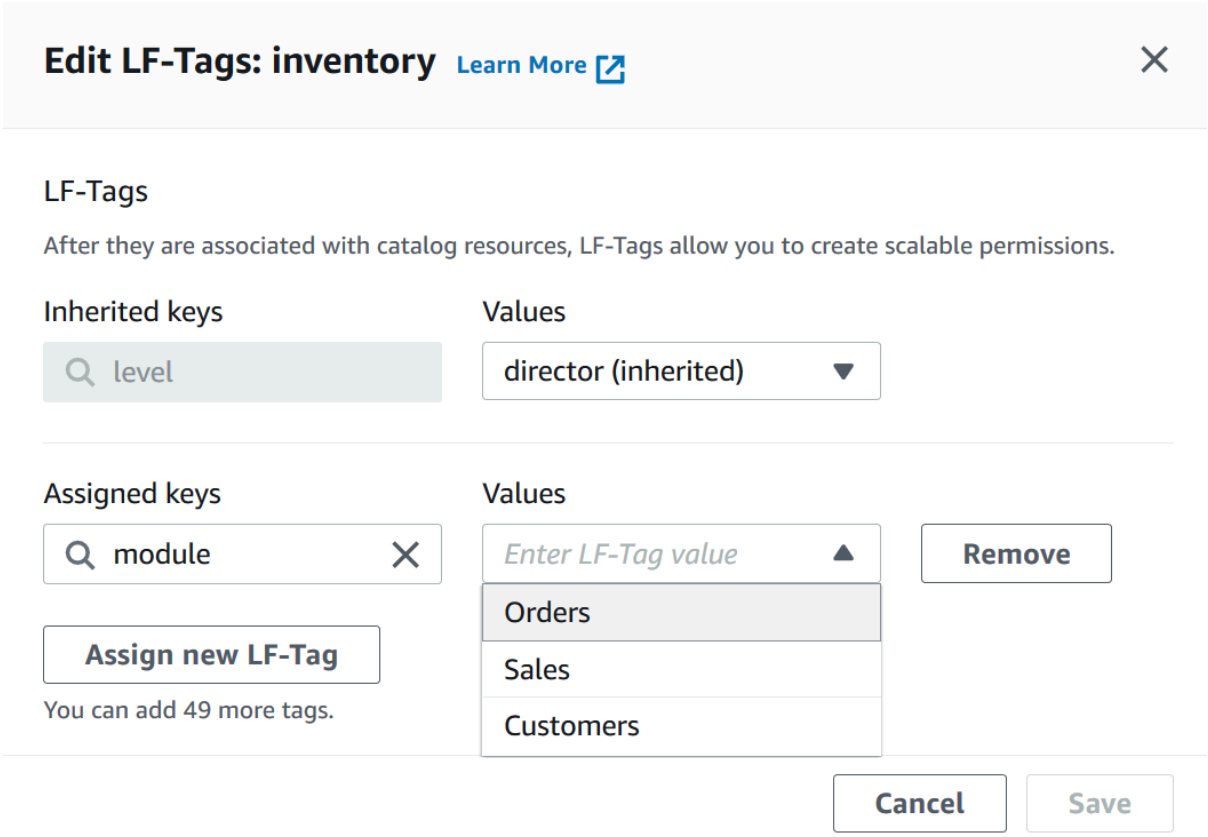
1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以符合先前所列需求的使用者身分登入。
2. 在導覽窗格的 [資料目錄] 下，執行下列其中一個動作：
   - 若要指派 LF 標籤給資料庫，請選擇 [資料庫]。
   - 若要將 LF 標籤指派給表格，請選擇「表格」。

3. 選擇資料庫或表格，然後在 [動作] 功能表上選擇 [編輯標記]。

   這時系統顯示 「編輯 LF-標籤：####」對話框。

   如果資料表從其包含的資料庫繼承 LF 標籤，視窗會顯示繼承的 LF 標籤。否則，它會顯示文字「沒有與資源相關聯的繼承 LF 標籤」。



4. (選擇性) 如果表格已繼承 LF 標籤，您可以針對「繼承的索引鍵」欄位旁的「值」清單選擇一個值來覆寫繼承的值。
5. 若要指派新的 LF 標籤，請執行下列步驟：

   a. 選擇「指派新 LF 標籤」。

   b. 在「指定的金鑰」欄位中，選擇 LF 標籤鍵，然後在「值」欄位中選擇一個值。

   c. (選擇性) 再次選擇「指派新 LF 標籤」以指定額外的 LF 標籤。
6. 選擇儲存。

AWS CLI

若要將 LF 標籤指定給資料目錄資源

- 執行 add-lf-tags-to-resource 命令。

  下列範例會將 LF 標籤module=orders指派給資料庫orders中的資料表。erp它使用引數的快捷語--lf-tags法。的CatalogID屬性--lf-tags是選擇性的。如果未提供，則會假設資源的目錄 ID (在本例中為表格)。

```
aws lakeformation add-lf-tags-to-resource --resource '{ "Table":
 {"DatabaseName":"erp", "Name":"orders"}}' --lf-tags
 CatalogId=111122223333,TagKey=module,TagValues=orders
```

  如果命令成功，則輸出如下。

```
{
    "Failures": []
}
```

  下一個範例會將兩個 LF 標籤指派給sales資料表，並使用引數的 JSON 語法。--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"]}]'
```

  下一個範例會將 LF 標籤指定level=director給total資料表的資料行。sales

```
aws lakeformation add-lf-tags-to-resource --resource '{ "TableWithColumns":
 {"DatabaseName":"erp", "Name":"sales", "ColumnNames":["total"]}}' --lf-tags
 TagKey=level,TagValues=director
```

更新資源的 LF 標籤

更新資料目錄資源的 LF 標籤的步驟 ()AWS CLI

- 使用add-lf-tags-to-resource指令，如先前程序所述。

使用與現有 LF 標籤相同的索引鍵新增 LF 標籤，但使用不同的值會更新現有值。

從資源中刪除 LF 標籤

若要移除資料目錄資源的 LF 標籤 ()AWS CLI

- 執行 `remove-lf-tags-from-resource` 命令。

  如果資料表具有覆寫繼承自父項資料庫的值的 LF 標籤值，則從資料表中移除該 LF 標籤會還原繼承的值。此行為也適用於覆寫繼承自資料表之索引鍵值的資料行。

  下`total`列範例會從資料表的資料行移`level=director`除 LF 標籤。`sales`的`CatalogID`屬性`--lf-tags`是選擇性的。如果未提供，則會假設資源的目錄 ID (在本例中為表格)。

```
aws lakeformation remove-lf-tags-from-resource
--resource ' { "TableWithColumns":
{ "DatabaseName": "erp",  "Name": "sales",  "ColumnNames":[ "total"]}}'
--lf-tags  CatalogId=111122223333,TagKey=level,TagValues=director
```

檢視指定給資源的 LF 標籤

您可以檢視指定給資料目錄資源的 LF 標籤。您必須擁有 LF 標籤的`DESCRIBE`或`ASSOCIATE`權限才能檢視它。

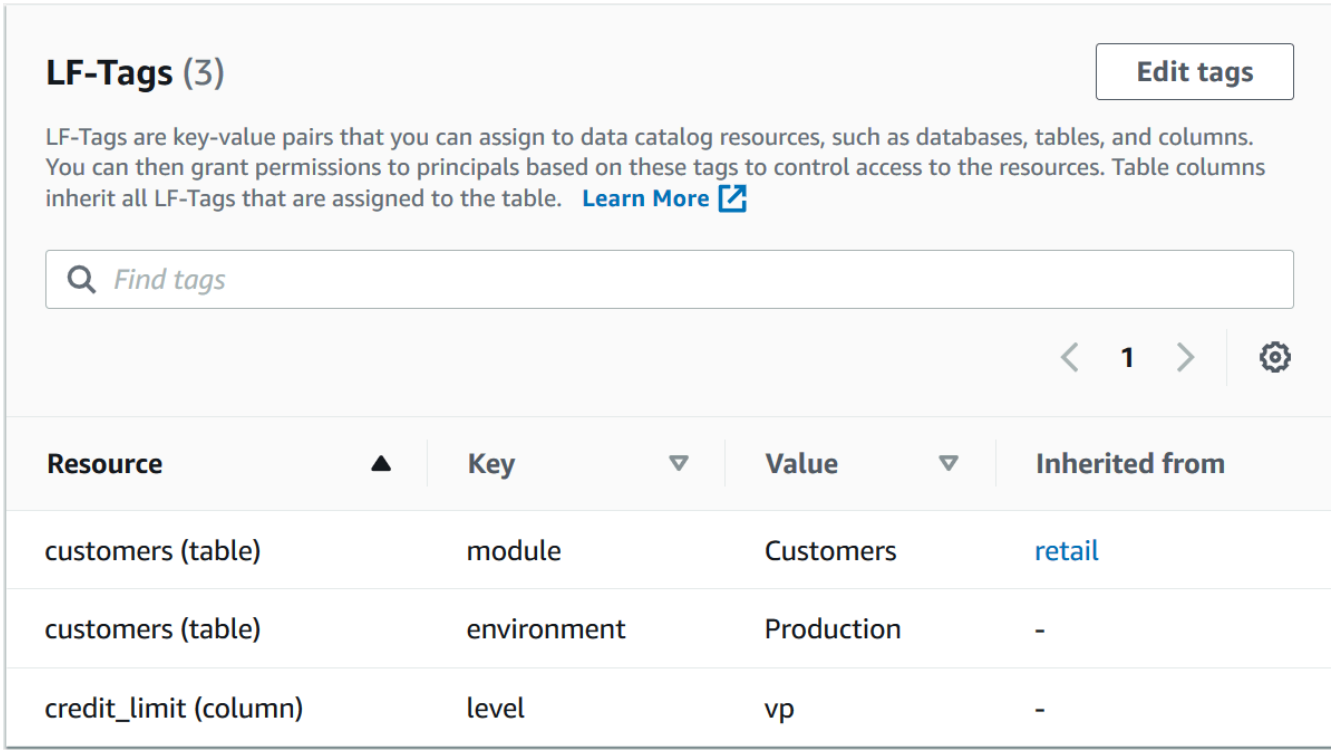
Console

若要檢視指派給資源的 LF 標籤 (主控台)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、資源擁有者或已獲得資源 Lake Formation 權限的使用者身分登入。
2. 在導覽窗格的「資料目錄」標題下，執行下列其中一項作業：
   - 若要檢視指派給資料庫的 LF 標籤，請選擇 [資料庫]。
   - 若要檢視指派給表格的 LF 標籤，請選擇「表格」。
3. 在「表格或資料庫」頁面上，選擇資料庫或表格的名稱。然後在詳細信息頁面上，向下滾動到 LF 標籤部分。

下面的屏幕截圖顯示了分配給一個customers表，其包含在數據庫中的 LF 標籤。retailmoduleLF 標籤是從數據庫繼承。此credit_limit欄具有指定的 level=vp LF 標籤。



AWS CLI

若要檢視指定給資源的 LF 標籤 ()AWS CLI

- 輸入與以下相似的命令。

```
aws lakeformation get-resource-lf-tags --show-assigned-lf-tags --
resource '{ "Table": {"CatalogId":"111122223333", "DatabaseName":"erp",
 "Name":"sales"}}'
```

命令會傳回下列輸出：

```
{
    "TableTags": [
        {
            "CatalogId": "111122223333",
            "TagKey": "module",
```

```
            "TagValues": [
                "sales"
            ]
        },
        {
            "CatalogId": "111122223333",
            "TagKey": "environment",
            "TagValues": [
                "development"
            ]
        }
    ],
    "ColumnTags": [
        {
            "Name": "total",
            "Tags": [
                {
                    "CatalogId": "111122223333",
                    "TagKey": "level",
                    "TagValues": [
                        "director"
                    ]
                }
            ]
        }
    ]
}
```

此輸出僅顯示已明確指派的 LF 標籤，而不是繼承。如果要查看所有列（包括繼承的 LF 標籤）上的所有 LF 標籤，請省略該選項。`--show-assigned-lf-tags`

檢視 LF 標籤指定給的資源

您可以檢視指定特定 LF 標籤鍵的所有「資料目錄」資源。為此，您需要以下 Lake Formation 權限：

- `Describe`或者`Associate`在 LF 標籤上。
- `Describe`或對資源的任何其他 Lake Formation 許可。

此外，您還需要下列 AWS Identity and Access Management (IAM) 許可：

- `lakeformation:SearchDatabasesByLFTags`

- `lakeformation:SearchTablesByLFTags`

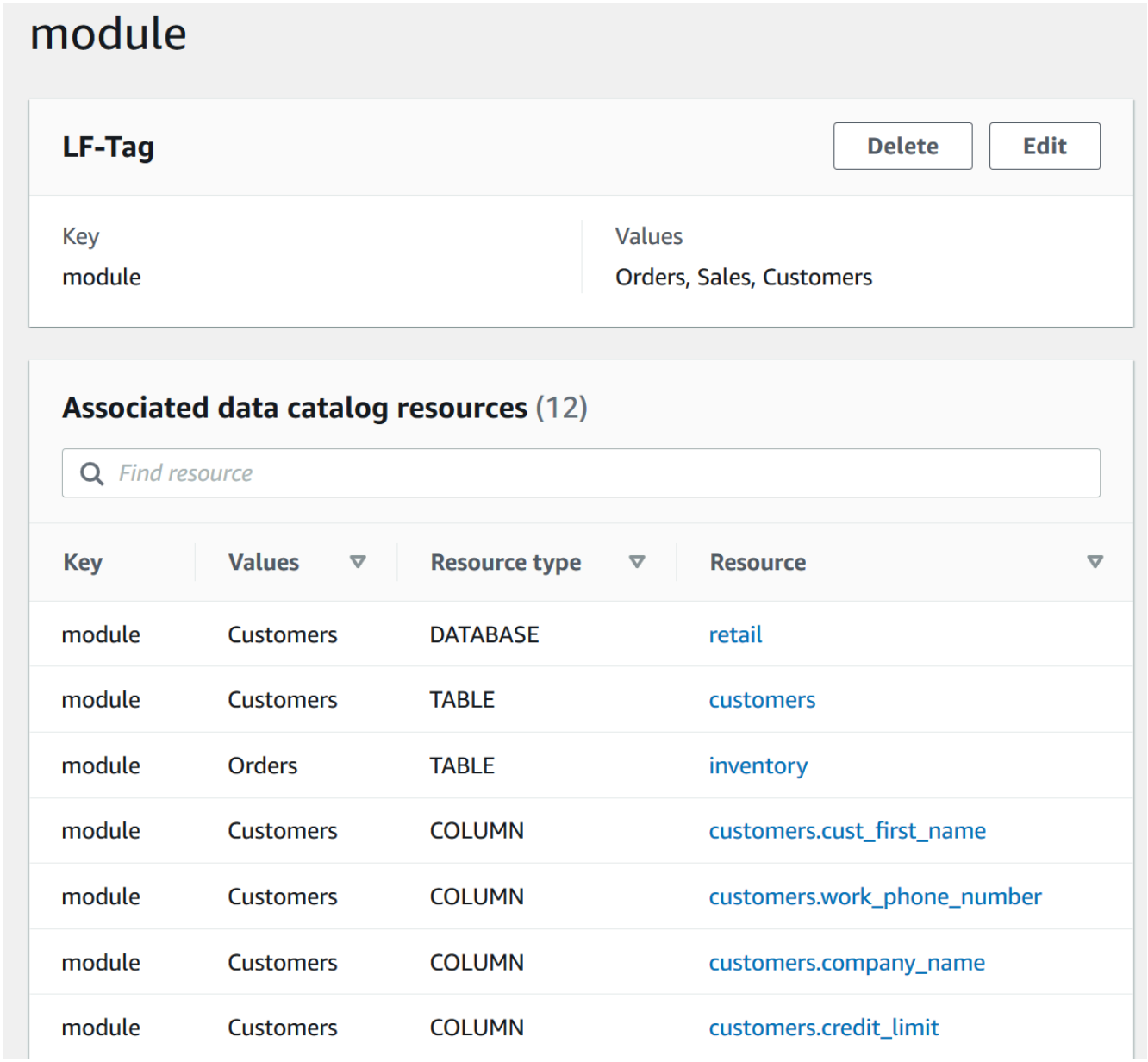
Console

若要檢視 LF 標籤指派給的資源 (主控台)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員或符合先前所列需求的使用者身分登入。
2. 在功能窗格的 [權限] 和 [LF 標籤和權限] 下，選擇 [LF 標籤]。
3. 選擇 LF 標籤鍵 (而不是金鑰名稱旁邊的選項按鈕)。

   LF 標籤詳細資訊頁面會顯示已指定 LF 標籤的資源清單。



AWS CLI

若要檢視 LF 標籤指定給的資源，請執行下列步驟：

- 執行`search-tables-by-lf-tags`或`search-databases-by-lf-tags`命令。

  Example

  下列範例會列出已指派 `level=vp` LF 標籤的資料表和資料行。對於列出的每個表和列，所有為表或列指定的 LF 標籤都會輸出，而不僅僅是搜索表達式。

```
aws lakeformation search-tables-by-lf-tags --expression
 TagKey=level,TagValues=vp
```

如需所需許可的詳細資訊，請參閱Lake Formation 角色和 IAM 許可參考。

LF 標籤的生命週期

1. LF 標籤的創造者邁克爾創建一個 LF 標籤。module=Customers
2. 邁克爾授予 Associate LF 標籤的數據工程師愛德華多。授予Associate隱含授予。Describe
3. 邁克爾授予在桌子Super上Custs給愛德華多與授予選項，使愛德華多可以分配 LF-標籤表。如需詳細資訊，請參閱 將 LF 標籤指定給資料目錄資源。
4. 愛德華多 LF 標籤module=customers分配給表。Custs
5. 邁克爾向數據工程師桑德拉（偽代碼）提出以下贈款。

   ```
   GRANT (SELECT, INSERT ON TABLES) ON TAGS module=customers TO Sandra WITH GRANT OPTION
   ```
6. 桑德拉向數據分析師瑪麗亞提供以下資助。

   ```
   GRANT (SELECT ON TABLES) ON TAGS module=customers TO Maria
   ```

   瑪麗亞現在可以在Custs表上運行查詢。

> **另請參閱**
>
> - 元數據訪問控制

基於 Lake Formation 標籤的訪問控制與基於 IAM 屬性的訪問控制的比較

屬性型存取控制 (ABAC) 是一種授權策略，可根據屬性來定義許可。在中 AWS，這些屬性稱為標籤。您可以將標籤附加到 IAM 資源，包括 IAM 實體 (使用者或角色) 和 AWS 資源。您可以為您的 IAM 主體建立單一 ABAC 政策或是一組政策。這些 ABAC 政策可以設計成在主體的標籤與資源標籤相符時允許操作。ABAC 在成長快速的環境中相當有幫助，並能在政策管理變得繁瑣時提供協助。

雲端安全和治理團隊使用 IAM 為所有資源定義存取政策和安全許可，包括 Amazon S3 儲存貯體、Amazon EC2 執行個體，以及您可以透過 ARN 參考的任何資源。IAM 政策會定義資料湖資源的廣

泛 (粗粒度) 許可，例如允許或拒絕 Amazon S3 儲存貯體、前綴層級或資料庫層級的存取。如需 IAM ABAC 的詳細資訊，請參閱 ABAC 的用途為何？ AWS在 IAM 使用者指南中。

例如，您可以使用 `project-access` 標籤鍵建立三個角色。將第一個角色的鍵值設為 `Dev`，第二個角色的鍵值設為 `Marketing`，並將第三個角色的鍵值設為 `Support`。將具有適當值的標籤指派給資源。然後，您可以使用單一政策，在角色和資源針對 `project-access` 使用相同值標記時允許存取。

資料控管團隊使用 Lake Formation 來定義特定資料湖資源的精細權限。LF 標籤會指派給資料目錄資源 (資料庫、表格和欄)，並授與主參與者。具有 LF 標籤的主參與者符合資源的 LF 標籤可以存取該資源。Lake Formation 許可是 IAM 許可的次要許可。例如，如果 IAM 許可不允許使用者存取資料湖，則 Lake Formation 不會將該資料湖中任何資源的存取權授予該使用者，即使主體和資源具有相符的 LF-標籤也一樣。

基於 Lake Formation 標籤的訪問控制（LF-TBAC）與 IAM ABAC 一起使用，為您的 Lake Formation 數據和資源提供額外的許可級別。

- Lake Formation TBAC 許可通過創新擴展。管理員不再需要更新現有政策來允許存取新的資源。例如，假設您將 IAM ABAC 策略與`project-access`標籤搭配使用，以提供對 Lake Formation 內特定資料庫的存取權。使用 LF-TBAC，LF 標籤`Project=SuperApp`被分配給特定的表或列，並且相同的 LF 標籤被授予該項目的開發人員。透過 IAM，開發人員可以存取資料庫，而 LF-TBAC 許可授予開發人員進一步存取資料表中特定資料表或資料行的權限。如果將新表格新增至專案，Lake Formation 管理員只需要將標籤指派給新資料表，即可讓開發人員存取該表格。
- Lake Formation TBAC 需要較少的 IAM 政策。由於您使用 IAM 政策授予對湖泊形成資源的高層級存取權限，以及 Lake Formation TBAC 來管理更精確的資料存取，因此您建立的 IAM 政策較少。
- 使用 Lake Formation TBAC，團隊可以快速改變和成長。這是因為新資源會自動根據屬性授與許可。例如，如果有新的開發人員加入專案，您可以輕鬆地將 IAM 角色與使用者建立關聯，然後將所需的 LF 標籤指派給使用者，藉此授予此開發人員存取權。您不需要變更 IAM 政策即可支援新專案或建立新的 LF 標籤。
- 使用 Lake Formation TBAC 可以獲得更好的粒度權限。IAM 政策授予對頂級資源的存取權，例如資料目錄資料庫或資料表。使用 Lake Formation TBAC，您可以授與包含特定資料值之特定資料表或資料行的存取權。

> **Note**
>
> IAM 標籤與 LF 標籤不一樣。這些標籤不可互換。LF 標籤用於授予 Lake Formation 許可，IAM 標籤用於定義 IAM 政策。

授予、撤銷和列出 LF 標籤值權限

您可以將 LF 標籤的Alter權Drop限授與主參與者，以管理 LF 標籤值運算式。您也可以將 LF 標籤的Grant with LF-Tag expressions權限授Describe與主體Associate，以檢視 LF 標籤並將其指派給資料目錄資源 (資料庫、資料表和欄)。將 LF 標籤指定給「資料目錄」資源時，您可以使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法來保護這些資源。如需詳細資訊，請參閱 基於 Lake Formation 標籤的訪問控制。

您可以使用授與選項授與這些權限，以便其他主參與者可以授與這些權限。Grant with LF-Tag expressions、Describe和Associate權限將在中說明新增 LF 標籤建立者。

您可以將 LF 標籤的Describe和Associate權限授予外部 AWS 帳戶。然後，該帳戶中的資料湖管理員可以將這些權限授與帳戶中的其他主體。外部帳戶中的資料湖管理員授與Associate權限的主參與者可以將 LF 標籤指派給您與其帳戶共用的資料目錄資源。

授予外部帳戶時，您必須包括授予選項。

您可以通過使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ ）AWS CLI授予 LF 標籤的權限。

主題

- 使用控制台列出 LF 標籤權限
- 使用控制台授予 LF 標籤權限
- 使用授與、撤銷和列出 LF 標籤權限 AWS CLI

如需更多資訊，請參閱管理 LF 標籤以進行中繼資料存取控制及基於 Lake Formation 標籤的訪問控制。

使用控制台列出 LF 標籤權限

您可以使用 Lake Formation 控制台查看 LF-標籤上授予的權限。您必須是 LF 標籤建立者、資料湖管理員，或者擁有 LF 標籤的Describe或Associate權限才能查看它。

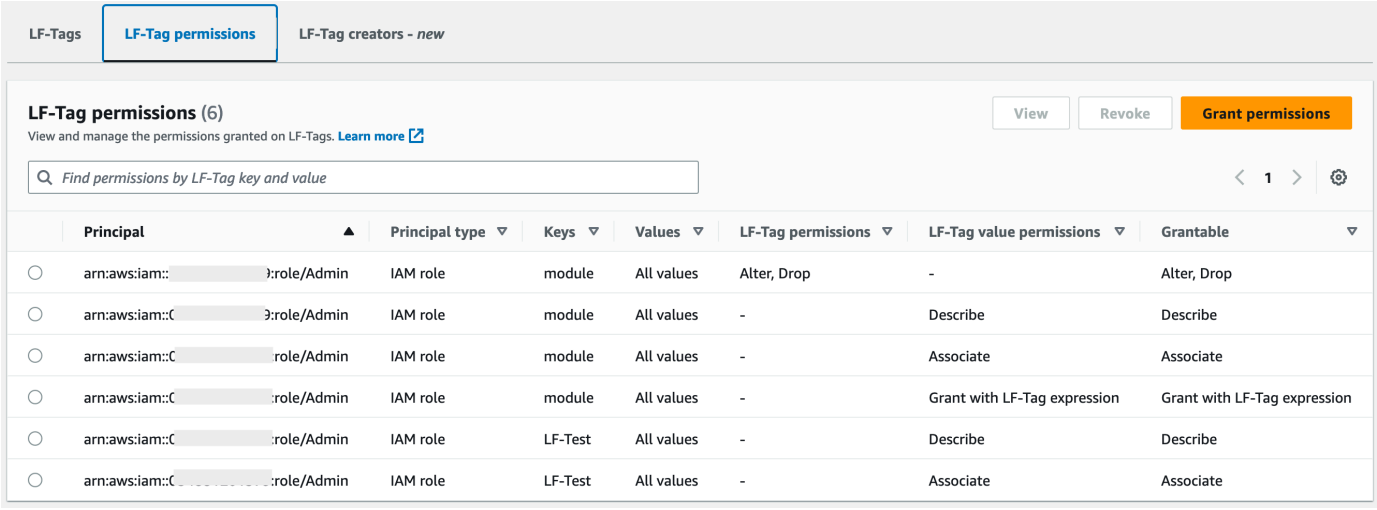
列出 LF 標籤權限（控制台）

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以 LF 標籤建立者、資料湖管理員或已授與 LF 標籤Describe權限的Drop使用者身分登入。Alter Associate

2. 在功能窗格的 [權限] 下，選擇 [LF-標籤和權限]，然後選擇 [LF 標籤權限] 區段。

   LF 標籤權限區段會顯示包含主參與者、標籤索引鍵、值和權限的資料表。



使用控制台授予 LF 標籤權限

下列步驟說明如何使用 Lake Formation 主控台上的 Grant LF-tag 權限頁面授予 LF-tag 權限的權限。該頁面分為以下幾個部分：

- 權限類型 — 要授予的權限類型。
- 主參與者 — 要授與權限的使用者、角色或 AWS 帳號。
- LF 標籤-LF 標籤授予權限。
- 權限 — 要授予的權限。

打開授予 LF 標籤權限頁面

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以 LF 標籤建立者、資料湖管理員身分登入，或以使用者 LF 標籤權限或 LF 標籤索引鍵值配對權限的身分登入已授與選項。`Grant`

2. 在導航窗格中，選擇 LF-標籤和權限，選擇 LF 標籤權限部分。
3. 選擇 授予許可。

指定權限類型

在 [權限類型] 區段中，選擇權限類型。

LF 標籤權限

選擇 LF 標籤權限，以允許主參與者更新 LF 標籤值或刪除 LF 標籤。

LF 標籤鍵值對權限

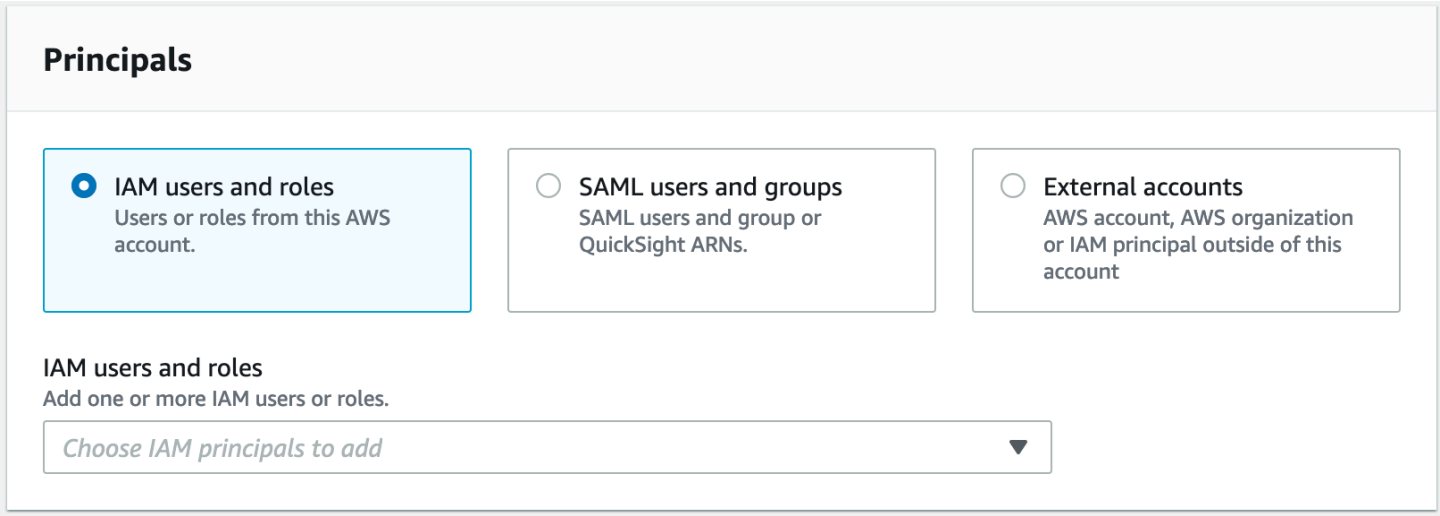
選擇 LF 標籤索引鍵值配對權限，以允許主體將 LF 標籤指派給資料目錄資源、檢視 LF 標籤和值，以及將以 LF 標籤為基礎的資料目錄資源權限授與主參與者。

以下各節中可用的選項取決於「權限」類型。

指定主參與者

> **Note**
>
> 您無法將 LF 標籤權限 (`Alter`和`Drop`) 授與外部帳戶或其他帳戶中的主體。

在「主參與者」區段中，選擇主參與者類型並指定要授與權限的主參與者。



IAM 使用者和角色

從 IAM 使用者和角色清單中選擇一或多個使用者或角色。

SAML 使用者和群組

對於 SAML 和 Amazon QuickSight 使用者和群組，請為透過 SAML 聯合的使用者或群組輸入一或多個 Amazon 資源名稱 (ARN)，或針對 Amazon 使用者或群組輸入 ARN。 QuickSight在每個 ARN 之後按 Enter 鍵。

如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。

> **Note**
>
> 僅支持 Amazon QuickSight 企業版與 Amazon QuickSight 的 Lake Formation 整合。

外部帳戶

對於AWS 帳戶，請輸入一或多個有效的 AWS 帳號 ID。在每個 ID 之後按 Enter 鍵。

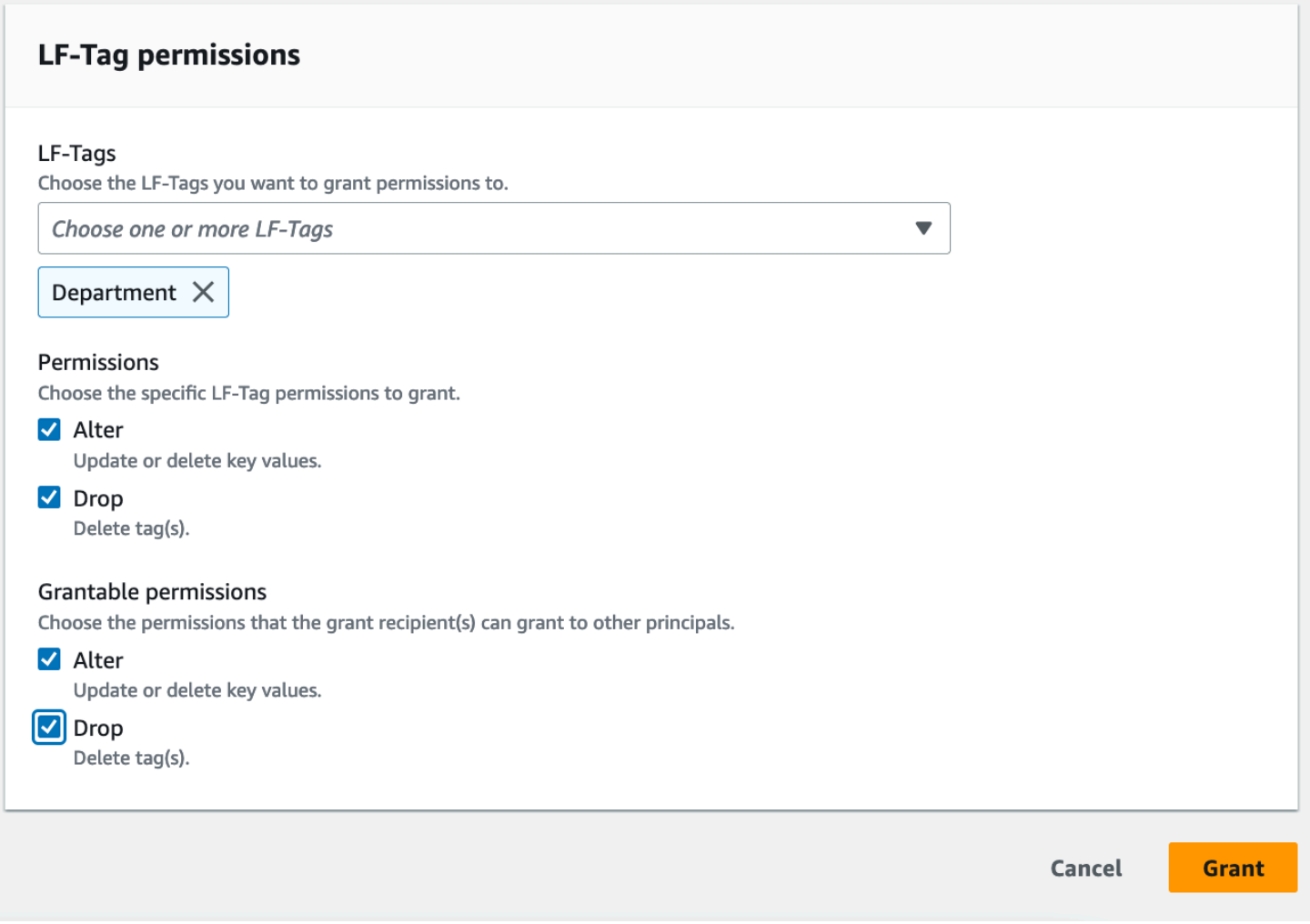
組織 ID 由「o-」後跟 10 到 32 個小寫字母或數字組成。

組織單位 ID 以「ou-」開頭，後面接著 4 到 32 個小寫字母或數字 (包含 OU 的根目錄識別碼)。該字符串後跟第二個「-」破折號和 8 到 32 個其他小寫字母或數字。

針對 IAM 主體，請輸入 IAM 使用者或角色的 ARN。

指定 LF 標籤

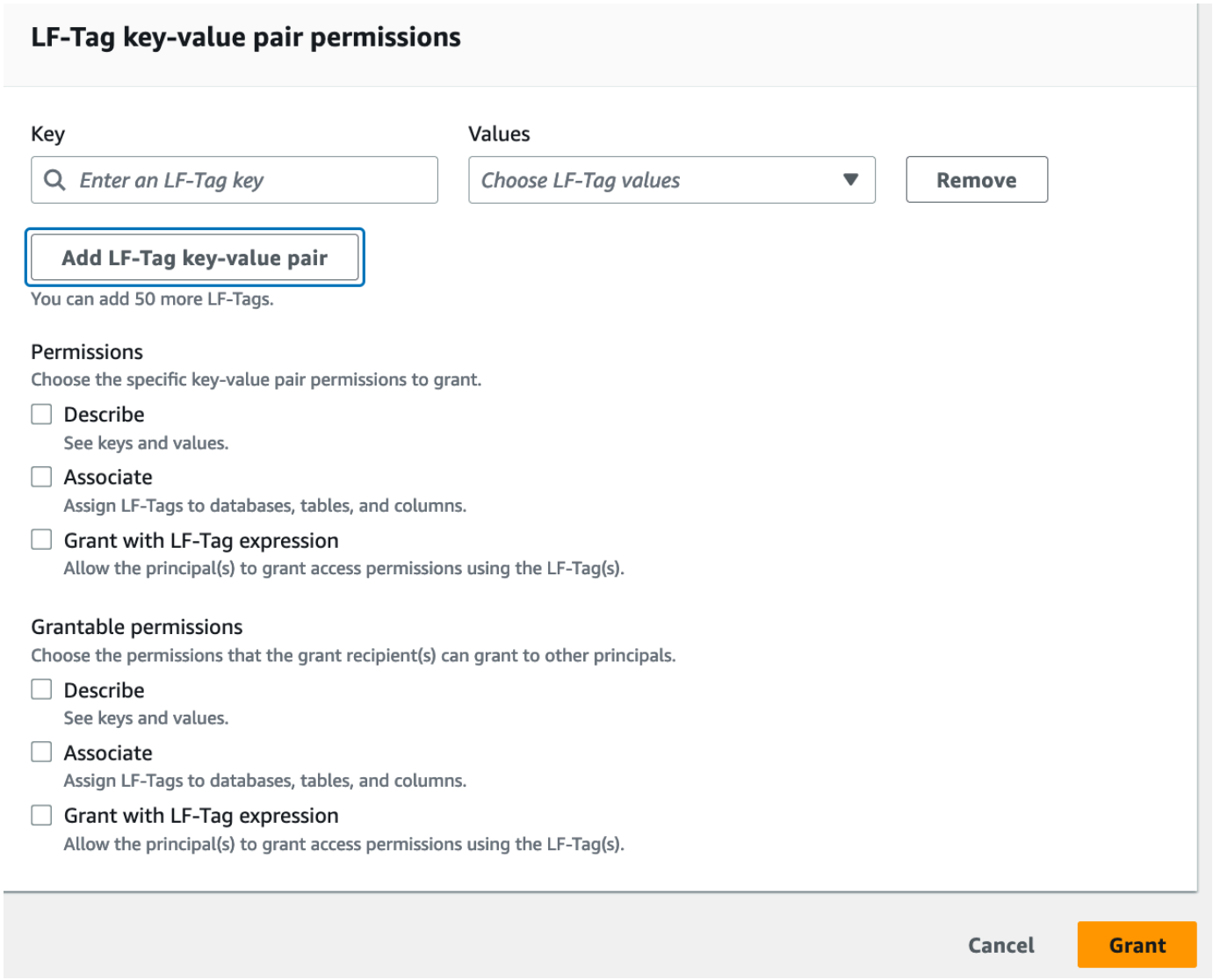
若要授與 LF 標籤的權限，請在 LF 標籤權限區段中，指定要授與權限的 LF 標籤。



- 使用下拉式清單選擇一或多個 LF 標籤。

指定 LF 標籤鍵值對

1. 若要授與 LF 標籤索引鍵值配對的權限，(您必須先選擇選擇 LF 標籤索引鍵值組權限做為權限類型) 選擇 [新增 LF 標籤索引鍵值組] 以顯示第一列欄位用於指定 LF-tag 索引鍵和值。



2. 將游標置於「關鍵字」欄位中，選擇性地開始鍵入以縮小選取清單範圍，然後選取 LF 標籤鍵。
3. 在「值」清單中，選取一個或多個值，然後按 Tab 鍵，或在欄位外按一下或點選以儲存選取的值。

> **Note**
>
> 如果「值」清單中的其中一列具有焦點，則按 Enter 可選取或清除該勾選方塊。

所選值會在「值」清單下方顯示為並排顯示。選擇 ✖ 以移除值。選擇「移除」以移除整個 LF 標籤。

4. 若要新增其他 LF 標籤，請再次選擇「新增 LF 標籤」，然後重複前兩個步驟。

指定權限

本節顯示 LF 標籤權限或 LF 標籤值權限，以您在上一個步驟中選擇的權限類型為基礎。

根據您選擇授與的權限類型，選取 LF 標籤權限或 LF 標籤索引鍵值組權限，以及可授予的權限。

1. 在 LF 標籤權限下，選取要授與的權限。

   隱含授予 [卸除] 和 [變更] 授與 [描述]

   您必須授與所有標籤值的「變更」和「刪除」權限。
2. 在 LT 標籤鍵值權限下，選取要授與的權限。

   隱含授予關聯性會授予「描述」。選擇「使用 LF 標籤運算式授與」，以允許授與收件者使用 LF-TBAC 方法授與或撤銷「資料目錄」資源的存取權限。
3. (選擇性) 在可授與權限下，選取授與收件者可以授與其帳戶中其他主體的權限。 AWS
4. 選擇 Grant (授予)。

使用授與、撤銷和列出 LF 標籤權限 AWS CLI

您可以使用 AWS Command Line Interface （ AWS CLI ）授予，撤銷和列出 LF-標籤的權限。

列出 LF 標籤權限（ ）AWS CLI

- 輸入`list-permissions`指令。您必須是 LF 標籤建立者、資料湖管理員，或具有 LF 標籤的`DropAlterDescribe`、`Associate`、、`Grant with LF-Tag permissions`權限才能查看它。

  下列命令會要求您擁有權限的所有 LF 標籤。

  ```
  aws lakeformation list-permissions --resource-type LF_TAG
  ```

  以下是資料湖管理員的範例輸出，管理員會看到授與所有主體的所有 LF 標籤。非系統管理使用者只會看到授與他們的 LF 標籤。從外部帳戶授予的 LF 標籤權限會顯示在個別的結果頁面上。要查看它們，請重複該命令並使用上一個命令運行返回的令牌提供`--next-token`參數。

  ```
  {
      "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"
}
```

您可以列出特定 LF 標籤鍵的所有授權。下列命令會傳回 LF module 標籤上授予的所有權限。

```
aws lakeformation list-permissions --resource-type LF_TAG --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

您也可以列出授與特定 LF 標籤之特定主參與者的 LF 標籤值。提供--principal引數時，您必須提供引--resource數。因此，命令只能有效地要求授與特定 LF 標籤索引鍵之特定主體的值。下面的命令顯示如何為主體datalake_user1和 LF 標籤鍵執行此操作。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":["*"]}}'
```

下列為範例輸出。

```
{
    "PrincipalResourcePermissions": [
        {
            "Principal": {
                "DataLakePrincipalIdentifier": "arn:aws:iam::111122223333:user/
datalake_user1"
            },
            "Resource": {
                "LFTag": {
                    "CatalogId": "111122223333",
                    "TagKey": "module",
                    "TagValues": [
                        "Orders",
                        "Sales"
                    ]
                }
            },
            "Permissions": [
                "ASSOCIATE"
            ],
            "PermissionsWithGrantOption": []
        }
    ]
```

```
}
```

授予 LF-標籤的權限 ( ) AWS CLI

1. 輸入與以下相似的命令。這個例子授予用戶datalake_user1對 LF 標籤的Associate權限與密鑰。module它會授與檢視及指派該機碼所有值的權限，如星號 (*) 所示。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

隱含授與Associate權限授與Describe權限。

下一個示例授Associate予外部 AWS 帳戶 1234-5678-9012 在 LF 標籤與密鑰，與授予選項。module它授予僅查看和分配值sales和的權限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. 隱含授與GrantWithLFTagExpression權限授與Describe權限。

下一個範例會GrantWithLFTagExpression使用授予選項授與金鑰module的 LF 標籤上的使用者。它授與僅使用值sales和的權限來檢視和授與資料目錄資源的權限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. 下一個範例會使用授予選項授予使用者在 LF 標籤上使用金鑰module的Drop權限。它授予刪除 LF 標籤的權限。要刪除 LF 標籤，您需要對該鍵的所有值的權限。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333 --permissions "DROP"
 --permissions-with-grant-option "DROP" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

4. 下一個範例會使用授Alter權選項授予使用者在 LF 標籤上使用金鑰module的權限。它授予刪除 LF 標籤的權限。要更新 LF 標籤，您需要對該密鑰的所有值的權限。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=111122223333 --permissions "ALTER"
 --permissions-with-grant-option "ALTER" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

撤銷 LF-標籤的權限（ ）AWS CLI

- 輸入與以下相似的命令。此範例會使用使用Associate者的金鑰module撤銷 LF 標籤的權限。datalake_user1

```
aws lakeformation revoke-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
datalake_user1 --permissions "ASSOCIATE" --resource '{ "LFTag":
 {"CatalogId":"111122223333","TagKey":"module","TagValues":["*"]}}'
```

使用 LF-TBAC 方法授與資料湖權限

您可以將 LF 標籤上的DESCRIBE和 L ASSOCIATE ake Formation 權限授與主體，以便他們可以檢視 LF 標籤，並將它們指派給資料目錄資源 (資料庫、資料表、檢視和資料行)。將 LF 標籤指定給資料目錄資源時，您可以使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法來保護這些資源。如需詳細資訊，請參閱 基於 Lake Formation 標籤的訪問控制。

首先，只有資料湖管理員可以授與這些權限。如果資料湖管理員使用授與選項授與這些權限，則其他主參與者可以授與這些權限。DESCRIBE和ASSOCIATE權限將在中說明基於 Lake Formation 標籤的訪問控制最佳實踐和考量。

您可以將 LF 標籤的DESCRIBE和ASSOCIATE權限授予外部 AWS 帳戶。然後，該帳戶中的資料湖管理員可以將這些權限授與帳戶中的其他主體。外部帳戶中的資料湖管理員授與ASSOCIATE權限的主參與者可以將 LF 標籤指派給您與其帳戶共用的資料目錄資源。

授予外部帳戶時，您必須包括授予選項。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface ()AWS CLI授予 LF 標籤的權限。

主題

- 授與資料目錄權限

另請參閱

- 授予、撤銷和列出 LF 標籤值權限
- 管理 LF 標籤以進行中繼資料存取控制
- 基於 Lake Formation 標籤的訪問控制

授與資料目錄權限

使用 Lake Formation 主控台，或使用 AWS CLI 以湖泊形成標籤為基礎的存取控制 (LF-TBAC) 方法，授與「資料目錄」資料庫、資料表、檢視和欄的「Lake Formation」權限。

Console

下列步驟說明如何使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法和 Lake Formation 主控台上的 [授與資料湖權限] 頁面來授與權限。該頁面分為以下幾個部分：

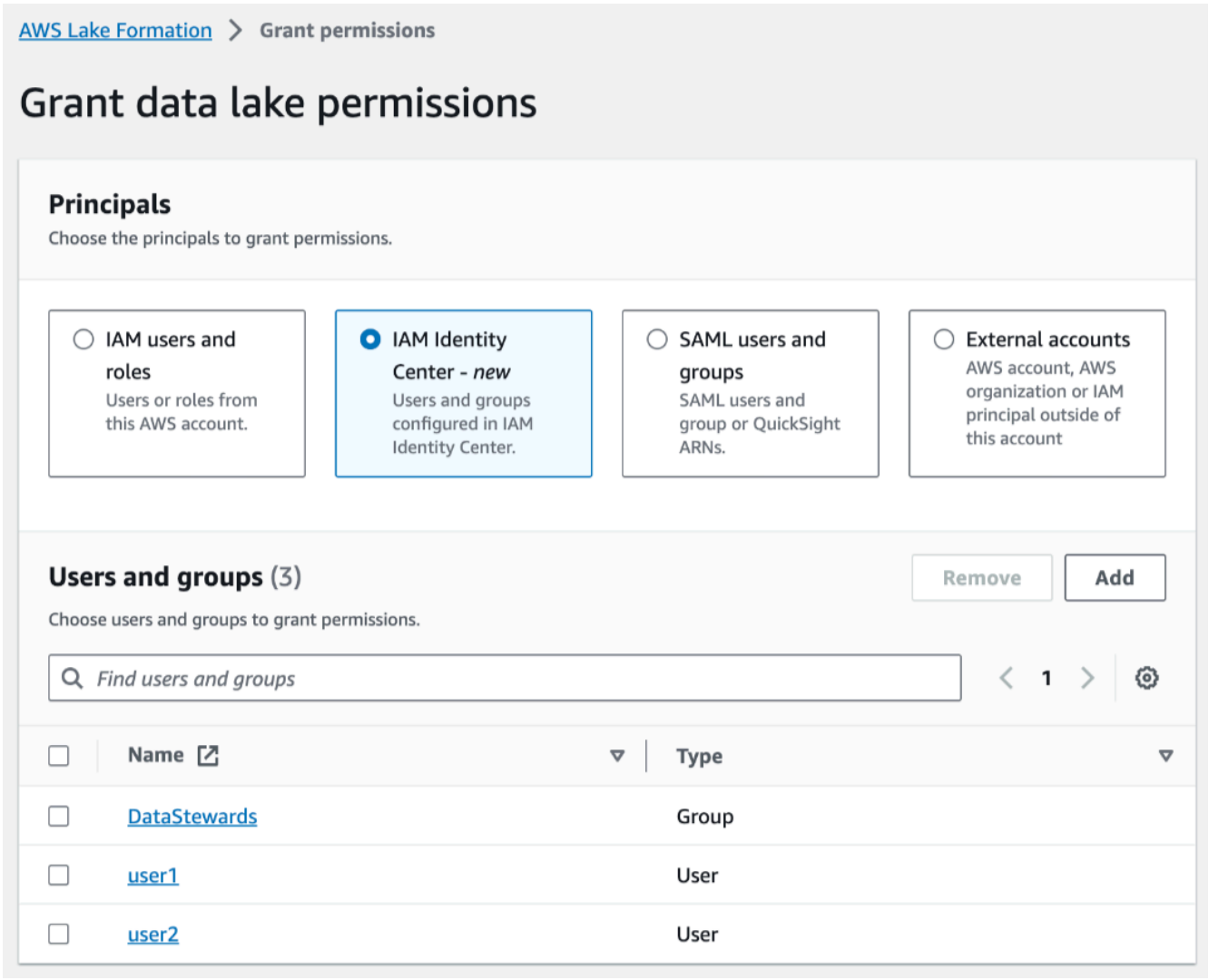
- 主參與者 — 要授與權限的使 AWS 帳戶 用者、角色和。
- LF 標籤或目錄資源 — 要授與權限的資料庫、表格或資源連結。
- 權限-要授予的 Lake Formation 許可權。

1. 開啟 [授與資料湖權限] 頁面。

   在 https://console.aws.amazon.com/lakeformation/ 開啟 AWS Lake Formation 主控台，然後以資料湖管理員身分或已透過 LF-TBAC 授與「資料目錄」資源的「Lake Formation 格式」權限的使用者身分登入，並使用授予選項。

   在功能窗格的 [權限] 下，選擇 [資料湖權限]。然後選擇授予。
2. 指定主參與者。

   在「主參與者」區段中，選擇主參與者類型，然後指定要授與權限的主參與者。



IAM 使用者和角色

從 IAM 使用者和角色清單中選擇一或多個使用者或角色。

IAM Identity Center

選擇一個或多個使用者，或從使用者和群組清單中選擇。

SAML 使用者和群組

對於 SAML 和 Amazon QuickSight 使用者和群組，請為透過 SAML 聯合的使用者或群組輸入一或多個 Amazon 資源名稱 (ARN)，或針對 Amazon 使用者或群組輸入 ARN。QuickSight 在每個 ARN 之後按 Enter 鍵。

如需有關如何建構 ARN 的資訊，請參閱Lake Formation 授予和撤銷 AWS CLI 命令。

> **Note**
>
> 僅支持 Amazon QuickSight 企業版與 Amazon QuickSight 的 Lake Formation 整合。

外部帳戶

針對AWS 帳戶、 AWS 組織或 IAM 主體，輸入 IAM 使用者或角色的一或多個有效 AWS 帳戶 ID、組織 ID、組織單位 ID 或 ARN。在每個 ID 之後按 Enter 鍵。

組織 ID 由「o-」後跟 10 到 32 個小寫字母或數字組成。

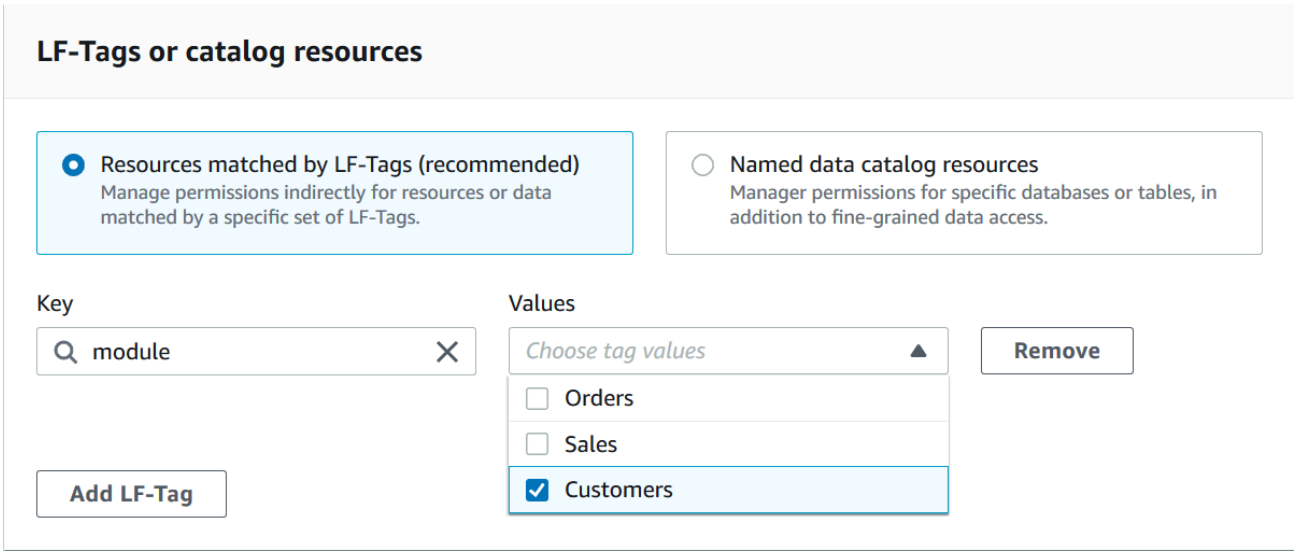
組織單位 ID 以「ou-」開頭，後面接著 4 到 32 個小寫字母或數字 (包含 OU 的根目錄識別碼)。該字符串後跟第二個「-」破折號和 8 到 32 個其他小寫字母或數字。

3. 指定 LF 標籤。

確定已選取 [LF-標籤符合的資源] 選項。選擇「新增 LF 標籤」。

1. 選擇 LF 標籤鍵和值。

如果您選擇多個值，則會建立含有運算子的 LF 標籤運算式OR。這表示如果任何 LF 標籤值與指定給資料目錄資源的 LF 標籤相符，您將獲得該資源的權限。



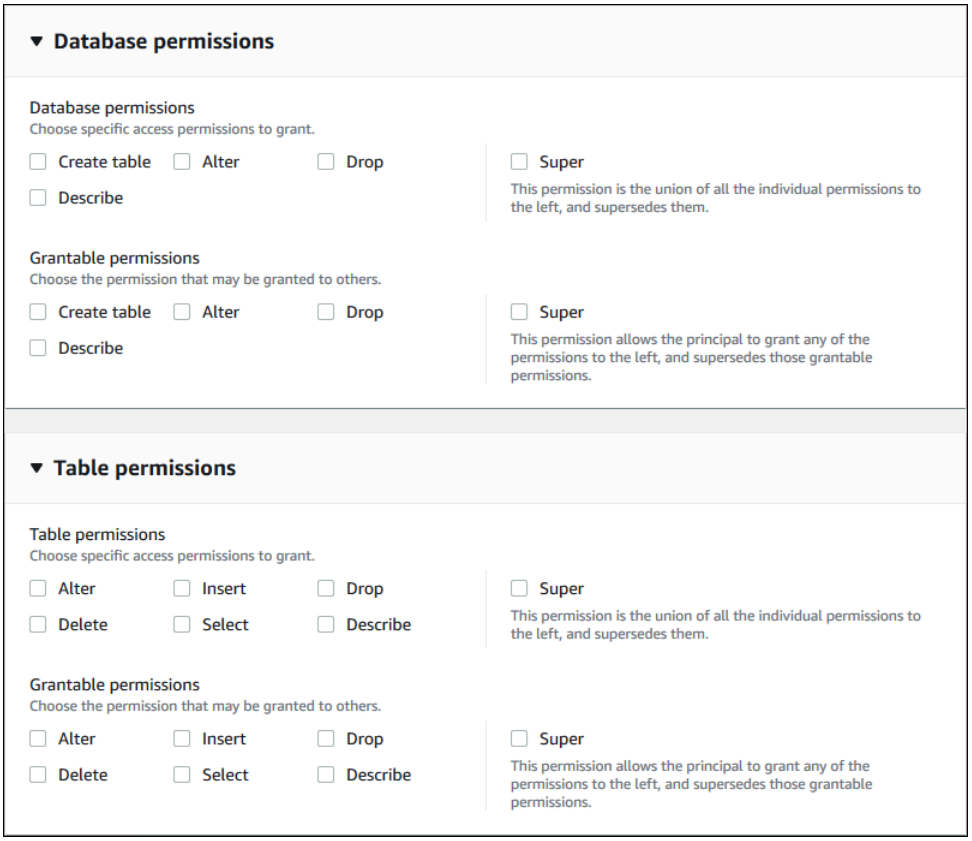
2. (選擇性) 再次選擇「新增 LF 標籤」以指定另一個 LF 標籤。

如果您指定多個 LF 標籤，則會使用運算子建立 LF 標籤運算式。AND僅當資源為 LF 標籤運算式中的每個 LF 標籤指派相符的 LF 標籤時，才會授與資料目錄資源的權限。

4. 指定權限。

   指定您要在相符資料目錄資源上授與主參與者的權限。相符資源是指派 LF 標籤的資源，符合授與主體的 LF 標籤運算式之一。

   您可以指定要授與相符資料庫、相符資料表和相符檢視的權限。



   在 [資料庫權限] 底下，選取要授與相符資料庫主體的資料庫權限。

   在 [資料表權限] 底下，選取要授與相符資料表和檢視表上主參與者的資料表或檢視權限。

   您也可以從 [表格] **Drop** 權限中選擇**Select**、**Describe**和權限，以套用至檢視。

5. 選擇 Grant (授予)。

AWS CLI

您可以使用 AWS Command Line Interface (AWS CLI) 和以湖泊形成標籤為基礎的存取控制 (LF-TBAC) 方法，授與「資料目錄」資料庫、資料表和資料行的 Lake Formation 權限。

使用 AWS CLI 和 LF-TBAC 方法授與資料湖權限

- 使用 `grant-permissions` 命令。

  Example

  下列範例會將 LF-tag 運算式 "`module=*`" (LF 標籤索引鍵`module`的所有值) 授與使用者。`datalake_user1`該使用者將擁有所有相符資料庫的`CREATE_TABLE`權限 — 已指派 LF 標籤的資料庫以及任何值`module`。

  ```
  aws lakeformation grant-permissions --principal
   DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/
  datalake_user1 --permissions "CREATE_TABLE" --resource '{ "LFTagPolicy":
   {"CatalogId":"111122223333","ResourceType":"DATABASE","Expression":
  [{"TagKey":"module","TagValues":["*"]}]}}'
  ```

  Example

  下一個範例會將 LF 標籤運算式 "`(level=director) AND (region=west OR region=south)`" 授與使用者。`datalake_user1`該使用者在相符資料表 (已同時指`DROP`派和 (`region=west`或`region=south`) 的資料表上`ALTER`，具有授與選項的、`level=director`和權限。`SELECT`

  ```
  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

  下一個範例會將 LF 標籤運算式 "`module=orders`" 授與 AWS 帳戶 1234-5678-9012。然後，該帳戶中的資料湖管理員可以將 "`module=orders`" 運算式授與其帳戶中的主體。然後，這些

主體將擁有帳戶 1111-2222-3333 所擁有的相符資料庫的CREATE_TABLE權限，並使用具名的資源方法或 LF-TBAC 方法與帳戶 1234-5678-9012 共用。

```
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"]}]}}'
```

權限範例案例

下列案例有助於示範如何設定權限以確保中資料的存取安全 AWS Lake Formation。

雪莉是資料管理員。她想為自己的公司建立一個資料湖 AnyCompany。目前，所有資料都存放在 Amazon S3 中。John 是行銷經理，需要寫入客戶購買資訊的存取權 (包含在中s3://customerPurchases)。一位營銷分析師迭戈今年夏天加入約翰。John 需要能夠授予 Diego 存取權，以便在不涉及雪莉的情況下對資料執行查詢。

Mateo, 從財務, 需要訪問查詢會計數據 (例如,s3://transactions). 他想要查詢資料庫 (Finance_DB) 財務團隊使用的資料表中的交易資料。他的Finance_DB經理, 阿納姆, 可以給他訪問. 儘管他不應該能夠修改會計數據，但他需要將數據轉換為適合預測的格式（模式）的能力。這些數據將被存儲在一個單獨的存儲桶（s3://financeForecasts）中，他可以修改。

總結一下：

- 雪莉是資料湖管理員。
- John 需要CREATE_DATABASE和CREATE_TABLE權限才能在「資料目錄」中建立新的資料庫和資料表。
- John 也需要SELECT他所建立之表格的INSERT、和DELETE權限。
- Diego 需要資料表的SELECT權限才能執行查詢。

的員工 AnyCompany 執行下列動作來設定權限。此案例中顯示的 API 作業會顯示簡化的語法，以提高清晰度。

1. 雪莉在 Amazon S3 路徑註冊了包含客戶購買信息的 Lake Formation。

```
RegisterResource(ResourcePath("s3://customerPurchases"), false, Role_ARN )
```

2. 雪莉授予 John 存取包含客戶購買資訊的 Amazon S3 路徑。

```
GrantPermissions(John, S3Location("s3://customerPurchases"),
 [DATA_LOCATION_ACCESS]) )
```

3. 雪莉授予 John 建立資料庫的權限。

```
GrantPermissions(John, catalog, [CREATE_DATABASE])
```

4. 約翰創建數據庫John_DB。John 自動擁有該資料庫的CREATE_TABLE權限，因為他建立了該資料庫。

```
CreateDatabase(John_DB)
```

5. 約翰創建John_Table指向的表s3://customerPurchases。因為他建立了資料表，所以他擁有該資料表的所有權限，而且可以授與該資料表的權限。

```
CreateTable(John_DB, John_Table)
```

6. 約翰允許他的分析師迭戈訪問該表John_Table。

```
 GrantPermissions(Diego, John_Table, [SELECT])
```

7. 約翰允許他的分析師, 迭戈, 訪問s3://customerPurchases/London/. 因為 Shirley 已經註冊s3://customerPurchases，所以其子文件夾已在 Lake Formation 註冊。

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, [DATA_LOCATION_ACCESS], [],
S3Location("s3://customerPurchases/London/") )
```

8. 約翰允許他的分析師，迭戈，在數據庫中創建表John_DB。

```
GrantDataLakePrivileges( 123456789012/datalake, Diego, John_DB, [CREATE_TABLE],
[] )
```

9. Diego 會John_DB在中建立資料表ALTER，s3://customerPurchases/London/並自動取得DROPSELECT、INSERT、和DELETE權限。

```
CreateTable( 123456789012/datalake, John_DB, Diego_Table )
```

Lake Formation 中的數據過濾和細胞級安全

當您授與「資料目錄」表格的 Lake Formation 權限時，您可以納入資料篩選規格，以限制查詢結果中某些資料的存取，以及與 Lake Formation 整合的引擎。Lake Formation 使用資料篩選來實現欄層級安全性、資料列層級安全性和儲存格層級安全性。如果來源資料包含巢狀結構，您可以在巢狀資料行上定義並套用資料篩選器。

主題

資料篩選概觀

借助 Lake Formation 的數據過濾功能，您可以實現以下級別的數據安全性。

資料欄層級安全

授與具有資料行層級安全性 (欄篩選) 之「資料目錄」表格的權限，可讓使用者僅檢視其在資料表中有權存取的特定欄和巢狀資料行。考慮一個`persons`用於大型多區域通信公司的多個應用程序的表格。透過欄篩選授與「資料目錄」表格的權限，可以限制不在 HR 部門工作的使用者看到個人身分識別資訊 (PII)，例如社會安全號碼或出生日期。您也可以定義安全性原則，並僅授與部分巢狀資料欄子結構的存取權。

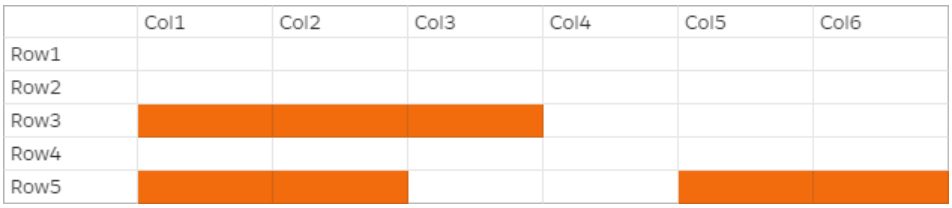
資料列層級安全性

授與具有列層級安全性 (資料列篩選) 之「資料目錄」表格的權限，可讓使用者僅檢視其在表格中有權存取的特定資料列。篩選是以一個或多個資料行的值為基礎。定義列篩選運算式時，您可以包含巢狀資料行結構。例如，如果通訊公司的不同區域辦公室有自己的人力資源部門，您可以將 HR 員工可以查看的人員記錄限制為只顯示其區域中員工的記錄。

儲存格層級安全

儲存格層級安全性結合了資料列篩選和欄篩選，提供高度彈性的權限模型。如果您以網格的方式檢視表格的列與欄，則使用儲存格層級安全性，您可以在兩個維度中的任何位置限制對網格個別元素 (儲存格)

的存取。也就是說，您可以根據該行限制對不同列的訪問。下圖說明了這一點，其中受限制的欄會被著色。



繼續 persons 表格的範例，您可以在儲存格層級建立資料篩選器，以限制對街道位址欄的存取 (如果該列的國家/地區欄設定為「UK」)，但如果該列的國家/地區欄設定為「US」，則允許存取街道位址欄。

篩選器僅適用於讀取作業。因此，您只能使用過濾器授予 SELECT Lake Formation 許可。

巢狀欄的儲存格層級安全性

Lake Formation 可讓您在巢狀資料行上定義和套用具有儲存格層級安全性的資料篩選器。但是，整合式分析引擎如 Amazon Athena、Amazon EMR 和亞 Amazon Redshift Spectrum 支援對 Lake Formation 管理的嵌套表格執行查詢，並具有資料列層級和欄層級安全性。

如需限制的詳細資訊，請參閱資料篩選限制。

Lake Formation 的數據過濾器

您可以建立資料篩選條件，來實作資料欄層級、資料列層級和儲存格層級安全性。當您授與表格的 SELECT Lake Formation 權限時，您可以選取資料篩選器。如果您的表格包含巢狀資料行結構，您可以透過包含或排除子資料欄來定義資料篩選，並在巢狀屬性上定義資料列層級篩選運算式。

每個資料篩選均屬於「資料目錄」中的特定表格。資料篩選器包含下列資訊：

- 篩選器名稱
- 與篩選器相關聯之表格的目錄 ID
- 資料表名稱
- 包含資料表的資料庫名稱
- 資料欄規格 — 要包含或排除在查詢結果中的`struct`資料行和巢狀資料欄 (含有資料類型) 的清單。
- 列篩選運算式 — 指定要包含在查詢結果中之列的運算式。在某些限制下，運算式會使用 PartiQL 語言中的`WHERE`子句語法。若要指定所有列，請在主控台的資料列層級存取權下選擇 [存取所有列]，或在 API 呼叫`AllRowsWildcard`中使用。

如需資料列篩選運算式支援之項目的詳細資訊，請參閱資料列篩選運算式中的 PartiQL 支援。

您取得的篩選層級取決於資料篩選條件的填入方式。

- 如果您指定「所有資料行」萬用字元並提供資料列篩選條件運算式，則只會建立資料列層級安全性 (資料列篩選)。
- 當您包含或排除特定資料欄和巢狀資料欄，並使用全資料列萬用字元指定「所有資料列」時，您只會建立資料行層級安全性 (資料行篩選)。
- 如果包含或排除特定資料行並提供資料列篩選條件運算式，則會建立儲存格層級安全性 (儲存格篩選)。

來自 Lake Formation 主控台的下列螢幕擷取畫面顯示了執行儲存格層級篩選的資料篩選器。對於對資料`orders`表的查詢，它會限制對資`customer_name`料行的存取，而查詢結果只會傳回資料行包含「pharma」的資料`product_type`列。

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

請注意使用單引號來括住字串常值。'pharma'

您可以使用 Lake Formation 主控台來建立此資料篩選器，也可以將下列要求物件提供給 CreateDataCellsFilter API 作業。

```
{
     "Name": "restrict-pharma",
     "DatabaseName": "sales",
     "TableName": "orders",
     "TableCatalogId": "111122223333",
     "RowFilter": {"FilterExpression": "product_type='pharma'"},
     "ColumnWildcard": {
         "ExcludedColumnNames": ["customer_name"]
     }
}
```

您可以根據需要為表格建立任意數量的資料篩選器。為了這樣做，您需要具有表格上授予選項的SELECT權限。依預設，Data Lake 管理員具有在該帳戶中的所有表格上建立資料篩選器的權限。在將資料表的權限授與主參與者時，通常只會使用可能的資料篩選器子集。例如，您可以為資料篩選器的資料orders表建立第二個 row-security-only 資料篩選器。參照前面的螢幕擷取畫面，您可以選擇 [存取所有欄] 選項，並包含的資料列篩選器運算式product_type<>pharma。此資料篩選器的名稱可以是no-pharma。它限制了對列設置為「藥」的product_type所有行的訪問。

此資料篩選器之 CreateDataCellsFilter API 作業的要求物件如下。

```
{
     "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"]
}
```

然後，您可以將具有restrict-pharma資料篩選器的資料orders表授SELECT與系統管理使用者，並SELECT在具有no-pharma資料篩選器的資料orders表上授與非系統管理使用者。對於醫療保健領域的使用者，您可以授予orders表格SELECT上所有資料列和欄的完整存取權 (無資料篩選器)，或者授與另一個資料篩選器，以限制對定價資訊的存取權限。

在資料篩選中指定資料行層級和資料列層級安全性時，您可以包含或排除巢狀資料行。在下列範例中，使用限定資料行名稱 (以雙引號括住) 來指定欄`product.offer`位的存取權。這對於巢狀欄位非常重要，以避免資料行名稱包含特殊字元時發生錯誤，並維持與頂層資料行層級安全性定義的回溯相容性。

```
{
     "Name": "example_dcf",
     "DatabaseName": "example_db",
     "TableName": "example_table",
     "TableCatalogId": "111122223333",
     "RowFilter": { "FilterExpression": "customer.customerName <> 'John'" },
     "ColumnNames": ["customer", "\"product\".\"offer\""]
}
```

另請參閱

- 管理資料篩選

資料列篩選運算式中的 PartiQL 支援

您可以使用 PartiQL 資料類型、運算子和彙總的子集來建構資料列篩選運算式。Lake Formation 不允許在過濾器表達式中使用任何用戶定義或標準的 PartiQL 函數。您可以使用比較運算子來比較資料行與常數 (例如，`views >= 10000`)，但無法將資料行與其他資料行進行比較。

資料列篩選器運算式可以是簡單運算式或複合運算式。運算式的總長度必須小於 2048 個字元。

簡單的表達

一個簡單的表達式將是格式： `<column name > <comparison operator ><value >`

- 資料欄名稱

  它可以是最上層資料欄、分割區資料欄或資料表結構定義中的巢狀資料欄，而且必須屬於下支援的資料類型列資料欄。
- 比較運算子

  以下是支持的運算符：`=, >, <, >=, <=, <>,!=, BETWEEN, IN, LIKE, NOT, IS [NOT] NULL`

- 所有字符串比較和LIKE模式匹配都區分大小寫。您不能在分區列上使用 IS [NOT] NULL 運算符。
- 欄值

  「欄」值必須符合資料行名稱的資料類型。

複合表達式

複合運算式的格式為：`( <simple expression >) <AND/OR >(<simple expression >)`。複合表達式可以使用邏輯運算符進一步組合AND/OR。

支援的資料類型

參照包含不受支援資料類型之資料 AWS Glue Data Catalog 表的資料列篩選器會導致錯誤。下列是對應至資料類型的資料表資料行和常數所支援的 Amazon Redshift 資料類型：

- `STRING, CHAR, VARCHAR`
- `INT, LONG, BIGINT, FLOAT, DECIMAL, DOUBLE`
- `BOOLEAN`
- `STRUCT`

如需 Amazon Redshift 中資料類型的詳細資訊，請參閱 Amazon Redshift 資料庫開發人員指南中的資料類型。

列篩選運算式

Example

以下是具有欄之資料表的有效資料列篩選運算式範例： `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

以下是具有巢狀資料欄之資料表之資料列篩選運算式的有效範例：`year > 2010 and customer.customerId <> 1`

定義巢狀資料列層級運算式時，不應參考資料分割資料行下的巢狀欄位。

字符串常量必須用單引號括起來。

保留的關鍵字

如果您的資料列篩選運算式包含 PartiQL 關鍵字，您就會收到剖析錯誤，因為資料欄名稱可能會與關鍵字衝突。發生這種情況時，請使用雙引號逸出列名。保留關鍵字的一些例子是「第一」，「最後」，「asc」，「缺少」。如需保留關鍵字的清單，請參閱 PartiQL 規範。

PartiQL 參考

如需有關 PartiQL 的詳細資訊，請參閱https://partiql.org/。

使用儲存格層級篩選查詢資料表所需的權限

若要對具有儲存格層級篩選的資料表執行查詢，需要下列 AWS Identity and Access Management (IAM) 權限。

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "lakeformation:StartQueryPlanning",
                "lakeformation:GetQueryState",
                "lakeformation:GetWorkUnits",
                "lakeformation:GetWorkUnitResults"
            ],
            "Resource": "*"
        }
    ]
}
```

有關 Lake Formation 權限的更多信息，請參閱Lake Formation 角色和 IAM 許可參考。

管理資料篩選

若要實作資料行層級、資料列層級和儲存格層級安全性，您可以建立和維護資料篩選器。每個資料篩選均屬於一個「資料目錄」表格。您可以為資料表建立多個資料篩選器，然後在授與資料表的權限時使用其中一或多個篩選器。您也可以在具有資料`struct`類型的巢狀資料行上定義和套用資料篩選器，讓使用者只能存取巢狀資料行的子結構。

您需要具有授予選項的`SELECT`權限才能建立或檢視資料篩選器。若要允許您帳戶中的主體檢視和使用資料篩選器，您可以授與該篩選器的`DESCRIBE`權限。

> **Note**
>
> Lake Formation 不支持授`Describe`予從另一個帳戶共享的數據過濾器的權限。

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 來管理資料篩選器。

如需資料篩選器的資訊，請參閱〈Lake Formation 的數據過濾器

建立資料篩選器

您可以為每個「資料目錄」表格建立一個或多個資料篩選。

為「資料目錄」表格 (主控台) 建立資料篩選的步驟

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、目標資料表擁有者或目標資料表擁有 Lake Formation 權限的主參與者身分簽署。
2. 在導覽窗格的 [資料目錄] 下，選擇 [資料篩選器]。
3. 在 [資料篩選] 頁面上，選擇 [建立新篩選器]。
4. 在 [建立資料篩選] 對話方塊中，輸入下列資訊：

   - 資料篩選器名稱
   - 目標資料庫 — 指定包含表格的資料庫。
   - 目標資料表
   - 欄層級存取 — 將此設定保留為 [存取所有欄]，以僅指定資料列篩選。選擇「包含欄」或「排除欄」以指定欄或儲存格篩選，然後指定要包含或排除的欄。

巢狀資料欄 — 如果您要在包含巢狀資料欄的資料表上套用篩選器，您可以在資料篩選中明確指定巢狀結構資料行的子結構。

當您將 SELECT 權限授與此檔案管理員上的主體時，執行下列查詢的主體只會看到的資料，`customer.customerName`而不`customer.customerId`會看到的資料。

```
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'
```

當您授與customer資料行的權限時，主參與者會收到資料行和資料行下巢狀欄位的存取權(customerName和customerID)。

- 列篩選運算式 — 輸入篩選運算式以指定列或儲存格篩選。如需支援的資料類型和運算子，請參閱資料列篩選運算式中的 PartiQL 支援。選擇存取所有資料列以授與所有資料列的存取權。

  您可以在資料列篩選運算式中包含巢狀資料行的部分資料行結構，以篩選包含特定值的資料列。

  當主參與者被授與具有資料列篩選器運算式之資料表的權限`Select * from example_nestedtable where customer.customerName <>'John'`，且資料行層級存取權設定為對所有資料行的存取權時，查詢結果只會顯示`customerName <>'John'`評估為 true 的資料列。

下列螢幕擷取畫面顯示實作儲存格篩選的資料篩選器。在對資料orders表進行查詢時，它會拒絕對資customer_name料行的存取，而且只會顯示資料行中有「pharma」的資料列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. 選擇 Create filter (建立篩選條件)。

若要在巢狀欄位上使用儲存格篩選原則建立資料篩選器

本節使用下列範例結構描述來顯示如何建立資料儲存格篩選：

```
[
    { 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. 在 [建立資料篩選] 上，輸入資料篩選的名稱。
2. 接下來，使用下拉式清單選擇資料庫名稱和表格名稱。
3. 在「資料欄層級存取權」段落中，選擇「包含的資料欄」，然後選取巢狀資料欄 () `customer.customerName`。
4. 在「列層級存取」區段中，選擇「存取所有列」選項。
5. 選擇 Create filter (建立篩選條件)。

   當您授與此篩選器的`SELECT`權限時，主參與者會取得資料行中所有資料列的`customerName`存取權。
6. 接下來，為同一個數據庫/表定義另一個數據過濾器。
7. 在「資料欄層級存取權」段落中，選擇「包含的資料欄」，然後選取另一個巢狀資料欄 () `customer.customerid`。
8. 在「列層級存取權」區段中，選擇「篩選列」，然後輸入資料列篩選運算式 () `customer.customerid <> 5`。
9. 選擇 Create filter (建立篩選條件)。

   當您授與此篩選器的`SELECT`權限時，主參與者會接收存取`customerName`、和`customerId`欄位中所有資料列的存取權，但資料行中值為 5 的`customerId`儲存格除外。

授與資料篩選權限

您可以將資料篩選器的`SELECT`、`DESCRIBE`和 `DROP` Lake Formation 權限授與主體。

首先，只有您可以檢視為表格建立的資料篩選器。若要啟用另一個主體檢視資料篩選器，並使用資料篩選器授與「資料目錄」權限，您必須：

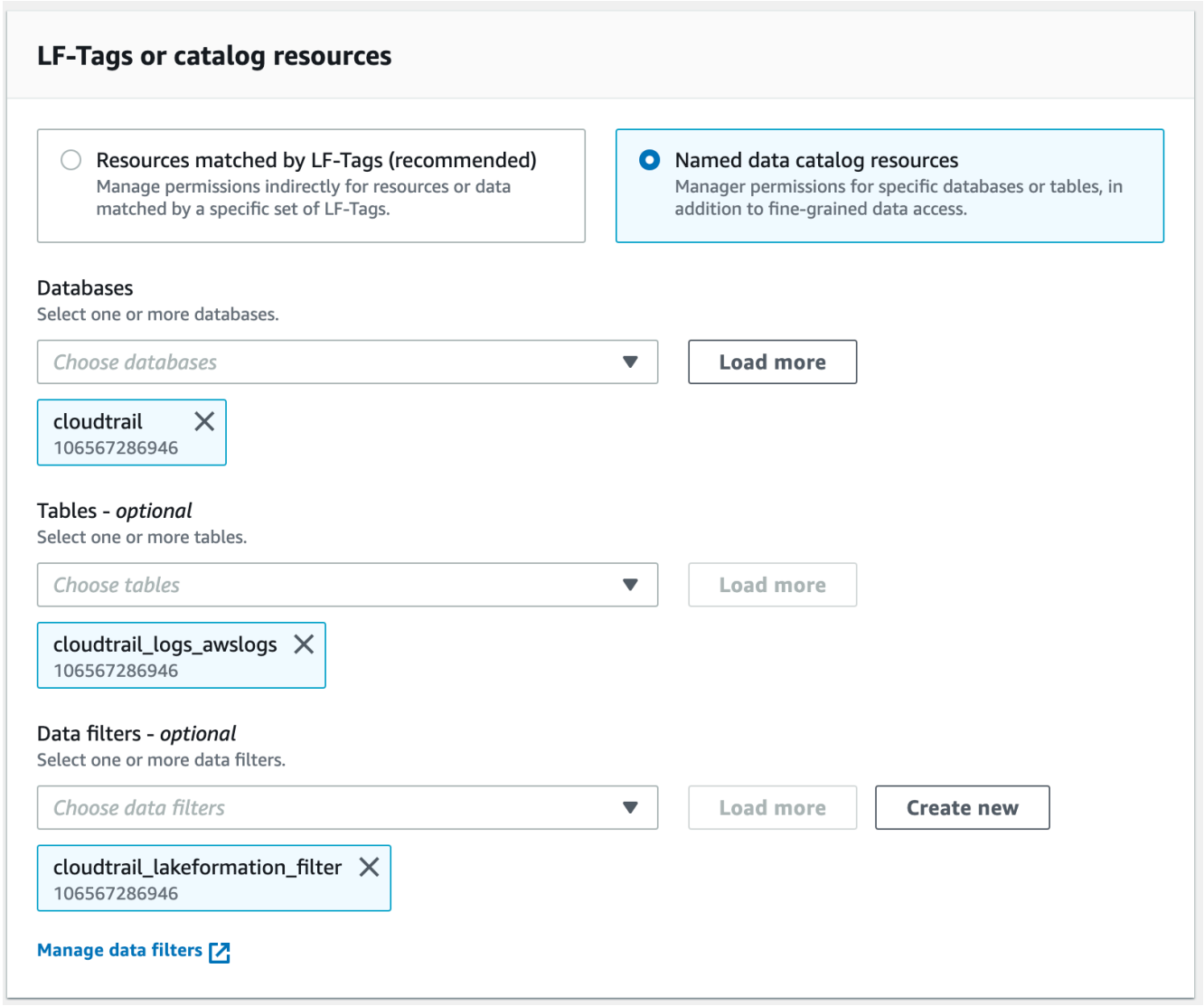
- 使用授SELECT與選項將資料表授與主參與者，然後將資料篩選套用至授權。
- 將資料篩選器的DESCRIBE或DROP權限授與主參與者。

您可以將SELECT權限授予外部 AWS 帳戶。然後，該帳戶中的資料湖管理員可以將該權限授與帳戶中的其他主體。授與外部帳戶時，您必須包含授予選項，以便外部帳戶的管理員可以進一步將權限重疊顯示給其帳戶中的其他使用者。授予帳戶中的主體時，授予授權選項是選擇性的。

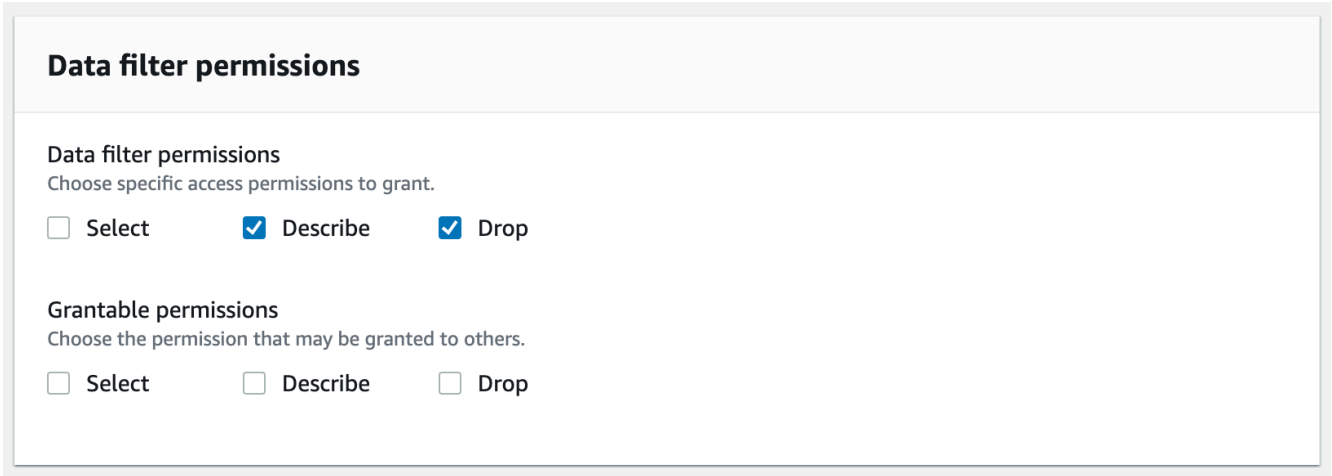
您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 授予和撤銷資料篩選器的權限。

Console

1. 登錄到 AWS Management Console 並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在功能窗格的 [權限] 下，選擇 [資料湖權限]。
3. 在 [權限] 頁面的 [資料權限] 區段中，選擇 [授權]。
4. 在 [授與資料權限] 頁面上，選擇要授與權限的主體。
5. 在 LF 標籤或目錄資源區段中，選擇具名資料目錄資源。然後選擇您要授與權限的資料庫、資料表和資料篩選器。



6. 在 [資料篩選器權限] 區段中，選擇您要授與所選主參與者的權限。



AWS CLI

- 輸入grant-permissions指令。DataCellsFilter為resource引數指定，並DROP為引Permissions數指定DESCRIBE或，以及 (選擇性) 的PermissionsWithGrantOption引數。

  下列範例會在資料篩選器datalake_user1上將授DESCRIBE與選項授與給使用者restrict-pharma，該篩選器屬於 AWS 帳戶 1111-2222-3333 中資料sales庫中的資料orders表。

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

  以下是文件的內容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"]
}
```

授與資料篩選器提供的資料權限

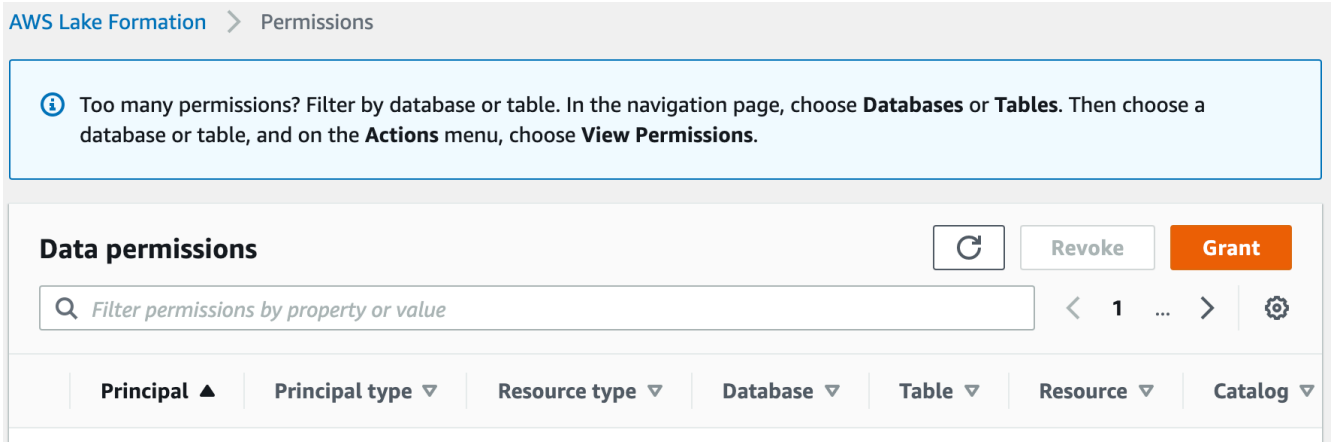
資料篩選器代表資料表內的資料子集。若要提供主參與者的資料存取SELECT權，必須將權限授與給這些主參與者。有了這個權限，主參與者可以：

- 在與其帳戶共用的資料表清單中檢視實際的資料表名稱。
- 在共用資料表上建立資料篩選器，並授與使用者使用這些資料篩選器的權限。

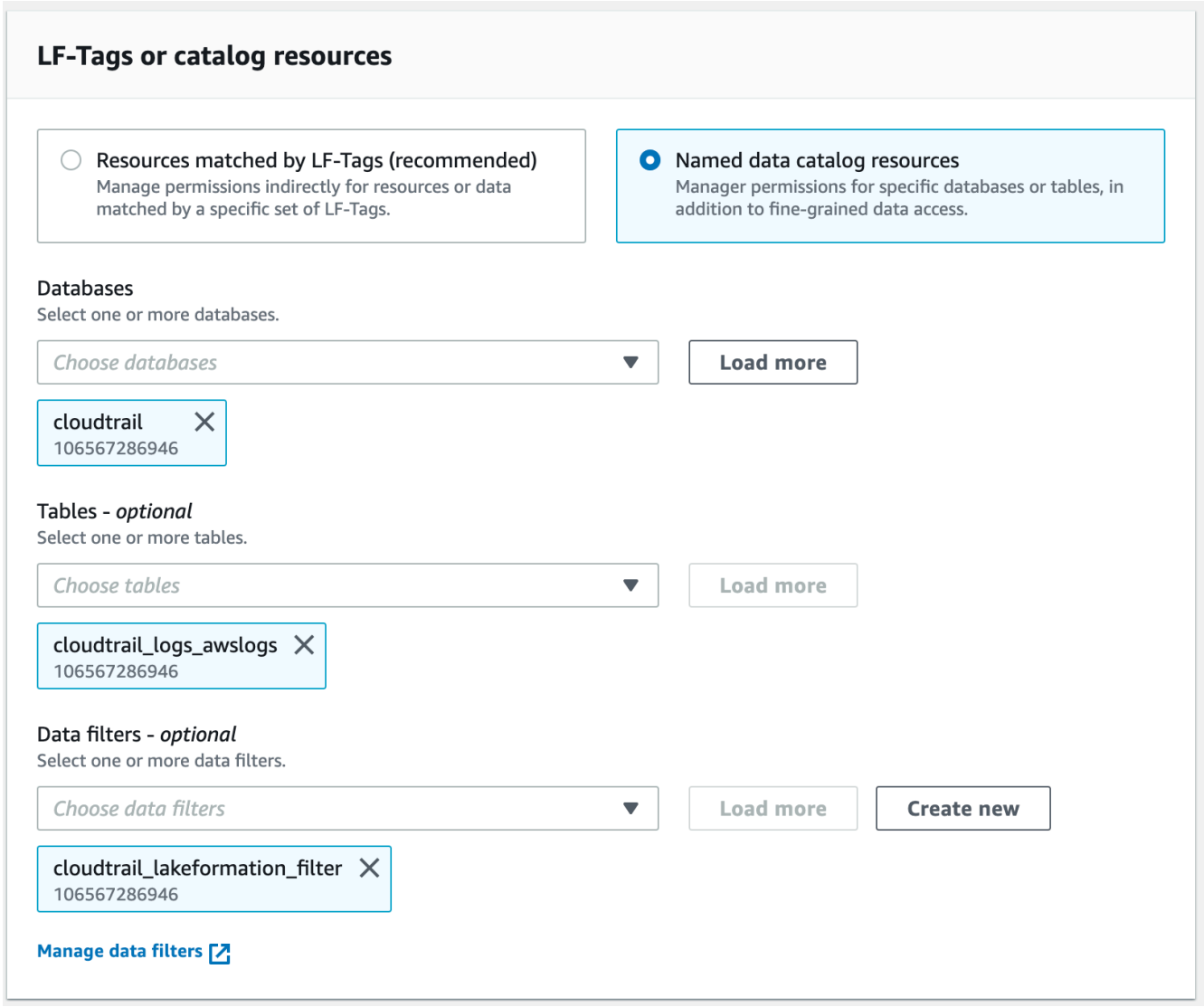
Console

若要授與選取權限

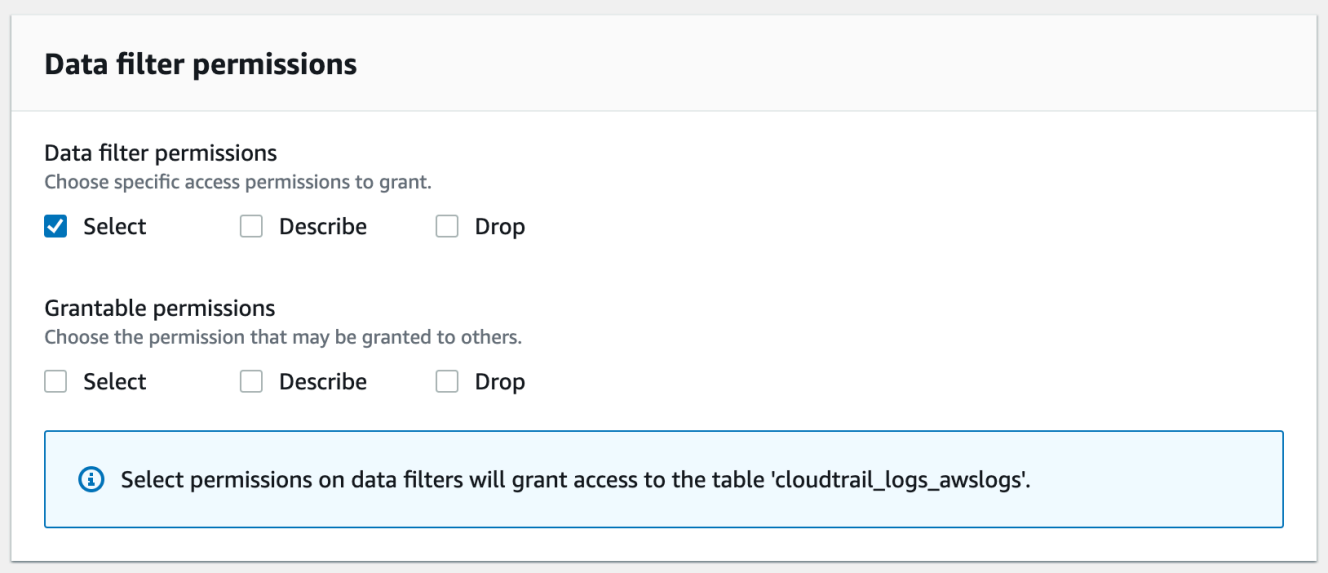
1. 轉到「Lake Formation」控制台中的「權限」頁面，然後選擇「授予」。



2. 選取您要提供存取權的主參與者，然後選取具名資料目錄資源。



3. 若要提供篩選器所代表之資料的存取權，請選擇 [資料篩選器權限] 下的 [選取]。



CLI

輸入grant-permissions指令。DataCellsFilter為資源引數指定，並SELECT為「權限」引數指定。

下列範例會將授SELECT與選項授與資料給使用者 datalake_user1restrict-pharma，該篩選器屬於中資料sales庫中的資料orders表 AWS 帳戶 1111-2222-3333。

```
aws lakeformation grant-permissions --cli-input-json file://grant-params.json
```

以下是文件的內容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"]
```

```
}
```

檢視資料篩選

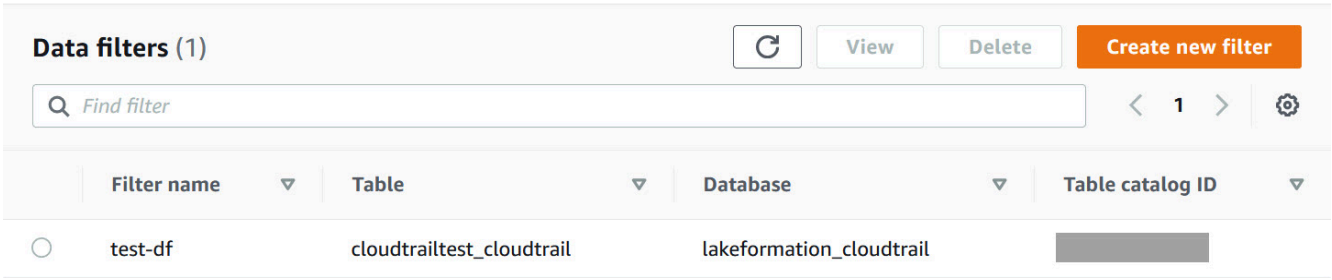
您可以使用 Lake Formation 控制 AWS CLI台或 Lake Formation API 來查看數據過濾器。

若要檢視資料篩選器，您必須是 Data Lake 管理員或擁有資料篩選器的必要權限。

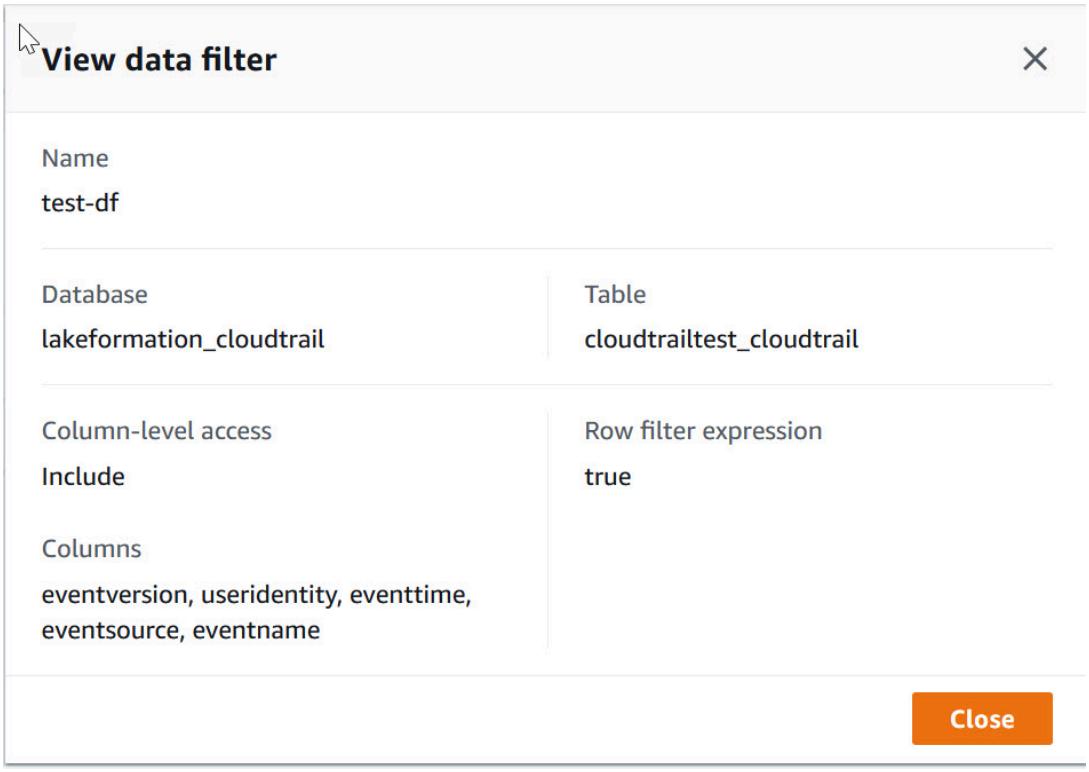
Console

1. 登錄到 AWS Management Console 並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在導覽窗格的 [資料目錄] 下，選擇 [資料篩選器]。

   此頁面會顯示您有權存取的資料篩選器。



3. 若要檢視資料篩選詳細資訊，請選擇資料篩選，然後選擇 [檢視]。出現一個新窗口，其中包含數據過濾器的詳細信息。



AWS CLI

輸入指`list-data-cells-filter`令並指定表格資源。

下列範例會列出資料`cloudtrailtest_cloudtrail`表的資料篩選器。

```
aws lakeformation list-data-cells-filter --table '{ "CatalogId":"123456789012",
"DatabaseName":"lakeformation_cloudtrail", "Name":"cloudtrailtest_cloudtrail"}'
```

API/SDK

使用 `ListDataCellsFilter` API 並指定表格資源。

下列範例使用 Python 列出資料`myTable`表的前 20 個資料篩選器。

```
response = client.list_data_cells_filter(
    Table = {
        'CatalogId': '111122223333',
        'DatabaseName': 'mydb',
        'Name': 'myTable'
    },
    MaxResults=20
```

```
)
```

列出資料篩選權限

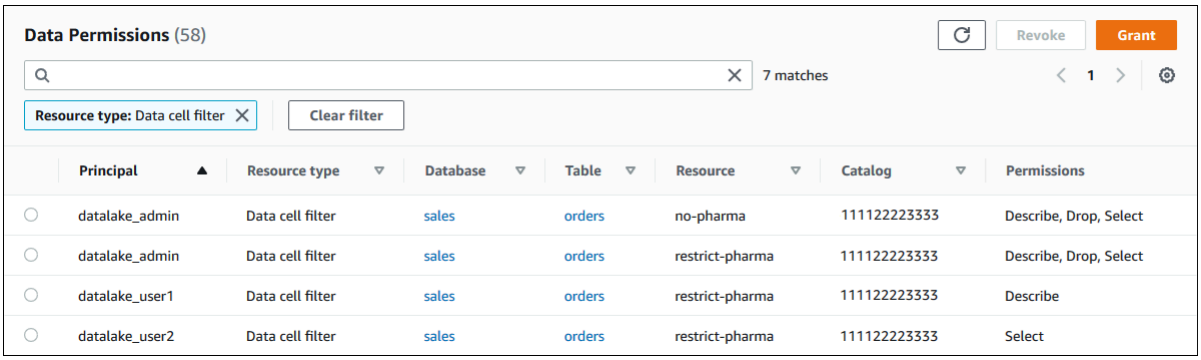
您可以使用 Lake Formation 主控台來檢視資料篩選器授予的權限。

若要檢視資料篩選器的權限，您必須是 Data Lake 管理員或擁有資料篩選器的必要權限。

Console

1. 登錄到 AWS Management Console 並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在功能窗格的 [權限] 下，選擇 [資料權限]。
3. 在 [資料權限] 頁面上，按一下或點選搜尋欄位，然後在 [內容] 功能表上選擇 [資源類型]。
4. 在 [資源類型] 功能表上，選擇 [資源類型:資料儲存格篩選]。

   列出您具有權限的資料篩選器。您可能必須水平捲動才能看到「權限」和「可授權」資料行。



AWS CLI

- 輸入`list-permissions`指令。`DataCellsFilter`為`resource`引數指定，並`DROP`為引`Permissions`數指定`DESCRIBE`或，以及 (選擇性) 的`PermissionsWithGrantOption`引數。

  下列範例會列出`DESCRIBE`資料篩選器上具有授與選項的權限`restrict-pharma`。結果僅限於針對 AWS 帳戶 1111-2222-3333 中`sales`資料庫中的主體`datalake_user1`和資料`orders`表所授與的權限。

```
aws lakeformation list-permissions --cli-input-json file://list-params.json
```

以下是文件的內容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"]
}
```

查看 Lake Formation 中的數據庫和表權限

您可以檢視授與「資料目錄」資料庫或表格的 Lake Formation 權限。您可以通過使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ AWS CLI ）來做到這一點。

使用主控台，您可以從 [資料庫] 或 [表格] 頁面開始，或從 [資料權限] 頁面檢視權限。

> **Note**
>
> 如果您不是資料庫管理員或資源擁有者，則只有在具有授與選項的資源具有 Lake Formation 權限時，才能檢視其他主體對資源擁有的權限。
> 除了必要的 Lake Formation 許可外，您還需要 AWS Identity and Access Management (IAM) 許可glue:GetDatabasesglue:GetDatabaseglue:GetTables、glue:GetTable、和glue:ListPermissions。

檢視資料庫的權限 (主控台，從「資料庫」頁面開始)

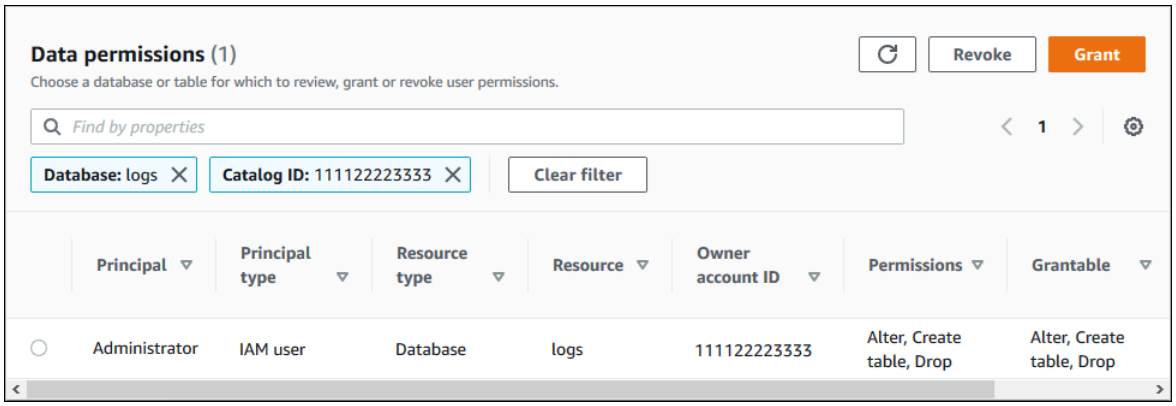
1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、資料庫建立者身分登入，或使用授與選項的資料庫具有任何 Lake Formation 權限的使用者身分登入。

2. 在導覽窗格中，選擇 Databases (資料庫)。
3. 選擇資料庫，然後在 [動作] 功能表上選擇 [檢視權限]。

> **Note**
>
> 如果您選擇資料庫資源連結，Lake Formation 會顯示資源連結的權限，而不會顯示資源連結的目標資料庫上的權限。

[資料] 權限頁面會列出資料庫的所有 Lake Formation 權限。資料庫擁有者的資料庫名稱和目錄 ID (AWS 帳戶 ID) 會在搜尋方塊下顯示為標籤。圖標表示篩選器已套用至僅適用於該資料庫的清單權限。您可以關閉拼貼或選擇「清除濾鏡」來調整濾鏡。



檢視資料庫的權限 (主控台，從 [資料權限] 頁面開始)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、資料庫建立者身分登入，或使用授與選項的資料庫具有任何 Lake Formation 權限的使用者身分登入。
2. 在導覽窗格中，選擇 [資料權限]。
3. 將游標置於頁面頂端的搜尋方塊中，然後在出現的 [內容] 功能表上選擇 [資料庫]。
4. 在出現的「資料庫」功能表上，選擇資料庫。

> **Note**
>
> 如果您選擇資料庫資源連結，Lake Formation 會顯示資源連結的權限，而不會顯示資源連結的目標資料庫上的權限。

[資料] 權限頁面會列出資料庫的所有 Lake Formation 權限。資料庫名稱會在搜尋方塊下顯示為圖標。圖標表示篩選器已套用至僅適用於該資料庫的清單權限。您可以關閉動態磚或選擇 [清除篩選器] 來移除篩選器。

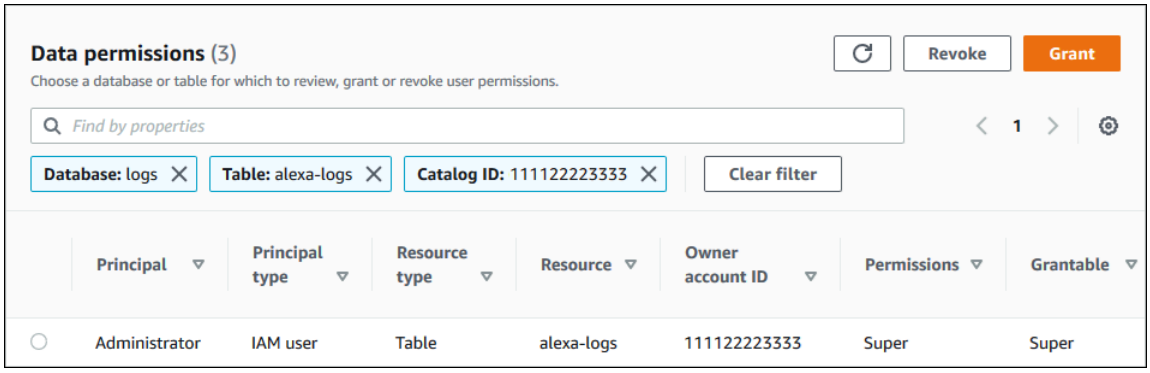
若要檢視資料表的權限 (主控台，從 [表格] 頁面開始)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員、表格建立者身分登入，或使用授與選項在表格上具有任何 Lake Formation 權限的使用者身分登入。
2. 在導覽窗格中，選擇 Tables (資料表)。
3. 選擇表格，然後在 [動作] 功能表上選擇 [檢視權限]。

> **Note**
>
> 如果您選擇表格資源連結，Lake Formation 會顯示資源連結的權限，而不會顯示在資源連結的目標表格上。

[資料] 權限頁面會列出表格的所有 Lake Formation 權限。表格名稱、包含表格之資料庫的資料庫名稱，以及表格擁有者的目錄 ID (AWS 帳戶 ID) 會顯示為搜尋方塊下方的標籤。標籤表示篩選器已套用至僅該表格的清單權限。您可以透過關閉標籤或選擇「清除濾鏡」來調整濾鏡。



若要檢視資料表的權限 (主控台，從 [資料權限] 頁面開始)

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

以資料湖管理員、表格建立者身分登入，或使用授與選項在表格上具有任何 Lake Formation 權限的使用者身分登入。

2. 在導覽窗格中，選擇 [資料權限]。
3. 將游標置於頁面頂端的搜尋方塊中，然後在出現的 [內容] 功能表上選擇 [資料庫]。
4. 在出現的「資料庫」功能表上，選擇資料庫。

> **Important**
>
> 如果您想要檢視從外部帳戶與您的 AWS 帳戶共用之資料表的權限，您必須在包含該表格的外部帳戶中選擇資料庫，而不是資料庫的資源連結。

[資料] 權限頁面會列出資料庫的所有 Lake Formation 權限。

5. 再次將游標置於搜尋方塊中，然後在出現的 [內容] 功能表上選擇 [表格]。
6. 在顯示的「表格」功能表上，選擇表格。

[資料] 權限頁面會列出表格的所有 Lake Formation 權限。包含表格之資料庫的表格名稱和資料庫名稱會在搜尋方塊下顯示為並排。圖標表示篩選器已套用至僅適用於該資料表的清單權限。您可以關閉拼貼或選擇「清除濾鏡」來調整濾鏡。

若要檢視資料表的權限 (AWS CLI)

- 輸入`list-permissions`指令。

下列範例會列出從外部帳戶共用之資料表的權限。`CatalogId`屬性是外部 AWS 帳戶的帳戶 ID，而資料庫名稱是指包含資料表之外部帳戶中的資料庫。

```
aws lakeformation list-permissions  --resource-type TABLE --resource '{ "Table":
 {"DatabaseName":"logs", "Name":"alexa-logs", "CatalogId":"123456789012"}}'
```

使用 Lake Formation 控制台撤銷權限

您可以使用主控台撤銷所有類型的 Lake Formation 權限 — 資料目錄權限、原則標記權限、資料篩選權限和位置權限。

撤銷資源（控制台）上的 Lake Formation 權限

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員或已獲得資源授與權限的使用者身分登入。
2. 在導覽窗格的 [權限] 下，選擇 [資料湖權限]、[LF 標籤和權限] 或 [資料位置]。
3. 選取權限或位置，然後選擇 [撤銷]。
4. 在開啟的對話方塊中，選擇「撤銷」。

Lake Formation 的跨賬戶數據共享

Lake Formation 跨帳戶功能可讓使用者在多個 AWS 組織間安全地共用分散式資料湖 AWS 帳戶，或直接與其他帳戶中的 IAM 主體共用分散式資料湖，從而提供對資料型錄中繼資料和基礎資料的精細存取。大型企業通常使用多個帳戶 AWS 帳戶，其中許多帳戶可能需要存取由單一管理的資料湖 AWS 帳戶。使用者和 AWS Glue 擷取、轉換和載入 (ETL) 工作可以查詢和聯結多個帳戶之間的資料表，並且仍可利用 Lake Formation 資料表層級和資料行層級資料保護。

當您將資料目錄資源的 Lake Formation 權限授與外部帳戶或直接授與其他帳戶中的 IAM 主體時，Lake Formation 會使用 AWS Resource Access Manager (AWS RAM) 服務來共用資源。如果受權者帳號與授與者帳號位於相同的組織中，則共用資源可立即供受權者使用。如果受權者帳號不在同一個組織中，則 AWS RAM 會傳送邀請給受權者帳戶，以接受或拒絕資源授與。然後，若要使共用資源可用，受權者帳戶中的資料湖管理員必須使用 AWS RAM 主控台或 AWS CLI 接受邀請。

Lake Formation 支援以混合式存取模式與外部帳戶共用資料目錄資源。混合式存取模式提供彈性 Lake Formation 可選擇性地啟用 AWS Glue Data Catalog.
透過混合式存取模式，您現在擁有一個增量路徑，可讓您為一組特定使用者設定 Lake Formation 權限，而不會中斷其他現有使用者或工作負載的權限原則。

如需詳細資訊，請參閱 混合存取模式。

直接跨帳戶共享

授權的主體可以與外部帳戶中的 IAM 主體明確共用資源。當帳戶擁有者想要控制外部帳戶中的哪些人可以存取資源時，此功能非常有用。IAM 主體收到的許可將是直接授權的聯合，以及串聯至主體的帳戶層級授權。收件者帳戶的資料湖管理員可以檢視直接跨帳戶授與，但無法撤銷權限。接收資源共用的主參與者無法與其他主參與者共用資源。

共用資料目錄資源的方法

透過單一 Lake Formation 授與作業，您可以授與下列資料目錄資源的跨帳戶權限。

- 一個數據庫
- 一個單獨的表（具有可選的列過濾）
- 幾個選擇的表
- 資料庫中的所有資料表 (使用「所有資料表」萬用字元)

有兩個選項可與另一個帳戶中的其他主體 AWS 帳戶 或 IAM 主體共用您的資料庫和表格。

- 基於 Lake Formation 標籤的訪問控制（LF-TBAC）（推薦）

  以 Lake Formation 標籤為基礎的存取控制是一種授權策略，可根據屬性定義權限。您可以使用以標籤為基礎的存取控制，與外部 IAM 主體、組織和組織單位 (OU) 共用資料目錄資源 (資料庫 AWS 帳戶、表格和欄)。在 Lake Formation，這些屬性被稱為 LF-標籤。如需詳細資訊，請參閱使用以 Lake Formation 標籤為基礎的存取控制來管理資料湖。

  > **Note**
  >
  > 授與資料目錄權限的 LF-TBAC 方法，用 AWS Resource Access Manager 於跨帳戶授與。Lake Formation 現在支持使用 LF-TBAC 方法向組 Organizations 和組織單位授予跨帳戶權限。
  > 若要啟用此功能，您需要將跨帳戶版本設定更新為版本 3。
  > 如需詳細資訊，請參閱 更新跨帳戶資料共用版本設定。

- Lake Formation 命名資源

  使用具名資源方法的 Lake Formation 跨帳戶資料共用，可讓您將 Lake Formation 權限以及「資料目錄」表格和資料庫的授與選項授與外部 AWS 帳戶、IAM 主體、組織或組織單位。授權作業會自動共用這些資源。

> **Note**
>
> 您也可以允許 AWS Glue 爬蟲使用 Lake Formation 認證存取其他帳戶中的資料存放區。如需詳細資訊，請參閱 AWS Glue 開發人員指南中的跨帳戶檢索。

Athena 和 Amazon Redshift Spectrum 等整合式服務需要資源連結才能在查詢中包含共用資源。如需有關資源連結的詳細資訊，請參閱資源連結在 Lake Formation 中如何運作。

如需考量和限制，請參閱跨帳戶資料共用最佳做法與考量。

主題

相關主題

- Lake Formation 許可權概述
- 存取和檢視共用資料目錄表格和資料庫
- 建立資源連結
- 跨帳戶存取疑難排解

必要條件

您的 AWS 帳戶才能與另一個帳戶中的其他帳戶或主參與者共用 Data Catalog 資源 (資料庫和表格)，以及存取與帳戶共用的資源之前，必須符合下列先決條件。

一般跨帳戶資料共用需求

- 若要以混合式存取模式共用資料目錄資料庫和表格，您需要將跨帳戶版本設定更新為第 4 版。
- 在授與資料目錄資源的跨帳戶權限之前，您必須撤銷該資源之`IAMAllowedPrincipals`群組的所有 Lake Formation 權限。如果呼叫主體具有跨帳戶存取資源的權限，且資源上存在`IAMAllowedPrincipals`權限，則 Lake Formation 會擲回`AccessDeniedException`。

僅當您在 Lake Formation 模式下註冊基礎資料位置時，此要求才適用。如果您以混合模式註冊資料位置，則IAMAllowedPrincipals群組權限可存在於共用資料庫或資料表上。

- 對於包含您想要共用之資料表的資料庫，您必須防止新資料表具有預設授權Super給IAMAllowedPrincipals。在 Lake Formation 主控台上，編輯資料庫並關閉僅對此資料庫中的新資料表使用 IAM 存取控制，或輸入以下 AWS CLI 指令，以資料庫名稱取代database。如果基礎資料位置是以混合式存取模式註冊，則不需要變更此預設設定。在混合式存取模式下，Lake Formation 可讓您針對 Amazon S3 和相同資源選擇性地強制執行 Lake Formation 許可和 AWS Glue IAM 許可政策。

```
aws glue update-database --name database --database-input
 '{"Name":"database","CreateTableDefaultPermissions":[]}'
```

- 若要授與跨帳戶許可，授與者必須在AWS Glue和 AWS RAM 服務上擁有必要的 AWS Identity and Access Management (IAM) 許可。受 AWS 管理的原則會AWSLakeFormationCrossAccountManager授與必要的權限。

  使用接收資源共用之帳號中的資料湖管理員 AWS RAM 必須具有下列其他策略。它允許管理員接受 AWS RAM 資源共用邀請。它還允許管理員啟用與組織的資源共享。

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

- 如果您要與 AWS Organizations 組織單位共用資料目錄資源，則必須在中啟用與組織共用 AWS RAM。

如需有關如何啟用與組織共用的資訊，請參閱《AWS RAM 使用指南》中的「啟用與 AWS 組織共用」。

您必須擁有啟用與組織共用的`ram:EnableSharingWithAwsOrganization`權限。

- 若要直接與另一個帳戶中的 IAM 主體共用資源，您需要將跨帳戶版本設定更新為第 3 版。此設定位於 [資料目錄設定] 頁面上。如果您使用的是版本 1，請參閱更新設定的指示更新跨帳戶資料共用版本設定。
- 您無法與其他帳戶共用以 AWS Glue 服務受管金鑰加密的資料目錄資源。您只能共用使用客戶加密金鑰加密的資料目錄資源，且接收資源共用的帳號必須具有資料目錄加密金鑰的權限，才能解密物件。

使用 LF-TBAC 要求進行跨帳戶資料共用

- 若要 AWS Organizations 與組織單位 (OU) 共用資料目錄資源，您需要將跨帳戶版本設定更新為第 3 版。
- 若要與跨帳戶版本設定的第 3 版共用資料目錄資源，授與者需要在您帳戶的 AWS 受管政策`AWSLakeFormationCrossAccountManager`中定義 IAM 許可。
- 如果您使用的是第 1 版或第 2 版的跨帳戶版本設定，則必須具有啟用 LF-TBAC 的資料目錄資源策略 (`glue:PutResourcePolicy`)。如需詳細資訊，請參閱 使用AWS Glue和 Lake Formation 管理跨帳戶權限。
- 如果您目前正在使用資AWS Glue料目錄資源策略來共用資源，並且想要使用第 3 版的交叉帳戶版本設定授與跨帳戶權限，則必須使用 `glue:PutResourcePolicy` API 作業在「資料目錄設定」中新增權限，如使用AWS Glue和 Lake Formation 管理跨帳戶權限區段所示。`glue:ShareResource`如果您的帳戶沒有使用資AWS Glue料目錄資源策略 (版本 1 和第 2 版使用`glue:PutResourcePolicy`權限) 來授與跨帳戶存取權限，則不需要此原則。

```
{
      "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"
      ]
    }
```

- 如果您的帳戶已使用資AWS Glue料目錄資源策略建立跨帳戶共用，且您目前正在使用具名資源方法或具有跨帳戶設定版本 3 的 LF-TBAC 來共用資源 (用於共用 AWS RAM 資源)，則必須在呼叫 API 作業'true'時將EnableHybrid引數設定為。glue:PutResourcePolicy如需詳細資訊，請參閱使用AWS Glue和 Lake Formation 管理跨帳戶權限。

存取共用資源的每個帳號都需要進行設定

- 如果您要與共用資源 AWS 帳戶，則取用者帳戶中至少有一個使用者必須是資料湖管理員，才能檢視共用資源。如需如何建立資料湖管理員的資訊，請參閱建立資料湖管理員。

  資料湖管理員可以將共用資源的 Lake Formation 權限授與帳戶中的其他主體。在資料湖管理員授與資源的權限之前，其他主體無法存取共用資源。
- Athena 和 Redshift 頻譜等整合式服務需要資源連結才能在查詢中包含共用資源。主參與者必須在其「資料目錄」中建立資源連結，連至另一個 AWS 帳戶共用資源。如需有關資源連結的詳細資訊，請參閱資源連結在 Lake Formation 中如何運作。
- 直接與 IAM 主體共用資源時，若要使用 Athena 查詢資料表，主體需要建立資源連結。若要建立資源連結，主體需要 Lake Formation CREATE_TABLE 或CREATE_DATABASE權限，以及glue:CreateTable或 glue:CreateDatabase IAM 權限。

  如果生產者帳戶與相同或另一個主體共用相同資料庫下的不同資料表，則該主體可以立即查詢資料表。

> **Note**
>
> 對於資料湖管理員以及資料湖管理員已授與權限的主參與者而言，共用資源會顯示在「資料目錄」中，就像它們是本機 (擁有的) 資源一樣。擷取、轉換和載入 (ETL) 工作可以存取共用資源的基礎資料。
> 對於共用資源，Lake Formation 主控台上的 [資料表] 和 [資料庫] 頁面會顯示擁有者的帳號 ID。
> 存取共用資源的基礎資料時，會在共用資源收件者的帳號和資源擁有者的帳號中產生 CloudTrail 記錄事件。 CloudTrail 事件可以包含存取資料之主體的 ARN，但只有當收件者帳戶選擇在記錄檔中包含主體 ARN 時才會顯示。如需詳細資訊，請參閱 跨帳戶 CloudTrail 記錄。

更新跨帳戶資料共用版本設定

不時 AWS Lake Formation 更新跨帳戶資料共用設定，以區分對 AWS RAM 使用情況所做的變更，並支援跨帳戶資料共用功能的更新。當 Lake Formation 這樣做時，它會創建一個新版本的跨帳戶版本設置。

跨帳戶版本設置之間的主要區別

如需跨帳戶資料共用如何在不同跨帳戶版本設定下運作的詳細資訊，請參閱下列章節。

> **Note**
>
> 若要與其他帳戶共用資料，授與者必須具有`AWSLakeFormationCrossAccountManager`受管 IAM 政策許可。這是所有版本的先決條件。
> 更新跨帳戶版本設定不會影響收件者對共用資源的權限。這適用於從版本 1 更新至版本 2、版本 2 至版本 3，以及版本 1 更新至版本 3 時。更新版本時，請參閱下列注意事項。

第一版

具名資源方法：將每個跨帳戶 Lake Formation 權限授予對應至一個 AWS RAM 資源共享。使用者 (授與者角色或主參與者) 不需要其他權限。

LF-TBAC 方法：跨帳戶 Lake Formation 許可權授予不用於共享數據。 AWS RAM 使用者必須擁有`glue:PutResourcePolicy`權限。

更新版本的好處：初始版本-不適用。

更新版本時的注意事項：初始版本-不適用

2 版

具名資源方法：透過將多個跨帳戶權限授與與一個 AWS RAM 資源共用對應，以最佳化 AWS RAM 資源共用數目。使用者不需要其他權限。

LF-TBAC 方法：跨帳戶 Lake Formation 許可權授予不用於共享數據。 AWS RAM 使用者必須擁有`glue:PutResourcePolicy`權限。

更新版本的好處：透過最佳的 AWS RAM 容量使用率進行擴充的跨帳戶設定。

更新版本時的考量事項：想要授與跨帳戶 Lake Formation 權限的使用者必須具有受`AWSLakeFormationCrossAccountManager` AWS 管理策略中的權限。否則，您必須擁

有ram:AssociateResourceShare和ram:DisassociateResourceShare權限才能與其他帳戶成功共用資源。

第 3 版

具名資源方法：透過將多個跨帳戶權限授與與一個 AWS RAM 資源共用對應，以最佳化 AWS RAM 資源共用數目。使用者不需要其他權限。

LF-TBAC 方法：Lake Formation 用 AWS RAM 於跨帳戶補助金。用戶必須添加膠水：ShareResource 聲明的glue:PutResourcePolicy權限。收件者必須接受來自的資源共用邀請 AWS RAM。

更新版本的好處：支援下列功能：

- 允許與外部帳戶中的 IAM 主體明確共用資源。

  如需詳細資訊，請參閱 授與和撤銷資料目錄資源的權限。
- 使用 LF-TBAC 方法對組織或組織單位 (OU) 啟用跨帳戶共用。
- 免除維護跨帳戶授權額外 AWS Glue 原則的額外負荷。

更新版本時的考量：當您使用 LF-TBAC 方法共用資源時，如果授與者使用的版本低於版本 3，且收件者使用的是版本 3 或更高版本，則授與者會收到下列錯誤訊息：「無效的跨帳戶授與請求。消費者帳戶已選擇加入跨帳戶版本：v3。請更新CrossAccountVersionDataLakeSetting至最低版本 v3 (服務: AmazonDataCatalog; 狀態碼:400; 錯誤代碼: InvalidInputException)」. 但是，如果授與者使用版本 3，而收件者使用的是版本 1 或版本 2，則使用 LF 標籤的跨帳戶授權會順利完成。

使用指定資源方法進行的跨帳戶授權在不同版本之間相容。即使授與者帳戶使用的是舊版本 (版本 1 或 2)，且收件者帳戶使用的是較新的版本 (版本 3 或更高版本)，跨帳戶存取功能也能順暢運作，而不會出現任何相容性問題或錯誤。

若要直接與其他帳戶中的 IAM 主體共用資源，只有授與者需要使用版本 3。

使用 LF-TBAC 方法進行的跨帳戶授權要求使用者在帳號中具有 AWS Glue Data Catalog 資源策略。當您更新到版本 3 時，LF-TBAC 授予使用。 AWS RAM若要允許以跨帳戶為 AWS RAM 基礎的授權成功，您必須將glue:ShareResource陳述式新增至現有的資料目錄資源策略，如使用 AWS Glue和 Lake Formation 管理跨帳戶權限區段所示。

第 4 版

授與者需要版本 4 或更高版本，才能以混合存取模式共用資料目錄資源。

最佳化 AWS RAM 資源共用

跨帳戶的新版本（第 2 版及以上版本）可以最佳地利用 AWS RAM 能力來最大限度地提高跨帳戶使用率。當您與外部 AWS 帳戶 或 IAM 主體共用資源時，Lake Formation 可能會建立新的資源共用，或將資源與現有共用產生關聯。通過與現有股份建立關聯，Lake Formation 減少了消費者需要接受的資源共享邀請的數量。

透過 TBAC 啟用 AWS RAM 共用，或直接將資源共用給主體

若要直接與其他帳戶中的 IAM 主體共用資源，或為組織或組織單位啟用 TBAC 跨帳戶共用，您需要將跨帳戶版本設定更新為第 3 版。如需有關資 AWS RAM 源限制的更多資訊，請參閱跨帳戶資料共用最佳做法與考量。

更新跨帳戶版本設定所需的權限

如果跨帳戶權限授與者具有AWSLakeFormationCrossAccountManager受管 IAM 政策許可，則跨帳戶權限授與者角色或主體不需要額外的權限設定。不過，如果跨帳戶授與者未使用受管政策，則授與者角色或主體應該具有下列 IAM 許可，以便新版跨帳戶授與成功。

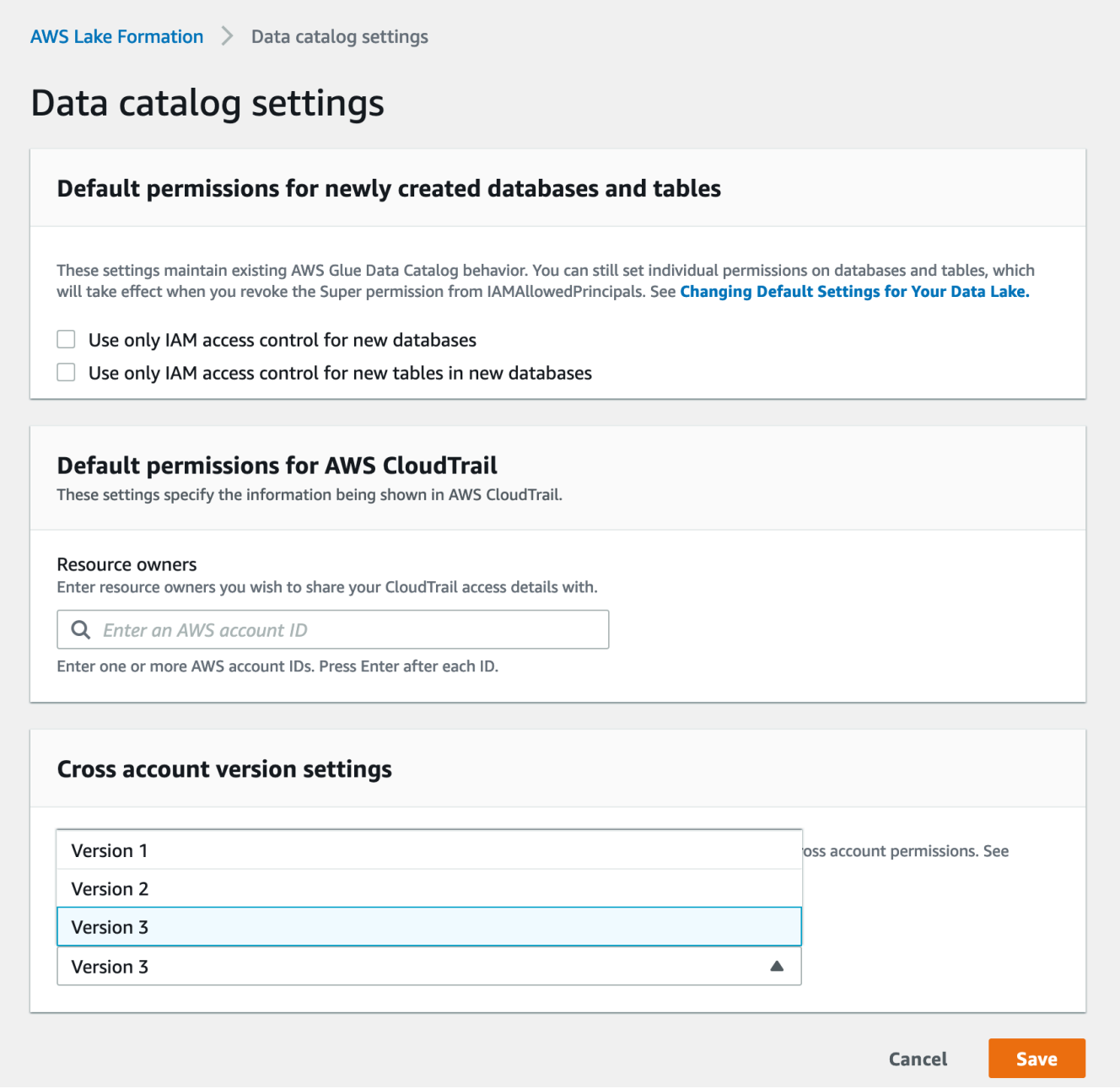
```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor1",
      "Effect": "Allow",
      "Action": [
         "ram:AssociateResourceShare",
         "ram:DisassociateResourceShare",
         "ram:GetResourceShares"
       ],
     "Resource": "*",
     "Condition": {
       "StringLike": {
         "ram:ResourceShareName": "LakeFormation*"
        }
      }
    }
  ]
}
```

若要啟用新版本

請依照下列步驟，透過 AWS Lake Formation主控台或更新跨帳戶版本設定 AWS CLI。

Console

1. 在 [資料目錄設定] 頁面上的 [跨帳戶版本設定] 下選擇 [版本 2]、[版本 3] 或 [版本 4]。如果您選擇第 1 版，Lake Formation 將使用預設的資源共享模式。



2. 選擇儲存。

AWS Command Line Interface (AWS CLI)

使用`put-data-lake-settings` AWS CLI 指令設定`CROSS_ACCOUNT_VERSION`參數。接受的值為 1、2、3 和 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**
>
> 選擇版本 2 或版本 3 之後，所有新的具名資源補助都會經過新的跨帳戶授權模式。若要以最佳方式使用現有跨帳戶共用的 AWS RAM 容量，我們建議您撤銷舊版本的授權，然後在新模式下重新授予。

跨外部帳戶 AWS 帳戶 或 IAM 主體共用資料目錄表格和資料庫

本節包含如何對外部帳戶、IAM 主體、組織或組織單位啟用 Data Catalog 資料表和資料庫跨 AWS 帳戶許可的指示。授權作業會自動共用這些資源。

主題

- 使用基於標籤的訪問控制共享數據
- 使用指定的資源方法進行跨帳戶資料共用

使用基於標籤的訪問控制共享數據

生產者/授權人帳戶需要設定

1. 定義 LF 標籤。如需建立 LF 標籤的指示，請參閱。創建 LF-標籤
2. 將 LF 標籤指派給目標資源。如需詳細資訊，請參閱 將 LF 標籤指定給資料目錄資源。
3. 將 LF 標籤權限授予外部帳戶。如需詳細資訊，請參閱 使用控制台授予 LF 標籤權限。

   此時，取用者資料湖系統管理員應該能夠在 [權限]、[系統管理角色和工作 LF-tag] 下，找到透過受權者帳戶 L ake Formation 主控台共用的原則標記。
4. 將資料權限授與外部/受權人帳戶。
   a. 在功能窗格的 [權限] 下的 [資料湖權限] 下，選擇 [授與]。
   b. 對於主體，請選擇外部帳戶，然後輸入主體的目標 AWS 帳戶 ID 或 IAM 角色，或輸入主體 (主體 ARN) 的 Amazon 資源名稱 (ARN)。
   c. 對於 LF 標籤或目錄資源，請選擇要與消費者帳戶共用的 LF 標籤的索引鍵和值 (索引鍵**Confidentiality**和值)。 `public`
   d. 對於「權限」，在「由 LF 標籤匹配的資源（推薦）」下選擇「添加 LF 標籤」。
   e. 選取要與受權者帳戶 (索引鍵和值`public`) 共用之標籤的索引鍵`Confidentiality`和值。
   f. 對於資料庫權限，請選取 [資料庫權限] 下的 [說明] 以授與資料庫層級的存取權限
   g. 消費者資料湖系統管理員應該能夠在 Lake Formation 主控台 https://console.aws.amazon.com/lakeformation/ 的「權限」、「系統管理角色和工作」、「LF- tag」下，找到透過消費者帳戶共用的原則標籤。
   h. 選取 [可授與權限] 下的 [說明]，讓消費者帳戶可以將資料庫層級權限授與其使用者。

      由於資料湖管理員必須將共用資源的權限授與受權者帳戶中的主體，因此必須始終以授與選項授與跨帳戶權限授與。

      > **Note**
      >
      > 收到直接跨帳戶授權的主參與者不會有「可授與」權限選項。

   i. 對於資料表和資料行權限，請選取表格權限下的選取並說明。
   j. 選取 [可授與權限] 下的 [選取並說明]。
   k. 選擇 Grant (授予)。

在收款/受授權人帳戶上設置

1. 當您與其他帳號共用資源時，該資源仍屬於生產者帳號，而且在 Athena 主控台中看不到。若要在 Athena 主控台中顯示資源，您需要建立指向共用資源的資源連結。如需建立資源連結的指示，請參閱建立共用資料目錄表格的資源連結和 建立共用資料目錄資料庫的資源連結
2. 您需要在使用者帳戶中建立一組個別的 LF 標籤，才能在共用資源連結時使用以 LF 標籤為基礎的存取控制。創建所需的 LF 標籤並將其分配給共享數據庫/表和資源鏈接。
3. 將這些 LF 標籤的許可授與受權者帳戶中的 IAM 主體。

使用指定的資源方法進行跨帳戶資料共用

您可以將權限直接授與其他 AWS 帳戶中的主參與者，或授與外部 AWS 帳戶 或 AWS Organizations。將 Lake Formation 權限授予組 Organizations 或組織單位等同於將權限授予該組織或組織單位 AWS 帳戶 中的每個人。

當您將權限授與外部帳戶或組織時，必須包含「可授與的權限」 選項。只有外部帳戶中的資料湖管理員才能存取共用資源，直到管理員將共用資源的權限授與外部帳戶中的其他主體為止。

> **Note**
>
> 從外部帳戶直接授與 IAM 主體時，不支援可授予許可權限選項。

遵循中使用指定的資源方法授與資料庫權限的指示，使用指定的資源方法授與跨帳戶權限。

授與與您帳戶共用的資料庫或資料表的權限

將屬於其他 AWS 帳號的資料目錄資源與您的 AWS 帳戶共用後，身為資料湖管理員，您可以將共用資源的權限授與帳戶中的其他主體。但是，您無法將資源的權限授與其他 AWS 帳號或組織。

您可以使用 AWS Lake Formation 控制台、API 或 AWS Command Line Interface (AWS CLI) 授予權限。

授與共用資料庫的權限 (具名資源方法、主控台)

- 請遵循中的說明進行使用指定的資源方法授與資料庫權限 在 LF 標籤或目錄資源下的「資料庫」清單中，請確定您在外部帳戶中選取資料庫，而不是資料庫的資源連結。

如果您在資料庫清單中沒有看到資料庫，請確定您已接受資料庫的 AWS Resource Access Manager (AWS RAM) 資源共用邀請。如需詳細資訊，請參閱 接受來自的資源共用邀請 AWS RAM。

此外，對於CREATE_TABLE和ALTER權限，請按照中的說明進行操作授予資料位置權限 (相同帳戶)，並確保在 [註冊帳戶位置] 欄位中輸入擁有帳戶 ID。

授與共用資料表的權限 (具名資源方法、主控台)

- 請遵循中的說明進行使用指定的資源方法授與資料表權限 在 LF 標籤或目錄資源下的「資料庫」清單中，請確定您在外部帳戶中選取資料庫，而不是資料庫的資源連結。

  如果您在表格清單中看不到該表格，請確定您已接受表格的 AWS RAM 資源共用邀請。如需詳細資訊，請參閱 接受來自的資源共用邀請 AWS RAM。

  此外，若要取ALTER得權限，請遵循中的指示授予資料位置權限 (相同帳戶)，並務必在 [註冊帳戶位置] 欄位中輸入擁有帳戶 ID。

若要授與共用資源的權限 (LF-TBAC 方法、主控台)

- 請遵循中的說明進行授與資料目錄權限 在 LF 標籤或目錄資源區段中，授與外部帳戶授與您帳戶的確切 LF 標籤運算式，或該運算式的子集。

  例如，如果外部帳戶以資料湖管理員身分module=customers AND environment=production將 LF 標籤運算式授與您的帳戶，您可以environment=production將該運算式授與module=customers或授與帳戶中的主體。您只能授與透過 LF-tag 運算式在資源上授予的相同或一部分 Lake Formation 權限 (例如ALTER、等)。SELECT

若要授與共用資料表的權限 (已命名的資源方法， AWS CLI)

- 輸入與以下相似的命令。在此範例中：
  - 你的 AWS 帳戶編號是
  - 擁有該表格且授與您帳戶的帳戶的帳戶是 1234-5678-9012。
  - 共用資料表上的SELECT權限正在授予pageviews使用者datalake_user1。該使用者是您帳戶中的主體。

- 資料pageviews表位於資料庫中，該analytics資料庫是由帳戶 1234-5678-9012 所擁有。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
 --permissions "SELECT" --resource '{ "Table": {"CatalogId":"123456789012",
 "DatabaseName":"analytics", "Name":"pageviews"}}'
```

請注意，必須在resource引數的CatalogId屬性中指定擁有帳戶。

授與資源連結權限

請遵循下列步驟來 AWS Lake Formation 授與 AWS 帳戶中主參與者的一或多個資源連結的權限。

建立資源連結後，只有您可以檢視和存取它。這假設資料庫未啟用此資料庫中新資料表的「僅使用 IAM 存取控制」。) 若要允許您帳號中的其他主參與者存取資源連結，請至少授與DESCRIBE權限。

> **Important**
>
> 授與資源連結的權限不會授與目標 (連結) 資料庫或表格的權限。您必須分別授與目標的權限。

您可以使用 Lake Formation 控制台，API 或 AWS Command Line Interface （ AWS CLI ）來授予權限。

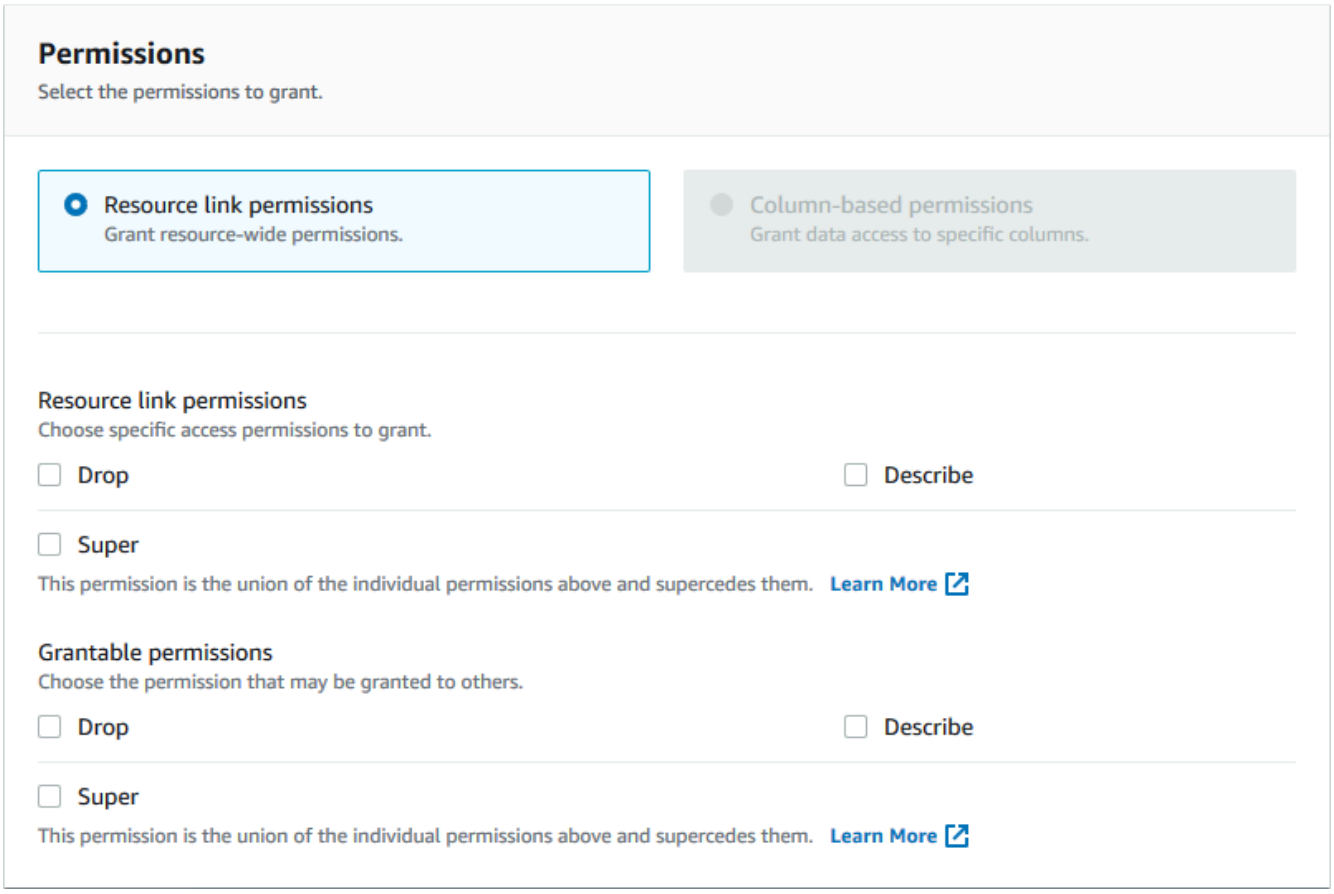
console

使用 Lake Formation 主控台授與資源連結權限

1. 執行以下任意一項：
   - 對於資料庫資源連結，請遵循使用指定的資源方法授與資料庫權限. 中的步驟執行下列步驟：
     1. 開啟 [授與資料湖權限] 頁面。
     2. 指定資料庫。指定一或多個資料庫資源連結。
     3. 指定主參與者。
   - 對於表格資源連結，請遵循中使用指定的資源方法授與資料表權限的步驟執行下列操作：
     1. 開啟 [授與資料湖權限] 頁面。

2. 分析表格。指定一或多個表格資源連結。
3. 指定主參與者。

2. 在「權限」下，選取要授與的權限。選擇性地選取可授與的權限。



3. 選擇 Grant (授予)。

AWS CLI

若要使用授與資源連結權限 AWS CLI

- 執行命令`grant-permissions`，將資源連結指定為資源。

  Example

  此範例會授與`DESCRIBE` AWS 帳號 1111-2222-3333 資料庫`incidents-link`中表格資源連結`issues`的使`datalake_user1`用者。

```
aws lakeformation grant-permissions --principal
 DataLakePrincipalIdentifier=arn:aws:iam::111122223333:user/datalake_user1
```

```
--permissions "DESCRIBE" --resource '{ "Table": {"DatabaseName":"issues",
"Name":"incidents-link"}}'
```

> 另請參閱:
>
> - 建立資源連結
> - Lake Formation 權限參考

存取共用資料表的基礎資料

假設 AWS 帳戶 A 與帳戶 B 共用「資料目錄」資料表 — 例如，透過將資料表上的授權選項授予SELECT與帳戶 B，若要讓帳戶 B 中的主體能夠讀取共用資料表的基礎資料，則必須符合下列條件：

- 帳戶 B 中的資料湖管理員必須接受共用。（如果帳戶 A 和 B 位於同一組織中，或者授予是使用基於 Lake Formation 標籤的訪問控制方法進行的，則不需要這樣做。）
- 資料湖管理員必須將帳戶 A 授與在共用資料表上授予的 Lake Formation SELECT 權限重新授與主體。
- 主體必須在資料表、包含該資料庫的資料庫以及帳戶 A 資料目錄上具有下列 IAM 許可。

> Note
>
> 在下列 IAM 政策中：
>
> - <account-id-A>以帳戶 A 的 AWS 帳號 ID 取代。
> - 以<region>有效的區域取代。
> - 以<database>包含共用資料表的帳戶 A 中的資料庫名稱取代。
> - <table>以共用資料表的名稱取代。

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

另請參閱:

- 接受來自的資源共用邀請 AWS RAM

跨帳戶 CloudTrail 記錄

Lake Formation 針對資料湖中的所有跨帳戶資料存取提供集中式稽核追蹤。當收件者 AWS 帳戶存取共用資料表中的資料時，Lake Formation 會將 CloudTrail 事件複製到擁有帳戶的 CloudTrail 記錄中。

複製的事件包括透過整合式服務 (例如 Amazon Athena 和 Amazon Redshift Spectrum) 對資料進行查詢，以及依任AWS Glue務存取資料。

CloudTrail 同樣複製資料目錄資源上跨帳戶作業的事件。

身為資源擁有者，如果您在 Amazon S3 中啟用物件層級日誌記錄，您可以執行查詢，將 S3 CloudTrail 事件與 Lake Formation CloudTrail 事件聯結在一起，以判斷已存取 S3 儲存貯體的帳戶。

主題

在跨帳戶 CloudTrail 記錄中包含主體身分

依預設，新增至共用資源收件者記錄檔並複製到資源擁有者日誌的跨帳戶 CloudTrail 事件僅包含外部帳戶 AWS 主體的主體 ID，而不是主體 (主體 ARN) 的人類可讀 Amazon 資源名稱 (ARN)。在信任範圍內共用資源時 (例如同一組織或團隊)，您可以選擇加入以在 CloudTrail 事件中包含主參與者 ARN。然後，資源擁有者帳號可以追蹤存取其擁有資源之收件者帳號中的主參與者。

> **Important**
>
> 身為共用資源收件者，若要在您自己的 CloudTrail 記錄檔中查看事件中的主體 ARN，您必須選擇與擁有者帳戶共用主體 ARN。
> 如果透過資源連結進行資料存取，則會在共用資源收件者帳號中記錄兩個事件：一個用於資源連結存取，另一個用於目標資源存取。資源連結存取的事件不包括主參與者 ARN。目標資源存取的事件不包括沒有選擇加入的主參與者 ARN。資源連結存取事件不會複製到擁有者帳號。

以下是預設跨帳戶 CloudTrail 事件的摘錄 (不加入)。執行資料存取的帳戶是 1111-2222-3333。這是顯示在呼叫帳號和資源擁有者帳號中的記錄。在跨帳戶案例中，Lake Formation 會在兩個帳戶中填充日誌。

```
{
    "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"
    },
...
}
```

身為共用資源用戶，當您選擇加入包含主參與者 ARN 時，摘錄會變成以下內容。此`lakeFormationPrincipal`欄位代表透過亞馬遜雅典娜、亞馬 Amazon Redshift Spectrum 或任務執行查詢的最終角色或AWS Glue使用者。

```
{
    "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"
    },
...
}
```

選擇加入跨 CloudTrail 帳戶記錄中包含主要 ARN

1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以`Administrator`使用者或使用 `Administrator Access` IAM 政策的使用者身分登入。
2. 在導覽窗格中，選擇設定。

3. 在 [資料目錄設定] 頁面的 [預設權限] AWS CloudTrail區段中，對於資源擁有者，輸入一或多個 AWS 資源擁有者帳號 ID。

   在每個帳戶 ID 之後按 Enter 鍵。

4. 選擇儲存。

   現在，共用資源收件者和資源擁有者的記錄檔中儲存的跨帳號 CloudTrail 事件都包含主參與者 ARN。

查詢 Amazon S3 跨帳戶存取的 CloudTrail 日誌

身為共用資源擁有者，您可以查詢 S3 CloudTrail 日誌以判斷已存取 Amazon S3 儲存貯體的帳戶 (前提是您已在 Amazon S3 中啟用物件層級記錄功能)。這僅適用於您在 Lake Formation 註冊的 S3 位置。如果共用資源取用者選擇在 Lake Formation CloudTrail 記錄檔中包含主要 Ran，您可以決定存取值區的角色或使用者。

使用執行查詢時 Amazon Athena，您可以在工作階段名稱屬性上加入 Lake Formation CloudTrail CloudTrail 事件和 S3 事件。查詢也可以在或上`eventName="GetDataAccess"`篩選 Lake Formation 事件和 S3 事`eventName="Get Object"`件`eventName="Put Object"`。

以下摘錄自 Lake Formation 跨帳戶 CloudTrail 事件，其中存取已註冊 S3 位置中的資料。

```
{
  "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"
   }
}
```

索`lakeFormationRoleSessionName`引鍵值可以與 S3 CloudTrail 事件`principalId`金鑰中的工作階段名稱加入。`AWSLF-00-GL-111122223333-B8JSAjo5QA`以下是 S3 CloudTrail 事件的摘錄。它顯示了會話名稱的位置。

```
{
```

```
   "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"
     },
   ..............
   ..............
}
```

工作階段名稱的格式如下：

```
AWSLF-<version-number>-<query-engine-code>-<account-id->-<suffix>
```

version-number

此格式的版本，目前00。如果作業階段名稱格式發生變更，則下一個版本將會是01。

query-engine-code

指示存取資料的實體。目前的值為：

GL	AWS Glue工作
AT	Athena
RE	Amazon Redshift Spectrum

account-id

要求 Lake Formation 憑證的 AWS 帳戶 ID。

suffix

隨機生成的字符串。

使用AWS Glue和 Lake Formation 管理跨帳戶權限

您可以使用AWS Glue或 AWS Lake Formation授與跨帳戶存取資料目錄資源和基礎資料。

在中AWS Glue，您可以透過建立或更新資料目錄資源策略來授與跨帳戶權限。在 Lake Formation 中，您可以使用 Lake Formation 成權限模型和 `Grant Permissions` API 操作來授予跨帳戶GRANT/REVOKE權限。

> **Tip**
>
> 我們建議您僅依賴 Lake Formation 權限來保護您的資料湖。

您可以使用 Lake Formation 控制台或 AWS Resource Access Manager （AWS RAM）控制台查看 Lake Formation 跨帳戶補助。不過，這些主控台頁面不會顯示資AWS Glue料目錄資源原則所授與的跨帳戶權限。同樣地，您可以使用AWS Glue主控台的 [設定] 頁面檢視 [資料目錄] 資源策略中的跨帳戶授與，但該頁面不會顯示使用 Lake Formation 授與的跨帳戶權限。

為了確保您在查看和管理跨帳戶權限時不會錯過任何授權，Lake Formation 並AWS Glue要求您執行以下操作，以指示您已知道並允許 Lake Formation 和的跨帳戶授予。AWS Glue

使用資AWS Glue料目錄資源策略授與跨帳戶權限時

如果您的帳戶 (授與者帳號或生產者帳號) 沒有進行任何用於共用 AWS RAM 資源的跨帳戶授權，您可以在中一如往常儲存資料目錄資源策略。AWS Glue但是，如果已經完成涉及 AWS RAM 資源共用的授權，您必須執行下列其中一項作業，以確保成功儲存資源策略：

- 當您在主控台的 「設定」 頁面上儲存資源策略時，AWS Glue主控台會發出警示，指出政策中的權限除了使用 Lake Formation 主控台授予的任何權限之外，還會發出警示。您必須選擇 [繼續] 才能儲存原則。
- 使用 `glue:PutResourcePolicy` API 作業儲存資源策略時，必須將`EnableHybrid`欄位設定為 'TRUE' (類型 = 字串)。下面的代碼示例演示了如何在 Python 中做到這一點。

```
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')
```

如需詳細資訊，請參閱開發人員指南中的PutResourcePolicy 動作 (Python: put_ resource)。AWS Glue

使用 Lake Formation 命名資源方法授予跨帳戶權限時

如果您的帳戶（生產者帳戶）中沒有資料目錄資源策略，則 Lake Formation 跨帳戶授予您照常進行。但是，如果有資料目錄資源策略存在，您必須在其中新增下列陳述式，以允許跨帳戶授權在使用指定的資源方法建立時成功。請以<region>有效的地區名稱和<account-id>您的 AWS 帳戶 ID (生產者帳號 ID) 取代。

```
    {
```

```
    "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"
    ]
  }
```

沒有這個額外的聲明，Lake Formation 授予成功，但被阻止 AWS RAM，收件人帳戶無法訪問授予的資源。

> **Important**
>
> 使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 方法進行跨帳戶授與時，您必須擁有至少具有中指定權限的資料目錄資源策略。必要條件

> **另請參閱:**
>
> - 元數據訪問控制（對於具名資源方法與基於 Lake Formation 標籤的訪問控制（LF-TBAC）方法的討論）。
> - 檢視共用資料目錄表格和資料庫
> - 在AWS Glue 開發人員指南中使用AWS Glue主控台上的資料目錄設定
> - 在AWS Glue 開發人員指南中授與跨帳戶存取權 (適用於範例資料目錄資源策略)

使用 GetResourceShares API 作業檢視所有跨帳戶授權

如果您的企業同時使用 AWS Glue Data Catalog 資源策略和 Lake Formation 授予跨帳戶權限，則在一個位置查看所有跨帳戶授予的唯一方法是使用 `glue:GetResourceShares` API 操作。

當您使用指定的資源方法跨帳戶授予 Lake Formation 權限時， AWS Resource Access Manager (AWS RAM) 會建立 AWS Identity and Access Management (IAM) 資源政策並將其儲存在您的 AWS

帳戶中。該策略授予訪問資源所需的權限。 AWS RAM 為每個跨帳號授權建立個別的資源策略。您可以使用 `glue:GetResourceShares` API 作業來檢視所有這些政策。

> **Note**
>
> 此作業也會傳回資料目錄資源原則。但是，如果您在 [資料目錄] 設定中啟用中繼資料加密，且您沒有 AWS KMS 金鑰的權限，則作業將不會傳回資料目錄資源原則。

檢視所有跨帳戶撥款

- 輸入以下 AWS CLI 命令。

```
aws glue get-resource-policies
```

以下是當您db1將資料庫t中資料表的權限授與 AWS 帳號 1111-2222-3333 時 AWS RAM 建立並儲存的資源策略範例。

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

```
}
```

另請參閱：

- GetResourceShares 開發人員指南中的動作（Python：獲取資源政策）AWS Glue

存取和檢視共用資料目錄表格和資料庫

對於資料湖管理員以及已授與權限的主參與者，與您 AWS 帳戶共用的資源會顯示在「資料目錄」中，就像這些資源是您帳戶中的資源一樣。主控台會顯示擁有該資源的帳號。

您可以使用 Lake Formation 控制台查看與您的帳戶共享的資源。您也可以使用 AWS Resource Access Manager (AWS RAM) 主控台，同時檢視與帳戶共用的資源，以及使用具名資源方法與其他 AWS 帳號共用的資源。

Important

當有人使用具名的資源方法將資料目錄資源的跨帳戶權限授與您的帳戶或 AWS 組織時，Lake Formation 會使用 AWS Resource Access Manager (AWS RAM) 服務來共用資源。如果您的帳號與授與帳號位於相同的 AWS 組織中，您可立即使用共用資源。
但是，如果您的帳戶不在同一個組織中，則 AWS RAM 會傳送邀請至您的帳戶，以接受或拒絕資源共用。然後，若要使共用資源可用，您帳戶中的資料湖管理員必須使用 AWS RAM 主控台或 CLI 接受邀請。
如果有 AWS RAM 資源共享邀請等待被接受，則 Lake Formation 控制台會顯示警報。只有獲得授權檢視 AWS RAM 邀請的使用者才會收到警示。

另請參閱:

- 跨 AWS 帳戶共用資料目錄表格和資料庫
- Lake Formation 的跨賬戶數據共享
- 存取共用資料表的基礎資料
- 元數據訪問控制(如需具名資源方法與用於共用資源的 LF-TBAC 方法的相關資訊。)

主題

接受來自的資源共用邀請 AWS RAM

如果資料目錄資源與您的 AWS 帳戶共用，且您的帳號與共用帳號不在同一個 AWS 組織中，則除非您接受來自 AWS Resource Access Manager (AWS RAM) 的資源共用邀請，否則您無法存取共用資源。身為資料湖管理員，您必須先查詢擱置中 AWS RAM 的邀請，然後接受邀請。

您可以使用 AWS RAM 主控台、API 或 AWS Command Line Interface (AWS CLI) 來檢視和接受邀請。

若要從 AWS RAM (主控台) 檢視和接受資源共用邀請

1. 確保您具有檢視和接受資源共用邀請的必要 AWS Identity and Access Management (IAM) 許可。

   如需資料湖管理員建議的 IAM 政策的相關資訊，請參閱the section called “資料湖管理員權限”。
2. 請遵循使用者指南中「接受和拒絕邀請」中的AWS RAM 指示進行。

若要檢視並接受來自 AWS RAM (AWS CLI) 的資源共用邀請

1. 確保您具有檢視和接受資源共用邀請的必要 AWS Identity and Access Management (IAM) 許可。

   如需資料湖管理員建議的 IAM 政策的相關資訊，請參閱the section called “資料湖管理員權限”。
2. 輸入下列命令以檢視擱置的資源共用邀請。

```
aws ram get-resource-share-invitations
```

輸出格式應類似以下內容。

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

請注意的狀態PENDING。

3. 將resourceShareInvitationArn金鑰的值複製到剪貼簿。
4. 將值貼到以下指令中，進行取代<invitation-arn>，然後輸入指令。

```
aws ram accept-resource-share-invitation --resource-share-invitation-
arn <invitation-arn>
```

輸出格式應類似以下內容。

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

請注意的狀態ACCEPTED。

檢視共用資料目錄表格和資料庫

您可以使用 Lake Formation 控制台或 AWS CLI 查看與您的帳戶共享的資源。您也可以使用 AWS Resource Access Manager (AWS RAM) 主控台或 CLI 來檢視與您帳戶共用的資源，以及與其他 AWS 帳戶共用的資源。

使用 Lake Formation 控制台查看共享資源

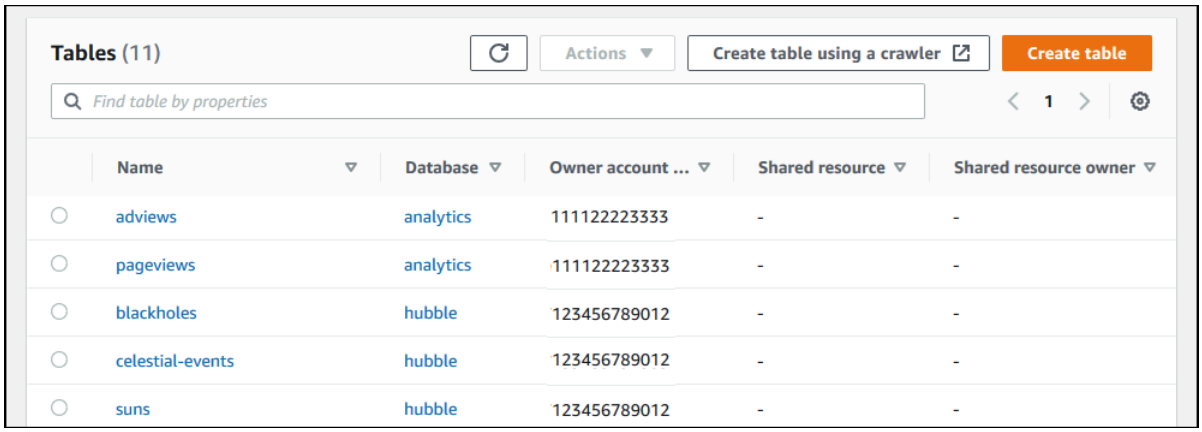
1. 開啟 Lake Formation 主控台，網址為 https://console.aws.amazon.com/lakeformation/。

   以資料湖管理員或已授與共用資料表權限的使用者身分登入。

2. 若要檢視與您的 AWS 帳戶共用的資源，請執行下列其中一個動作：

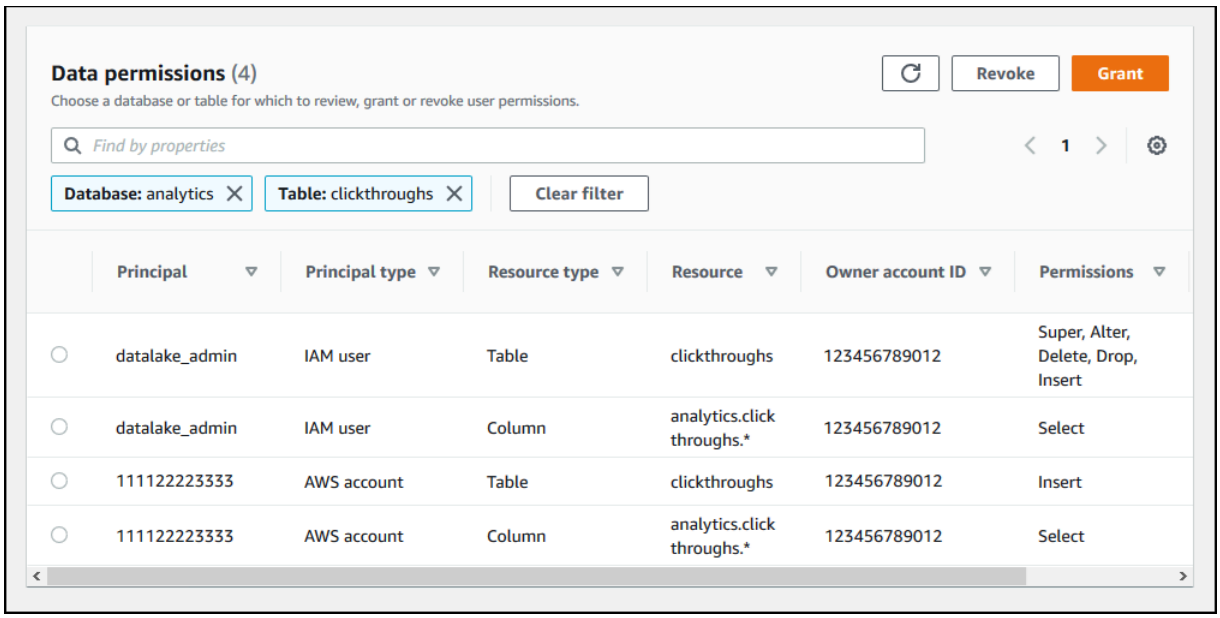
   - 若要檢視與帳戶共用的資料表，請在功能窗格中選擇 [表格]。
   - 若要檢視與帳戶共用的資料庫，請在功能窗格中選擇 [資料庫]。

   主控台會顯示您帳戶中並與您帳戶共用的資料庫或表格清單。對於與您的帳戶共用的資源，主控台會在「擁有者 AWS 帳號 ID」欄 (下列螢幕擷取畫面的第三欄) 下方顯示擁有者的帳號 ID。



3. 若要檢視您與其他 AWS 帳戶或組織共用的資源，請在導覽窗格中選擇 [資料權限]。

   您共用的資源會列在 [資料權限] 頁面上，外部帳號會顯示在 [主參與者] 欄中，如下圖所示。



若要使用 AWS RAM 主控台檢視共用資源

1. 確保您具有使用查看共用資源的必要 AWS Identity and Access Management (IAM) 許可 AWS RAM。

   您至少必須具有權限`ram:ListResources`。此權限包含在 AWS 受管政策中。`AWSLakeFormationCrossAccountManager`

2. 請登入 AWS Management Console 並開啟 AWS RAM 主控台，網址為 https://console.aws.amazon.com/ram。

3. 執行以下任意一項：

   - 若要查看您共用的資源，請在導覽窗格的 [由我共用] 下，選擇 [共用資源]。
   - 若要查看與您共用的資源，請在導覽窗格的 [與我共用] 下，選擇 [共用資源]。

建立資源連結

資源連結是資料目錄物件，它們是中繼資料資料庫和表格的連結，通常是來自其他 AWS 帳戶的共用資料庫和表格。它們有助於跨帳戶存取所有 AWS 區域的資料湖中的資料。

> **Note**
>
> Lake Formation 支援跨 AWS 區域查詢資料目錄資料表。您可以在指向不同區域的共用資料庫和表格的 AWS 區域中建立資源連結，從任何區域存取「資料目錄」資料庫和表格。

主題

資源連結在 Lake Formation 中如何運作

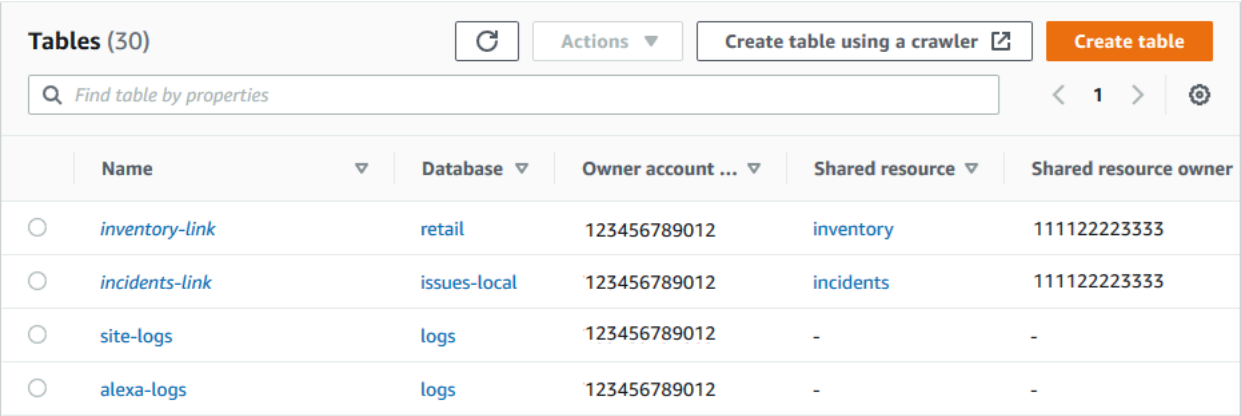
資源連結是指向本機或共用資料庫或表格的連結的資料目錄物件。建立資料庫或表格的資源連結後，無論您要使用資料庫或表格名稱，都可以使用資源連結名稱。除了您擁有的表格或與您共用的表格外，表格資源連結會由傳回，`glue:GetTables()`並以項目的形式顯示在 Lake Formation 主控台的「表格」(T ables) 頁面上。資源鏈接到數據庫的行為以類似的方式。

建立資料庫或表格的資源連結可讓您執行下列作業：

- 為「資料目錄」中的資料庫或表格指定其他名稱。如果不同的 AWS 帳戶共用具有相同名稱的資料庫或資料表，或者帳戶中的多個資料庫具有相同名稱的資料表，此功能特別有用。
- 在指向另一個區域中的資料庫和表格的 AWS 區域中建立資源連結，從任何區域存取「資料目錄」資料庫和表格。您可以使用 Athena、Amazon EMR 在任何區域使用這些資源連結執行查詢，並執行 AWS Glue ETL Spark 任務，而不必複製來源資料或 Glue 資料型錄中的中繼資料。
- 使用整合式 AWS 服務 (例如 Amazon Athena 和 Amazon Redshift Spectrum) 來執行存取共用資料庫或資料表的查詢。某些整合式服務無法跨帳戶直接存取資料庫或資料表。不過，他們可以存取您帳戶中的資源連結，連至其他帳戶中的資料庫和表格。

> **Note**
>
> 您不需要建立資源連結來參照AWS Glue擷取、轉換和載入 (ETL) 指令碼中的共用資料庫或表格。但是，為了避免多個 AWS 帳戶共用具有相同名稱的資料庫或表格時產生歧義，您可以建立並使用資源連結，或在呼叫 ETL 作業時指定目錄 ID。

下列範例顯示「Lake Formation」主控台「表格」頁面，其中列出了兩個資源連結。資源連結名稱一律以斜體顯示。每個資源連結會連同其連結共用資源的名稱和擁有者一起顯示。在此範例中， AWS 帳戶 1111-2222-3333 中的資料湖系統管理員共用了帳戶 1234-5678-9012 的`inventory`和`incidents`資料表。然後，該帳號中的使用者建立了這些共用資料表的資源連結。



以下是資源連結的注意事項和限制：

- 需要資源連結，才能讓 Athena 和 Redshift Spectrum 等整合式服務查詢共用資料表的基礎資料。這些整合式服務中的查詢會根據資源連結名稱建構。
- 假設針對包含資料庫關閉「僅針對此資料庫中的新資料表使用 IAM 存取控制」設定，則只有建立資源連結的主體才能檢視和存取該資源連結。若要讓您帳號中的其他主參與者存取資源連結，請授與該資源連結的`DESCRIBE`權限。若要讓其他人卸除資源連結，請授與該連結的`DROP`權限。資料湖管理員可以存取帳戶中的所有資源連結。若要刪除由另一個主參與者建立的資源連結，資料湖管理員必須先授與自己資源連結的`DROP`權限。如需詳細資訊，請參閱 Lake Formation 權限參考。

  > **Important**
  >
  > 授與資源連結的權限不會授與目標 (連結) 資料庫或表格的權限。您必須分別授與目標的權限。

- 若要建立資源連結，您需要 Lake Formation `CREATE_TABLE` 或`CREATE_DATABASE`權限，以及`glue:CreateTable`或 `glue:CreateDatabase` AWS Identity and Access Management (IAM) 權限。
- 您可以建立本機 (擁有的) 資料目錄資源的資源連結，以及與 AWS 帳戶共用的資源。
- 建立資源連結時，不會執行檢查以查看目標共用資源是否存在，或者您是否具有資源的跨帳號權限。這可讓您以任何順序建立資源連結和共用資源。
- 如果刪除資源連結，則不會捨棄連結的共用資源。如果您卸除共用資源，則不會刪除該資源的資源連結。

- 您可以建立資源連結鏈結。但是，這樣做沒有任何價值，因為 API 只遵循第一個資源鏈接。

> 另請參閱：
>
> - 授與和撤銷資料目錄資源的權限

建立共用資料目錄表格的資源連結

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 建立任何 AWS 區域中共用資料表的資源連結。

若要建立共用資料表 (主控台) 的資源連結

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以具有資料庫 Lake Formation CREATE_TABLE 權限的主體身分登入，以包含資源連結。
2. 在導覽窗格中，選擇 [表格]，然後選擇 [建立] > [資源連結]。
3. 在 [建立資源連結] 頁面上，提供下列資訊：

   資源連結名稱

   輸入與表格名稱相同規則的名稱。名稱可以與目標共用資料表相同。

   資料庫

   本機資料目錄中包含資源連結的資料庫。

   共用表格擁有者區域

   如果您要在其他區域中建立資源連結，請選取目標共用資料表的區域。

   共用資料表

   從清單中選取共用資料表，或輸入本機 (擁有的) 或共用表格名稱。

   該列表包含與您的帳戶共享的所有表格。記下每個資料表所列出的資料庫和擁有者帳戶 ID。如果您沒有看到已與帳戶共享的表格，請檢查以下事項：

   - 如果您不是資料湖管理員，請檢查資料湖管理員是否已授與您資料表上的 Lake Formation 權限。

- 如果您是資料湖管理員，且您的帳戶與授與帳戶不在同一個 AWS 組織中，請確定您已接受表格的 AWS Resource Access Manager (AWS RAM) 資源共用邀請。如需詳細資訊，請參閱 接受來自的資源共用邀請 AWS RAM。

共用資料表的資料庫

如果您從清單中選取共用資料表，則此欄位會在外部帳戶中填入共用資料表的資料庫。否則，請在外部帳戶中輸入本機資料庫 (用於本機資料表的資源連結) 或共用資料表的資料庫。

共享表格擁有者

如果您從清單中選取共用資料表，則此欄位會填入共用資料表的擁有者帳戶 ID。否則，請輸入您的 AWS 帳號 ID (用於本機資料表的資源連結) 或共用該表格的 AWS 帳號 ID。

4. 選擇 [建立] 以建立資源連結。

然後，您可以在「表格」頁面的「名稱」欄下檢視資源連結名稱。

5. (選擇性) 將資源連結的 Lake Formation DESCRIBE 權限授與必須能夠檢視連結並存取目標表格的主參與者。

但是，授與資源連結的權限並不會授與目標 (連結) 資料庫或表格的權限。您必須個別授與目標資料庫的權限，表格/資源連結才能在 Athena 中顯示。

若要在相同區域中建立共用資料表的資源連結 (AWS CLI)

1. 輸入與以下相似的命令。

```
aws glue create-table --database-name myissues --table-input
 '{"Name":"my_customers","TargetTable":
{"CatalogId":"111122223333","DatabaseName":"issues","Name":"customers"}}'
```

這個命令會建立名為共用資料表`my_customers`的資源連結`customers`，該資源連結位於 AWS 帳戶 1111-2222-3333 的資料庫`issues`中。資源連結儲存在本機資料庫中`myissues`。

2. (選擇性) 將資源連結的 Lake Formation DESCRIBE 權限授與必須能夠檢視連結並存取目標表格的主參與者。

但是，授與資源連結的權限並不會授與目標 (連結) 表格的權限。您必須個別授與目標資料庫的權限，表格/資源連結才能在 Athena 中顯示。

若要在不同區域中建立共用資料表的資源連結 (AWS CLI)

1. 輸入與以下相似的命令。

```
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"
        }
    }
}‘
```

此命令會建立一個名為 `rl_useast1salestb_ireland`「歐洲 (愛爾蘭) 區域」的資源連結至共用資料表`useast1_salestb`，該資源連結位於美國東部 (維吉尼亞北部) 區域 AWS 帳戶 444455556666 的資料庫`useast1_salesdb`中。資源連結儲存在本機資料庫中`ireland_db`。

2. 將 Lake Formation DESCRIBE 權限授與必須能夠透過連結檢視連結並存取連結目標的主參與者。

但是，授與資源連結的權限並不會授與目標 (連結) 表格的權限。您必須個別授與目標資料表的權限，表格/資源連結才能在 Athena 中顯示。

另請參閱：

- 資源連結在 Lake Formation 中如何運作
- DESCRIBE

建立共用資料目錄資料庫的資源連結

您可以使用 AWS Lake Formation 主控台、API 或 AWS Command Line Interface (AWS CLI) 建立共用資料庫的資源連結。

若要建立共用資料庫的資源連結 (主控台)

1. 請在以下位置開啟 AWS Lake Formation 主控台。 https://console.aws.amazon.com/lakeformation/ 以資料湖管理員或資料庫建立者身分登入。

   資料庫建立者是已獲得 Lake Formation CREATE_DATABASE 權限的主體。

2. 在功能窗格中，選擇 [資料庫]，然後選擇 [建立] > [資源] 連結。

3. 在 [建立資源連結] 頁面上，提供下列資訊：

   資源連結名稱

   輸入與資料庫名稱相同規則的名稱。名稱可以與目標共用資料庫相同。

   共用資料庫擁有者區

   如果要在其他「區域」中建立資源連結，請選取目標共用資料庫的「區域」。

   共用資料庫

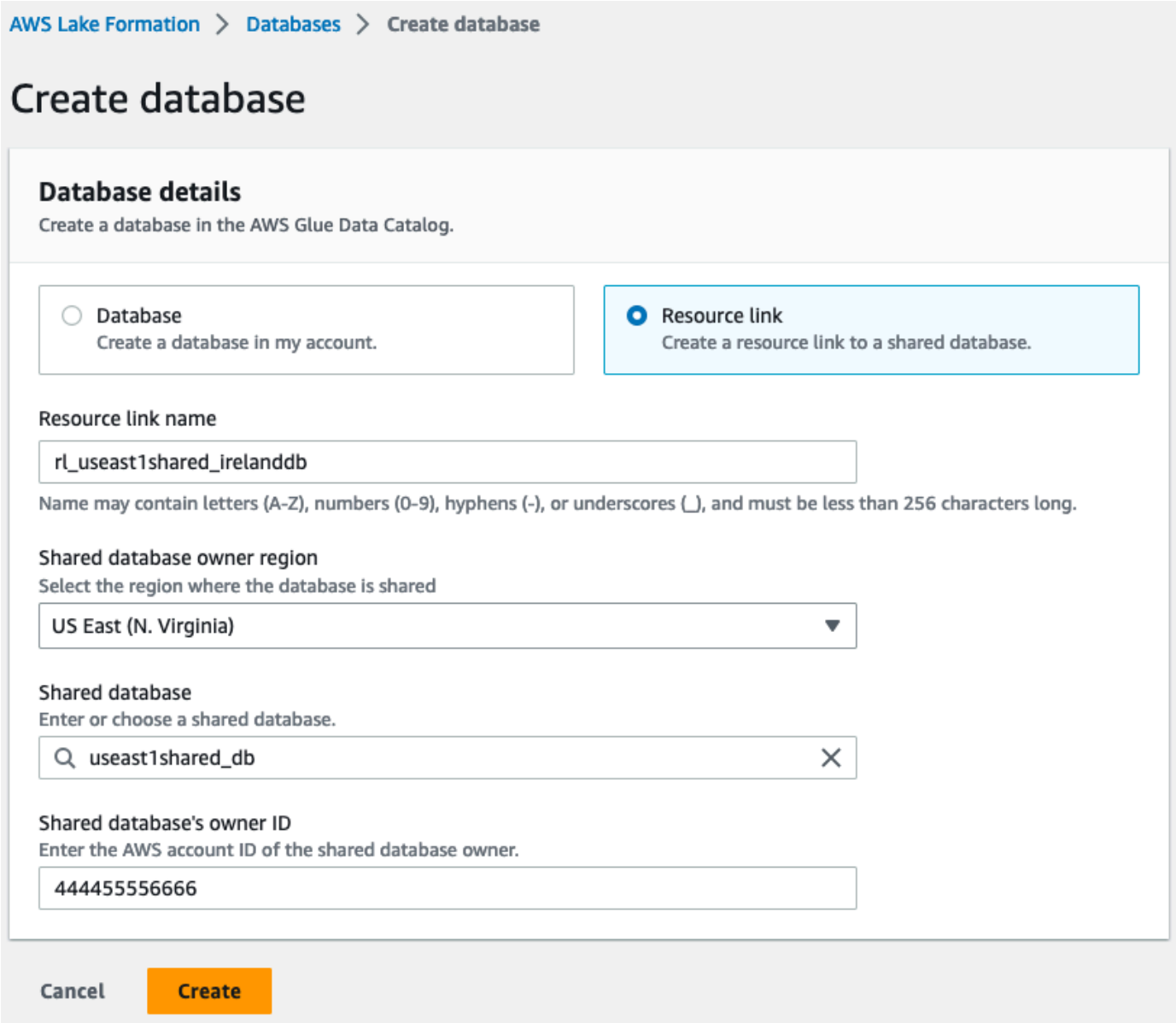
   從清單中選擇資料庫，或輸入本機 (擁有) 或共用資料庫名稱。

   該列表包含與您的帳戶共享的所有數據庫。請記下每個資料庫隨附的擁有者帳戶 ID。如果您沒有看到已與帳戶共用的資料庫，請檢查下列項目：

   - 如果您不是資料湖管理員，請檢查資料湖管理員是否已授與您資料庫的 Lake Formation 權限。
   - 如果您是資料湖管理員，且您的帳戶與授與帳戶不在同一個 AWS 組織中，請確定您已接受資料庫的 AWS Resource Access Manager (AWS RAM) 資源共用邀請。如需詳細資訊，請參閱 接受來自的資源共用邀請 AWS RAM。

   共用資料庫擁有

   如果您從清單中選取共用資料庫，則此欄位會填入共用資料庫的擁有者帳戶 ID。否則，請輸入您的 AWS 帳號 ID (用於本機資料庫的資源連結) 或共用資料庫的 AWS 帳戶 ID。



4. 選擇 [建立] 以建立資源連結。

   然後，您可以在「資料庫」頁面的「名稱」欄下檢視資源連結名稱。

5. (選擇性) 將資源連結的 Lake Formation DESCRIBE 權限授與來自歐洲 (愛爾蘭) 區域的主參與者 (必須能夠檢視連結並存取目標資料庫)。

   但是，授與資源連結的權限並不會授與目標 (連結) 資料庫或表格的權限。您必須個別授與目標資料庫的權限，表格/資源連結才能在 Athena 中顯示。

在相同區域中建立共用資料庫的資源連結 (AWS CLI)

1. 輸入與以下相似的命令。

```
aws glue create-database --database-input '{"Name":"myissues","TargetDatabase":
{"CatalogId":"111122223333","DatabaseName":"issues"}}'
```

這個命令會建立一個名為共用資料庫myissues的資源連結issues，該資源連結位於 AWS 帳號 1111-2222-3333 中。

2. (選擇性) 將 Lake Formation DESCRIBE 權限授與資源連結上必須能夠檢視連結並存取目標資料庫或表格的主參與者。

但是，授與資源連結的權限並不會授與目標 (連結) 資料庫或表格的權限。您必須個別授與目標資料庫的權限，表格/資源連結才能在 Athena 中顯示。

若要在不同區域中建立共用資料庫的資源連結 (AWS CLI)

1. 輸入與以下相似的命令。

```
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"
       }
    }
}'
```

此命令會rl_useast1shared_irelanddb在歐洲 (愛爾蘭) 區域的 AWS 帳戶 111122223333 中建立名為的資源連結到共用資料庫useast1shared_db，該資源連結位於美國東部 (維吉尼亞北部) 區域的 AWS 帳戶 444455556666。

2. 將 Lake Formation DESCRIBE 權限授予來自歐洲 (愛爾蘭) 區域的主體，必須能夠透過連結檢視連結並存取連結目標。

另請參閱：

- 資源連結在 Lake Formation 中如何運作
- DESCRIBE

AWS GlueAPI 中的資源連結處理

下表說明資AWS Glue料目錄 API 如何處理資料庫和表格資源連結。對於所有 Get* API 操作，只有調用者具有獲得返回權限的數據庫和表。此外，透過資源連結存取目標資料庫或表格時，您必須在目標和資源連結上同時具有 AWS Identity and Access Management (IAM) 和 Lake Formation 權限。在資源鏈接上需要的 Lake Formation 許可是DESCRIBE。如需詳細資訊，請參閱 DESCRIBE。

資料庫 API 作業

API 操作	資源連結處理
CreateDatabase	如果資料庫是資源連結，則會建立指向指定目標資料庫的資源連結。
UpdateDatabase	如果指定的資料庫是資源連結，請跟隨連結並更新目標資料庫。如果必須修改資源連結才能連結至其他資料庫，您必須將其刪除並建立新的資料庫。
DeleteDatabase	刪除資源連結。它不會刪除鏈接的（目標）數據庫。
GetDatabase	如果呼叫者對目標具有權限，請按照連結傳回目標的屬性。否則，它返回鏈接的屬性。
GetDatabases	返回數據庫列表，包括資源鏈接。對於結果集中的每個資源連結，作業會跟隨連結以取得連結目標的內容。您必須指定 ResourceShareType = ALL 才能查看與您的帳戶共用的資料庫。

表格 API 作業

API 操作	資源連結處理
CreateTable	如果資料庫是資源連結，請跟隨資料庫連結，並在目標資料庫中建立表格。如果表格是資源連結，則作業會在指定的資料庫中建立資源連結。不支援透過資料庫資源連結建立表格資源連結。
UpdateTable	如果表格或指定的資料庫是資源連結，則會更新目標表格。如果資料表和資料庫都是資源連結，則作業會失敗。
DeleteTable	如果指定的資料庫是資源連結，請跟隨連結並刪除目標資料庫中的表格或表格資源連結。如果表格是資源連結，則作業會刪除指定資料庫中的表格資源連結。刪除表格資源連結並不會刪除目標資料表。
BatchDeleteTable	與 DeleteTable 相同。
GetTable	如果指定的資料庫是資源連結，請跟隨資料庫連結，並從目標資料庫傳回表格或表格資源連結。否則，如果表格是資源連結，則作業會跟隨連結並傳回目標資料表屬性。
GetTables	如果指定的資料庫是資源連結，請跟隨資料庫連結，並從目標資料庫傳回表格和表格資源連結。如果目標資料庫是另一個 AWS 帳戶的共用資料庫，則作業只會傳回該資料庫中的共用資料表。它不跟隨目標數據庫中的表格資源鏈接。否則，如果指定的資料庫是本機 (擁有的) 資料庫，則作業會傳回本機資料庫中的所有表格，並跟隨每個表格資源連結以傳回目標資料表屬性。
SearchTables	返回表格和表格的資源鏈接。它不會跟隨鏈接返回目標表屬性。您必須指定 ResourceShareType = 才ALL能查看與您帳戶共用的資料表。
GetTableVersion	與 GetTable 相同。
GetTableVersions	與 GetTable 相同。
DeleteTableVersion	與 DeleteTable 相同。
BatchDeleteTableVersion	與 DeleteTable 相同。

分割區 API 作業

API 操作	資源連結處理
CreatePartition	如果指定的資料庫是資源連結，請跟隨資料庫連結，並在目標資料庫的指定表格中建立一個分割區。如果表格是資源連結，則作業會跟隨資源連結，並在目標表格中建立分割區。不支援透過表格資源連結和資料庫資源連結建立分割區。
BatchCreatePartition	與 CreatePartition 相同。
UpdatePartition	如果指定的資料庫是資源連結，請跟隨資料庫連結，並更新目標資料庫中指定表格中的分割區。如果表格是資源連結，則作業會跟隨資源連結，並更新目標資料表中的分割區。不支援透過表格資源連結和資料庫資源連結來更新分割區。
DeletePartition	如果指定的資料庫是資源連結，請跟隨資料庫連結，並刪除目標資料庫中指定表格中的分割區。如果表格是資源連結，則作業會跟隨資源連結，並刪除目標表格中的分割區。不支援透過表格資源連結和資料庫資源連結刪除分割區。
BatchDeletePartition	與 DeletePartition 相同。
GetPartition	如果指定的資料庫是資源連結，請跟隨資料庫連結，並從指定的表格傳回分割區資訊。否則，如果表格是資源連結，則作業會跟隨連結並傳回分割區資訊。如果表和數據庫都是資源鏈接，則返回一個空的結果集。
GetPartitions	如果指定的資料庫是資源連結，請跟隨資料庫連結，並傳回指定表格中所有分割區的分割區資訊。否則，如果表格是資源連結，則作業會跟隨連結並傳回分割區資訊。如果表和數據庫都是資源鏈接，則返回一個空的結果集。
BatchGetPartition	與 GetPartition 相同。

用戶定義函數 API 操作

API 操作	資源連結處理
(所有 API 作業)	如果資料庫是資源連結，請跟隨資源連結，並在目標資料庫上執行作業。

另請參閱：

- 資源連結在 Lake Formation 中如何運作

跨區域存取表格

Lake Formation 支援跨 AWS 區域查詢資料目錄資料表。您可以使用 Amazon Athena、Amazon EMR 和 AWS Glue ETL，在指向來源資料庫和表格的其他區域建立資源連結，從其他區域存取某個區域中的資料。透過跨區域表格存取權，您可以跨區域存取資料，而無需將基礎資料或中繼資料複製到「資料目錄」中。

例如，您可以將生產者帳戶中的資料庫或表格共用給區域 A 中的消費者帳戶。接受區域 A 中的資源共用邀請後，取用者帳戶的資料湖管理員可以在區域 A 中建立共用資源的資源連結。使用者帳戶可以在區域 A 中將共用資源的權限授與該帳戶中的 IAM 主體，並可以授與區域 A 中的資源連結。來自 B 區的共享數據。

您也可以在區域 A 的生產者帳戶中託管 Amazon S3 資料來源，並在區域 B 的中央帳戶中註冊資料位置。您可以在中央帳戶中建立資料目錄資源、設定 Lake Formation 權限，以及與您帳戶中的消費者或區域 B 的外部帳戶共用資料。跨區域功能允許使用者使用資源連結從區域 C 存取這些資料目錄表格。

使用此功能，您可以在跨區域的 Apache Hive 中繼存放區中查詢聯合資料庫，也可以在執行查詢時，將本機區域中的資料表與另一個區域中的資料表聯結在一起。

Lake Formation 通過跨區域表訪問支持以下功能：

- 基於 LF 標籤的訪問控制
- 精細的存取控制權限
- 在具有適當權限的共用資料庫或資料表上寫入作業
- 帳戶層級的跨帳戶資料共用，並直接與 IAM 主管層級

具有Create_Database和Create_Table權限的非管理使用者可以建立跨區域資源連結。

> **Note**
>
> 您可以在任何區域中建立跨區域資源連結，並存取資料，而無需套用 Lake Formation 權限。對於未向 Lake Formation 註冊的 Amazon S3 中的來源資料，存取由 Amazon S3 的 IAM 許可政策和 AWS Glue 動作決定。

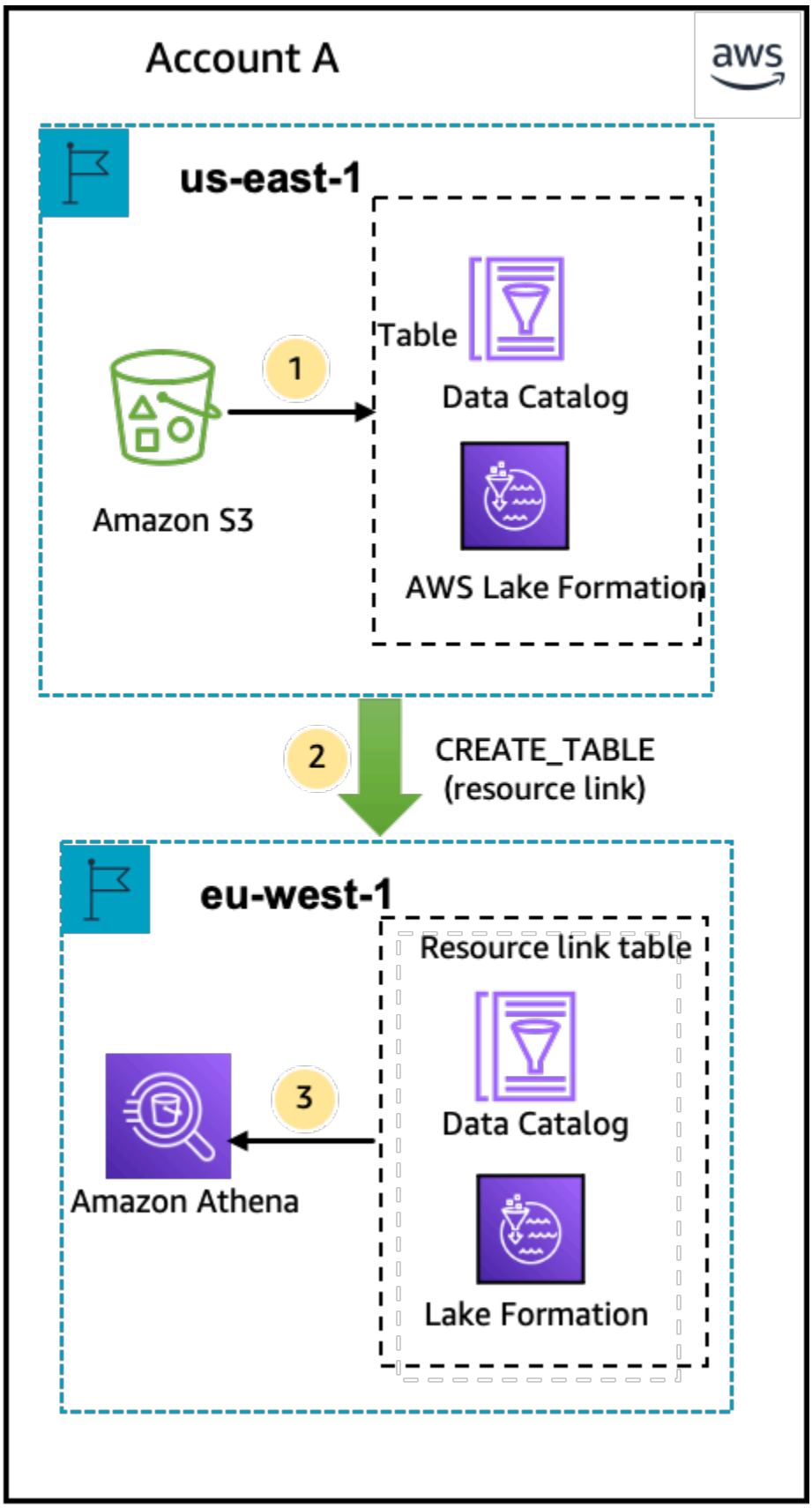
如需限制的詳細資訊，請參閱跨區域資料存取限制。

工作流程

下圖顯示從相同 AWS 帳戶和外部帳戶跨 AWS 區域存取資料的工作流程。

存取同一 AWS 帳戶內共用資料表的工作流程

在下圖中，資料會與美國東部 (維吉尼亞北部) 區域內相同 AWS 帳戶中的使用者共用，而使用者會查詢來自歐洲 (愛爾蘭) 區域的共用資料。


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

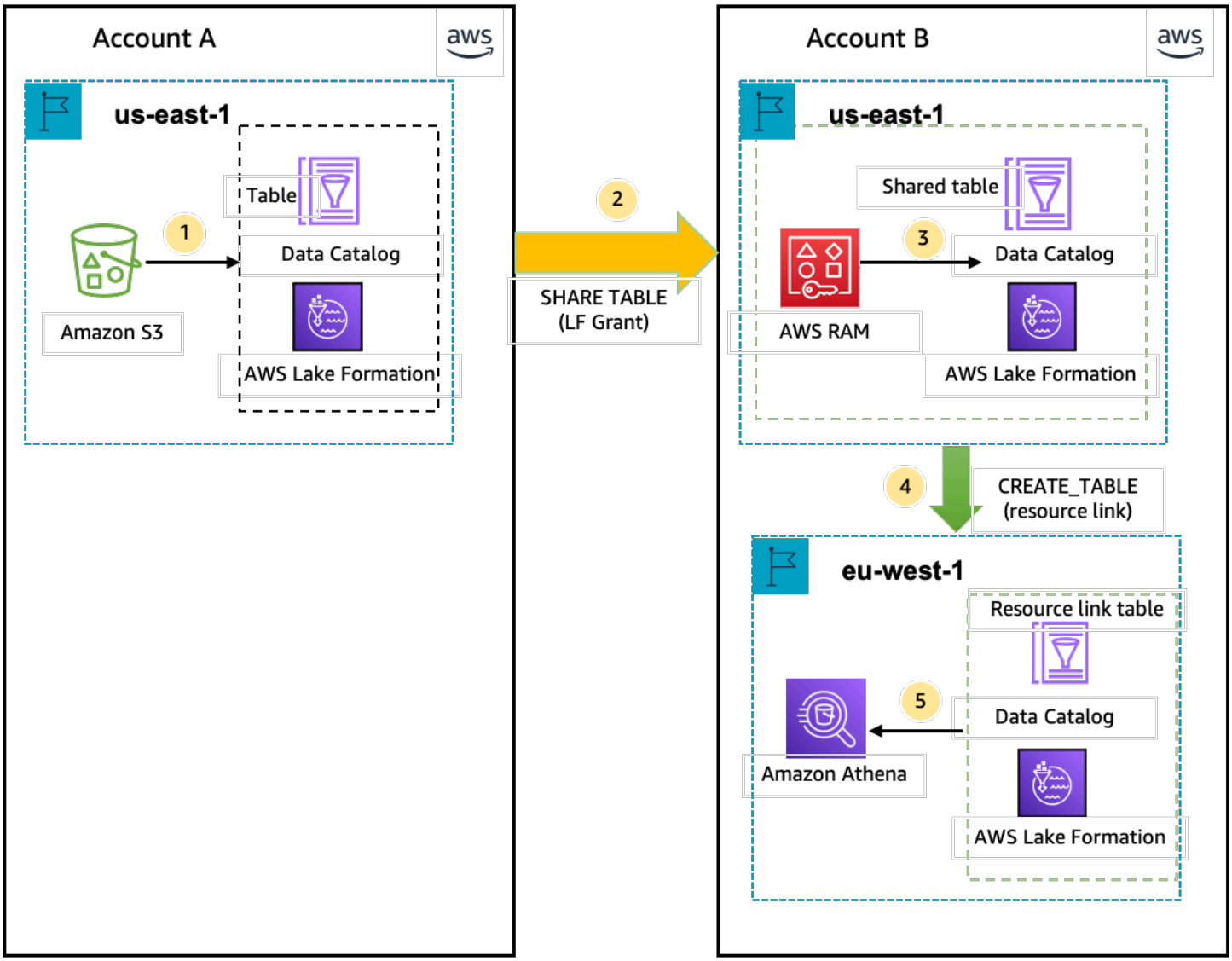
資料湖管理員會執行下列活動 (步驟 1-2)：

1. 資料湖管理員使用資料目錄資料庫和表格設定 AWS 帳戶，並在美國東部 (維吉尼亞北部) 區域的 Lake Formation 註冊 Amazon S3 資料位置。

   將資料目錄資源 (圖表中的產品表格) 的`Select`權限授與相同帳戶中的主參與者 (使用者)。
2. 在歐洲 (愛爾蘭) 區域中建立資源連結，指向美國東部 (維吉尼亞北部) 區域中的來源表格。`DESCRIBE`授與從歐洲 (愛爾蘭) 區域到主體之資源連結的權限。
3. 使用者使用 Athena 從歐洲 (愛爾蘭) 區域查詢資料表。

存取與外部 AWS 帳戶共用資料表的工作流程

在下圖中，生產者帳戶 (帳戶 A) 託管 Amazon S3 儲存貯體、註冊資料位置，並與美國東部 (維吉尼亞北部) 區域的消費者帳戶 (帳戶 B) 和來自歐洲 (愛爾蘭) 區域的消費者帳戶 (帳戶 B) 的使用者共用資料目錄表。



1. 資料湖管理員使用資料目錄資源和在美國東部 (維吉尼亞北部) 區域的 Lake Formation 註冊的 Amazon S3 資料位置設定帳戶 (生產者帳戶)。 AWS
2. 生產者帳戶的資料湖管理員會將「資料目錄」表格共用至取用者帳戶。
3. 消費者帳戶的資料湖管理員接受美國東部 (維吉尼亞北部) 區域的資料共用邀請，並將共用資料表的`Select`權限授與來自相同區域的主體。
4. 消費者帳戶的資料湖管理員會在歐洲 (愛爾蘭) 區域建立資源連結，指向美國東部 (維吉尼亞北部) 區域中的目標共用資料表，並授與來自歐洲 (愛爾蘭) 區域之資源連結的使用者`DESCRIBE`權限。
5. 使用者使用 Athena 查詢來自歐洲 (愛爾蘭) 區域的資料。

設定跨區域表格存取

若要存取不同區域的資料，您需要先在註冊 Amazon S3 資料位置的區域中設定資料目錄資料庫和表格。您可以與帳戶或其他帳戶中的主參與者共用資料目錄資料庫和表格。然後，您需要建立資料湖管理員，這些管理員可以建立資源連結，指向使用者查詢資料的區域中的目標共用資料位置。

若要查詢來自不同區域的同一帳戶共用資料

在此段落中，目標共用資料表「區域」稱為「區域 A」，而使用者會從區域 B 執行查詢。

1. 地區 A 中的帳戶設定 (您在此建立和共用資料)

   資料湖管理員需要完成下列動作：

   a. 註冊 Amazon S3 資料位置。

      如需詳細資訊，請參閱 將 Amazon S3 位置新增至您的資料湖。

   b. 在帳戶中建立資料庫和資料表。這也可以由具有建立資料庫和資料表權限的非系統管理使用者來完成。

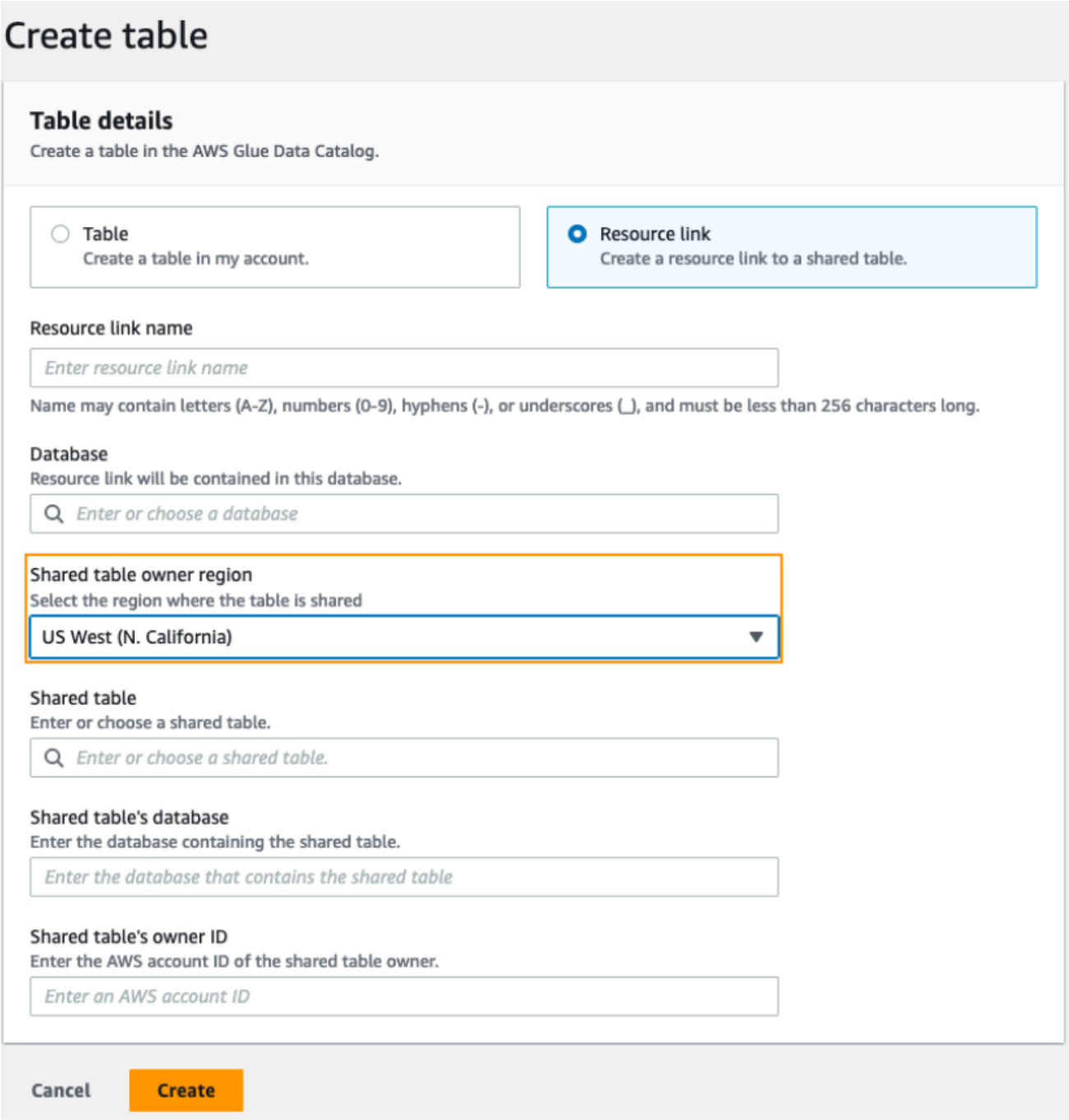
   c. 使用將資料表的資料權限授與主參與`Grantable permissions`者。

      若要取得更多資訊，請參閱授與和撤銷資料目錄資源的權限。

2. 地區 B 中的帳戶設定 (您存取資料的位置)

   資料湖管理員需要完成下列動作：

   a. 在區域 B 中建立資源連結，指向「地區 A」中的目標共用資料表。在「建立」表格畫面上指定共用資料表擁有者區域。



如需建立資料庫和表格之資源連結的指示，請參閱建立資源連結。

b. 授Describe予區域 B 中資源連結上 IAM 主體的權限。

如需授與資源連結權限的詳細資訊，請參閱授與資源連結權限。

區域 B 中的 IAM 主體可以使用 Athena 透過連結查詢目標資料表。

從不同區域存取跨帳戶資料

1. 生產者/授權人帳戶設定

   資料湖管理員需要完成下列動作：

   a. 在地區A中設定生產者/授權人帳戶
   b. 在區域 A 註冊 Amazon S3 資料位置。
   c. 創建數據庫和表。這可以由具有建立資料表權限的非系統管理使用者完成。
   d. 使用區域 A 中的資料表，將資料權限授與使用者/受權者帳戶。`Grantable permissions`

   如需詳細資訊，請參閱 跨外部帳戶 AWS 帳戶 或 IAM 主體共用資料目錄表格和資料庫。

2. 消費者/授權人帳戶設定

   資料湖管理員需要完成下列動作：

   a. 接受來自 AWS RAM 地區 A 的資源共用邀請。
   b. 在區域 B 中建立指向共用資料表的資源連結。區域 B 是用戶希望查詢表的地方。
   c. 將共用資料表的資料許可授與區域 A 中的 IAM 主體

   > **Note**
   >
   > 您必須將權限授與共用資料表的相同區域中的共用資料表。

   d. 將權限授與地區 B 中資源連結上的主參與者。

   區域 B 中消費者帳戶中的主體接著使用 Athena 從區域 B 查詢共用資料表。

資料共用 AWS Lake Formation

您可以使用 AWS Lake Formation 資料共用功能授與和管理存放在 Amazon S3 以外位置的資料，以及存放在非其他位置的中繼資料的許可 AWS Glue Data Catalog。透過資料共用功能，您可以在 Amazon Redshift 中設定和管理資料集的許可，而無需將資料遷移到 Amazon S3。您也可以使用「資料目錄」聯合功能連接至外部中繼存放區。

之後，您可以透過定義精細的存取控制原則，使用 Lake Formation 來管理中央資料目錄中的資料和存取權限。資料湖管理員可以將權限授與帳戶內的其他 IAM 主體，也可以將權限授與資料目錄資源上的跨帳戶。IAM 主體可以使用 Amazon Redshift Spectrum 和亞馬 Amazon Athena 查詢共用資料。

Lake Formation 提供下列方法來共用資料，以及管理外部資料集和外部中繼存放區的權限：

- 將湖泊形成與 Amazon Redshift 資料共用整合 — 使用 Lake Formation 集中管理 Amazon Redshift 資料庫的資料庫、表格、欄和資料列層級存取權限，並限制使用者存取資料清單中的物件。
- 連線 AWS Glue Data Catalog 至外部中繼存放區 — 使用 Lake Formation Connect AWS Glue Data Catalog 至外部中繼存放區，以管理 Amazon S3 中資料集的存取許可。不需要將中繼資料 AWS Glue Data Catalog 移轉至。
- 將湖泊形成與 AWS Data Exchange 整合 — Lake Formation 支援透過授權存取您的資料 AWS Data Exchange。如果您對 Lake Formation 資料的授權有興趣，請參閱AWS Data Exchange 使用者指南AWS Data Exchange中的內容。

主題

管理 Amazon Redshift 數據照顧中的數據許可

使用此功能 AWS Lake Formation，您可以在 Amazon Redshift 的資料記錄中安全地管理資料。Amazon Redshift 是雲端中的全受管 PB 級資料倉儲服務。 AWS 使用資料共用功能，Amazon Redshift 可協助您在各 AWS 帳戶處共用資料。如需有關 Amazon Redshift 資料共用的詳細資訊，請參閱 Amazon Redshift 中的資料共用概觀。

在 Amazon Redshift 中，生產者叢集管理員會建立資料清單，並與資料湖管理員共用。 step-by-step 如需建立資料湖管理員的指示，請參閱建立資料湖管理員。

在您（資料湖管理員）接受資料保護之後，您必須為特定的資料保護建立資料 AWS Glue Data Catalog 庫。這樣您就可以使用 Lake Formation 權限控制對其的訪問。Lake Formation 會將每個資料清單對應至對應的「資料目錄」資料庫。這些資料庫會在資料目錄中顯示為聯合資料庫。

當資料庫指向「資料目錄」外部的圖元時，即稱為聯合資料庫。Amazon Redshift 資料清單中的表格和檢視會在資料目錄中以個別表格的形式列出。您可以與相同帳戶內的選取 IAM 主體和 SAML 使用者共用聯合資料庫，或與 Lake Formation 的另一個帳戶共用聯合資料庫。您也可以包含列與欄篩選運算式，以限制對特定資料的存取。如需詳細資訊，請參閱 資料篩選概觀。

若要讓使用者能夠存取 Amazon Redshift 資料照護裝置，您必須執行下列動作：

1. 更新資料目錄設定以啟用 Lake Formation 權限。
2. 接受來自 Amazon Redshift 生產者叢集管理員的資料清單邀請，並在 Lake Formation 中註冊資料保護。

   完成此步驟後，您可以管理 Lake Formation 成資料目錄中的資料清單。
3. 建立聯合資料庫並定義該資料庫的權限。
4. 將權限授與資料庫和資料表上的使用者。您可以與相同帳戶或其他帳戶中的使用者共用整個資料庫或資料表子集。

如需限制的詳細資訊，請參閱Amazon Redshift 數據共享限制。

主題

- 在 Amazon Redshift 數據庫上設置許可的先決條件
- 為 Amazon Redshift 數據庫設置許可
- 查詢聯合資料庫

在 Amazon Redshift 數據庫上設置許可的先決條件

更新預設資料目錄設定

若要啟用「資料目錄」資源的 Lake Formation 權限，建議您停用 Lake Formation 中的預設「資料目錄」設定。如需詳細資訊，請參閱 變更預設權限模型或使用混合存取模式。

更新權限

除了資料湖管理員許可 (`AWSLakeFormationDataAdmin`) 之外，還需要下列許可才能在 Lake Formation 中接受 Amazon Redshift 資料識別：

- glue:PassConnection on aws:redshift
- redshift:AssociateDataShareConsumer
- redshift:DescribeDataSharesForConsumer
- redshift:DescribeDataShares

資料湖管理員 IAM 使用者隱含具有下列權限。

- 資料位置存取
- 建立資料庫 (_I)
- 湖泊照明:註冊資源

為 Amazon Redshift 數據庫設置許可

本主題說明接受資料查看邀請、建立聯合資料庫及授與權限所需遵循的步驟。您可以使用 Lake Formation 控制台或 AWS Command Line Interface （ AWS CLI ）。本主題中的範例顯示相同帳戶中的生產者叢集、資料目錄和資料用戶。

要了解有關 Lake Formation 跨帳戶功能的更多信息，請參閱Lake Formation 的跨賬戶數據共享。

若要設定資料清單的權限

1. 檢閱資料查看邀請並接受。

   Console

   1. 以資料湖管理員身分登入湖泊形成主控台，網址為 https://console.aws.amazon.com/lakeformation/。導覽至「資料共用」頁面。
   2. 檢閱您獲得授權存取的資料庫。「狀態」(S tatus) 欄會指出您目前的資料存取參與狀態。「擱置中」狀態表示您已新增至資料存放區，但您尚未接受或拒絕邀請。
   3. 若要回應資料查看邀請，請選取資料清單名稱，然後選擇 [檢閱邀請]。在接受或拒絕資料清單中，檢閱邀請詳細資料。選擇「接受」以接受邀請，或選擇「拒絕」以拒絕邀請。如果拒絕邀請，則無法訪問數據保護。

   AWS CLI

   下列範例顯示如何檢視、接受及註冊邀請。將 AWS 帳戶 ID 取代為有效的 AWS 帳戶 ID。將其取data-share-arn代為參考資料識別的實際 Amazon 資源名稱 (ARN)。

1. 檢視擱置中的邀請。

```
aws redshift describe-data-shares \
 --data-share-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds' \
```

2. 接受一個數據標題。

```
 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. 在 Lake Formation 帳戶中註冊數據存儲器。使用 RegisterResourceAPI 操作在 Lake Formation 中註冊數據清單。 `DataShareArn`是的輸入參數`ResourceArn`。

> **Note**
>
> 這是一個強制性的步驟。

```
aws lakeformation register-resource \
 --resource-arn 'arn:aws:redshift:us-
east-1:111122223333:datashare:abcd1234-1234-ab12-cd34-1a2b3c4d5e6f/
federatedds'
```

2. 建立資料庫。

接受資料查看邀請之後，您需要建立一個資料庫，該資料庫指向與該資料保護關聯的 Amazon Redshift 資料庫。您必須是資料湖管理員才能建立資料庫。

Console

1. 從「邀請」窗格中選取資料清單，然後選擇「設定資料庫詳細資訊」。
2. 在設定資料庫詳細資訊中，輸入資料保護的唯一名稱和識別碼。您可以使用此識別碼在中繼資料階層 (資料庫名稱 .schema.Table) 內部對應資料指令。
3. 選擇下一步，將權限授與共用資料庫和表格上的其他使用者。

AWS CLI

使用下列範例程式碼建立資料庫，該資料庫指向使用與 Lake Formation 共用的 Amazon Redshift 資料庫。 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. 授予權限。

建立資料庫之後，您可以將權限授與帳戶中的使用者或外部 AWS 帳戶 和組織。您將無法在對應至 Amazon Redshift 資料識別的聯合資料庫上授與寫入資料許可 (插入、刪除) 和中繼資料權限 (更改、刪除、建立)。如需授與權限的詳細資訊，請參閱管理 Lake Formation 權限。

> **Note**
>
> 身為資料湖管理員，您只能檢視聯合資料庫中的表格。若要執行任何其他動作，您需要授與自己這些資料表的更多權限。

Console

1. 在「授與權限」畫面上，選取要授與權限的使用者。
2. 選擇 Grant (授予)。

AWS CLI

您可以使用下列範例來授與資料庫和資料表權限 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"
                     ]
```

```
}
```

查詢聯合資料庫

授與許可後，使用者可以使用 Amazon Redshift 登入並開始查詢聯合資料庫。使用者現在可以使用本機資料庫名稱來參考 SQL 查詢中的 Amazon Redshift 資料識別。在 Amazon Redshift 中，透過資料清單共用的公用結構描述中的客戶資料表將具有如資料目錄`public.customer`中所建立的對應表格。

1. 在使用 Amazon Redshift 查詢聯合資料庫之前，叢集管理員使用下列命令從資料目錄資料庫建立資料庫：

```
CREATE DATABASE sharedcustomerdb FROM ARN
 'arn:aws:glue:<region>:111122223333:database/tahoedb' WITH DATA CATALOG SCHEMA
 tahoedb
```

2. 叢集管理員會授與資料庫的使用權限。

```
GRANT USAGE ON DATABASE sharedcustomerdb TO IAM:user;
```

3. 您 (聯合使用者) 現在可以登入 SQL 工具來查詢資料表。

```
Select * from sharedcustomerdb.public.customer limit 10;
```

如需詳細資訊，請參閱 Amazon Redshift 管理指南 AWS Glue Data Catalog中的查詢。

管理使用外部中繼存放區之資料集的權限

使用 AWS Glue Data Catalog 中繼資料聯合 (資料目錄聯合)，您可以將資料目錄連接到存放 Amazon S3 資料中繼資料的外部中繼資料的外部中繼資料，並使用 AWS Lake Formation安全地管理資料存取許可。您不需要將中繼資料從外部中繼存放區移轉至資料目錄。

資料目錄提供集中的中繼資料儲存庫，可讓跨不同系統的資料管理和探索更加輕鬆。當您的組織管理資料目錄中的資料時，您可以用 AWS Lake Formation 來控制對 Amazon S3 中資料集的存取。

> **Note**
>
> 目前，我們僅支援 Apache Hive（版本 3 及更高版本）中繼儲存庫聯盟。

若要設定資料目錄聯盟，我們HiveMetastore在中提供名為 GlueDataCatalogFederation- 的 AWS Serverless Application Model (AWS SAM) 應用程式 AWS Serverless Application Repository。

參考實 GitHub 作是以開放原始碼專案的形式在AWS Glue Data Catalog 聯合-Hive 中繼存放區中提供。

AWS SAM 應用程式會建立並部署下列資源，以便將資料目錄連線至 Hive 中繼存放區所需的資源：

- AWS Lambda 函數 — 主控在資料目錄和 Hive 中繼存放區之間進行通訊的聯合服務的實作。 AWS Glue 調用此 Lambda 函數從蜂巢中繼存儲中檢索元數據對象。
- Amazon API Gateway-Hive 中繼存儲的連接端點，充當代理，將所有調用路由到 Lambda 函數。
- IAM 角色 — 具有建立資料目錄和 Hive 中繼存放區之間連線所需權限的角色。
- AWS Glue 連線 — 一 Amazon API Gateway 種存放端點的 AWS Glue 連線類型，以及要呼叫 Amazon API Gateway 端點的 IAM 角色。

當您查詢資料表時， AWS Glue 服務會對 Hive 中繼儲存區進行執行階段呼叫，並擷取中繼資料。Lambda 函數充當蜂巢中繼存儲和數據目錄之間的轉換器。

建立連線之後，為了將 Hive 中繼資料與資料目錄同步處理中繼資料，您需要使用 Hive 中繼存放區連線詳細資料在資料目錄中建立聯合資料庫，並將此資料庫對應至 Hive 資料庫。當資料庫指向「資料目錄」外部的圖元時，即稱為聯合資料庫。

您可以使用以標籤為基礎的存取控制和聯合資料庫上的具名資源方法來套用 Lake Formation 權限，並在多 AWS 帳戶個組織單位 (OU) 之間共用。 AWS Organizations您也可以直接與其他帳戶的 IAM 主體共用聯合資料庫。

您可以使用外部 Hive 資料表上的 Lake Formation 資料篩選器，在資料行層級、資料列層級和儲存格層級定義細微的權限。您可以使用 Amazon Athena，Amazon Redshift 或 Amazon EMR 查詢 Lake Formation 託管的外部蜂巢表。

如需跨帳戶資料共用和資料篩選的詳細資訊，請參閱：

- Lake Formation 的跨賬戶數據共享
- Lake Formation 中的數據過濾和細胞級安全

資料目錄中繼資料同盟高階步

1. 您可以建立具有適當權限的 IAM 使用者和角色來部署 AWS SAM 應用程式和建立聯合資料庫。

2. 您可以透過選取使用外部 Hive 中繼存放區的資料集`Enable Data Catalog federation`選項，向 Lake Formation 註冊 Amazon S3 資料位置。
3. 您可以設定 AWS SAM 應用程式設定 (AWS Glue 連線名稱、Hive 中繼存放區的 URL 以及 Lambda 函數參數)，然後部署 AWS SAM 應用程式。
4. AWS SAM 應用程式會部署連接外部 Hive 中繼存放區與資料目錄所需的資源。
5. 若要在 Hive 資料庫和資料表上套用 Lake Formation 權限，您可以使用 Hive 中繼存放區連線詳細資料在資料目錄中建立資料庫，並將此資料庫對應至 Hive 資料庫。
6. 將聯合資料庫的權限授與您帳戶或其他帳戶中的主體。

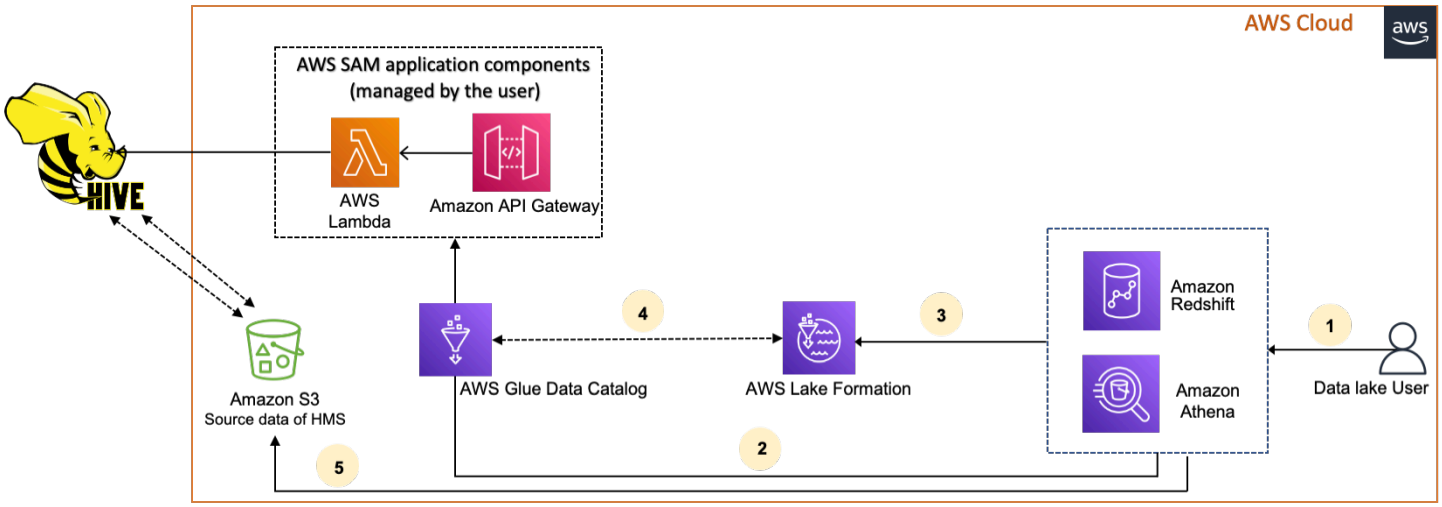
> **Note**
>
> 您可以將資料目錄連線到外部 Hive mestastore、建立聯合資料庫，以及在 Hive 資料庫和資料表上執行查詢和 ETL 指令碼，而無需套用 Lake Formation 權限。對於未向 Lake Formation 註冊的 Amazon S3 中的來源資料，存取由 Amazon S3 的 IAM 許可政策和 AWS Glue 動作決定。

如需限制的詳細資訊，請參閱Hive 中繼資料儲存資料共用考量和限制。

主題

工作流程

下圖顯示了連接到外部 Hive 中 AWS Glue Data Catalog 繼存儲的工作流程。



1. 主體使用整合式服務 (例如 Athena 或 Redshift 頻譜) 提交查詢。
2. 整合式服務會呼叫中繼資料的資料目錄，進而呼叫後方可用的 Hive 中繼存放區端點 Amazon API Gateway，並接收中繼資料要求的回應。
3. 整合式服務會將要求傳送至 Lake Formation，以驗證資料表資訊和憑證以存取資料表。
4. Lake Formation 授權請求並將臨時憑據出售給集成的應用程序，從而允許數據訪問。
5. 整合式服務會使用從 Lake Formation 接收到的臨時登入資料，從 Amazon S3 讀取資料，並將結果分享給主體。

將資料目錄連線到 Hive 中繼存放區的先決條件

若要連線 AWS Glue Data Catalog 至外部 Apache Hive 中繼存放區並設定資料存取權限，您需要完成下列需求：

> **Note**
>
> 我們建議 Lake Formation 管理員部署應用 AWS SAM 程式，而且只有有權限的使用者使用 Hive 中繼存放區連線來建立對應的聯合資料庫。

1. 建立 IAM 角色。

   若要部署 AWS SAM 應用程式

   - 建立具有部署資源 (Lambda 函數 Amazon API Gateway、IAM 角色和 AWS Glue 連線) 所需權限的角色，以建立與 Hive 中繼存放區的連線。

若要建立聯合資料庫

資源需要下列權限：

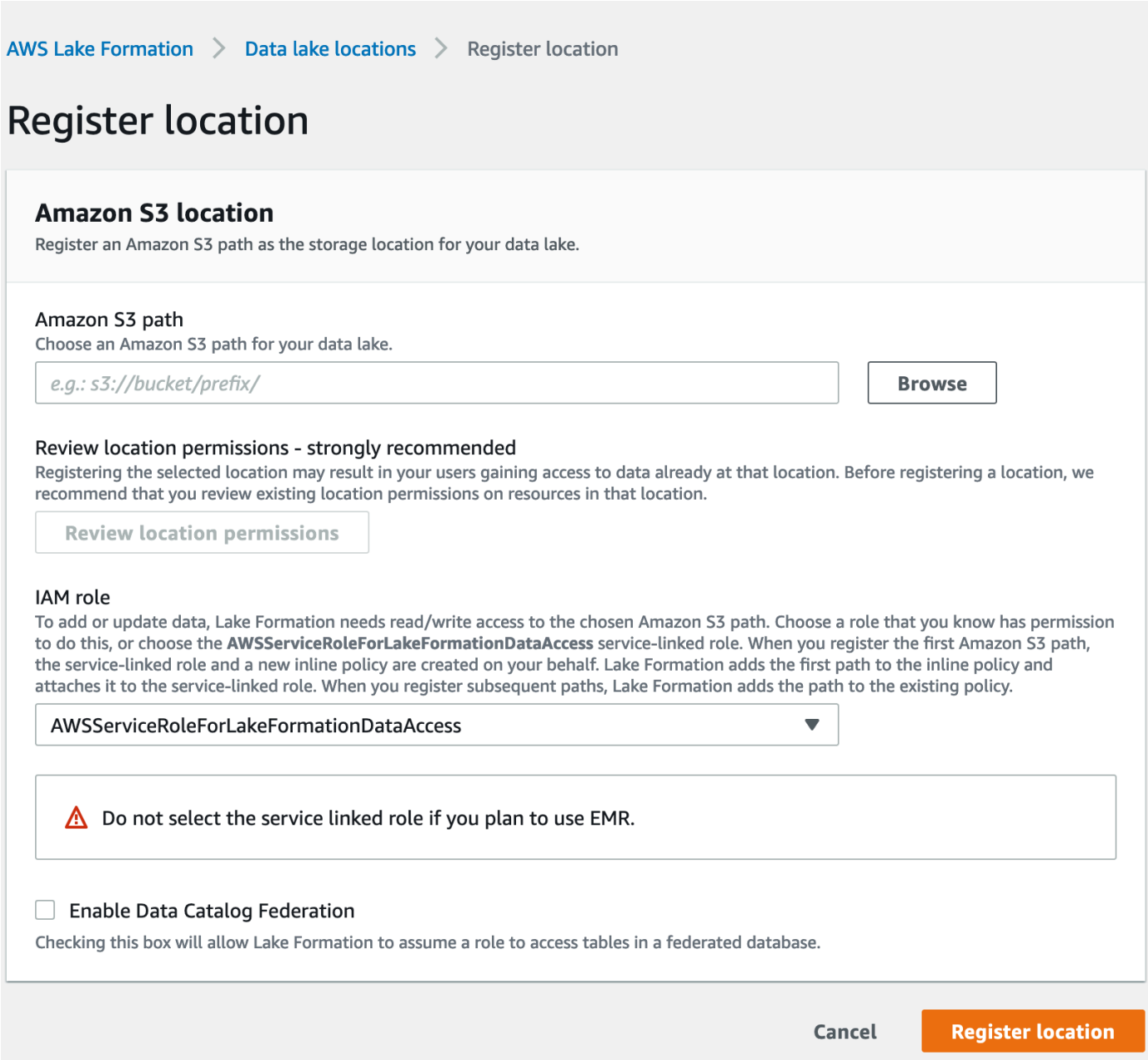
- `glue:CreateDatabase on resource arn:aws:glue:region:account-id:database/gluedatabasename`
- `glue:PassConnection on resource arn:aws:glue:region:account-id:connection/hms_connection`

2. 向 Lake Formation 註冊 Amazon S3 位置。

   若要使用 Lake Formation 來管理和保護資料湖中的資料，您必須註冊具有 Hive 中繼存放區中表資料的 Amazon S3 位置與 Lake Formation。通過這樣做，Lake Formation 可以將憑據出售給 AWS 分析服務，例如 Athena，Redshift 頻譜和 Amazon EMR。

   如需註冊 Amazon S3 位置的詳細資訊，請參閱將 Amazon S3 位置新增至您的資料湖。

   註冊 Amazon S3 位置時，請選取啟用資料目錄聯合核取方塊，以允許 Lake Formation 擔任角色來存取聯合資料庫中的表格。



若要取得有關向 Lake Formation 註冊資料位置的更多資訊，請參閱[為您的資料湖設定 Amazon S3 位置](#)。

3. 使用正確的 Amazon EMR 版本。

   要使用 Amazon EMR 與聯合蜂巢中繼存儲數據庫，你需要有蜂巢版本 3.x 或更高版本和 Amazon EMR 版本 6.x 或更高。

將數據目錄連接到外部 Hive 中繼存儲

若要連接 AWS Glue Data Catalog 到 Hive 中繼存放區，您需要部署名為 GlueDataCatalogFederation-HiveMetastore 的 AWS SAM 應用程式。它創建與數據目錄連接外部 Hive 中繼存儲所需的資源。您可以在中存取 AWS SAM 應用程式 AWS Serverless Application Repository。

AWS SAM 應用程式會使用 Lambda 函數為 Amazon API Gateway 後方的 Hive 中繼存放區建立連線。該 AWS SAM 應用程序使用統一資源標識符（URI）作為來自用戶的輸入，並將外部 Hive 中繼存儲連接到數據目錄。當使用者在 Hive 資料表上執行查詢時，資料目錄會呼叫 API Gateway 端點。端點會叫用 Lambda 函數以擷取 Hive 資料表的中繼資料。

將資料目錄連線至 Hive 中繼存放區並設定權限

1. 部署 AWS SAM 應用程式。

    1. 登入 AWS Management Console 並開啟 AWS Serverless Application Repository.
    2. 選擇在導覽窗格中的 Available applications (可用的應用程式)。
    3. 選擇公用應用程式。
    4. 選取選項 Show apps that create custom IAM roles or resource policies (顯示建立自訂 IAM 角色或資源政策的應用程式)。
    5. 在搜索框中，輸入名稱 GlueDataCatalogFederation-HiveMetastore。
    6. 選擇 GlueDataCatalogFederation-HiveMetastore 應用程序。
    7. 在「應用程式設定」 下，為 Lambda 函數輸入下列所需的最低設定：
        - 應用程式名稱- AWS SAM 應用程式的名稱。
        - GlueConnectionName-連接的名稱。
        - HiveMetastoreURI-您的配置單元存儲主機的 URI。
        - LambdaMemory-Lambda 內存量，以 MB 為單位，從 128-10240。預設值為 1024。
        - LambdaTimeout-Lambda 叫用執行階段上限 (以秒為單位)。預設值為 30。
        - VPC SecurityGroupIds 和 VPC SubnetIds-Hive 中繼存放區所在之 VPC 的資訊。
    8. 選擇 I acknowledge that this app creates custom IAM roles and resource policies (我認可此應用程式建立自訂的 IAM 角色和資源政策)。如需詳細資訊，請選擇 Info (資訊) 連結。
    9. 在 Application settings (應用程式設定) 部分的右下方，選擇 Deploy (部署)。部署完成後，Lambda 函數會出現在 Lambda 主控台的 Resources (資源) 區段中。

應用程式已部署至 Lambda。它的名稱前面加上無伺服器回購- 表示應用程式是從. AWS Serverless Application Repository選取應用程式會帶您前往「資源」頁面，其中會列出已建置之應用程式的每個資源。這些資源包括允許資料目錄和 Hive 中繼存放區之間進行通訊的 Lambda 函數、 AWS Glue 連線以及資料庫聯合所需的其他資源。

2. 在「資料目錄」中建立聯合資料庫。

   建立 Hive 中繼存放區的連線之後，您可以在資料目錄中建立指向外部 Hive 中繼存放區資料庫的聯合資料庫。您需要在資料目錄中為每個 Hive 中繼存放區資料庫連線至資料目錄建立對應的資料庫。

   Lake Formation console

   1. 在 [資料共用] 頁面上，選擇 [共用資料庫] 索引標籤，然後選擇 [建立資料庫]。
   2. 對於連接名稱，從下拉菜單中選擇您的 Hive 中繼存儲連接的名稱。
   3. 輸入資料庫的唯一資料庫名稱和聯合來源識別碼。這是您在查詢資料表時在 SQL 陳述式中使用的名稱。名稱最多可包含 255 個字元，而且在您的帳戶中必須是唯一的。
   4. 選擇建立資料庫。

   AWS CLI

   ```
   aws glue create-database \
   '{
    "CatalogId": "<111122223333>",
     "database-input": {
       "Name":"<fed_glue_db>",
       "FederatedDatabase":{
           "Identifier":"<hive_db_on_emr>",
           "ConnectionName":"<hms_connection>"
        }
      }
    }'
   ```

3. 檢視聯合資料庫中的表格。

   建立聯合資料庫之後，您可以使用 Lake Formation 主控台或. AWS CLI

Lake Formation console

1. 從 [共用資料庫] 索引標籤中選取資料庫名稱。
2. 在「資料庫」頁面上，選擇檢視表格。

AWS CLI

下列範例說明如何擷取連線定義、資料庫名稱，以及資料庫中的部分或所有資料表。將資料目錄的 ID 取代為您用來建立資料庫的有效 AWS 帳戶 ID。hms_connection以連線名稱取代。

```
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. 授予權限。

建立資料庫之後，您可以將權限授與帳戶中的其他 IAM 使用者和角色，或授 AWS 帳戶 與外部和組織。您將無法授與聯合資料庫的寫入資料權限 (插入、刪除) 和中繼資料權限 (變更、刪除、建立)。如需授與權限的詳細資訊，請參閱管理 Lake Formation 權限。

5. 查詢聯合資料庫。

授與許可後，使用者可以使用 Athena 和 Amazon Redshift 登入並開始查詢聯合資料庫。使用者現在可以使用本機資料庫名稱來參考 SQL 查詢中的 Hive 資料庫。

Amazon Athena 查詢語法範例

以先前建立的本機資料庫名稱取fed_glue_db代。

```
Select * from fed_glue_db.customers limit 10;
```

其他資源

以下部落格文章包含在 Hive 中繼存放區資料庫和資料表上設定 Lake Formation 權限的詳細說明，並使用 Athena 進行查詢。我們還說明了一個跨帳戶共享用例，其中生產者帳戶 A 中的 LLake Formation ms 主體共享聯合 Hive 數據庫和使用 LF-tag 到消費者帳戶 B 的表。

- 使 AWS Lake Formation 用權限查詢您的 Apache 蜂巢中繼存儲

中的安全性 AWS Lake Formation

雲安全 AWS 是最高的優先級。身為 AWS 客戶，您可以從資料中心和網路架構中獲益，該架構專為滿足對安全性最敏感的組織的需求而打造。

安全是 AWS 與您之間共同承擔的責任。共同責任模型 將此描述為雲端的安全和雲端內的安全：

- 雲端的安全性 — AWS 負責保護在 AWS 雲端中執行 AWS 服務的基礎架構。 AWS 還為您提供可以安全使用的服務。在 AWS 合規計畫中，第三方稽核員會定期測試並驗證我們的安全功效。若要深入瞭解適用於的規範遵循計劃 AWS Lake Formation，請參閱合規方案的AWS 服務範圍。
- 雲端中的安全性 — 您的責任取決於您使用的 AWS 服務。您也必須對其他因素負責，包括資料的機密性、您公司的要求和適用法律和法規。

本文件可協助您瞭解如何在使用 Lake Formation 時應用共同的責任模型。下列主題說明如何設定 Lake Formation 以符合您的安全性和合規性目標。您還將學習如何使用其他 AWS 服務來幫助您監控和保護您的 Lake Formation 資源。

主題

- Lake Formation 的數據保護
- 基礎架構安全性 AWS Lake Formation
- 預防跨服務混淆代理人
- 安全性事件登入 AWS Lake Formation

Lake Formation 的數據保護

AWS 共同責任模式適用於 AWS Lake Formation 的資料保護。如此模型中所述， AWS 負責保護執行所有 AWS 雲端. 您負責維護在此基礎設施上託管內容的控制權。您也同時負責所使用 AWS 服務 的安全組態和管理任務。如需資料隱私權的詳細資訊，請參閱資料隱私權常見問答集。如需有關歐洲資料保護的相關資訊，請參閱 AWS 安全性部落格上的 AWS 共同的責任模型和 GDPR 部落格文章。

基於資料保護目的，我們建議您使用 AWS IAM Identity Center 或 AWS Identity and Access Management (IAM) 保護 AWS 帳戶 登入資料並設定個別使用者。如此一來，每個使用者都只會獲得授與完成其任務所必須的許可。我們也建議您採用下列方式保護資料：

- 每個帳戶均要使用多重要素驗證 (MFA)。

- 使用 SSL/TLS 與 AWS 資源進行通訊。我們需要 TLS 1.2 並建議使用 TLS 1.3。
- 使用設定 API 和使用者活動記錄 AWS CloudTrail。
- 使用 AWS 加密解決方案以及其中的所有默認安全控制 AWS 服務。
- 使用進階的受管安全服務 (例如 Amazon Macie)，協助探索和保護儲存在 Amazon S3 的敏感資料。
- 如果您在透過命令列介面或 API 存取時需要經 AWS 過 FIPS 140-2 驗證的加密模組，請使用 FIPS 端點。如需有關 FIPS 和 FIPS 端點的更多相關資訊，請參閱聯邦資訊處理標準 (FIPS) 140-2 概觀。

我們強烈建議您絕對不要將客戶的電子郵件地址等機密或敏感資訊，放在標籤或自由格式的文字欄位中，例如名稱欄位。這包括當您使用控制台，API 或 AWS SDK AWS 服務 使用 Lake Formation 或其他工作時。 AWS CLI您在標籤或自由格式文字欄位中輸入的任何資料都可能用於計費或診斷日誌。如果您提供外部伺服器的 URL，我們強烈建議請勿在驗證您對該伺服器請求的 URL 中包含憑證資訊。

靜態加密

AWS Lake Formation 支援下列方面的資料加密：

- 亞馬遜簡單儲存服務 (Amazon S3) 資料湖中的資料。

  Lake Formation 支持使用 AWS Key Management Service（ AWS KMS ）的數據加密。資料通常會透過AWS Glue擷取、轉換和載入 (ETL) 工作將資料寫入資料湖。如需如何加密AWS Glue工作所寫入之資料的詳細資訊，請參閱《AWS Glue 開發人員指南》中的加密編目器、作業和開發端點所寫入的資料。
- 這是 Lake Formation 儲存描述資料湖中資料的中繼資料表的位置。 AWS Glue Data Catalog

  如需詳細資訊，請參閱AWS Glue 開發人員指南中的加密資料目錄。

若要將 Amazon S3 位置新增為資料湖中的儲存，請使用註冊該位置 AWS Lake Formation。然後，您可以使用 Lake Formation 權限，對指向此位置的 AWS Glue Data Catalog 物件以及位置中的基礎資料進行精細的存取控制。

Lake Formation 支援註冊包含加密資料的 Amazon S3 位置。如需更多詳細資訊，請參閱 註冊加密的 Amazon S3 位置。

基礎架構安全性 AWS Lake Formation

作為受管服務， AWS Lake Formation 受 Amazon Web Services：安 AWS 全流程概觀白皮書中所述的全球網路安全程序保護。

您可以使用 AWS 已發布的 API 調用通過網絡訪問 Lake Formation。用戶端必須支援 Transport Layer Security (TLS) 1.0 或更新版本。建議使用 TLS 1.2 或更新版本。用戶端也必須支援具備完美轉送私密 (PFS) 的密碼套件，例如臨時 Diffie-Hellman (DHE) 或橢圓曲線臨時 Diffie-Hellman (ECDHE)。現代系統 (如 Java 7 和更新版本) 大多會支援這些模式。

此外，請求必須使用存取金鑰 ID 和與 IAM 主體相關聯的私密存取金鑰來簽署。或者，您可以使用 AWS Security Token Service (AWS STS) 以產生暫時安全憑證以簽署請求。

預防跨服務混淆代理人

混淆代理人問題屬於安全性問題，其中沒有執行動作許可的實體可以強制具有更多許可的實體執行該動作。在中 AWS，跨服務模擬可能會導致混淆的副問題。在某個服務 (呼叫服務) 呼叫另一個服務 (被呼叫服務) 時，可能會發生跨服務模擬。可以操縱呼叫服務來使用其許可，以其不應有存取許可的方式對其他客戶的資源採取動作。為了防止這種情況發生， AWS 提供的工具可協助您透過已授予您帳戶中資源存取權的服務主體來保護所有服務的資料。

若要限制 AWS Lake Formation 為資源提供另一項服務的許可，我們建議在資源政策中使用 aws:SourceArn 和 aws:SourceAccount 全域條件內容索引鍵。如果同時使用全域條件內容索引鍵，則在相同政策陳述式中使用 aws:SourceAccount 值和 aws:SourceArn 值中的帳戶時，必須使用相同的帳戶 ID。

目前，Lake Formation 僅支持aws:SourceArn以下格式：

```
arn:aws:lakeformation:aws-region:account-id:*
```

下列範例說明如何在 Lake Formation 中使用aws:SourceArn和aws:SourceAccount全域條件上下文索引鍵，以防止混淆的副問題。

```
{
  "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:*"
        }
      }
    }
  ]
}
```

安全性事件登入 AWS Lake Formation

AWS Lake Formation 與該服務集成在一起 AWS CloudTrail，該服務可提供用戶，角色或 AWS 服務在 Lake Formation 中採取的行動記錄。 CloudTrail 捕獲所有 API 呼叫 Lake Formation 作為事件。捕獲的呼叫包括來自 Lake Formation 控制台的呼叫 AWS Command Line Interface，以及對 Lake Formation API 操作的代碼調用。

如需有關 Lake Formation 中事件記錄的詳細資訊，請參閱使用記錄 AWS Lake Formation API 調用 AWS CloudTrail。

> **Note**
>
> `GetTableObjects`、`UpdateTableObjects`和`GetWorkUnitResults`是大量資料平面作業。目前未記錄對這些 API 的呼叫 CloudTrail。如需有關中資料平面作業的詳細資訊 CloudTrail，請參閱《AWS CloudTrail 使用指南》中的記錄追蹤的資料事件。
> Lake Formation 的變化以支持其他 CloudTrail 活動將記錄在的文件歷史記錄 AWS Lake Formation。

整合第三方服務 Lake Formation

與整合AWS Lake Formation可讓第三方服務安全地存取其 Amazon S3 資料湖中的資料。您可以使用 Lake Formation 作為授權引擎，透過整合式 AWS 服務 (例如 Amazon 雅典娜、亞馬遜 EMR 和 Redshift 頻譜) 來管理或強制執行對資料湖的許可。Lake Formation 為整合服務提供了兩種選擇：

1. Lake Formation 應用程式整合設定：Lake Formation 可以根據有效許可，將 AWS STS 令牌形式的臨時登入資料分配到已註冊的 Amazon S3 位置，以便授權的應用程式可以代表使用者存取資料。
2. 集中強制執行：Lake Formation 查詢 API 操作會從 Amazon S3 擷取資料，並根據有效許可篩選結果。與查詢 API 作業整合的引擎或應用程式可以依賴 Lake Formation 來評估呼叫身分的權限，並根據這些權限安全地篩選資料。第三方查詢引擎只能查看和對過濾的數據進行操作。

主題

使用 Lake Formation 應用程式整合

Lake Formation 允許第三方服務與 Lake Formation 整合，並透過使用和GetTemporaryGluePartitionCredentials操作代表其使用GetTemporaryGlueTableCredentials者暫時存取 Amazon S3 資料。這允許第三方服務使用與其他 AWS 分析服務使用的相同授權和憑證自動售貨功能。本節說明如何使用這些 API 作業與第三方查詢引擎整合Lake Formation。

預設會停用這些 API 作業。有兩個選項可以授權 Lake Formation 集成應用程序：

- 設定每次呼叫應用程式整合 API 作業時都會驗證的 IAM 工作階段標籤

  如需詳細資訊，請參閱 啟用第三方查詢引擎的權限，以呼叫應用程式整合 API 作業。
- 啟用允許外部引擎在具有完整表格存取權的 Amazon S3 位置存取資料的選項

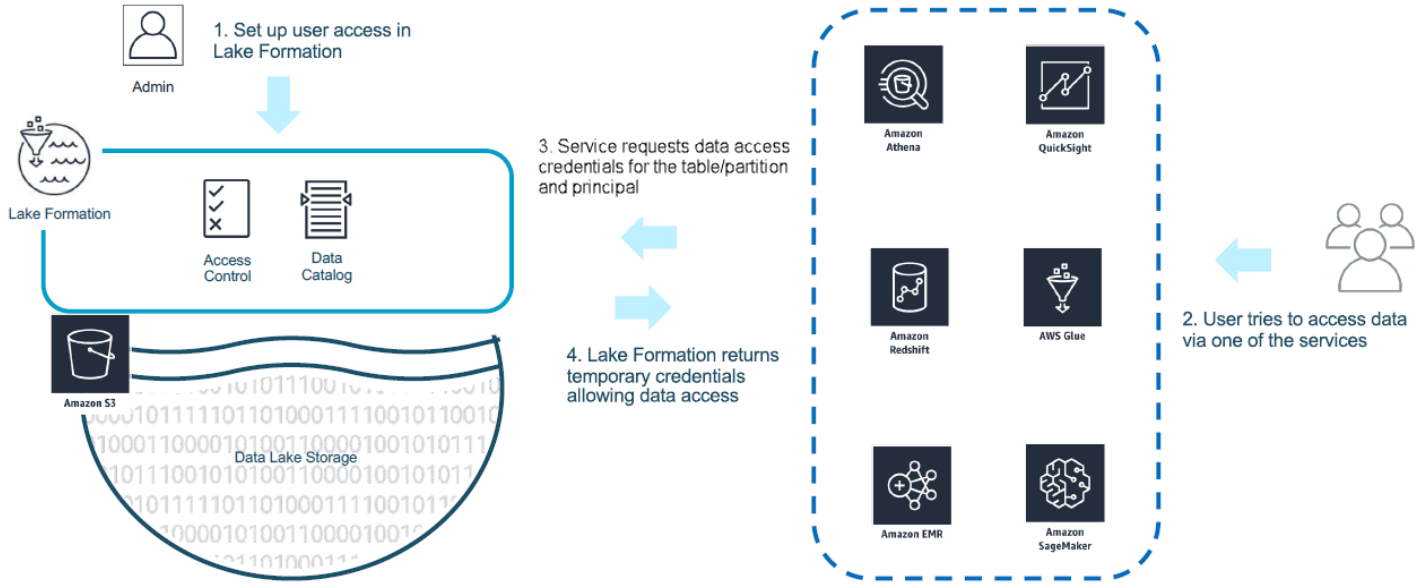
  如果使用者具有完整資料表存取權，此選項可讓查詢引擎和應用程式取得沒有 IAM 工作階段標籤的認證。它提供查詢引擎和應用程式的效能優勢，以及簡化資料存取。Amazon EC2 上的亞馬遜 EMR 能夠利用此設置。

  如需詳細資訊，請參閱 完整表格存取的應用程式整合 。

主題

湖平整應用程序集成如何工作

本節說明如何使用應用程式整合 API 作業，將協力廠商應用程式 (查詢引擎) 與整合Lake Formation。



1. Lake Formation管理員執行下列活動：
   - 透過提供 IAM 角色 (用於自動售貨登入資料) 向 Lake Formation 註冊 Amazon S3 位置，該角色具有存取 Amazon S3 位置內資料的適當許可
   - 註冊第三方應用程式，以便能夠呼叫 Lake Formation 的憑證自動販賣 API 作業。請參閱the section called “註冊第三方查詢引擎”
   - 授與使用者存取資料庫和資料表的權限

     例如，如果您想要發佈使用者工作階段資料集，其中包含一些包含個人識別資訊 (PII) 的資料行，以限制存取，您可以為這些欄指派名為「分類」的 LF-TBAC 標記，其值為「敏感」。接下來，您要定義允許商務分析師存取使用者工作階段資料的權限，但排除標記為「分類 = 敏感」的欄。
2. 主體 (使用者) 將查詢提交至整合式服務。

3. 集成的應用程序將請求發送到 Lake Formation，要求提供表信息和憑據以訪問表格。
4. 如果查詢主體獲得存取資料表的授權，Lake Formation 會將認證傳回給整合式應用程式，以允許資料存取。

> **Note**
>
> 自動售貨憑據時，Lake Formation 不會訪問基礎數據。

5. 整合式服務會從 Amazon S3 讀取資料，根據收到的政策篩選欄，然後將結果傳回主體。

> **Important**
>
> Lake Formation憑據自動售貨 API 操作可以使用明確拒絕故障（故障關閉）模型的分佈式強制執行。這引入了客戶，第三方服務和 Lake Formation 之間的三方安全模型。整合式服務受到信任，可正確強制執行Lake Formation權限 (分散式強制執行)。

整合式服務負責根據篩選資料傳回給使用者Lake Formation之前從傳回的政策篩選從 Amazon S3 讀取的資料。整合式服務遵循故障關閉模型，這表示如果無法強制執行必要Lake Formation的權限，則查詢就必須失敗。

湖泊形成應用程序集成中的角色和責任

角色	責任
客戶	• 啟用 Lake Formation 應用程式整合設定 (請參閱the section called “註冊第三方查詢引擎”)。 • 向 Lake Formation 明確註冊已批准的第三方（請參閱the section called “註冊第三方查詢引擎”）。 • 使用 Lake Formation 權限測試和驗證第三方解決方案。 • 監控和審核 Lake Formation 憑證自動售貨 API 操作的第三方使用情況。
第三方	• 公開記錄每個軟體修訂版本的支援功能，並提供正確啟用的指示。 • 在調用 Lake Formation 憑證自動售貨 API 操作時準確地宣傳支持的功能（根據文檔）。

角色	責任
	• 安全地存儲和處理付款憑據，以避免憑證洩漏和權限提升。 • 根據支援的功能強制執行權限，並僅將篩選的資料傳回給使用者 • 無法正確強制執行所需權限時，查詢失敗
AWS Lake Formation	• 正確衍生並傳回指定主參與者的有效權限。 • call-by-call 根據 API 作業驗證協力廠商支援的功能。 • 只有當引擎的廣告功能與目錄資源上定義的功能相符時，才傳回範圍下的 IAM 登入資料，否則會傳回錯誤。

Lake Formation應用程式整合 API 作業工作流程

以下是應用程序集成 API 操作的工作流程：

1. 用戶使用集成的第三方查詢引擎提交查詢或數據請求。查詢引擎會採用代表使用者或使用者群組的 IAM 角色，並擷取呼叫應用程式整合 API 作業時要使用的受信任登入資料。
2. 查詢引擎會呼叫`GetUnfilteredTableMetadata`，如果是分區資料表，則查詢引擎會呼叫`GetUnfilteredPartitionsMetadata`以從「資料目錄」擷取中繼資料和原則資訊。
3. Lake Formation 執行對請求的授權。如果使用者對資料表沒有適當的權限，AccessDeniedException則會擲回。
4. 作為請求的一部分，查詢引擎會傳送它支援的篩選。有跡象表明，可以在一個數組中被發送兩個標誌：列 _ 權限和 CELL_FILTER_ 權限。如果查詢引擎不支援這些功能中的任何一項，且該功能的資料表上存在原則，則會擲回 a PermissionTypeMismatchException，且查詢失敗。這是為了避免數據洩漏。
5. 返回的響應包含以下內容：
   - 表的整個模式，以便查詢引擎可以使用它來解析存儲中的數據。
   - 使用者可存取的授權資料行清單。如果授權的資料行清單是空的，表示使用者具有`DESCRIBE`權限，但沒有`SELECT`權限，而且查詢失敗。
   - 旗標`IsRegisteredWithLakeFormation`，表示 Lake Formation 是否可以將認證分配給此資源資料。如果傳回 false，則應使用客戶的登入資料來存取 Amazon S3。
   - `CellFilters`如果有任何應該被應用到數據行的列表。此清單包含用於評估每一列的欄和表示式。只有當 CELL_FILTER_ PERFORCE 作為請求的一部分發送，並且對調用用戶的表格有數據過濾器時，才應填充此選項。

6. 擷取中繼資料後，查詢引擎會呼叫GetTemporaryGlueTableCredentials或取GetTemporaryGluePartitionCredentials得 AWS 登入資料，以從 Amazon S3 位置擷取資料。
7. 查詢引擎會從 Amazon S3 讀取相關物件，根據步驟 2 中收到的政策篩選資料，然後將結果傳回給使用者。

的應用程式整合 API 作業Lake Formation包含用於設定與協力廠商查詢引擎整合的其他內容。您可以在憑證自動售貨 API 操作部分中查看操作詳細信息。

註冊第三方查詢引擎

在第三方查詢引擎可以使用應用程式整合 API 作業之前，您必須明確啟用查詢引擎的權限，才能代表您呼叫 API 作業。這是通過幾個步驟完成：

1. 您需要指定需要權限的 AWS 帳戶和 IAM 工作階段標記，才能透過 AWS Lake Formation主控台 AWS CLI 或 API/SDK 呼叫應用程式整合 API 操作。
2. 當第三方查詢引擎在您的帳戶中擔任執行角色時，查詢引擎必須附加在 Lake Formation 註冊的工作階段標記，代表第三方引擎。 Lake Formation使用此標籤來驗證請求是否來自核准的引擎。如需工作階段標籤的詳細資訊，請參閱 IAM 使用者指南中的工作階段標籤。
3. 設定第三方查詢引擎執行角色時，您必須在 IAM 政策中具有以下最低權限集：

```
{
  "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. 在查詢引擎執行角色上設定角色信任原則，以對哪些工作階段標籤金鑰值配對可附加至此角色具有良好的存取控制。在下列範例中，此角色只允許附加工作階段標籤金鑰"LakeFormationAuthorizedCaller"和工作階段標籤值"engine1"，且不允許其他工作階段標籤金鑰值配對。

```
{
    "Sid": "AllowPassSessionTags",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/query-execution-role"
    },
    "Action": "sts:TagSession",
    "Condition": {
    "StringLike": {
        "aws:RequestTag/LakeFormationAuthorizedCaller": "engine1"        }
    }
}
```

LakeFormationAuthorizedCaller呼叫 STS: AssumeRole API 作業以擷取要使用的查詢引擎的認證時，工作階段標記必須包含在 AssumeRole 要求中。傳回的臨時認證可用於提出Lake Formation應用程式整合 API 要求。

Lake Formation應用程式整合 API 作業要求呼叫主體為 IAM 角色。IAM 角色必須包含具有已註冊之預定值的工作階段標記。Lake Formation此標籤Lake Formation允許驗證用於調用應用程序集成 API 操作的角色是否允許這樣做。

啟用第三方查詢引擎的權限，以呼叫應用程式整合 API 作業

請遵循下列步驟，以允許第三方查詢引擎透過 AWS Lake Formation主控台 AWS CLI 或 API/SDK 呼叫應用程式整合 API 作業。

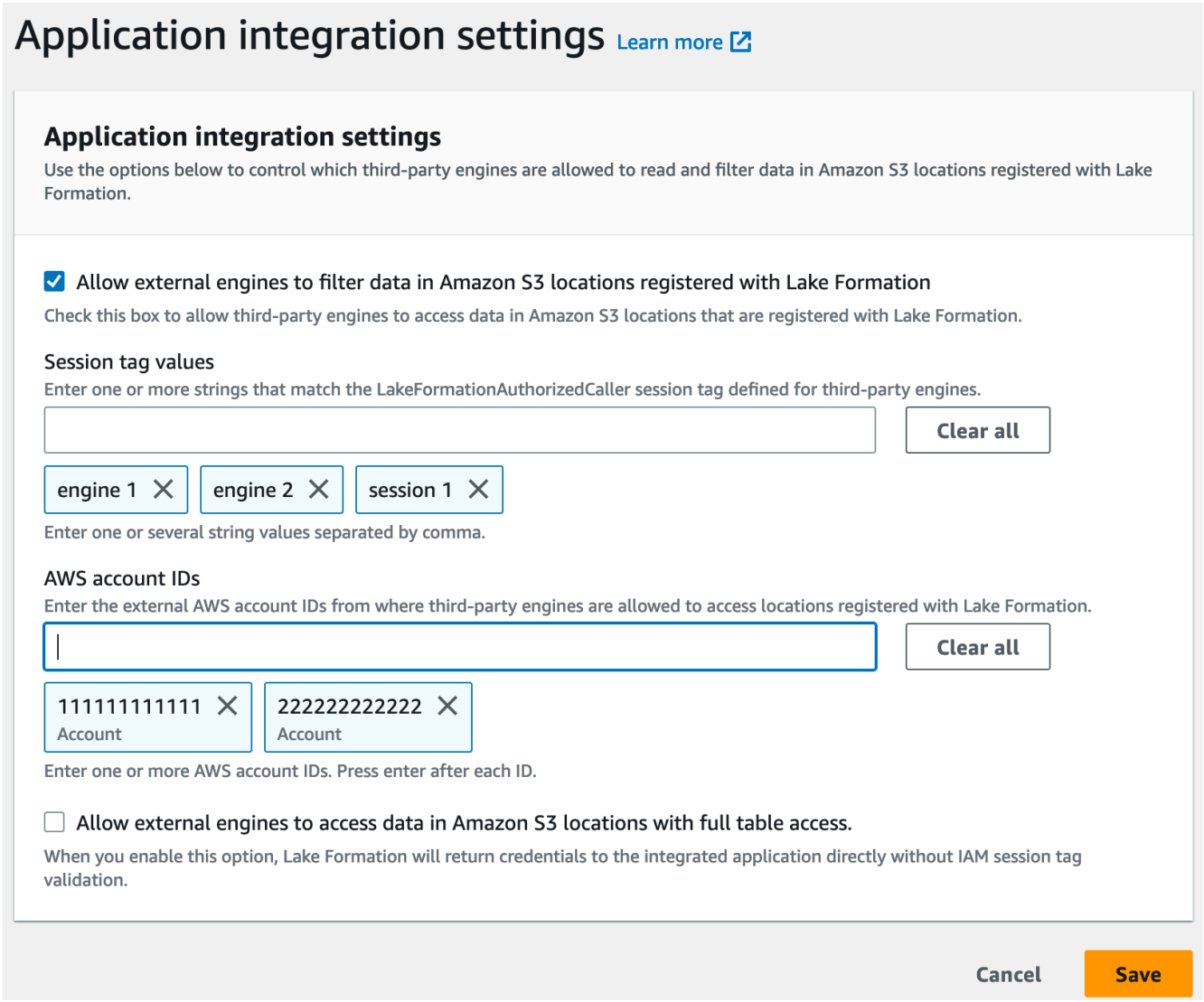
Console

若要註冊您的帳戶以進行外部資料篩選：

1. 登錄到 AWS Management Console，並打開 Lake Formation 控制台 https://console.aws.amazon.com/lakeformation/。
2. 在左側導覽中，展開 [系統管理]，然後選擇 [應用程式整合設定]。

3. 在「應用程式整合設定」頁面上，選擇「允許外部引擎篩選註冊過的 Amazon S3 位置中的資料」選項Lake Formation。
4. 輸入您為第三方引擎建立的工作階段標籤。如需有關工作階段標籤的資訊，請參閱AWS Identity and Access Management 使用者指南中的在 AWS STS 中傳遞工作階段標
5. 輸入可使用第三方引擎存取未篩選中繼資料資訊的使用者帳戶 ID，以及目前帳戶中資源的資料存取認證。

   您也可以使用 AWS 帳戶 ID 欄位來設定跨帳戶存取權限。



CLI

使用 `put-data-lake-settings` CLI 指令設定下列參數。

使用此 AWS CLI 命令時，有三個欄位需要設定：

- allow-external-data-filtering — (布林值) 表示協力廠商引擎可以存取目前帳戶中未篩選的中繼資料資訊和資料存取認證。
- external-data-filtering-allow-list— (陣列) 使用第三方引擎時，可存取目前帳戶中未篩選的中繼資料資訊和資料存取憑證的帳戶 ID 清單。
- authorized-sessions-tag-value-list— (陣列) 授權工作階段標籤值 (字串) 的清單。如果 IAM 角色登入資料已附加授權的索引鍵值配對，則如果工作階段標記包含在清單中，則會授與工作階段存取已設定帳戶中資源上未篩選的中繼資料資訊和資料存取登入資料的存取權。授權的工作階段標籤金鑰定義為*LakeFormationAuthorizedCaller*。
- AllowFullTableExternalDataAccess-(布林值) 當呼叫者具有完整資料存取權限時，是否允許第三方查詢引擎在沒有工作階段標籤的情況下取得資料存取認證。

例如：

```
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

使用 PutDataLakeSetting API 作業設定下列參數。

使用此 API 操作時，有三個字段需要配置：

- AllowExternalDataFiltering— (布林值) 指出協力廠商引擎是否可以存取目前帳戶中未篩選的中繼資料資訊和資料存取認證。
- ExternalDataFilteringAllowList— (陣列) 帳戶 ID 清單，可以使用第三方引擎存取未篩選的中繼資料資訊，以及目前帳戶中資源的資料存取憑證。
- AuthorizedSectionsTagValueList— (陣列) 授權標籤值 (字串) 的清單。如果 IAM 角色登入資料已附加授權標籤，則會授與工作階段存取未篩選的中繼資料資訊，以及已設定帳戶中資源的資料存取登入資料。授權的工作階段標籤金鑰定義為*LakeFormationAuthorizedCaller*。
- AllowFullTableExternalDataAccess-(布林值) 當呼叫者具有完整資料存取權限時，是否允許第三方查詢引擎在沒有工作階段標籤的情況下取得資料存取認證。

例如：

```
//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));
}
```

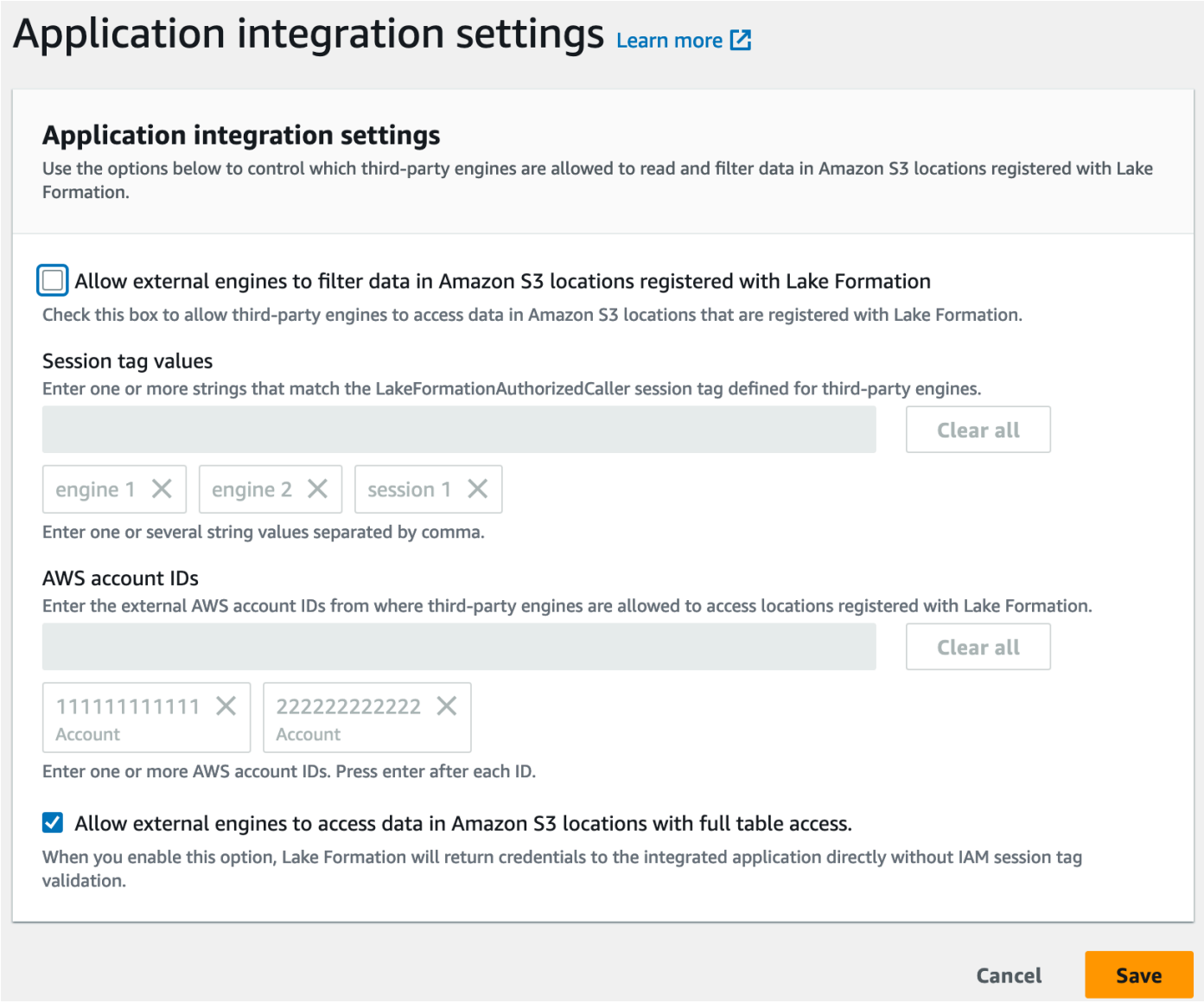
完整表格存取的應用程式整合

請遵循下列步驟，讓第三方查詢引擎在沒有 IAM 工作階段標籤驗證的情況下存取資料：

Console

1. 在 https://console.aws.amazon.com/lakeformation/ 上登錄到 Lake Formation 控制台。
2. 在左側導覽中，展開 [管理]，然後選擇 [應用程式整合設定]。
3. 在「應用程式整合設定」 頁面上，選擇允許外部引擎存取具有完整表格存取權的 Amazon S3 位置中的資料選項。

   啟用此選項時，Lake Formation 會直接將登入資料傳回至查詢應用程式，而不需要 IAM 工作階段標記驗證。



AWS CLI

使用 `put-data-lake-settings` CLI 指令設定`AllowFullTableExternalDataAccess`參數。

```
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
    }
}
```

與其他 AWS 服務合作

AWS Amazon Athena AWS Glue、Amazon Redshift Spectrum 和亞馬遜 EMR 等服務可用於安全地存取 AWS Lake Formation 向 Lake Formation 註冊的 Amazon S3 位置中的資料。使用 Lake Formation，您可以定義和管理. AWS Glue Data Catalog這些 AWS 服務中的每一項都是 Lake Formation 的受信任呼叫者，而 Lake Formation 可透過臨時登入資料存取 Amazon S3 中存放的資料。如需詳細資訊，請參閱 湖平整應用程序集成如何工作。

若要使用這些功能，Lake Formation 需要您先註冊 Amazon S3 位置，並將適當的許可指派給 IAM 主體，以存取表格、資料庫和 Amazon S3 位置。若要取得更多資訊，請參閱管理 Lake Formation 權限。

下表列出了亞馬遜雅典娜、亞馬遜 EMR 和亞馬 Amazon Redshift Spectrum 支援的 Lake Formation 許可類型 AWS Glue，以存取來自 AWS Glue 標準表和交易表 (Apache Iceberg、Apache Hudi 和 Linux 基礎三角洲湖) 的資料，其中包含存放在 Amazon S3 的資料以及資料目錄中的表中繼資料。

AWSAWS Glue 標準資料表和檢視表的服務和支援的權限類型

AWS 服務	資料表層級權限	資料行層級權限	資料列和儲存格層級權限
Athena	讀/寫存取	讀取存取權	讀取存取權
Athena Spark	不支援	不支援	不支援
已佈建叢集或亞馬遜無伺服器上的 Redshift 頻譜	讀/寫存取	讀取存取權	讀取存取權
阿帕奇星火在 Amazon EMR（EC2）	讀/寫存取	讀取存取權	讀取存取權
阿帕奇蜂巢在 Amazon EMR (EC2)	讀/寫存取	讀取存取權	不支援
阿帕奇星火在 EMR 無服務器	讀/寫存取	讀取存取權	讀取存取權

AWS 服務	資料表層級權限	資料行層級權限	資料列和儲存格層級權限
在 EMR 無伺服器上的阿帕奇蜂巢	不支援	不支援	不支援
Amazon EMR on EKS	不支援	不支援	不支援
AWS Glue ETL	讀/寫存取	不支援	不支援

考量與限制

- Athena Spark 不支援查詢具有 Lake Formation 權限的資料目錄資料表。
- 以 Athena SAML 為基礎的使用者可以啟用 SAML 2.0 型聯合，來讀取使用 Lake Formation 權限保護的資料來源。SAML 使用者可以將資料插入鑲木地板資料表。
- 在 EMR 無伺服器上的 Apache 星火不支援查詢資料目錄檢視。
- EMR 無伺服器上的 Apache Hive 不支援查詢具有 Lake Formation 權限的資料表。
- AWS Glue ETL 需要完整存取整個資料表，同時從基礎 Amazon S3 位置擷取資料。 AWS Glue 如果您在資料表上套用資料行層級權限，ETL 工作會失敗。

AWS 交易表格格式的服務和支援的權限類型

AWS 服務	Iceberg	Hudi	三角洲湖 (母語者)	三角洲湖（符號鏈接表）
Athena	支援使用表格、欄、列和儲存格層級權限讀取表格。寫入作業需要完整資料表存取權。	支援對具有資料表、欄、資料列和儲存格層級權限的資料表進行讀取和建立作業。不支援寫入作業。	Athena (引擎版本 3) 支援讀取具有資料表、欄、資料列和儲存格層級權限的原生 Delta Lake 資料表。不支援寫入作業。	Athena (引擎版本 3) 支援讀取具有資料表、欄、列和儲存格層級權限的符號連結 Delta Lake 資料表。不支援寫入作業。
已佈建叢集上的 Redshift 頻譜	支援使用表格、欄、列和儲存格層級權限讀取表	支援使用表格、欄、列和儲存格層級權限讀取表	不支援	支援透過具有資料表、欄、列和儲存格層級權限

AWS 服務	Iceberg	Hudi	三角洲湖 (母語者)	三角洲湖（符號鏈接表）
	格。不支援寫入作業。	格。不支援寫入作業。		的符號連結資訊清單讀取 Delta Lake 資料表。不支援寫入作業。
阿帕奇星火在 Amazon EMR（EC2）	支援使用表格、欄、列和儲存格層級權限讀取表格。寫入作業需要完整資料表存取權。	支援使用表格、欄、列和儲存格層級權限讀取表格。寫入作業需要完整資料表存取權。	支援使用表格、欄、列和儲存格層級權限讀取表格。不支援寫入作業。	支援使用表格、欄、列和儲存格層級權限讀取表格。寫入作業需要完整資料表存取權。
AWS Glue ETL	支援對具有資料表層級權限的表格進行讀取	支援對具有資料表層級權限的表格進行讀取	支援對具有資料表層級權限的表格進行讀取	支援對具有資料表層級權限的表格進行讀取

主題

- 使 AWS Lake Formation 用 Amazon Athena
- AWS Lake Formation 與 Amazon Redshift Spectrum 一起使用
- AWS Lake Formation 搭配使用 AWS Glue
- AWS Lake Formation 與 Amazon EMR 一起使用
- 使 AWS Lake Formation 用 Amazon QuickSight
- AWS Lake Formation 搭配 AWS CloudTrail 湖泊使用

使 AWS Lake Formation 用 Amazon Athena

Amazon Athena 是無伺服器的查詢服務，可協助您分析存放在 Amazon S3 中的結構化、半結構化和非結構化資料。您可以使用 Athena SQL 來查詢 CSV、JSON、實木複合地板和 Avro 資料格式的資料。Athena SQL 還支持表格格式，如阿帕奇蜂巢，阿帕奇胡迪和阿帕奇冰山。Athena 與資料集的中繼資料整合，可 AWS Glue Data Catalog 將資料集的中繼資料存放在 Amazon S3。Athena 可以使用 Lake Formation 定義並維護這些資料集的存取控制原則。

以下是一些常見的使用案例，您可以在 Athena 使用 Lake Formation。

- 使用 Lake Formation 權限從 Athena 存取資料目錄資源 (資料庫和表格)。您可以使用指定的資源方法或 LF 標籤來定義資料庫和資料表的權限。如需詳細資訊，請參閱：
  - 使用指定的資源方法授與資料庫權限
  - 基於 Lake Formation 標籤的訪問控制

> **Note**
>
> 僅當使用 Athena SQL 查詢來自 Amazon S3 的來源資料和資料目錄中的中繼資料時，Lake Formation 許可才適用。
> Athena Spark 不支援查詢具有 Lake Formation 權限的資料目錄資料表。Lake Formation m 權限支援資料庫和資料表的讀取和寫入作業。

> **Note**
>
> 當您使用 LF 標籤管理資料目錄資源的權限時，您無法套用資料篩選器。

- 透過在欄、列和儲存格層級授予權限，Lake Formation 的數據過濾器以保護 Amazon S3 資料湖中的表格，以控制查詢結果。請參閱 Amazon Athena 使用者指南中的分割區投影限制。
- 執行聯合查詢時，對 SAML 型 Athena 使用者可用的資料強制執行精細的存取控制。

  Athena JDBC 和 ODBC 驅動程式支援使用 SAML 型身分識別提供者 (IdP) 設定對資料來源的聯合存取。使用 Amazon 與 Lake Formation QuickSight 整合，搭配您現有的 IAM 角色或 SAML 使用者或群組，以視覺化方式呈現 Athena 查詢結果。

> **Note**
>
> 只有當您使用 JDBC 或 ODBC 驅動程式向 Athena 提交查詢時，SAML 使用者和群組的 Lake Formation 權限才適用。

如需詳細資訊，請參閱使用 Lake Formation 和 Athena JDBC 和 ODBC 驅動程式聯合存取 Athena。

> **Note**
>
> 目前，以下地區不支援在 Lake Formation 中授權存取 SAML 身分：
>
> - 中東 (巴林) – me-south-1

- 亞太區域 (香港) – ap-east-1
- 非洲 (開普敦) – af-south-1
- 中國 (寧夏) – cn-northwest-1
- 亞太區域 (大阪) - ap-northeast-3

- 用於Lake Formation 的跨賬戶數據共享查詢其他帳戶中的資料表。

> **Note**
>
> 如需使用 Lake Formation 權限時限制的詳細資訊`Views`，請參閱考量與限制。

Support 交易表格格式

套用 Lake Formation 權限可讓您保護 Amazon S3 資料湖中的交易資料。下表列出了 Athena 和 Lake Formation 權限支持的交易表格式。當 Athena 使用者執行查詢時，Lake Formation 會強制執行這些權限。

資料表格式	說明和允許的作業	Athena 支持的 Lake Formation 權限
Apache Hudi	用於簡化增量資料處理和資料管線開發的格式。 Athena 支援在 Amazon S3 資料集上使用 Apache Hudi 表格格式的建立和讀取操作，同時適用於寫入時複製 (CoW) 和讀取時合併 (MoR) 呼迪表類型。Athena 不支援 Hudi 資料表的寫入作業。 使用 Athena 查詢 Hudi 資料集。	用於使用資料表、欄、列和儲存格層級權限Lake Formation 中的數據過濾和細胞級安全來保護 Hudi 資料表的安全。
Apache Iceberg	一種開放式資料表格式，可將大量檔案集合作為資料表進行	支援資料表、欄、列和儲存格層級權限。目前，Lake

資料表格式	說明和允許的作業	Athena 支持的 Lake Formation 權限
	管理，並支援現代化的分析資料湖作業，例如記錄層級的插入、更新、刪除和時間旅行查詢。 如需有關雅典娜對冰山表格支援的詳細資訊，請參閱使用冰山表。	Formation 不支援管理寫入作業的權限，例如 VACUUMMERGE，以UPDATE及開OPTIMIZE啟資料表格式中的資料表。
Linux Foundation Delta Lake	三角洲湖是一個開放原始碼專案，有助於實作通常在 Amazon S3 或 Hadoop 分散式檔案系統 (HDFS) 上建置的現代資料湖架構。 Athena 支援使用符號連結式資訊清單資訊清單資料表定義建立的 Delta 湖泊資料表。 AWS Glue Data Catalog 如需詳細資訊，請參閱使用檢 AWS Glue 索器編目 Delta Lake 資料表。 Athena（引擎版本 3）支援讀取原生三角洲湖表。 如需詳細資訊，請參閱使用 AWS Glue 檢索器介紹原生 Delta Lake 表格支援。	符號連結表格和原生 Delta Lake 資料表支援資料表、資料行、資料列和儲存格層級權限。

其他資源

部落格文章、影片和工作坊

- 使用亞馬遜 Amazon Athena 查詢 Amazon S3 資料湖中的 Apache 胡迪資料集

- 使用 Amazon 雅典娜、亞馬遜 EMR 和建立 Apache 冰山資料湖 AWS Glue
- 使用 Athena 和阿帕奇冰山在 Amazon S3 上插入，更新，刪除
- 基於 LF 標籤的訪問控制 Lake Formation 研討會查詢數據湖。

AWS Lake Formation 與 Amazon Redshift Spectrum 一起使用

Amazon Redshift Spectrum 可讓您查詢和擷取 Amazon S3 資料湖中的資料，而無需將資料載入 Amazon Redshift 叢集節點。

Redshift 頻譜支持兩種註冊啟用 Lake Formation 的外部 AWS Glue 數據目錄的方法。

- 使用具有資料目錄權限的叢集附加 IAM 角色

  若要建立 IAM 角色，請遵循以下程序中概述的步驟。

  若要使用 AWS Glue Data Catalog 已啟用的功能建立 Amazon Redshift 的 IAM 角色 AWS Lake Formation
- 使用設定來管理外部 AWS Glue Data Catalog 資源存取權的聯合身分 IAM 身分

  Redshift 頻譜支援使用聯合 IAM 身分查詢 Lake Formation 資料表。IAM 身分可以是 IAM 使用者或 IAM 角色。如需 Redshift 頻譜中 IAM 身分聯合的詳細資訊，請參閱使用聯合身分管理 Amazon Redshift 對本機資源的存取和 Redshift 頻譜外部表格。

透過 Lake Formation 與 Redshift Spectrum 整合，您可以在資料向 LLake Formation Form 註冊後，在資料表上定義列、欄和儲存格層級的存取控制權限。

如需詳細資訊，請參閱搭 AWS Lake Formation配使用 Redshift 光譜。

Redshift 頻譜支援對 Lake Formation 管理的外部結構描述表進行讀取或SELECT查詢。

如需詳細資訊，請參閱建立 Redshift 頻譜的外部結構描述。

Support 交易資料表類型

此表格列出 Redshift 頻譜支援的交易表格格式，以及適用的 Lake Formation 權限。

支援的表格格式

資料表格式	說明和允許的作業	Redshift 頻譜中支持 Lake Formation 權限
Apache Hudi	用於簡化增量資料處理和資料管線開發的格式。 Redshift 頻譜支援在 Amazon S3 上使用阿帕奇胡迪複製 (CoW) 表格式的插入、刪除和更新寫操作。 如需詳細資訊，請參閱針對在 Apache Hudi 中管理的資料建立外部資料表。	用於使用資料表、欄、列和儲存格層級權限Lake Formation 中的數據過濾和細胞級安全來保護 Hudi 資料表的安全。
Apache Iceberg	一種開放式資料表格式，可將大量檔案集合作為資料表進行管理，並支援現代化的分析資料湖作業，例如記錄層級的插入、更新、刪除和時間旅行查詢。 如需詳細資訊，請參閱搭配 Amazon Redshift 使用 Apache 冰山表。	Redshift 頻譜支持阿帕奇冰山表進行查詢。
Linux Foundation Delta Lake	三角洲湖是一個開放原始碼專案，可協助實作通常在 Amazon S3 或 Hadoop 分散式檔案系統 (HDFS) 上建置的現代化資料湖架構。 Redshift 頻譜支援查詢三角洲湖資料表。如需詳細資訊，請參閱為在 Delta Lake 中管理的資料建立外部資料表。	支援資料表、欄、列和儲存格層級權限。

其他資源

部落格文章和工作坊

- 使用 Amazon Redshift Spectrum 啟用現代化資料架構的 AWS Lake Formation 同時，將資料湖集中控管
- 使用 Redshift 頻譜查詢 Amazon S3 數據湖中的寫入時複製（CoW）表

AWS Lake Formation 搭配使用 AWS Glue

資料工程師和 DevOps 專業人員 AWS Glue 與 Apache Spark 搭配擷取、轉換和載入 (ETL) 搭配使用，在 Amazon S3 中對其資料集執行轉換，並將轉換後的資料載入資料湖和資料倉儲，以進行分析、機器學習和應用程式開發。由於不同的團隊存取 Amazon S3 中的相同資料集，因此必須根據其角色授予和限制許可。

AWS Lake Formation 建立在其上 AWS Glue，並且服務以下列方式進行交互：

- Lake Formation 和AWS Glue共享相同的數據目錄。
- 以下 Lake Formation 控制台功能調用AWS Glue控制台：
  - 工作 — 如需詳細資訊，請參閱AWS Glue 開發人員指南中的新增工作。
  - 爬行者程式 — 如需詳細資訊，請參閱開發人員指AWS Glue 南中的使用爬行者程式編目表格。
- 使用 Lake Formation 藍圖時產生的AWS Glue工作流程是工作流程。您可以在 Lake Formation 主控台和主控台中檢視和管理這些工作流程。AWS Glue
- 機器學習轉型與 Lake Formation 一起提供，並建立在 AWS Glue API 操作之上。您可以在AWS Glue主控台上建立和管理機器學習轉換。如需詳細資訊，請參閱AWS Glue 開發人員指南中的 Machine Learning 轉換。

您可以使用 Lake Formation 精細的存取控制來管理現有的資料目錄資源和 Amazon S3 資料位置。

> **Note**
>
> AWS Glue ETL 需要完整存取整個資料表，同時從基礎 Amazon S3 位置擷取資料。 AWS Glue 如果您在資料表上套用資料行層級權限，ETL 工作會失敗。

Support 交易資料表類型

套用 Lake Formation 權限可讓您保護 Amazon S3 資料湖中的交易資料。下表列出了支持的交易表格格式 AWS Glue 和 Lake Formation 權限。Lake Formation 強制執行這些權限進行 AWS Glue 操作。

支援的表格格式

資料表格式	說明和允許的作業	支持 Lake Formation 權限 AWS Glue
Apache Hudi	用於簡化增量數據處理和數據管道開發的開放表格格式。 如需範例，請參閱中 AWS Glue的使用 Hudi 架構。	資料表層級權限適用於 Hudi 資料表。 如需詳細資訊，請參閱限制。
Apache Iceberg	一種開放式表格格式，可將大型檔案集合當作資料表來管理。 如需範例，請參閱中 AWS Glue的 < 使用冰山架構 > 。	資料表層級權限適用於冰山資料表。 如需詳細資訊，請參閱限制。
Linux Foundation Delta Lake	三角洲湖是一個開放原始碼專案，可協助實作通常在 Amazon S3 或 Hadoop 分散式檔案系統 (HDFS) 上建置的現代化資料湖架構。 如需範例，請參閱中的〈使用三角洲湖架構〉 AWS Glue。	Delta Lake 資料表可使用資料表層級權限。 如需詳細資訊，請參閱限制。

其他資源

部落格文章和儲存庫

- 使用 AWS Glue 連接器讀取和寫入具有 ACID 事務的 Apache 冰山表，並執行時間旅行
- 使用 AWS Glue 自訂連接器寫入 Apache 胡迪資料表

- AWS 雲形範本和火花程式碼範例的儲存庫 AWS Glue，可使用 Apache Hudi 和 Amazon S3 來分析串流資料。

AWS Lake Formation 與 Amazon EMR 一起使用

Amazon EMR 是一個靈活的 AWS 託管集群平台，你可以在支持的大數據框架，如 Hadoop 的地圖減少，星火，蜂巢，普雷斯托等運行任何自定義代碼。Organizations 也使用 Amazon EMR 跨高度分散式叢集執行批次和串流資料處理應用程式。在 Amazon EMR 上使用 Apache Spark，您可以在許可由 Lake Formation 管理的數據庫和表上運行數據轉換和自定義代碼。

部署 Amazon EMR 有三個選項：

- EC2 上的 EMR
- EMR Serverless
- Amazon EMR on EKS

如需詳細資訊，請參閱將 Amazon EMR 與 Lake Formation 整合或使用 EMR 無伺服器搭配 AWS Lake Formation 進行精細的存取控制

Support 交易表格格式

當您使用 Spark SQL 讀取和寫入資料時，Amazon EMR 版本 6.15.0 及更高版本包括對 Apache Hudi、Apache 冰山和三角洲湖資料表格式的湖泊 Lake Formation 表、列、欄和儲存格層級存取控制權限的支援。

有關限制，請參閱 Amazon EMR 與 Lake Formation 的注意事項。

支援的表格格式

資料表格式	說明和允許的作業	Amazon EMR 支持 Lake Formation 許可
Apache Hudi	用於簡化增量數據處理和數據管道開發的開放表格格式。 有關支持的操作列表，請參閱 Apache Hudi 和 Lake Formation。	Amazon EMR 使用 Apache Hudi 來支援資料表、資料列、資料欄和儲存格層級存取控制。

資料表格式	說明和允許的作業	Amazon EMR 支持 Lake Formation 許可
Apache Iceberg	一種開放式表格格式，可將大型檔案集合當作資料表來管理。 有關支持的操作列表，請參閱 Apache 冰山和 Lake Formation。	Amazon EMR 使用 Apache Iceberg 來支援資料表、資料列、資料欄和儲存格層級存取控制。
Linux Foundation Delta Lake	三角洲湖是一個開放原始碼專案，可協助實作通常在 Amazon S3 或 Hadoop 分散式檔案系統 (HDFS) 上建置的現代化資料湖架構。 有關支持的操作列表，請參閱三角洲湖泊和 Lake Formation。	Amazon EMR 透過三角洲湖表支援資料表、資料列、欄和儲存格層級存取控制。

其他資源

使用者指南、部落格文章和研討會

- 使用執行時間角色與 Amazon EMR 整合
- 快速入門與阿帕奇胡迪，阿帕奇冰山和三角洲湖與 Amazon EMR 在 EKS
- 搭配無伺服器使用台達湖 OSS

使 AWS Lake Formation 用 Amazon QuickSight

Amazon QuickSight 支持使用 Athena 探索由 Amazon S3 中 Lake Formation 許可管理的數據集。

Amazon 的標準版和企業版用戶都與 Lake Formation QuickSight 集成，但略有不同。

- 企業版 — 將精細的存取控制 (FGAC) 許可授予個別 Amazon QuickSight 使用者、群組和 IAM 角色，以存取資料庫和表格。

- 標準版 — 將權限授予 IAM 角色以存取資料庫和表格。

> **Note**
>
> 默認情況下，Amazon QuickSight 使用名為的角色`aws-quicksight-service-role-v0`。您也可以定義具有必要許可的自訂角色，讓 Amazon 能 QuickSight 夠存取 Athena。

如需詳細資訊，請參閱授權連線 AWS Lake Formation

其他資源

部落格文章

- 在中為 Amazon QuickSight 作者啟用精細的許可 AWS Lake Formation
- 使用 AWS Lake Formation 和 Amazon 安全地分析您的數據 QuickSight

AWS Lake Formation 搭配 AWS CloudTrail 湖泊使用

AWS CloudTrail Lake 支援使用 Amazon Athena 中 AWS Lake Formation的細微權限探索事件資料存放區。

> **Note**
>
> CloudTrail 湖只能通過 Amazon Athena查詢。

若要向 CloudTrail Lake Formation 註冊您的 Lake 活動資料存放區，請參閱聯合事件資料存放區。

使用記錄 AWS Lake Formation API 調用 AWS CloudTrail

AWS Lake Formation 與該服務集成在一起 AWS CloudTrail，該服務可提供用戶，角色或 AWS 服務在 Lake Formation 中採取的行動記錄。 CloudTrail 捕獲所有 Lake Formation API 調用作為事件。捕獲的呼叫包括來自 Lake Formation 控制台的呼叫 AWS Command Line Interface，以及對 Lake Formation API 操作的代碼調用。如果您建立追蹤，您可以啟用持續傳遞 CloudTrail事件到 Amazon S3 儲存貯體，包括 Lake Formation 的事件。如果您未設定追蹤，您仍然可以在 [事件歷程記錄] 中檢視 CloudTrail 主控台中最近的事件。使用收集的信息 CloudTrail，您可以確定向 Lake Formation 提出的請求，提出請求的 IP 地址，提出請求的人員，提出請求的時間以及其他詳細信息。

若要進一步了解 CloudTrail，請參閱使AWS CloudTrail 用者指南。

湖的形成信息 CloudTrail

CloudTrail 當您建立新 AWS 帳號時，預設為啟用。當活動在 Lake Formation 中發生時，該活動會與 CloudTrail 事件歷史記錄中的其他 AWS 服務事件一起記錄為事件。事件即為來自任何來源的單一請求，其中包含請求動作、動作日期和時間，以及請求參數的相關資訊。此外，每個事件或記錄項目都包含產生請求者的相關資訊。身分資訊可協助您判斷下列事項：

- 要求是使用根使用者登入資料還是 AWS Identity and Access Management (IAM) 使用者登入資料提出。
- 提出該請求時，是否使用了特定角色或聯合身分使用者的暫時安全憑證。
- 請求是否由其他 AWS 服務提出。

如需詳細資訊，請參閱CloudTrail 使 userIdentity 元素。

您可以檢視、搜尋和下載 AWS 帳戶的最近活動。如需詳細資訊，請參閱使用 CloudTrail 事件歷程記錄檢視事件。

為了獲得您 AWS 帳戶中的持續事件記錄，包括 Lake Formation 的活動，請創建一條線索。追蹤可 CloudTrail 將日誌檔交付到 Amazon S3 儲存貯體。根據預設，在主控台建立線索時，該線索會套用到所有 AWS 區域。追蹤記錄來自 AWS 分區中所有區域的事件，並將日誌檔傳送到您指定的 Amazon S3 儲存貯體。此外，您可以設定其他 AWS 服務，例如 Amazon Athena，進一步分析 CloudTrail 記錄檔中收集的事件資料並採取行動。 CloudTrail 也可以將日誌檔傳送到 Amazon CloudWatch 日誌和 CloudWatch 事件。

如需詳細資訊，請參閱下列內容：

- 建立追蹤的概觀
- CloudTrail 支援的服務與整合
- 設定 Amazon SNS 通知 CloudTrail
- 從多個區域接收 CloudTrail 日誌文件並從多個帳戶接收 CloudTrail 日誌文件

了解 Lake Formation 事件

所有 Lake Formation API 操作都由記錄 CloudTrail 並記錄在 AWS Lake Formation 開發人員指南中。例如，呼叫PutDataLakeSettingsGrantPermissions、和RevokePermissions動作會在 CloudTrail 記錄檔中產生項目。

下列範例顯示 CloudTrail 動GrantPermissions作的事件。項目包括授與權限的使用者 (datalake_admin)、授與權限的主體 (datalake_user1)，以及授與的權限 (CREATE_TABLE)。此項目也會顯示授權失敗，因為resource引數中未指定目標資料庫。

```
{
    "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"
}
```

下一個範例顯示GetDataAccess動作的 CloudTrail 記錄項目。主參與者不會直接呼叫此 API。相反地，GetDataAccess每當主體或整合 AWS 服務要求臨時登入資料以存取在 Lake Formation 註冊的資料湖位置中的資料時，就會記錄這些資料。

```
{
    "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"
    },
...
}
```

另請參閱

- 跨帳戶 CloudTrail 記錄

Lake Formation 的最佳做法，考慮因素和限制

使用本節可快速找到中的最佳做法、考量事項和限制 AWS Lake Formation。

如需您的服務資源或作業數目上限，請參閱服務配額 AWS 帳戶。

主題

跨帳戶資料共用最佳做法與考量

Lake Formation 跨帳戶功能可讓使用者在多個 AWS 組織間安全地共用分散式資料湖 AWS 帳戶，或直接與其他帳戶中的 IAM 主體共用分散式資料湖，從而提供對資料型錄中繼資料和基礎資料的精細存取。

使用 Lake Formation 跨帳戶資料共用時，請考慮下列最佳做法：

- 您可以對自己 AWS 帳戶中的校長進行的 Lake Formation 許可授予的數量沒有限制。但是，Lake Formation 使用 AWS Resource Access Manager （AWS RAM）容量進行您的帳戶可以使用指定的資源方法進行的跨帳戶贈款。若要最大化 AWS RAM 容量，請遵循指定資源方法的下列最佳做法：
  - 使用新的跨帳戶授權模式（在「跨帳號版本設定」下的第 3 版及以上版本）與外部 AWS 帳戶人員共用資源。如需詳細資訊，請參閱 更新跨帳戶資料共用版本設定。
  - 將 AWS 帳戶排列到組織中，並授與權限給組織或組織單位。對組織或組織單位的授權算作一個授權。

授與組織或組織單位也不需要接受授權的 AWS Resource Access Manager (AWS RAM) 資源共用邀請。如需詳細資訊，請參閱 存取和檢視共用資料目錄表格和資料庫。

- 請使用特殊的「所有資料表」萬用字元來授與資料庫中所有資料表的權限，而不是授與資料庫中許多個別資料表的權限。對所有資料表授與計為單一授權。如需詳細資訊，請參閱 授與和撤銷資料目錄資源的權限。

> **Note**
>
> 如需有關要求提高中資源共用數目限制的詳細資AWS 訊 AWS RAM，請參閱 AWS 一般參考.

- 您必須建立共用資料庫的資源連結，該資料庫才會顯示在 Amazon Athena 和 Amazon Redshift Spectrum 查詢編輯器中。同樣地，若要能夠使用 Athena 和 Redshift 頻譜查詢共用資料表，您必須建立資料表的資源連結。然後資源連結會顯示在查詢編輯器的表格清單中。

  您可以使用「所有資料表」萬用字元來授與資料庫中所有資料表的權限，而不是為許多個別資料表建立資源連結以進行查詢。然後，當您為該資料庫建立資源連結，並在查詢編輯器中選取該資料庫資源連結時，您就可以存取該資料庫中的所有表格以供查詢。如需詳細資訊，請參閱 建立資源連結。
- 當您直接與其他帳戶中的主體共用資源時，收件者帳戶中的 IAM 主體可能沒有建立資源連結的權限，無法使用 Athena 和 Amazon Redshift Spectrum 查詢共用資料表。資料湖管理員可以建立預留位置資料庫並將`CREATE_TABLE`權限授與`ALLIAMPrincipal`群組，而不是為每個共用的資料表建立資源連結。接著，收件者帳戶中的所有 IAM 主體都可以在預留位置資料庫中建立資源連結，並開始查詢共用資料表。

  請參閱中授與權限的 CLI 命令範例使用指定的資源方法授與資料庫權限。`ALLIAMPrincipals`
- Athena 和 Redshift 頻譜支援資料行層級的存取控制，但僅用於包含，而不是排除。AWS GlueETL 工作不支援資料行層級存取控制。
- 當資源與您的 AWS 帳號共用時，您只能將資源的權限授與您帳號中的使用者。您無法將資源的權限授與其他 AWS 帳號、組織 (甚至不是您自己的組織) 或`IAMAllowedPrincipals`群組。
- 您無法將資料庫授與`DROP`或`Super`授與外部帳戶。
- 刪除資料庫或資料表之前，請先撤銷跨帳戶權限。否則，您必須刪除中的孤立資源共用。 AWS Resource Access Manager

另請參閱

- 基於 Lake Formation 標籤的訪問控制最佳實踐和考量
- CREATE_TABLE在中，瞭Lake Formation 權限參考解更多跨帳戶存取規則與限制。

跨區域資料存取限制

Lake Formation 支持跨查詢數據目錄表 AWS 區域。您可以使用 Amazon Athena Amazon EMR 和 AWS Glue ETL 在指向來源資料庫和表格的其他區域建立資源連結，從其他區域存取區域中的資料。透過跨區域表格存取權，您可以跨區域存取資料，而無需將基礎資料或中繼資料複製到「資料目錄」中。

下列限制適用於跨區域表格存取。

- Lake Formation 不支援使用 Amazon Redshift Spectrum 從其他區域查詢資料目錄資料表。
- 在 Lake Formation 控制台中，數據庫和表視圖不顯示源區域數據庫/表名稱。
- 若要從另一個區域檢視共用資料庫下的表格清單，您必須先建立共用資料庫的資源連結，然後選取資源連結，然後選擇 [檢視表格]。
- 當您在 AWS 區域 該點中建立資源連結至在選擇加入區域中建立的共用資料庫和表格時，跨區域表格存取功能無法運作。

  如需詳細資訊，請參閱支援 AWS 區域 與服務頁面上的選擇加入區域。
- Lake Formation 不支援 SAML 使用者進行的跨區域資源連結呼叫。

資料目錄檢視考量和限制

在中 AWS Glue Data Catalog，檢視是虛擬資料表，其中的內容是由參照一或多個資料表的查詢來定義。您可以使用適用於亞馬遜雅典娜或亞馬 Amazon Redshift 的 SQL 編輯器建立最多 10 個表格的檢視。視圖的基礎參考表可以屬於相同的數據庫或相同內的不同數據庫 AWS 帳戶。

下列考量和限制適用於「資料目錄」檢視。

- Amazon Redshift 始終從具有字符串的表中創建具有 varchar 列的視圖。從其他引擎添加方言時，必須將字符串列轉換為具有明確長度的 varchar。
- 將資料湖權限授與資料庫`All views`內，將導致受權者擁有資料庫中所有資料表和檢視的權限。

- 您無法建立檢視表：
  - 引用其他視圖。
  - 當參考資料表是資源連結時。
  - 當參照表具有IAM_ALLOWED_GROUP主參與者權限時。
  - 當參考資料表位於另一個帳戶時。
  - 從外部蜂巢元存儲。

資料篩選限制

當您授與「資料目錄」表格的 Lake Formation 權限時，您可以納入資料篩選規格，以限制查詢結果中某些資料的存取，以及與 Lake Formation 整合的引擎。Lake Formation 使用資料篩選來實現欄層級安全性、資料列層級安全性和儲存格層級安全性。如果來源資料包含巢狀結構，您可以在巢狀資料行上定義資料並套用資料篩選器。

資料行層級篩選的注意事項和限制

有三種方法可以指定欄篩選：

- 通過使用數據過濾器
- 通過使用簡單的列過濾或嵌套列過濾。
- 通過使用標籤。

簡單的列過濾只是指定要包含或排除的列列表。Lake Formation 控制台，API 和 AWS CLI 支持簡單的列過濾。如需範例，請參閱Grant with Simple Column Filtering。

下列備註與限制適用於欄篩選：

- AWS Glue ETL 工作不支援欄篩選。如果將欄篩選套用至工作參照的任何資料表，工作就會失敗。
- 要使SELECT用授予選項和列篩選進行授予，您必須使用包含列表，而不是排除列表。如果沒有授予選項，您可以使用包含或排除清單。
- 若要授與資料SELECT行篩選的資料表，您必須已授與資料表SELECT上的 grant 選項，而且沒有任何資料列限制。您必須擁有所有列的存取權。
- 如果您SELECT使用 grant 選項和資料行篩選授與帳戶中的主參與者，則該主體在授與另一個主體時，必須針對相同的資料行或授與的資料行子集指定資料行篩選。如果您SELECT使用授與選項和

欄篩選授與外部帳戶，則外部帳戶中的資料湖管理員可以將所有欄授SELECT與其帳戶中的另一個主體。不過，即使SELECT在所有資料行上，該主參與者也只能看到授與外部帳戶的資料行。

- 您無法在分割區索引鍵上套用欄篩選。
- 具有資料表中資料行子集SELECT權限的主參與者無法授與該資料表的ALTERDROPDELETE、、或INSERT權限。對於具有ALTER、DROPDELETE、或資料表INSERT權限的主參與者而言，如果您授與資料行篩選的SELECT權限，則不會有任何作用。

下列備註和限制適用於巢狀資料欄篩選：

- 您可以在資料篩選中包括或排除五個層級的巢狀欄位。

  Example

  彩色 1 _1 _ 1 _ 密集 1 _1_1_1
- 您無法對分區欄內的巢狀欄位套用欄篩選。
- 如果您的資料表結構定義包含頂層欄名稱 (「customer」。 address」) 在資料篩選器中具有相同的巢狀欄位表示模式 (具有頂層欄名稱customer和巢狀欄位名稱的巢狀資料行與資料篩選器"customer"."address"中的指定方式相同)，您無法明確指定對頂層欄或巢狀欄位的存取權，因為兩者都address是使用包含/排除清單中的相同模式來表示。這是不明確的，如果您指定了頂層列或嵌套字段，則 Lake Formation 無法解決。
- 如果頂層欄或巢狀欄位在名稱中包含雙引號，則在資料儲存格篩選的包括和排除清單中指定巢狀欄位的存取權時，必須包含第二個雙引號。

  Example

  帶雙引號的嵌套列名稱示例-a.b.double"quote

  Example

  數據過濾器中的示例嵌套列表示- "a"."b"."double""quote"

儲存格層級篩選限制

請記住下列資料列層級和儲存格層級篩選的注意事項和限制。

- 巢狀資料行、檢視表和資源連結不支援儲存格層級安全性。

- 巢狀欄也支援頂層欄所支援的所有運算式。但是，在定義巢狀資料列層級運算式時，不應參考資料分割資料行下的巢狀欄位。
- 使用 Athena 引擎第 3 版或 Amazon Redshift Spectrum 時，所有區域均可使用儲存格層級安全性。對於其他服務，儲存格層級安全性僅適用於上述的支援地區區域。
- 不支援 SELECT INTO 陳述式。
- map資料列篩選運算式不支援和資料類型。array支持struct數據類型。
- 資料表上可定義的資料篩選器數目沒有限制，但資料表上單一主體的資料篩選SELECT權限限制為 100。
- 資料表中可包含在授權中的資料篩選器數目上限為 10 個。
- 若要使用資料列篩選運算式套用資料篩選器，您必須在所有表格欄上SELECT使用授與選項。對外部帳戶進行授權時，此限制不適用於外部帳戶中的管理員。
- 如果主參與者是群組的成員，且主參與者和群組都被授與資料列子集的權限，則主參與者的有效資料列權限就是主參與者的權限與群組權限的聯集。
- 下列資料行名稱在資料列層級和儲存格層級篩選的資料表中受到限制：
  - 感染者
  - oid
  - xmin
  - 分鐘
  - xmax
  - C 最大
  - 表情
  - 插入
  - 刪除
  - 入侵
  - 雷卡特獨特
- 如果您將所有資料列篩選運算式與其他具有述詞的篩選運算式同時套用至資料表，則所有資料列運算式會優於其他所有篩選運算式。
- 當資料列子集的權限授與外部帳戶，而外部 AWS 帳戶的資料湖管理員會將這些權限授與該帳戶中的主體時，主體的有效篩選述詞就是帳戶述詞與直接授與主體之任何述詞的交集。

  例如，如果帳戶具有述詞的資料列權限，dept='hr'且主參與者已個別授與權限country='us'，則主參與者只能存取具有dept='hr'和country='us'的資料列。

如需儲存格層級篩選的詳細資訊，請參閱Lake Formation 中的數據過濾和細胞級安全。

混合式存取模式考量與限制

混合式存取模式提供彈性 Lake Formation 可選擇性地啟用 AWS Glue Data Catalog.
透過混合式存取模式，您現在擁有一個增量路徑，可讓您為一組特定使用者設定 Lake Formation 權限，而不會中斷其他現有使用者或工作負載的權限原則。

下列考量與限制適用於混合式存取模式。

限制

- 更新 Amazon S3 位置註冊 — 您無法使用服務連結角色編輯在 Lake Formation 註冊的位置參數。
- 使用 LF-tag 時選擇加入選項 — 當您可以使用 LF-tag 授予 Lake Formation 權限時，您可以選擇連續步驟加入主體以連續步驟強制執行 Lake Formation 權限，方法是選擇附加了 Lf 標籤的資料庫和表格。
- 選擇加入主參與者 — 目前只有資料湖管理員角色可以選擇加入資源的主參與者。
- 選擇加入資料庫中的所有表格 — 在跨帳戶授權中，當您授予權限並選擇加入資料庫中的所有表格時，您還需要選擇資料庫才能使用權限。

考量事項

- 將在 Lake Formation 註冊的 Amazon S3 位置更新為混合存取模式 — 我們不建議將已在 Lake Formation 註冊的 Amazon S3 資料位置轉換為混合式存取模式，但可以這樣做。
- 以混合存取模式註冊資料位置時的 API 行為
  - CreateTable — 無論混合訪問模式標誌和選擇狀態如何，該位置都被視為已註冊在 Lake Formation m 中。因此，使用者需要資料位置權限才能建立資料表。
  - CreatePartition/BatchCreatePartitions/UpdatePartitions (當分區位置更新為指向使用混合式註冊的位置時) — 無論混合存取模式旗標和選擇使用狀態為何，Amazon S3 位置都會被視為向 Lake Formation 註冊。因此，使用者需要資料位置權限才能建立或更新資料庫。
  - CreateDatabase/UpdateDatabase (當資料庫位置更新為指向以混合存取模式註冊的位置時) — 無論混合式存取模式旗標和選擇加入狀態為何，該位置都會被視為已註冊於 Lake Formation。因此，使用者需要資料位置權限才能建立或更新資料庫。
  - UpdateTable (更新表格位置以指向以混合存取模式註冊的位置時) — 無論混合式存取模式旗標和選擇加入狀態為何，都會將該位置視為已註冊於 Lake Formation。因此，使用者需要資料位置權

限才能更新資料表。如果表格位置未更新或指向未在 Lake Formation 註冊的位置，則使用者不需要資料位置權限即可更新資料表。

Hive 中繼資料儲存資料共用考量和限制

使用 AWS Glue Data Catalog 中繼資料聯合 (資料目錄聯合)，您可以將資料目錄連接到存放 Amazon S3 資料中繼資料的外部中繼資料的外部中繼資料，並使用 AWS Lake Formation安全地管理資料存取許可。

下列考量和限制適用於從 Hive 資料庫建立的聯合資料庫：

考量事項

- AWS SAM 應用程式支援 — 您必須負責 AWS SAM 部署的應用程式資源 (以Amazon API Gateway 及 Lambda 函數) 的可用性。當使用者執行查詢時，請確定 AWS Glue Data Catalog 和 Hive 中繼存放區之間的連線正常運作。
- 配置單元存儲版本要求 — 您只能使用 Apache Hive 版本 3 及更高版本創建聯合數據庫。
- 映射的數據庫要求-每個 Hive 數據庫必須映射到 Lake Formation 的新數據庫。
- 資料庫層級聯合支援 — 您只能在資料庫層級連線到 Hive 中繼存放區。
- 聯合資料庫的權限 — 即使刪除來源資料表或資料庫，套用至聯合資料庫下的一或多個資料表上的權限仍然存在。重新建立來源資料庫或資料表時，您不需要重新授與權限。在來源刪除具有 Lake Formation 權限的聯合資料表時，Lake Formation 權限仍然可見，您可以視需要撤銷它們。

  如果使用者刪除聯合資料庫，其所有對應的權限都會遺失。重新建立具有相同名稱的相同資料庫，將不會復原 Lake Formation 權限。使用者必須再次設定新的權限。
- 聯合資料庫上的 IAM AllowedPrincipal 群組許可 — 根據`DataLakeSettings`，Lake Formation 可能會將所有資料庫和資料表的權限設定為名為`IAMAllowedPrincipal`的虛擬群組。`IAMAllowedPrincipal`指的是透過 IAM 主體政策和 AWS Glue 資源政策存取資料目錄資源的所有 IAM 主體。如果這些權限存在於資料庫或資料表上，則會授與所有主體存取資料庫或資料表的存取權。

  但是，Lake Formation 不允許對聯合數據庫下的表的`IAMAllowedPrincipal`權限。當您建立聯合資料庫時，請務必將`CreateTableDefaultPermissions`參數傳遞為空白清單。

  如需詳細資訊，請參閱 變更資料湖的預設設定。
- 在查詢中聯結表格 — 您可以將 Hive 中繼存放區表與資料目錄原生資料表聯結以執行查詢。

限制

- AWS Glue Data Catalog 與 Hive 中繼資料庫之間同步中繼資料的限制 — 建立 Hive 中繼存放區連線之後，您需要建立聯合資料庫，以將 Hive 中繼資料庫中繼資料與. AWS Glue Data Catalog當使用者執行查詢時，同盟資料庫下的資料表會在執行階段同步。
- 在聯合資料庫下建立新資料表的限制 — 您將無法在聯合資料庫下建立新資料表。
- 資料權限限制 — 無法 Support Hive 中繼存放區表格檢視的權限。

Amazon Redshift 數據共享限制

AWS Lake Formation 可讓您安全地管理來自 Amazon Redshift 的資料記憶體中的資料。Amazon Redshift 是雲端中的全受管 PB 級資料倉儲服務。 AWS 使用資料共用功能，Amazon Redshift 可協助您在各 AWS 帳戶處共用資料。如需有關 Amazon Redshift 資料共用的詳細資訊，請參閱 Amazon Redshift 中的資料共用概觀。

下列注意事項和限制適用於從 Amazon Redshift 資料庫建立的聯合資料庫：

- 對應的資料庫需求 — 每個 Amazon Redshift 資料指標都必須對應至 Lake Formation 中的新資料庫。在「資料目錄」資料庫中展開資料清單物件表現法時，需要這樣做才能維護唯一的表格名稱。
- 在聯合資料庫下建立新資料表的限制 — 您將無法在聯合資料庫下建立新資料表。
- 聯合資料庫的權限 — 即使刪除來源資料表或資料庫，套用至聯合資料庫下的一或多個資料表上的權限仍然存在。重新建立來源資料庫或資料表時，您不需要重新授與權限。當源代碼刪除具有 Lake Formation 權限的聯合表時，Lake Formation 權限仍然可見，您可以根據需要撤銷它們。

  如果使用者刪除聯合資料庫，其所有對應的權限都會遺失。重新建立具有相同名稱的相同資料庫，將不會復原 Lake Formation 權限。使用者必須再次設定新的權限。
- 聯合資料庫上的 IAM AllowedPrincipal 群組許可 — 根據`DataLakeSettings`，Lake Formation 可能會將所有資料庫和資料表的權限設定為名為`IAMAllowedPrincipal`的虛擬群組。`IAMAllowedPrincipal`指的是透過 IAM 主體政策和 AWS Glue 資源政策存取資料目錄資源的所有 IAM 主體。如果這些權限存在於資料庫或資料表上，則會授與所有主體存取資料庫或資料表的存取權。

  但是，Lake Formation 不允許對聯合數據庫下的表的`IAMAllowedPrincipal`權限。當您建立聯合資料庫時，請務必將`CreateTableDefaultPermissions`參數傳遞為空白清單。

  如需詳細資訊，請參閱 變更資料湖的預設設定。

- 資料篩選 — 在 Lake Formation 中，您可以透過資料行層級和資料列層級篩選，授與聯合資料庫下表格的權限。不過，您無法結合資料行層級和資料列層級篩選，以在聯合資料庫下的資料表上，以儲存格層級精細度限制存取。
- 區分大小寫識別碼 — 由 Lake Formation 管理的 Amazon Redshift 資料清理物件僅支援小寫的資料表名稱和資料行名稱。如果 Amazon Redshift 資料庫中的資料庫、表格和欄將使用 Lake Formation 共用和管理，請勿開啟資料庫、資料表和欄的區分大小寫識別碼。

如需在 Amazon Redshift 中使用資料倉時限制的詳細資訊，請參閱 Amazon Redshift 資料庫開發人員指南中的資料共用限制。

IAM 身分識別中心整合限制

您可以透過連線至身分識別提供者 (IdPs) AWS IAM Identity Center，並跨 AWS 分析服務集中管理使用者和群組的存取。您可以在 IAM Identity Center 中設定 AWS Lake Formation 為已啟用的應用程式，資料湖管理員可以將精細的許可授與授權的使用者和 AWS Glue Data Catalog 資源群組。

以下限制適用於 Lake Formation 與 IAM 身分中心的整合：

- 您無法在 Lake Formation 中將 IAM 身分中心使用者和群組指派為資料湖管理員或唯讀管理員。
- 如果您使用的是可以代表您承擔的 IAM 角色來加密和解密資料目錄，則 IAM Identity Center 使用者和群組 AWS Glue 可以查詢加密的資料目錄資源。 AWS 受管理金鑰不支援受信任的身分傳播。
- IAM 身分中心使用者和群組只能叫用 IAM 身分中心提供的`AWSIAMIdentityCenterAllowListForIdentityContext`政策中列出的 API 操作。
- Lake Formation 允許來自外部帳戶的 IAM 角色，代表 IAM Identity Center 使用者和群組用於存取資料目錄資源的承運人角色，但只能授與擁有帳戶內的資料目錄資源的許可。如果您嘗試將權限授與外部帳戶中資料目錄資源的 IAM Identity Center 使用者和群組，Lake Formation 會擲回下列錯誤-「主體不支援跨帳戶授與」。

基於 Lake Formation 標籤的訪問控制最佳實踐和考量

您可以建立、維護和指派 LF 標籤，以控制對資料目錄資料庫、資料表和欄的存取。

使用以 Lake Formation 標籤為基礎的存取控制時，請考慮下列最佳作法：

- 必須先預先定義所有 LF 標籤，才能將其指定給資料目錄資源或授與主參與者。

資料湖管理員可以透過建立具有所需 IAM 許可的 LF 標籤建立者來委派標籤管理任務。數據工程師和分析師決定 LF 標籤的特徵和關係。LF 標籤創建者然後創建並維護 LF-標籤在 Lake Formation。

- 您可以為資料目錄資源指定多個 LF 標籤。特定索引鍵只能指定一個值給特定資源。

  例如，您可以將module=Orders、region=Westdivision=Consumer、等指派給資料庫、資料表或欄。您無法指派module=Orders,Customers。

- 建立資源時，您無法將 LF 標籤指派給資源。您只能將 LF 標籤新增至現有資源。
- 您可以授予 LF 標籤表達式，而不僅僅是單個 LF 標籤，給主體。

  LF 標籤表達式看起來像下面的（在偽代碼中）。

  ```
  module=sales AND division=(consumer OR commercial)
  ```

  授與此 LF 標籤運算式的主參與者division=consumer只能存取已指派和或的資料目錄資源 (資料庫、資料表module=sales和資料行)。division=commercial如果您希望主體能夠存取具有module=sales或的資源division=commercial，請勿將兩者都包含在相同授權中。提出兩個贈款，一個用於module=sales，一個用於division=commercial。

  最簡單的 LF-標籤表達式由只有一個 LF 標籤，如。module=sales

- 在具有多個值的 LF 標籤上被授與權限的主體可以存取具有這些值之一的資料目錄資源。例如，如果授與使用者具有 key= module 和 value = 的 LF 標籤orders,customers，則該使用者可以存取已指派或的資源。module=orders module=customers
- 您需要具有Grant with LF-Tag expressions使用 LF-TBAC 方法授與資料目錄資源的資料權限的權限。資料湖管理員和 LF 標籤建立者隱含地接收此權限。具有Grant with LFTag expressions權限的主體可以使用以下命令授與資源的資料權限：
  - 指定的資源方法
  - LF-TBAC 方法，但只能使用相同的 LF 標記表達式

    例如，假設資料湖管理員進行下列授權 (以虛擬程式碼形式)。

    ```
    GRANT (SELECT ON TABLES) ON TAGS module=customers, region=west,south TO user1 WITH
     GRANT OPTION
    ```

    在這種情況下，user1可以使用 LF-TBAC 方法將資料表授與其他主體，但只能使用完整的 LF 標籤運SELECT算式。module=customers, region=west,south

- 如果同時具有 LF-TBAC 方法和具名資源方法的資源授與權限的主參與者，則主參與者對資源所擁有的權限就是這兩種方法所授與之權限的聯集。
- Lake Formation 支援跨帳戶授DESCRIBE與 ASSOCIATE LF 標籤，並使用 LF-TBAC 方法授與跨帳戶資料目錄資源的權限。在這兩種情況下，主體都是 AWS 帳戶 ID。

> **Note**
>
> Lake Formation 支持使用 LF-TBAC 方法的組織和組織單位跨帳戶贈款。若要使用此功能，您必須將跨帳戶版本設定更新為第 3 版。

如需詳細資訊，請參閱 Lake Formation 的跨賬戶數據共享。

- 在一個帳戶中建立的資料目錄資源只能使用在相同帳戶中建立的 LF 標籤加上標籤。在一個帳戶中建立的 LF 標籤無法與另一個帳戶的共用資源相關聯。
- 使用以 Lake Formation 標籤為基礎的存取控制 (LF-TBAC) 授與跨帳戶對資料目錄資源的存取權，需要新增帳戶的「資料目錄」資源策略。 AWS 如需詳細資訊，請參閱 必要條件。
- LF 標籤鍵和 LF 標籤值的長度不能超過 50 個字元。
- 可指定給資料目錄資源的 LF 標籤數目上限為 50。
- 下列限制為軟限制：
  - 可以建立的 LF 標籤數目上限為 1000。
  - 可為 LF 標籤定義的最大值數目為 1000。
- 標籤鍵和值會在儲存時轉換為所有小寫。
- LF 標籤只能將一個值指定給特定資源。
- 如果將多個 LF 標籤授與具有單一授權的主參與者，則主體只能存取具有所有 LF 標籤的資料目錄資源。
- AWS GlueETL 工作需要完整的表格存取權。如果 AWS Glue ETL 角色沒有資料表中所有資料行的存取權，工作將會失敗。您可以在資料行層級套用 LF 標籤，但可能會導致 AWS Glue ETL 角色遺失完整資料表存取權，而且工作失敗。
- 如果 LF-tag 運算式評估結果只能存取資料表資料行的子集，但是存在相符項目時授與的 Lake Formation 權限是需要完整資料行存取權限的其中一個權限`Alter`，也就是`Drop`、`Insert`或`Delete`，則不會授與這些權限。相反，只`Describe`被授予。如果授與的權限為 `All` (`Super`)，則僅授`Select`與並`Describe`被授與。

- 通配符不與 LF 標籤一起使用。若要將 LF 標籤指派給資料表的所有資料行，請將 LF 標籤指派給資料表，且資料表中的所有資料行都會繼承 LF 標籤。若要將 LF 標籤指派給資料庫中的所有資料表，請將 LF 標籤指派給資料庫，並且資料庫中的所有資料表都會繼承該 LF 標籤。

受管理資料壓縮的支援格式和限制

若要透過 Amazon Athena、Amazon EMR 和 AWS Glue ETL 任務等 AWS 分析服務提供更佳的讀取效能，請為資料型錄中的冰山資料表 AWS Glue Data Catalog 提供受管壓縮 (將小型 Amazon S3 物件壓縮為較大物件的程序)。

數據壓縮支持多種數據類型和壓縮格式，用於讀取和寫入數據，包括從加密表中讀取數據。

資料壓縮支援：

- 檔案類型：鑲木地板
- 資料類型：布林值、整數、長整數、浮點數、雙精度、字串、小數、日期、時間、時間戳記、字串、UUID、二進位
- 壓縮：zstd、gzip、snappy、未壓縮
- 加密：資料壓縮僅支援預設的 Amazon S3 加密 (SSE-S3) 和伺服器端 KMS 加密 (SSE-KMS)。
- BinPack 壓縮
- 結構描述演進
- 具有目標檔案大小的表格 (寫入。 target-file-size-bytes 在冰山配置中的屬性）包含範圍內 128MB 至 512 MB。
- 區域
  - 亞太區域 (東京)
  - 亞太區域 (首爾)
  - 亞太區域 (孟買)
  - 亞太區域 (新加坡)
  - 歐洲 (愛爾蘭)
  - 歐洲 (法蘭克福)
  - 美國東部 (維吉尼亞北部)
  - 美國東部 (俄亥俄)
  - 美國西部 (加利佛尼亞北部)
  - 南美洲 (聖保羅)

- 當儲存基礎資料的 Amazon S3 儲存貯體位於其他帳戶中時，您可以從 Data Catalog 所在的帳戶中執行壓縮程序。若要執行此程序，壓縮角色需要 Amazon S3 儲存貯體的存取權。

資料壓縮目前不支援：

- 檔案類型：阿夫羅、ORC
- 資料類型：固定
- 壓縮：brotli、lz4
- 在分割區規格演進時壓縮文件。
- 一般排序或堆疊順序排序
- 合併或刪除檔案：壓縮程序會略過具有刪除與檔案相關聯的資料檔案。
- 跨帳戶資料表壓縮：您無法在跨帳戶資料表上執行壓縮程序。
- 跨區域表格壓縮：您無法在跨區域資料表上執行壓縮。
- 在資源連結上啟用壓縮功能
- Amazon S3 儲存貯體的 VPC 端點

解決 Lake Formation

如果您在使用 AWS Lake Formation 時遇到問題，請參閱本節中的主題。

主題

一般性問題的故障診斷

使用這裡的信息來幫助您診斷和解決各種 Lake Formation 問題。

錯誤：上的 Lake Formation 權限不足 <Amazon S3 location>

嘗試在資源指向的 Amazon S3 位置上，在沒有資料位置許可的情況下建立或變更資料目錄資源。

如果資料目錄資料庫或表格指向 Amazon S3 位置，則當您授與 Lake Formation 權限時`ALTER`，`CREATE_TABLE`或者您也必須授與該位置的`DATA_LOCATION_ACCESS`權限。如果您要將這些權限授與外部帳戶或組織，則必須包含授與選項。

將這些權限授與外部帳戶後，該帳戶中的資料湖管理員接著必須將權限授與帳戶中的主體 (使用者或角色)。授與從其他帳戶接收的`DATA_LOCATION_ACCESS`權限時，您必須指定擁有者AWS 帳戶的目錄 ID (帳戶 ID)。所有者帳戶是註冊該地點的帳戶。

如需詳細資訊，請參閱 基礎資料存取控制 及 授與資料位置權限。

錯誤:「Glue API 的加密金鑰權限不足」

嘗試在沒有 AWS Identity and Access Management (IAM) 權限的情況下授予加密資料目錄的 AWS KMS 加密金鑰的 Lake Formation 許可。

使用清單的我 Amazon Athena 或 Amazon Redshift 查詢失敗

Lake Formation 不支持使用清單的查詢。

錯誤：「Lake Formation 權限不足：需要在目錄上創建標籤」

使用者/角色必須是資料湖管理員。

刪除無效資料湖管理員時發生錯誤

您應該同時刪除所有無效的資料湖管理員 (定義為資料湖管理員的已刪除 IAM 角色)。如果您嘗試分別刪除無效的資料湖管理員，Lake Formation 會擲回無效的主體錯誤。

跨帳戶存取疑難排解

使用此處的資訊可協助您診斷並修正跨帳戶存取問題。

主題

我授予了跨帳戶 Lake Formation 權限，但收件人看不到資源

- 收件者帳戶中的使用者是否為資料湖管理員？ 只有資料湖管理員可以在共用時看到資源。
- 您是否使用指定的資源方法與組織外部的帳號共用？ 如果是這樣，收件者帳戶的資料湖管理員必須接受 AWS Resource Access Manager (AWS RAM) 中的資源共用邀請。

  如需詳細資訊，請參閱 the section called “接受資 AWS RAM 源共用邀請”。
- 您是否在中使用帳戶層級 (資料目錄) 資源策略？AWS Glue 如果是，則如果您使用指定的資源方法，則必須在政策中包含特殊聲明， AWS RAM 以授權代表您共用政策。

  如需詳細資訊，請參閱 the section called “使用AWS Glue和 Lake Formation 管理跨帳戶權限”。
- 您是否擁有授予跨帳戶存取權所需的 AWS Identity and Access Management (IAM) 許可？

如需詳細資訊，請參閱 the section called “必要條件”。

- 您授與權限的資源不得授與任何 Lake Formation ms 權限給該IAMAllowedPrincipals群組。
- 帳戶層級deny政策中是否有關於資源的陳述？

收件者帳戶中的主體可以看到資料目錄資源，但無法存取基礎資料

收件者帳戶中的主體必須具有必要的 AWS Identity and Access Management (IAM) 許可。如需詳細資訊，請參閱 存取共用資料表的基礎資料。

錯誤:接受 AWS RAM 資源共用邀請時出現「關聯失敗，因為呼叫者未獲得授權」

將資源的存取權授與不同帳號後，當接收帳號嘗試接受資源共用邀請時，動作會失敗。

```
$ 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
        }
    ]
}
```

之所以發生錯誤，glue:PutResourcePolicy是因為AWS Glue當接收帳號接受資源共用邀請時叫用。若要解決此問題，請允許生產者/授與者帳戶所使用的假定角色glue:PutResourcePolicy執行動作。

錯誤：「未授權授予資源的權限」

嘗試授與另一個帳戶擁有的資料庫或資料表的跨帳戶權限。當資料庫或表格與您的帳戶共用時，身為資料湖管理員，您只能將其權限授與帳戶中的使用者。

錯誤：「無法擷取 AWS 組織資訊的存取」

您的帳戶是組 Organ AWS izations 管理帳戶，而且您沒有擷取組織資訊 (例如帳戶中的組織單位) 的必要權限。

如需詳細資訊，請參閱 Required permissions for cross-account grants。

錯誤：「<organization-ID>找不到組織」

嘗試與組織共用資源，但未啟用與組織共用。啟用與組織的資源共用。

如需詳細資訊，請參閱AWS RAM 使用指南中的啟用與 Organ AWS izations 共用。

錯誤：「Lake Formation 權限不足：非法組合」

使用者共用資料目錄資源，而 Lake Formation 權限已授與資源的`IAMAllowedPrincipals`群組。使用者必須`IAMAllowedPrincipals`先撤銷所有 Lake Formation 權限，才能共用資源。

ConcurrentModificationException 對外部帳戶的授助/撤銷請求

當用戶對 LF 標籤策略的主體進行多個並發授予和/或撤銷權限請求時，Lake Formation 會拋出。 ConcurrentModificationException用戶需要 catch 異常並重試失敗的授權/撤銷請求。使用`GrantPermissions`/`RevokePermissions`API 操作的批處理版本-並通過減少BatchGrantPermissions並發授BatchRevokePermissions予/撤銷請求的數量在一定程度上緩解此問題。

使用 Amazon EMR 存取透過跨帳戶共用的資料時發生錯誤

當您使用 Amazon EMR 存取從其他帳戶與您共用的資料時，部分 Spark 程式庫將嘗試呼叫 `Glue:GetUserDefinedFunctions` API 操作。由於 AWS RAM 受管理權限的版本 1 和 2 不支援此動作，因此您會收到下列錯誤訊息：

```
"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"
```

若要解決此錯誤，建立資源共用的資料湖管理員必須更新附加至資源共用的 AWS RAM 受管理權限。AWS RAM 受管許可第 3 版允許主體執行 `glue:GetUserDefinedFunctions` 動作。

如果您建立新的資源共用，Lake Formation 預設會套用最新版本的 AWS RAM 受管理權限，且您不需要採取任何動作。若要啟用現有資源共用的跨帳戶資料存取，您需要將 AWS RAM 受管理的權限更新為版本 3。

您可以檢視指 AWS RAM 派給中與您共用之資源的權限 AWS RAM。第 3 版中包含下列許可：

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
  AWSRAMPermissionGlueAllTablesReadWriteForCatalog
  AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

更新現有資源共用的 AWS RAM 受管理權限版本

您 (資料湖管理員) 可以依照AWS RAM 使用者指南中的指示將 AWS RAM 受管理的權限更新為較新版本，或者您可以撤銷該資源類型的所有現有權限並重新授予這些權限。如果您撤銷權限，請 AWS RAM 刪除與 AWS RAM 資源類型相關聯的資源共用。當您重新授與權限時， AWS RAM 會建立附加最新版 AWS RAM 受管理權限的新資源共用。

疑難排解藍圖和工作流程

使用此處的資訊可協助您診斷和修正藍圖和工作流程問題。

主題

- 我的藍圖失敗，顯示「User：<user-ARN>未授權執行：iam：PassRole 在資源上：<role-ARN>」
- 我的工作流失敗，出現「用戶：<user-ARN>未授權執行：iam：PassRole 在資源上：<role-ARN>」
- 我的工作流程中的爬蟲失敗，「資源不存在或請求者未授權訪問請求的權限」

- 我的工作流程中的爬蟲失敗，並顯示「調用 CreateTable 操作時發生錯誤（AccessDeniedException）...」

我的藍圖失敗，顯示「User：<user-ARN>未授權執行：iam：PassRole 在資源上：<role-ARN>」

沒有足夠權限無法傳遞所選角色的使用者嘗試建立藍圖。

更新使用者的 IAM 政策以傳遞角色，或要求使用者選擇具有所需密碼角色權限的其他角色。

如需詳細資訊，請參閱 the section called “Lake Formation 角色和 IAM 許可參考”。

我的工作流失敗，出現「用戶：<user-ARN>未授權執行：iam：PassRole 在資源上：<role-ARN>」

您為工作流程指定的角色沒有允許角色自行傳遞的內嵌原則。

如需詳細資訊，請參閱 the section called “(選擇性) 為工作流程建立 IAM 角色”。

我的工作流程中的爬蟲失敗，「資源不存在或請求者未授權訪問請求的權限」

一個可能的原因是傳遞的角色沒有足夠的權限，無法在目標資料庫中建立資料表。授與角色對資料庫的CREATE_TABLE權限。

我的工作流程中的爬蟲失敗，並顯示「調用 CreateTable 操作時發生錯誤（AccessDeniedException）...」

一個可能的原因是工作流程角色對目標儲存位置沒有資料位置權限。將資料位置權限授與角色。

如需詳細資訊，請參閱 the section called “DATA_LOCATION_ACCESS”。

的已知問題 AWS Lake Formation

檢閱的這些已知問題 AWS Lake Formation。

主題

- 篩選表格中繼資料的限制

- 重新命名排除的欄的問題
- 刪除 CSV 表格中的欄時發生問題
- 表分區必須在一個共同的路徑下添加
- 在建立工作流程期間建立資料庫時發生
- 刪除然後重新建立使用者時發生問題
- GetTables和 SearchTables API 不會更新IsRegisteredWithLakeFormation參數的值
- 資料目錄 API 作業不會更新IsRegisteredWithLakeFormation參數的值
- Lake Formation 操作不支持 AWS Glue 模式註冊表

篩選表格中繼資料的限制

AWS Lake Formation 資料行層級權限可用來限制對資料表中特定資料行的存取。當使用者使用主控台或類似 API 擷取有關資料表的中繼資料時`glue:GetTable`，資料表物件中的資料行清單只會包含他們有權存取的欄位。瞭解此中繼資料篩選的限制非常重要。

雖然 Lake Formation 為整合式服務提供有關資料行權限的中繼資料，但實際篩選查詢回應中的資料行是整合式服務的責任。支援資料行層級篩選的湖泊形成用戶端，包括 Amazon 雅典娜、亞馬遜 Redshift 頻譜和亞馬遜 EMR 會根據向 Lake Formation 註冊的欄許可來篩選資料。使用者將無法讀取他們不應該存取的任何資料。目前，AWS GlueETL 不支援資料行篩選。

> **Note**
>
> EMR 叢集並非由完全管理 AWS。因此，EMR 管理員有責任妥善保護叢集，以避免未經授權的資料存取。

某些應用程式或格式可能會將其他中繼資料 (包括欄名稱和類型) 儲存在`Parameters`地圖中作為表格屬性。這些屬性會以未修改的方式傳回，且任何具有任何資料行`SELECT`權限的使用者都可以存取。

例如，Avro 會將資料表結構定義的 JSON 表示法SerDe儲存在名為的資料表屬性中`avro.schema.literal`，可供所有具有資料表存取權的使用者使用此屬性。建議您避免在資料表屬性中儲存敏感資訊，並注意使用者可以瞭解 Avro 格式表格的完整結構描述。此限制特定於有關表格的中繼資料。

AWS Lake Formation 如果呼叫者沒有資料表中所有資料行的`SELECT`權限，則會在回應`glue:GetTable`或類似要求`spark.sql.sources.schema`時移除任何資料表屬性。這樣可以

防止使用者存取有關使用 Apache Spark 建立之資料表的其他中繼資料。在 Amazon EMR 上執行時，Apache Spark 應用程式仍可讀取這些表格，但可能不會套用某些最佳化，且不支援區分大小寫的資料行名稱。如果使用者擁有資料表中所有資料行的存取權，Lake Formation 會傳回未修改且具有所有資料表屬性的資料表。

重新命名排除的欄的問題

如果您使用資料行層級權限排除資料行，然後重新命名資料行，則不會再從查詢中排除該資料行，例如。`SELECT *`

刪除 CSV 表格中的欄時發生問題

如果您使用 CSV 格式建立「資料目錄」表格，然後從結構描述中刪除資料行，則查詢可能會傳回錯誤的資料，而且可能不會遵守資料行層級權限。

因應措施：改為建立新資料表。

表分區必須在一個共同的路徑下添加

Lake Formation 預計表格的所有分區都位於表格位置欄位中設定的共同路徑下。當您使用爬蟲程式將分割區新增至目錄時，這會順暢運作。但是，如果您手動新增分割區，且這些分割區不在父資料表中設定的位置下，則資料存取將無法運作。

在建立工作流程期間建立資料庫時發生

使用 Lake Formation 主控台從藍圖建立工作流程時，您可以建立目標資料庫 (如果目標資料庫不存在)。當您這麼做時，已登入的使用者會取得所建立之資料庫的`CREATE_TABLE`權限。不過，工作流程產生的爬行者程式會在嘗試建立資料表時擔任工作流程的角色。這會失敗，因為角色沒有資料庫的`CREATE_TABLE`權限。

因應措施：如果您在工作流程設定期間透過主控台建立資料庫，則在執行工作流程之前，必須將您剛建立之資料庫的`CREATE_TABLE`權限賦予與工作流程相關聯的角色。

刪除然後重新建立使用者時發生問題

以下情況導致返回的 Lake Formation 許可權錯誤：`lakeformation:ListPermissions`

1. 建立使用者並授予 Lake Formation 權限。
2. 刪除使用者。

3. 重新建立具有相同名稱的使用者。

ListPermissions傳回兩個項目，一個用於舊使用者，另一個用於新使用者。如果您嘗試撤銷授與舊使用者的權限，則會撤銷新使用者的權限。

GetTables和 SearchTables API 不會更新IsRegisteredWithLakeFormation參數的值

有一個已知的限制，即資料目錄 API 作業 (例如GetTables和SearchTables不會更新的值)IsRegisteredWithLakeFormation parameter，並傳回預設值 (此為 false)。建議您使用 GetTable API 來檢視的正確值IsRegisteredWithLakeFormation parameter。

資料目錄 API 作業不會更新IsRegisteredWithLakeFormation參數的值

資料目錄 API 作業 (例如GetTables和SearchTables不會更新IsRegisteredWithLakeFormation參數的值)，並傳回預設值 (即 false) 存在已知限制。建議您使用 GetTable API 來檢視IsRegisteredWithLakeFormation參數的正確值。

Lake Formation 操作不支持 AWS Glue 模式註冊表

Lake Formation 作業不支援包含StorageDescriptor要SchemaReference在架構註冊中使用的中的 AWS Glue 表格。

更新錯誤訊息

AWS Lake Formation 已將資源特定例外狀況更新為下列 API 作業的一般EntityNotFound錯誤訊息，以符合安全性和合規性目標。

- RevokePermissions
- GrantPermissions
- GetResourceLF 標籤
- GetTable
- GetDatabase

AWS Lake Formation API

> **Note**
>
> AWS Lake Formation 服務的更新 API 參考現已推出。

內容

- 字串模式

權限 API

[權限 API] 區段說明在 AWS Lake Formation中授與和撤銷權限所需的作業和資料類型。有關所有 API 操作和數據類型，請參閱 Lake Formation AWS Lake Formation API 參考指南。

作業

- GrantPermissions
- RevokePermissions
- BatchGrantPermissions
- BatchRevokePermissions
- GetEffectivePermissionsForPath
- ListPermissions
- GetDataLakePrincipal

資料類型

- Resource
- DatabaseResource
- TableResource
- TableWithColumnsResource
- DataCellsFilterResourcee
- DataLocationResource
- DataLakePrincipal
- PrincipalPermissions
- PrincipalResourcePermissions
- DetailsMap
- ColumnWildcard
- BatchPermissionsRequestEntry
- BatchPermissionsFailureEntry

資料湖設定 API

本節包含用於管理資料湖管理員的資料湖設定 API 作業和資料類型。

作業

- GetDataLakeSettings
- PutDataLakeSettings

資料類型

- DataLakeSettings

IAM 身分識別中心整合 API

本節包含與 IAM 身分中心建立和管理 Lake Formation 整合的作業。

作業

- CreateLakeFormationIdentityCenterConfiguration
- DeleteLakeFormationIdentityCenterConfiguration
- DescribeLakeFormationIdentityCenterConfiguration
- UpdateLakeFormationIdentityCenterConfiguration

資料類型

- ExternalFilteringConfiguration

混合式存取模式 API

「混合式存取模式 API」段落說明在中設定混合式存取模式所需的作業和資料類型 AWS Lake Formation。有關所有 API 操作和數據類型，請參閱 Lake Formation AWS Lake Formation API 參考指南。

作業

- CreateLakeFormationOptIn
- DeleteLakeFormationOptIn
- ListLakeFormationOptIns

資料類型

- Resource
- DatabaseResource
- TableResource
- 資源資訊
- LakeFormationOptInsInfo
- DataLocationResource

憑證自動販賣 API

[認證自動販賣 API] 區段說明與使用 AWS Lake Formation 服務以販售認證以及註冊和管理資料湖資源相關的作業和資料類型。

作業

- RegisterResource
- DeregisterResource
- ListResources
- GetUnfilteredTableMetadata
- GetUnfilteredPartitionsMetadata
- GetTemporaryGluePartitionCredentials
- GetTemporaryGlueTableCredentials
- UpdateResource

資料類型

- FilterCondition
- RowFilter
- ResourceInfo

標記 API

標記 API 部分描述了與授權策略相關的操作和數據類型，該策略定義了屬性或鍵值對標籤的權限模型。

作業

- ADDLF TagsToResource
- 拆卸式 F TagsFromResource
- GetResourceLF 標籤
- 列表標籤
- 創建標籤
- 獲取 F 標籤
- 更新 fTag
- 刪除標籤
- SearchTablesByLF 標籤
- SearchDatabasesByLF 標籤

資料類型

- LF TagKeyResource
- LF TagPolicyResource
- TaggedTable
- TaggedDatabase
- LFTAG
- LF TagPair

- LF TagError
- 欄位標籤

資料篩選器 API

資料篩選器 API 說明如何在中管理資料儲存格篩選器 AWS Lake Formation。

作業

- CreateDataCellsFilter
- DeleteDataCellsFilter
- ListDataCellsFilter
- GetDataCellsFilter
- UpdateDataCellsFilter

資料類型

- DataCellsFilter
- RowFilter

常見資料類型

Common Data Types 說明 AWS Lake Formation中的其他常見資料類型。

ErrorDetail 結構

包含錯誤的詳細資訊。

欄位

- `ErrorCode` – UTF-8 字串，長度不可小於 1 個位元組，也不可以超過 255 個位元組，需符合Single-line string pattern。

  此錯誤相關的程式碼。
- `ErrorMessage` – 描述字串，長度不可超過 2048 個位元組，需符合URI address multi-line string pattern。

描述錯誤的訊息。

字串模式

API 使用以下常規表達式來定義適用於各種字串參數和成員的有效內容：

- 單行字串模式 –「[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\t]*」
- URI 位址多行字串模式 –「[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*」
- 自訂字串模式 #3 — "^\w+\.\w+\.\w+$」
- 自訂字串模式 #4 — "^\w+\.\w+$」
- 自訂字串模式 #5 — "arn:aws:iam::[0-9]*:role/.*」
- 自訂字串模式 #6 — "arn:aws:iam::[0-9]*:user/.*」
- 自訂字串模式 #7 — "arn:aws:iam::[0-9]*:group/.*」
- 自訂字串模式 #8 – "arn:aws:iam::[0-9]*:saml-provider/.*"
- 自訂字串模式 #9 – "^([\p{L}\p{Z}\p{N}_.:\/=+\-@%]*)$"
- 自訂字串模式 #10 – "^([\p{L}\p{Z}\p{N}_.:*\/=+\-@%]*)$"
- 自訂字串模式 #11 – "[\p{L}\p{N}\p{P}]*"

支援地區

本節介紹了 Lake Formation 的支持 AWS 區域 和功能的信息。

一般可用性

如需 AWS 區域 支援者 AWS Lake Formation，請參閱區域提供的 AWS 服務清單。

如需每個區域和 Lake Formation 服務配額的 Lake Formation 服務端點清單，請參閱AWS Lake Formation 端點和配額。

AWS GovCloud (US)

如需「 AWS GovCloud (US) 區域」與「標準」之間差異的概觀 AWS 區域，請參閱的 AWS Lake Formation 不同之處 AWS GovCloud (US)。

交易與儲存最佳化

以下提供適用於 Lake Formation 的受控資料表、交易支援和儲存最佳化功能： AWS 區域

區域名稱	區域參數	端點
美國東部 (維吉尼亞北部)	us-east-1	湖泊形成. 美國東部 1.Amazonaws.com 湖泊形成-菲普斯美国东部-亚马逊
美國東部 (俄亥俄)	us-east-2	湖泊形成. 美東 2. 亞馬遜 湖泊形成-菲普斯美國東部 2. 亞馬遜
美國西部 (奧勒岡)	us-west-2	湖泊形成. 美國西部-亞馬遜 湖泊形成-菲普斯美国西部-亚马逊

區域名稱	區域參數	端點
亞太區域 (孟買)	ap-south-1	湖形成. 阿勒南 1. 亚马孙
亞太區域 (首爾)	ap-northeast-2	湖形成. 阿拉伯東北部 2. 亞馬遜
亞太區域 (新加坡)	ap-southeast-1	湖形成. 阿拉伯東南部-亞馬遜
亞太區域 (悉尼)	ap-southeast-2	湖形成. 阿拉伯東南部 2. 亞馬遜
亞太區域 (東京)	ap-northeast-1	湖形成. 阿拉伯東北部-1. 亞馬遜
歐洲 (法蘭克福)	eu-central-1	湖泊形成. 歐盟-中部-1. 亞馬遜
歐洲 (愛爾蘭)	eu-west-1	湖泊形成. 歐洲西部-阿馬遜
歐洲 (倫敦)	eu-west-2	湖泊形成. 歐洲西部-亞馬遜
歐洲 (斯德哥爾摩)	eu-north-1	歐盟北部湖泊形成 1. 亞馬遜
加拿大 (中部)	ca-central-1	湖泊形成 .ca-中央-阿馬遜
南美洲 (聖保羅)	sa-east-1	湖泊形成. SA-東部-阿馬遜

的文件歷史記錄 AWS Lake Formation

下表說明的文件的重要變更 AWS Lake Formation。

變更	描述	日期
更新的政策變更	記錄了和 AWSLakeFormationDataAdmin策略的更改（添加了語句 ID AWSLakeFormationCrossAccountManager和刪除多餘權限）。	2024年3月14日
更新了 Lake Formation 的設置	更新了「設定 AWS Lake Formation」區段中的步驟。	2024年2月7日
更新的政策變更	為服務連結角色的內嵌原則新增權限。如需詳細資訊，請參閱使用服務連結角色進行 Lake Formation。	2024年2月7日
更新的政策變更	記錄了對LakeFormationDataAccessServiceRolePolicy策略的更改。	2024年2月2日
綜合 Lake Formation 限制	為 Lake Formation 的限制和考量創建了統一的部分。有關更多信息，請參閱 Lake Formation 限制。	2023 年 12 月 15 日
增加了文檔冰山壓實	為了透過 Athena 和 Amazon EMR 和 AWS Glue ETL 任務等 AWS 分析服務提供更好的讀取效能，請為資料目錄中的冰山表 AWS Glue Data Catalog 提供受管壓縮 (將小型 Amazon S3 物件壓縮為較大物	2023年11月25 日

	件的程序)。如需詳細資訊，請參閱最佳化 Iceberg 資料表。	
已新增 IAM 身分中心整合的說明文件	IAM 身分中心整合可讓使用者和群組存取強制執行 Lake Formation 權限的資料目錄資源。如需詳細資訊，請參閱 IAM 身分中心整合。	2023年11月25 日
已新增資料目錄檢視的文件	您可以使用 SQL 編輯器 Amazon Athena 或 Amazon Redshift 在中創建參考多達 10 AWS Glue Data Catalog 個表的視圖。如需詳細資訊，請參閱建立檢視表。	2023年11月25 日
更新了策略更改	記錄了對AWSLakeFormationCrossAccountManager策略的更改。	2023 年 10 月 25 日
已新增混合存取模式的文件	混合式存取模式提供彈性 Lake Formation 可選擇性地為您的 AWS Glue Data Catalog. 透過混合式存取模式，您現在擁有一個增量路徑，可讓您為一組特定使用者設定 Lake Formation 權限，而不會中斷其他現有使用者或工作負載的權限原則。如需詳細資訊，請參閱混合存取模式。	2023 年 9 月 26 日
增加了用於創建阿帕奇冰山表的文檔	現在，您可以創建使用 Apache 鑲木地板數據格式 AWS Glue Data Catalog 與駐留在 Amazon S3 的數據格式的 Apache 冰山表。如需詳細資訊，請參閱建立 Iceberg 資料表。	2023 年 8 月 16 日

新增跨區域資料存取的文件	Lake Formation 支援跨 AWS 區域查詢資料目錄資料表。您可以使用 Athena、Amazon EMR 從其他區域存取某個區域中的資料，並在指向來源資料庫和表格的其他區域建立資源連結來執行 AWS Glue ETL。您可以將資料目錄連接到存放 Amazon S3 資料中繼資料的外部中繼資料的外部中繼存放區，並使用 AWS Lake Formation安全地管理資料存取許可。如需詳細資訊，請參閱跨區域存取資料表。	2023 年 6 月 30 日
重新組織的內容	重新整理了指南中的章節，以配合 Lake Formation 的使用者旅程。	2023 年 5 月 15 日
已新增 HMS 聯盟的說明文件	您可以將資料目錄連接到存放 Amazon S3 資料中繼資料的外部中繼資料的外部中繼存放區，並使用 AWS Lake Formation安全地管理資料存取許可。如需詳細資訊，請參閱管理使用外部中繼存放區之資料集的權限。	2023年4月15日
添加了 Amazon Redshift 數據共享的文檔	您現在可以使用 Lake Formation 許可安全地管理來自 Amazon Redshift 的資料記憶體中的資料。Lake Formation 支持通過許可訪問您的數據 AWS Data Exchange。如需詳細資訊，請參閱中的資料共用 AWS Lake Formation。	2022 年 11 月 30 日

Support 直接與校長跨帳戶資料共用	已新增直接與其他帳戶中 IAM 主體共用資料的相關資訊。如需詳細資訊，請參閱中的跨帳戶資料共 AWS Lake Formation用。	2022 年 11 月 10 日
Support 使用 T AWS RAM BAC 啟用的資料共用	已新增有關授與跨帳戶授與資料目錄權限使用 AWS Resource Access Manager 之 LF-TBAC 方法的相關資訊。	2022 年 11 月 10 日
增加了與其他服務合作的部分	已新增有關 Athena AWS Glue、Redshift 頻譜和 Amazon EMR 等 AWS 服務如何使用 Lake Formation 安全存取註冊在湖泊形成的 Amazon S3 位置中的資料的資料的資訊。如需更多資訊，請參閱與其他 AWS 服務合作。	2022 年 11 月 10 日
???	已新增使用 Amazon EMR 存取跨帳戶資料時，疑難排解錯誤的相關資訊。如需詳細資訊，請參閱 使用 Amazon EMR 存取透過跨帳戶共用的資料時發生錯誤。	2022 年 11 月 7 日
跨帳號資源共用的更新	新增跨帳號資源共用在 Lake Formation 中如何運作的說明。記錄了對AWSLakeFormationCrossAccountManager策略的更改。	2022 年 5 月 6 日
新教學課程	已新增建立受控管資料表、保護資料湖和共用資料湖的新教學課程。如需詳細資訊，請參閱「開始使用」一節。	2022 年 4 月 20 日

新, Lake Formation, 著陸, 頁面	已更新 Lake Formation 登陸頁面，包含教學課程連結，這些教學課程提供如何使用 Lake Formation 建置資料湖、擷取資料、共用和保護資料湖的 step-by-step指示。	2022 年 4 月 20 日
Support 憑證自動售貨	已新增認證自動販賣的相關資訊，該資訊支援 Lake Formation，以允許第三方服務透過憑證自動販賣 API 作業與 Lake Formation 整合。有關更多信息，請參閱憑證自動售貨機如何在 Lake Formation 中工作。	2022 年 2 月 28 日
Support 受控管的資料表和進階資料篩選	已新增有關受控資料表的資訊，這些資料表支援 ACID 交易、自動資料壓縮和時間行程查詢。已新增有關建立資料篩選以支援資料行層級安全性、資料列層級安全性和儲存格層級安全性的資訊。如需詳細資訊，請參閱 L ake Formation 和資料篩選中的控管表和湖 Lake Formation 中的儲存格層級安全性。	2021 年 11 月 30 日
Support VPC 介面端點	新增有關為 Lake Formation 建立虛擬私有雲端 (VPC) 介面端點的資訊，以便 VPC 與 Lake Formation 之間的通訊在網路中完全安全地進行。 AWS 如需詳細資訊，請參閱搭配 VPC 端點使用 Lake Formation。	2021 年 10 月 11 日

支援 VPC 端點政策	已新增有關在 Lake Formation 中支援 Virtual Private Cloud (VPC) 端 (VPC) 端點原則的相關資訊。如需詳細資訊，請參閱搭配 VPC 端點使用 Lake Formation。	2021 年 10 月 11 日
Support 以標籤為基礎的存取控制	以 Lake Formation 標籤為基礎的存取控制提供了一種新的、更具擴展性的方式，可透過使用 LF-tag 來管理對資料目錄資源和基礎資料的存取。如需詳細資訊，請參閱〈Lake Formation 標籤型存取控制〉。	2021 年 5 月 7 日
Amazon EMR 上資料篩選的新選擇加入要求。	已新增有關選擇加入以允許 Amazon EMR 篩選由 Lake Formation 管理之資料的要求資訊。如需詳細資訊，請參閱允許 Amazon EMR 上的資料篩選。	2020 年 10 月 9 日
Support 授與資料目錄資料庫的完整跨帳戶權限	已新增有關授與跨 AWS 帳戶之資料目錄資料庫完整 Lake Formation 權限的資訊，包括CREATE_TABLE ．如需詳細資訊，請參閱共用資料目錄資料庫。	2020 年 10 月 1 日

Support 透過 SAML 驗證的 Amazon Athena 使用者。	已新增有關透過 JDBC 或 ODBC 驅動程式連線並透過 SAML 身分識別提供者 (例如 Okta 和 Microsoft 作用中目錄聯合服務 (AD FS) 進行驗證的 Athena 使用者的支援資訊。如需詳細資訊，請參閱與 Lake Formation 的AWS 服務整合。	2020 年 9 月 30 日
透過加密的資料目錄 Support 跨帳戶存取	已新增有關在資料目錄加密時授與跨帳戶權限的資訊。如需詳細資訊，請參閱跨帳戶存取先決條件。	2020 年 7 月 30 日
Support 跨帳戶存取資料湖	已新增有關將 Data Catalog 資料庫和表格 AWS Lake Formation 權限授與外部 AWS 帳戶和組織，以及存取從外部帳戶共用之資料目錄物件的相關資訊。如需詳細資訊，請參閱跨帳戶存取。	2020 年 7 月 7 日
與 Amazon 集成 QuickSight	已新增有關如何授與 Lake Formation 許可給 Amazon QuickSight 企業版使用者的資訊，以便他們可以存取駐留在已註冊 Amazon S3 位置的資料集。如需詳細資訊，請參閱授與資料目錄權限。	2020 年 6 月 29 日
設定和入門章節的更新	重新組織並改進了「設置」和「入門」章節。更新資料湖管理員的建議 AWS Identity and Access Management (IAM) 許可。	2020 年 2 月 27 日

Support AWS Key Management Service	已新增有關 AWS Key Management Service (AWS KMS) 的 Lake Formation 支援如何簡化設定整合式服務，以便在已註冊的 Amazon Simple Storage Service (Amazon S3) 位置讀取和寫入加密資料的相關資訊。已新增有關如何註冊使用加密的 Amazon S3 位置的資訊 AWS KMS keys。如需詳細資訊，請參閱 the section called “將 Amazon S3 位置新增至您的資料湖”。	2020 年 2 月 27 日
藍圖和資料湖管理員 IAM 政策的更新	已釐清增量資料庫藍圖的輸入參數。更新資料湖管理員所需的 IAM 政策。	2019 年 12 月 20 日
安全章節重寫和升級章節修訂	改進了安全性和升級章節。	2019 年 10 月 29 日
超級權限取代全部權限	更新了「安全性」和「升級」章節，以反映`All`使用的取代權限`Super`。	2019 年 10 月 10 日
增補，更正和澄清	作出補充, 更正, 和基於反饋澄清. 修訂了安全章節。更新了「安全性」和「升級」章節，以反映群組的取代方`Everyone`式`IAMAllowedPrincipals` 。	2019 年 9 月 11 日
新的指南	這是初版的 AWS Lake Formation 開發人員指南。	2019 年 8 月 8 日

AWS 詞彙表

有關最新 AWS 術語，請參閱AWS 詞彙表 參考文獻中的AWS 詞彙表。

本文為英文版的機器翻譯版本，如內容有任何歧義或不一致之處，概以英文版為準。