

Lab: Integrate Databricks Unity Catalog with OneLake

Objective:

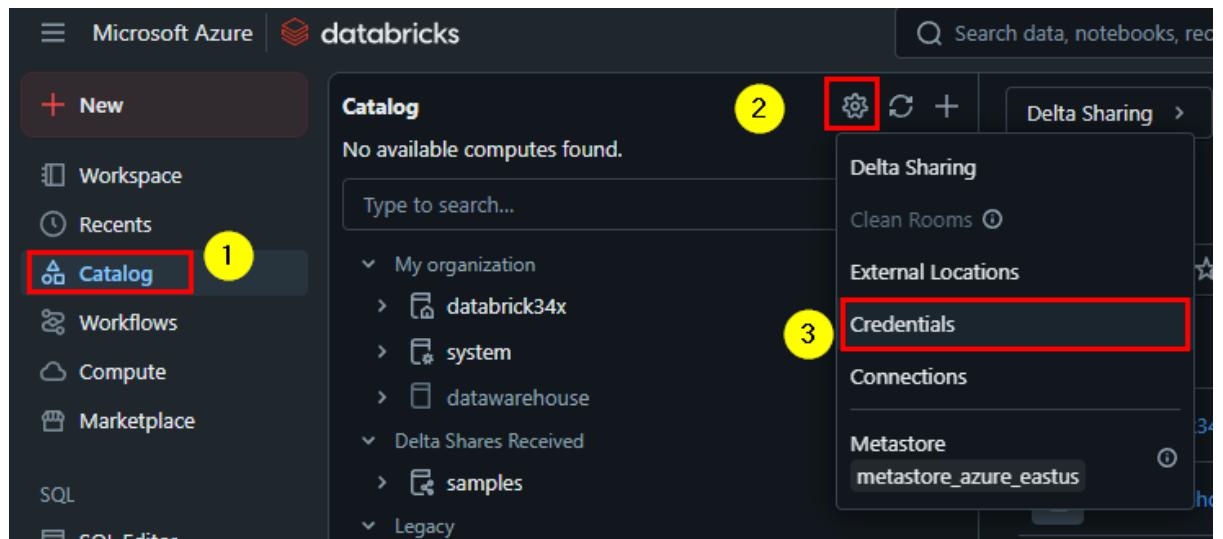
Understand the concept of integrating OneLake with the Unity Catalog.

Tasks:

1. Create a unity catalog in Databricks Workspace
2. Create external credentials and location
3. Set up your Cloud storage connection
4. Create a OneLake Shortcut

Task 1: Create credentials and external location

1. Open the Databricks workspace (<https://adb-3426738885164031.11.azuredatabricks.net/>) in a browser window.
2. Navigate to catalog click on the gear icon and select Credentials

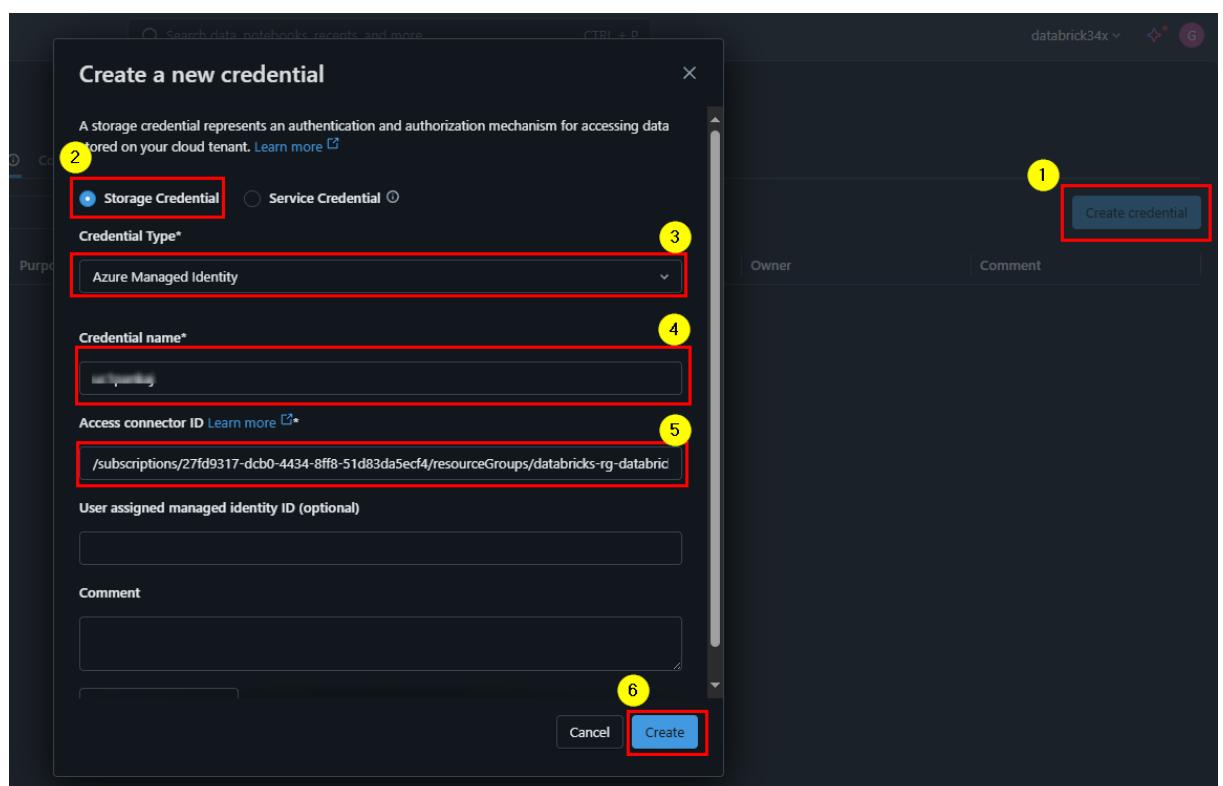


The screenshot shows the Databricks Catalog interface. On the left sidebar, the 'Catalog' item is highlighted with a red box and yellow circle number 1. At the top right, there is a gear icon highlighted with a red box and yellow circle number 2. On the right side, under the 'External Locations' section, the 'Credentials' item is highlighted with a red box and yellow circle number 3.

3. Click **Create credential** and provide the following details as shown:

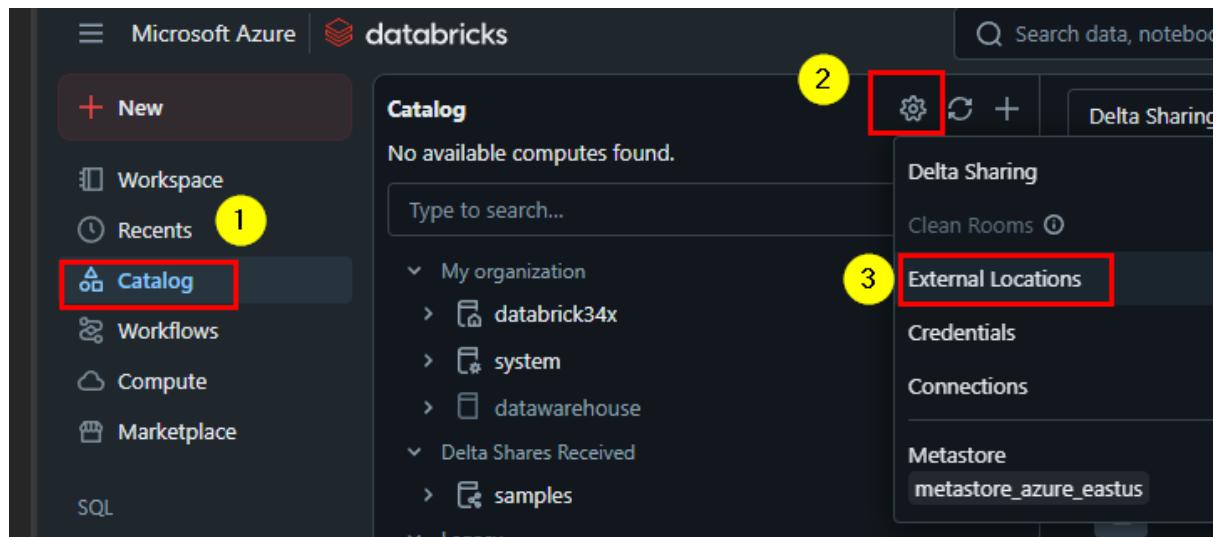
Access connector ID:

```
/subscriptions/27fd9317-dcb0-4434-8ff8-51d83da5ecf4/resourceGroups/RG-Genpact/providers/Microsoft.Databricks/accessConnectors/databricks-access-connector
```



The screenshot shows the 'Create a new credential' dialog box. The 'Storage Credential' radio button is selected, highlighted with a red box and yellow circle number 2. The 'Credential Type*' dropdown is set to 'Azure Managed Identity', highlighted with a red box and yellow circle number 3. The 'Credential name*' field contains 'databricks', highlighted with a red box and yellow circle number 4. The 'Access connector ID' field contains the URL '/subscriptions/27fd9317-dcb0-4434-8ff8-51d83da5ecf4/resourceGroups/databricks-rg-databr', highlighted with a red box and yellow circle number 5. The 'Create' button at the bottom right is highlighted with a red box and yellow circle number 6. On the right side of the dialog, there is a preview area showing the 'Owner' and 'Comment' columns, with the 'Create credential' button also highlighted with a red box and yellow circle number 1.

4. Create an external location



Catalog

No available computes found.

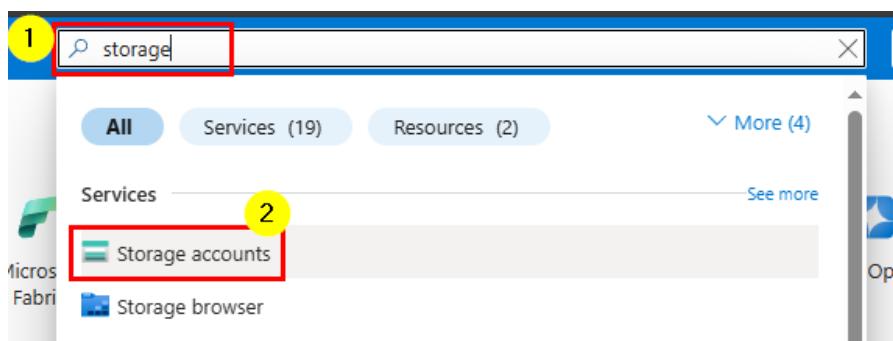
Type to search...

- My organization
 - databrick34x
 - system
 - datawarehouse
- Delta Shares Received
 - samples

External Locations

5. Create a container in Azure Data Lake storage

a. Search for Azure Storage:



storage

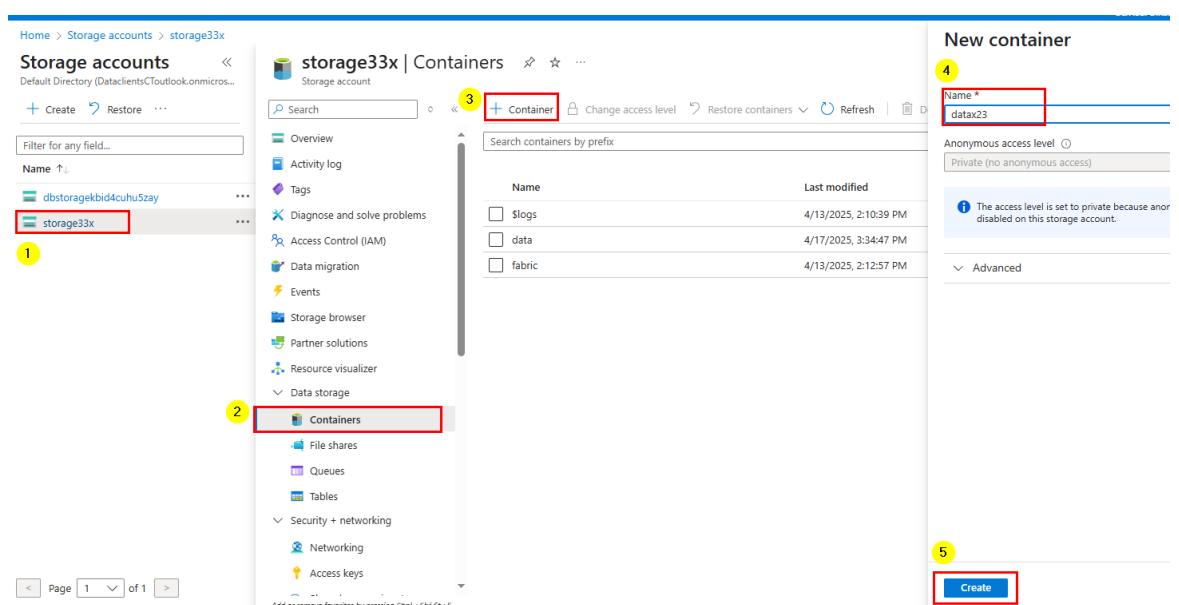
All Services (19) Resources (2) More (4)

Services

Storage accounts

Storage browser

b. Select **storage33x**, create a new container, and give a unique name.



storage33x | Containers

+ Container

Name	Last modified
Slogs	4/13/2025, 2:10:39 PM
data	4/17/2025, 3:34:47 PM
fabric	4/13/2025, 2:12:57 PM

New container

Name * data23

Anonymous access level Private (no anonymous access)

The access level is set to private because anonymous access is disabled on this storage account.

Create

- c. Construct your storage URL by replacing your container name

URL: abfss://<replace-with-your-containerName>@storage34x.dfs.core.windows.net/

Create a new external location

An external location allows you to access your data stored in cloud storage (e.g. Azure Data Lake Storage). You will need the cloud storage path and a paired credential (e.g. managed identity) which gives access to that path [Learn more](#)

External location name* uc1pankaj 1

URL* abfss://uc1pankaj@storage34x.dfs.core.windows.net/ 2

Storage credential* [Learn more](#) uc1pankaj (Managed Identity) 3

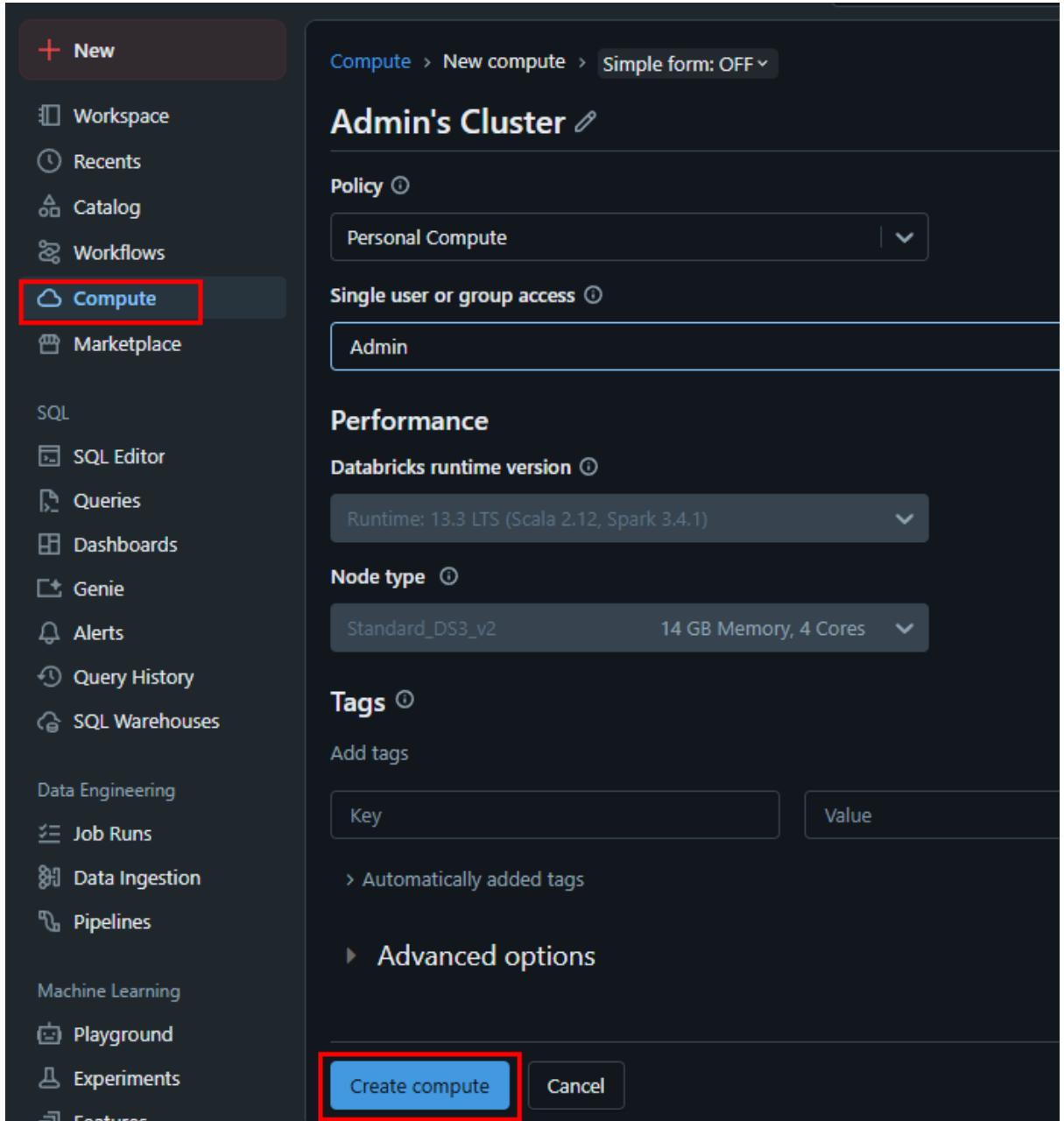
Comment

» Advanced Options 4

Cancel Create

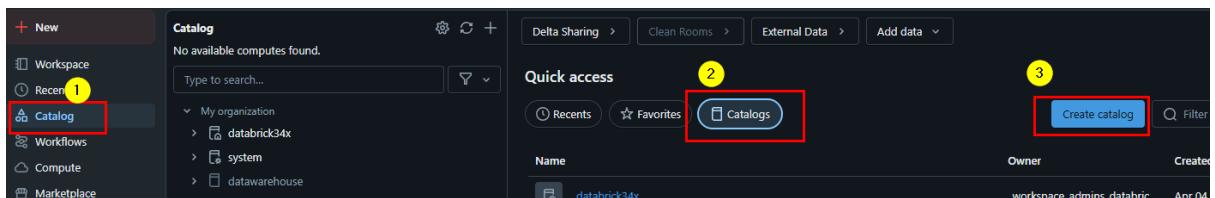
Task 2: Create a unity catalog in Databricks Workspace

- Create a new compute, navigate to compute and select **Create Compute**.



The screenshot shows the Databricks workspace interface. On the left, there is a sidebar with various options: Workspace, Recents, Catalog (which is highlighted with a red box), Workflows, Compute (which is also highlighted with a red box), Marketplace, SQL, SQL Editor, Queries, Dashboards, Genie, Alerts, Query History, SQL Warehouses, Data Engineering, Job Runs, Data Ingestion, Pipelines, Machine Learning, Playground, Experiments, and Features. On the right, the main area is titled "Admin's Cluster" and contains sections for Policy (set to Personal Compute), Single user or group access (set to Admin), Performance (Databricks runtime version set to 13.3 LTS (Scala 2.12, Spark 3.4.1)), Node type (Standard_DS3_v2, 14 GB Memory, 4 Cores), and Tags. At the bottom, there are "Create compute" and "Cancel" buttons, with the "Create compute" button highlighted by a red box.

- Select Catalog from the left vertical menu, select Catalogs, and click Create Catalog.



The screenshot shows the Databricks Catalog interface. On the left, there is a sidebar with options: New, Workspace, Recents (highlighted with a yellow circle), Catalog (highlighted with a red box), Workflows, Compute, and Marketplace. The main area is titled "Catalog" and shows a message "No available computes found." Below this is a search bar and a "Type to search..." input field. The "Quick access" bar includes "Recent" (highlighted with a yellow circle), "Favorites", and a "Catalogs" button (highlighted with a red box). The "Create catalog" button in the top right corner is also highlighted with a red box. The bottom of the screen shows a navigation bar with links for Delta Sharing, Clean Rooms, External Data, and Add data.

- Enter the catalog name and select the storage location to external location.

Create a new catalog

A catalog is the first layer of Unity Catalog's three-level namespace and is used to organize your data assets. [Learn more](#)

Catalog name* uc1pankaj 1

Type* 2 Standard

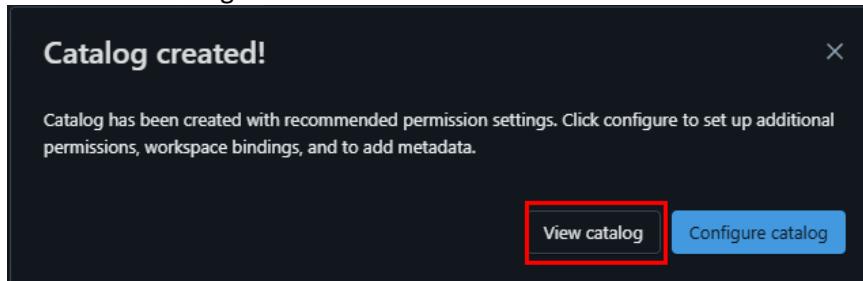
Storage location Cloud uc1pankaj sales

abfss://uc1pankaj@storage34x.dfs.core.windows.net/sales

Location in cloud storage where data for managed tables will be stored. If not specified, the location will default to the metastore root location. 3

Cancel Create

4. Select view catalog



5. Create a new notebook in Databricks inside the workspace
6. In the first cell enter the following code and create an empty cell and run it, if no cluster is created then create a new one.

```
use catalog <your-catalog-name>;
create schema if not exists sales;
use schema sales;
create table if not exists department
(
    deptcode int,
    deptname string,
    location string
);
```

```
INSERT INTO department VALUES
(10, 'FINANCE', 'EDINBURGH'),
(20, 'SOFTWARE', 'PADDINGTON');
```

```
Select * from department;
```

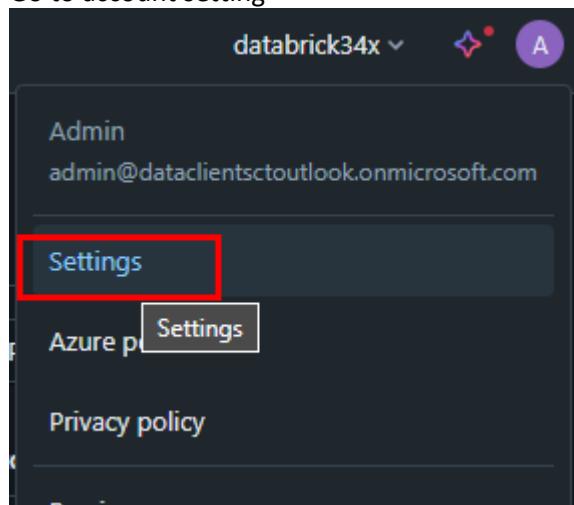
7. Create a new code cell and run the following code to create an external table. Replace your container name:

```
create external table trips_external
using delta
location 'abfss://<your-container-name>@storage34x.dfs.core.windows.net/trip'
as select * from samples.nyctaxi.trips;
```

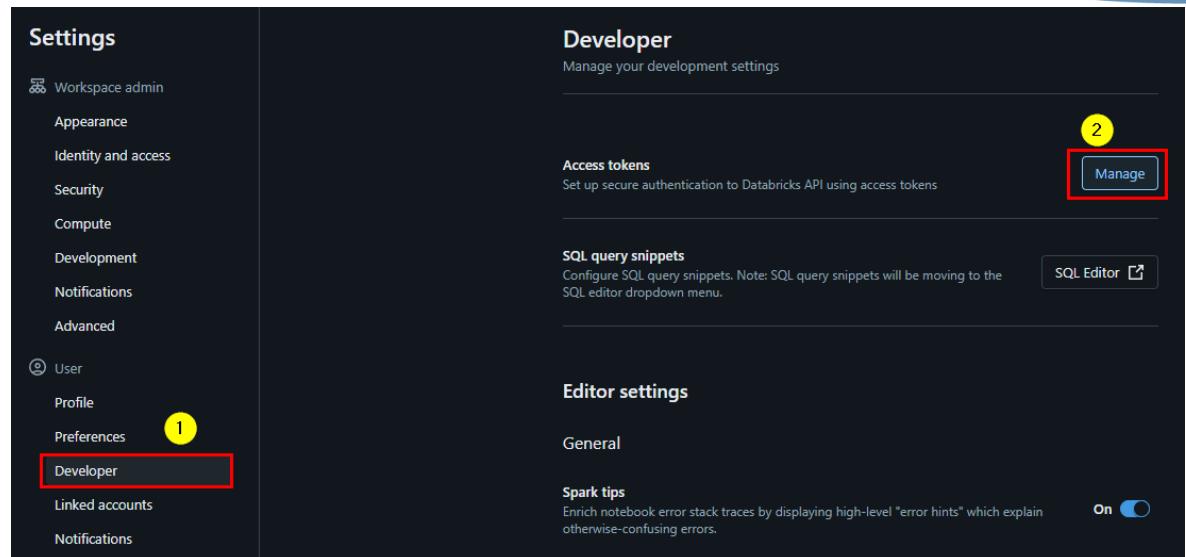
8. Terminate the compute resource.

9. Create a personal access token:

- Go to account setting



- Select Developer under User, and then Manage for Access token



Developer
Manage your development settings

Access tokens
Set up secure authentication to Databricks API using access tokens

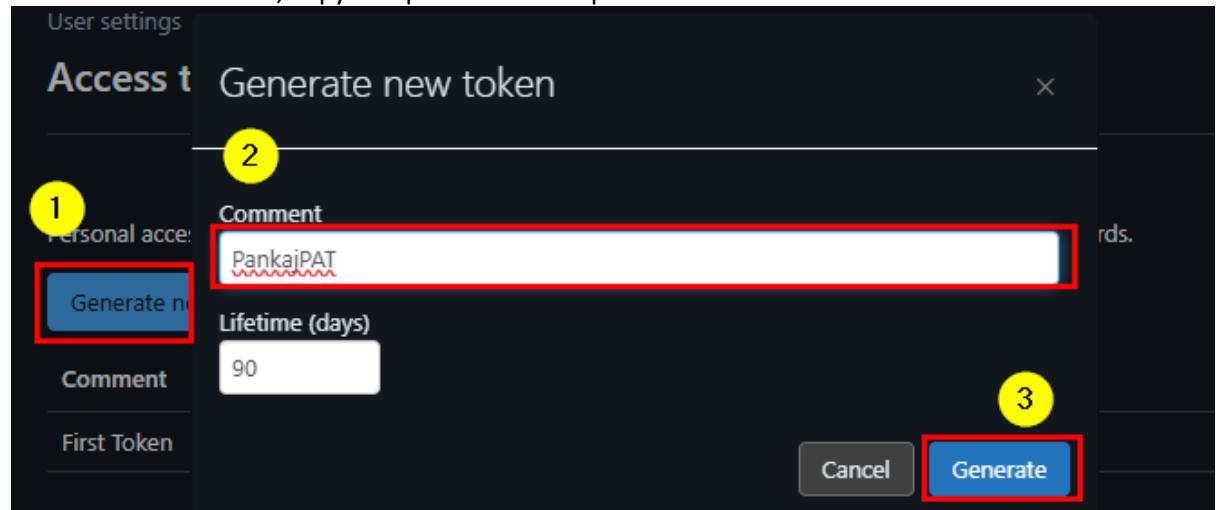
SQL query snippets
Configure SQL query snippets. Note: SQL query snippets will be moving to the SQL editor dropdown menu.

Editor settings

General

Spark tips
Enrich notebook error stack traces by displaying high-level "error hints" which explain otherwise-confusing errors. **On**

- c. Generate a new token, copy and past it in a notepad we need it later.



User settings

Access token **Generate new token** ×

Personal access token

Comment rds.

Generate new token Cancel **Generate**

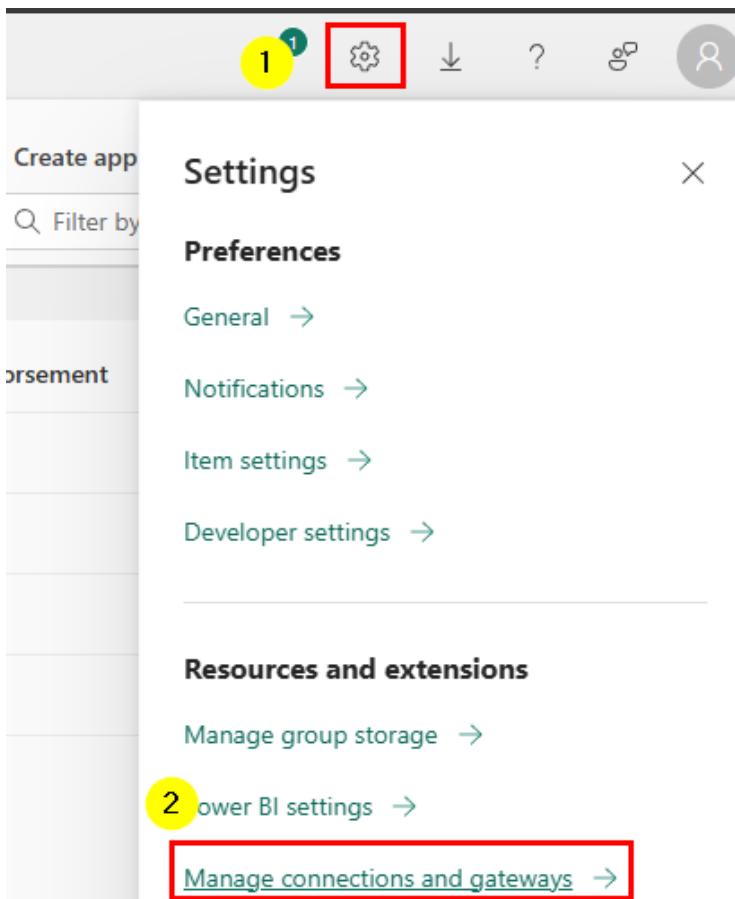
Lifetime (days)

Comment

First Token

Task 3: Set up your Cloud storage connection in Fabric

1. Switch back to the Microsoft Fabric portal and from setting open connections.



The screenshot shows the Microsoft Fabric portal's Settings page. At the top, there is a navigation bar with icons for notifications (yellow circle with '1'), settings (highlighted with a red box), download, help, and user profile. Below the navigation bar, the title 'Settings' is displayed. Under the 'Preferences' section, links for General, Notifications, Item settings, and Developer settings are shown. A horizontal line separates this from the 'Resources and extensions' section. In the 'Resources and extensions' section, links for Manage group storage and Manage connections and gateways (highlighted with a red box) are listed. A yellow circle with '2' is placed over the 'Manage connections and gateways' link.

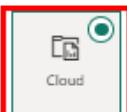
2. Enter the following details:

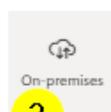
- a. Connection name: a unique name
- b. Connection type: ADLS Gen 2
- c. Storage URL: <https://storage33x.dfs.core.windows.net/>
- d. Full path: **Your Container Name**
- e. Authentication type: Key
- f. Account key:

**Us86NWDEUvmsB4UsUdwBi+BRTu1YSL1HA9IuQNgcnKoShZmab45Tf7Ulz/y9MbdJ5E
RIG7iEx2Jf+ASTOf+zMQ==**

New connection X

Currently, these cloud connections are not supported by Dataflows, Dataflows Gen2, and Datamarts. To create personal cloud connections for these experiences, please use the Dataflows or Datamarts editor in "Get Data".

1 

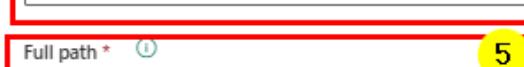
2 

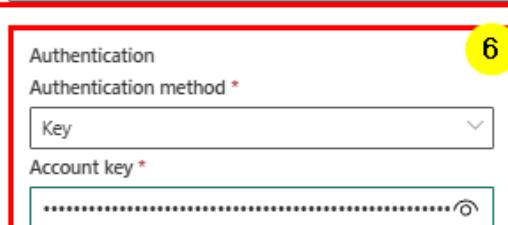
3 

4 

5 

6 

7 

8 

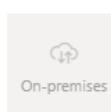
9 

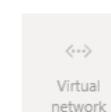
10 

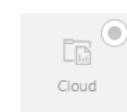
- When the connection is created, Open the newly created connection and copy the connection ID:

Settings X

Currently, these cloud connections are not supported by Dataflows, Dataflows Gen2, and Datamarts. To create personal cloud connections for these experiences, please use the Dataflows or Datamarts editor in "Get Data".

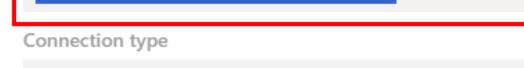
1 

2 

3 

4 

5 

6 

Task 4: Create a OneLake Shortcut

1. **Import the sync notebook** to your Fabric workspace. Download [This notebook](#)
2. **Configure the parameters** in the first cell of the notebook to integrate Unity Catalog tables.

The Databricks API, authenticated through the PAT token, is utilized for exporting Unity Catalog tables. The following snippet is used to configure the source (Unity Catalog) and destination (OneLake) parameters. Ensure to replace them with your own values.

```
# Databricks workspace
dbx_workspace = "<databricks_workspace_url>"
dbx_token = "<pat_token>"

# Unity Catalog
dbx_uc_catalog = "catalog1"
dbx_uc_schemas = '["sales"]'

# Fabric
fab_workspace_id = "<workspace_id>"
fab_lakehouse_id = "<lakehouse_id>"
fab_shortcut_connection_id = "<connection_id>"

# If True, UC table renames and deletes will be considered
fab_consider_dbx_uc_table_changes = True
```

3. Run all cells of the notebook to start synchronizing Unity Catalog Delta tables to OneLake using shortcuts. Once notebook is completed, shortcuts to Unity Catalog Delta tables are available in the lakehouse, SQL analytics endpoint, and semantic model.