

# Unity Catalog

# Learning Objectives

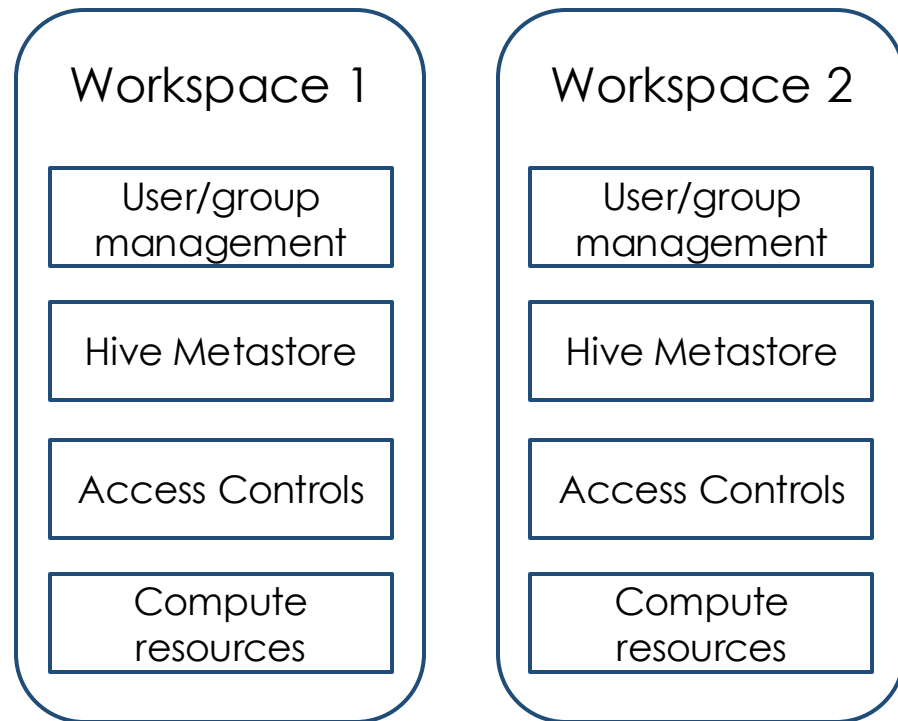
- ▶ What is Unity Catalog
- ▶ 3-level namespace
- ▶ Security model for governing data objects

# Unity Catalog

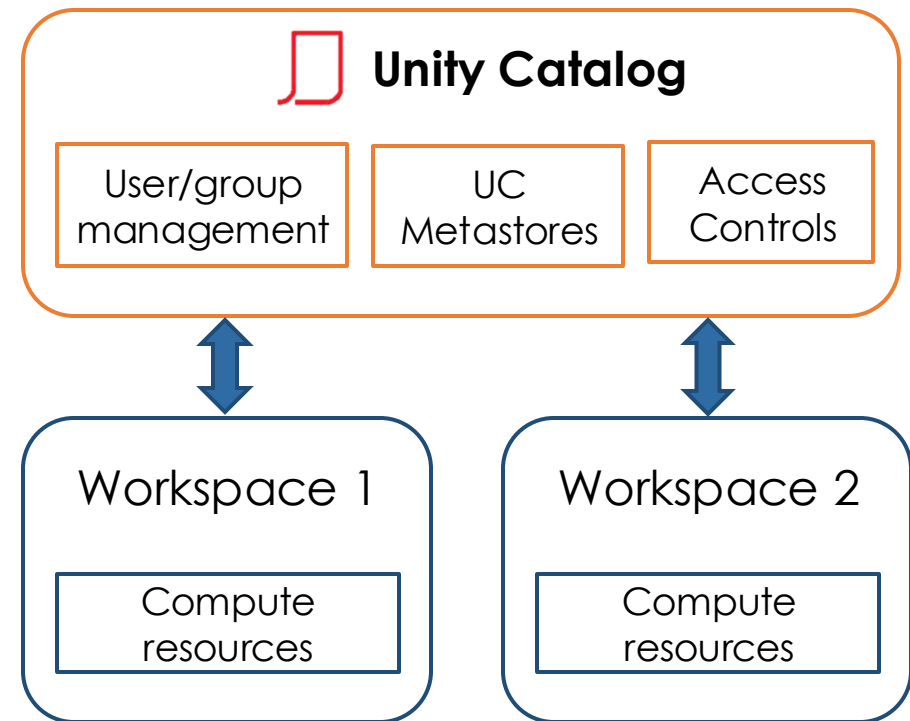
- ▶ Centralized governance solution across all your workspaces on any cloud.
- ▶ Unify governance for all data and AI assets
  - ▶ files, tables, machine learning models and dashboards
  - ▶ based on SQL

# Architecture

## Before Unity Catalog



## With Unity Catalog



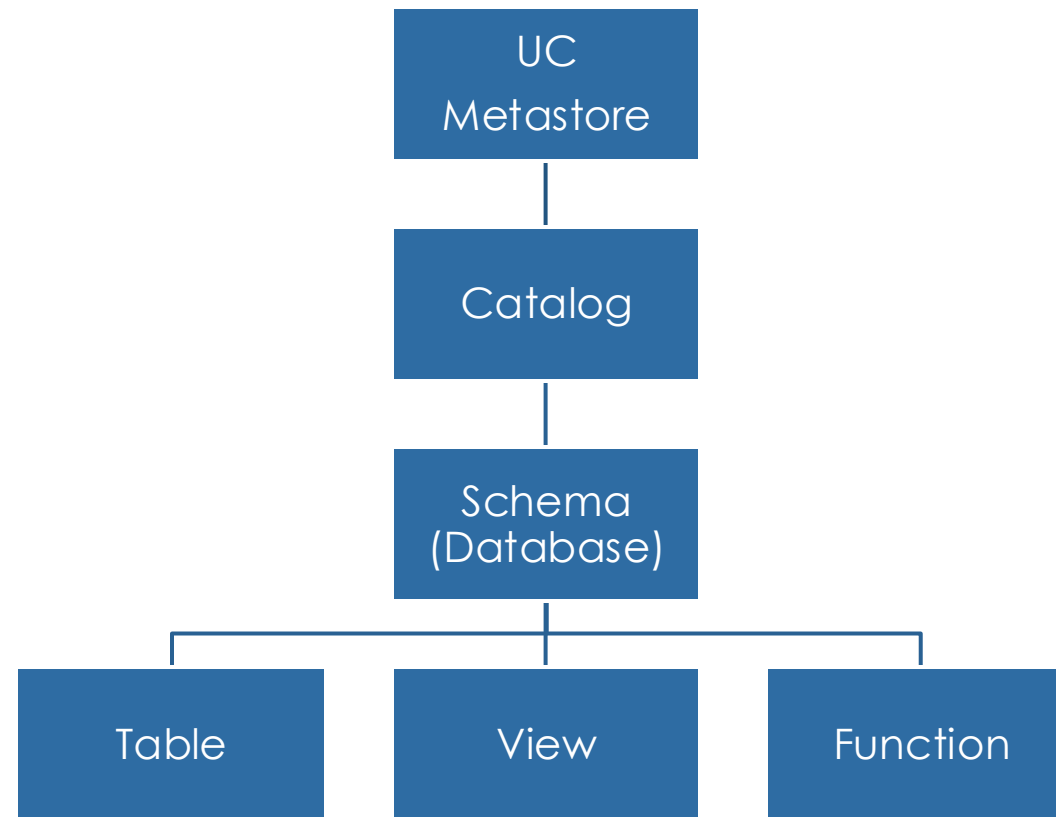
# UC 3-level namespace

```
SELECT * FROM schema.table
```

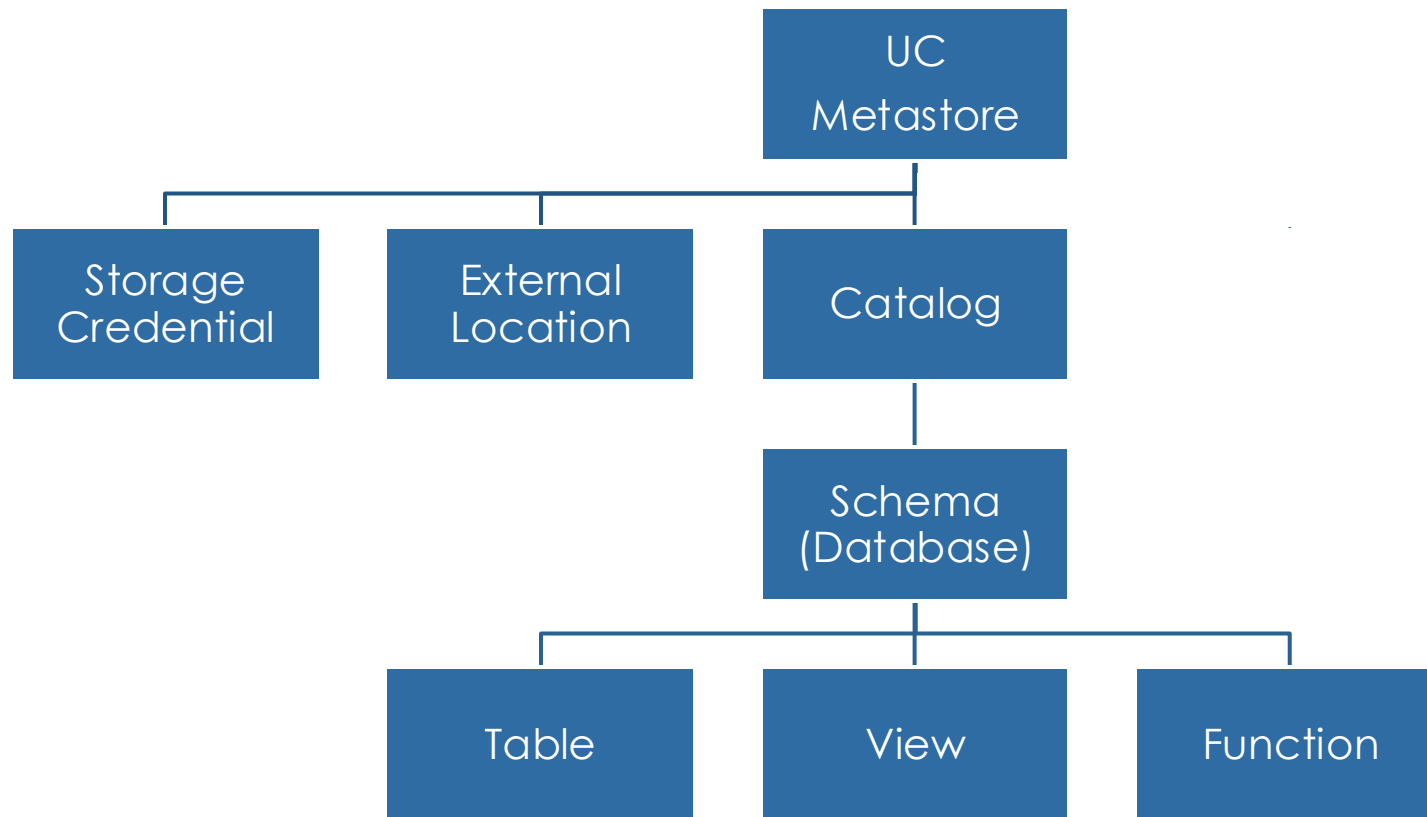


```
SELECT * FROM catalog.schema.table
```

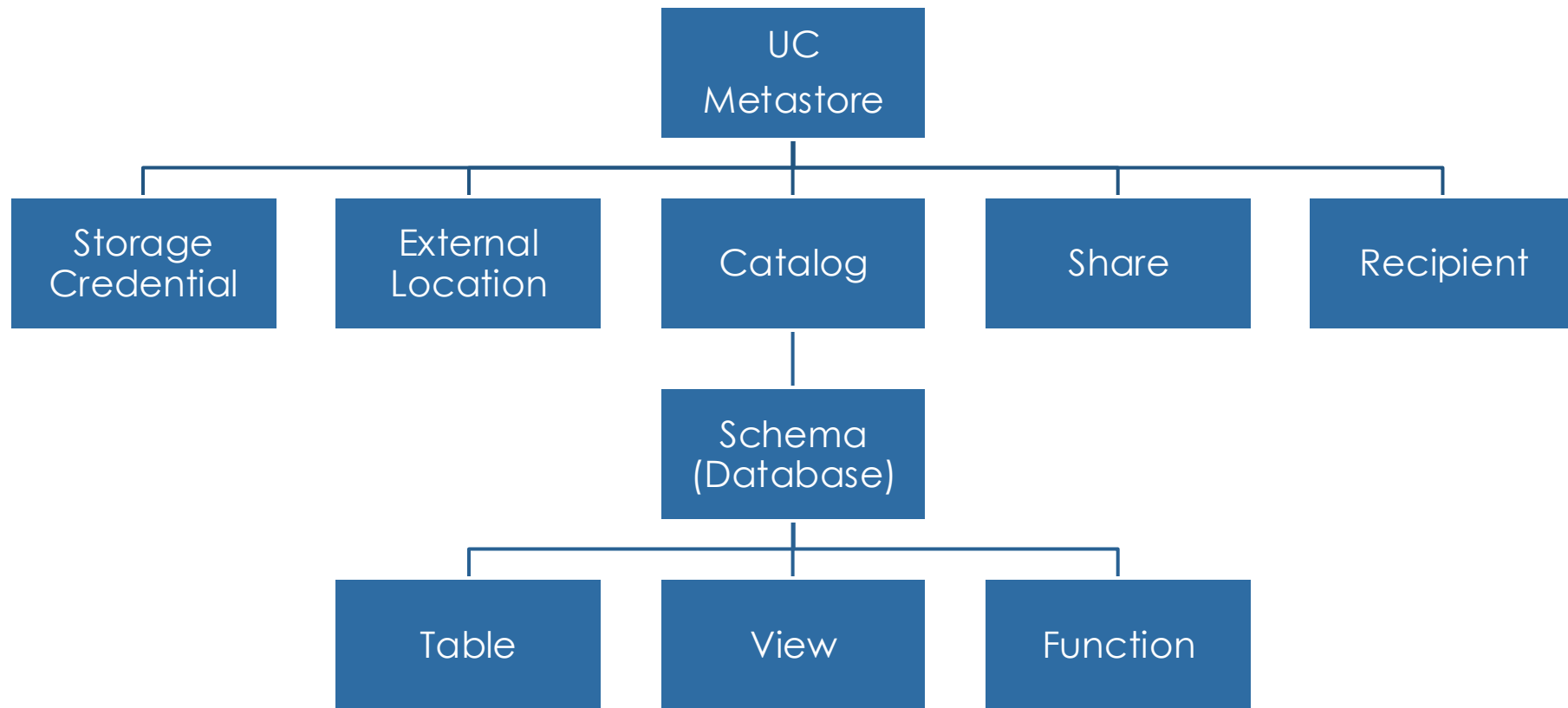
# UC hierarchy



# UC hierarchy



# UC hierarchy

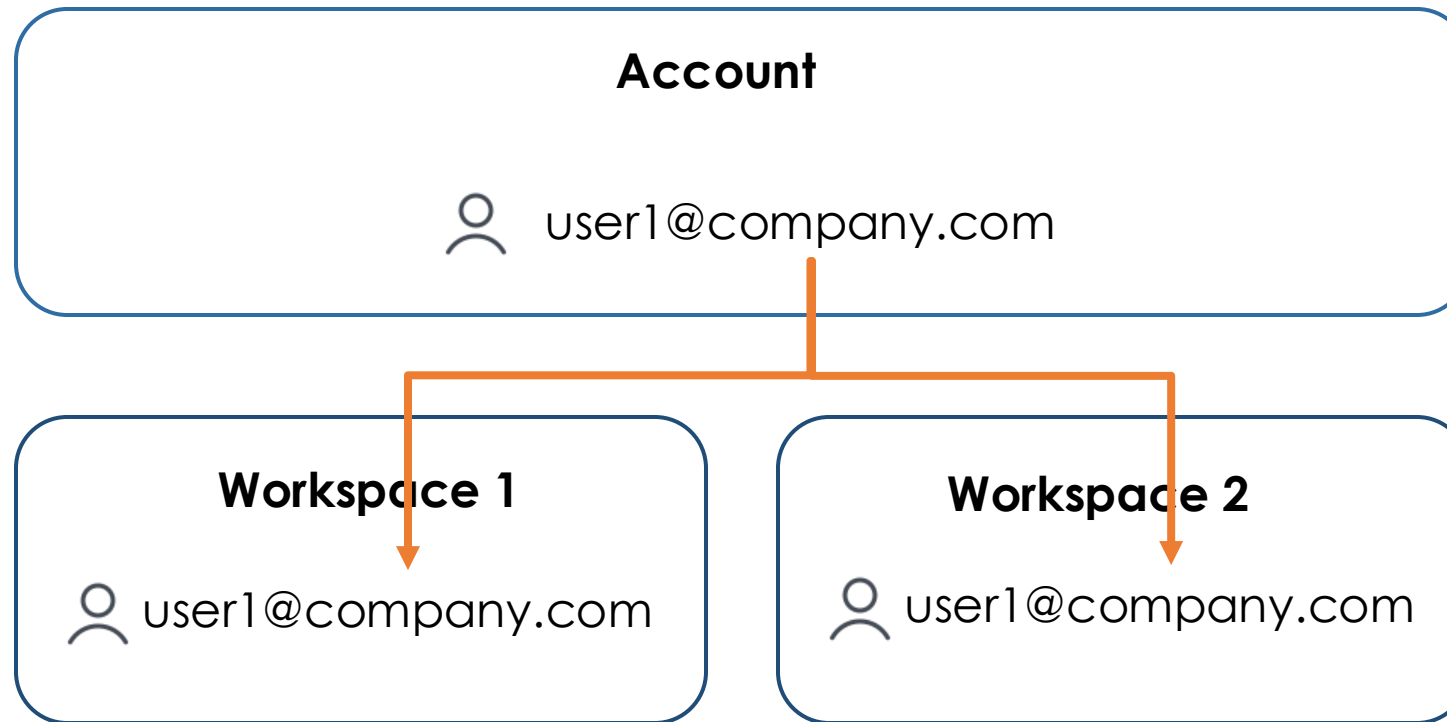




# Identities

- ▶ **Users:** identified by e-mail addresses
  - ▶ Account administrator
- ▶ **Service Principles:** identified by Application IDs
  - ▶ Service Principles with administrative privilege
- ▶ **Groups:** grouping Users and Service Principles
  - ▶ Nested groups

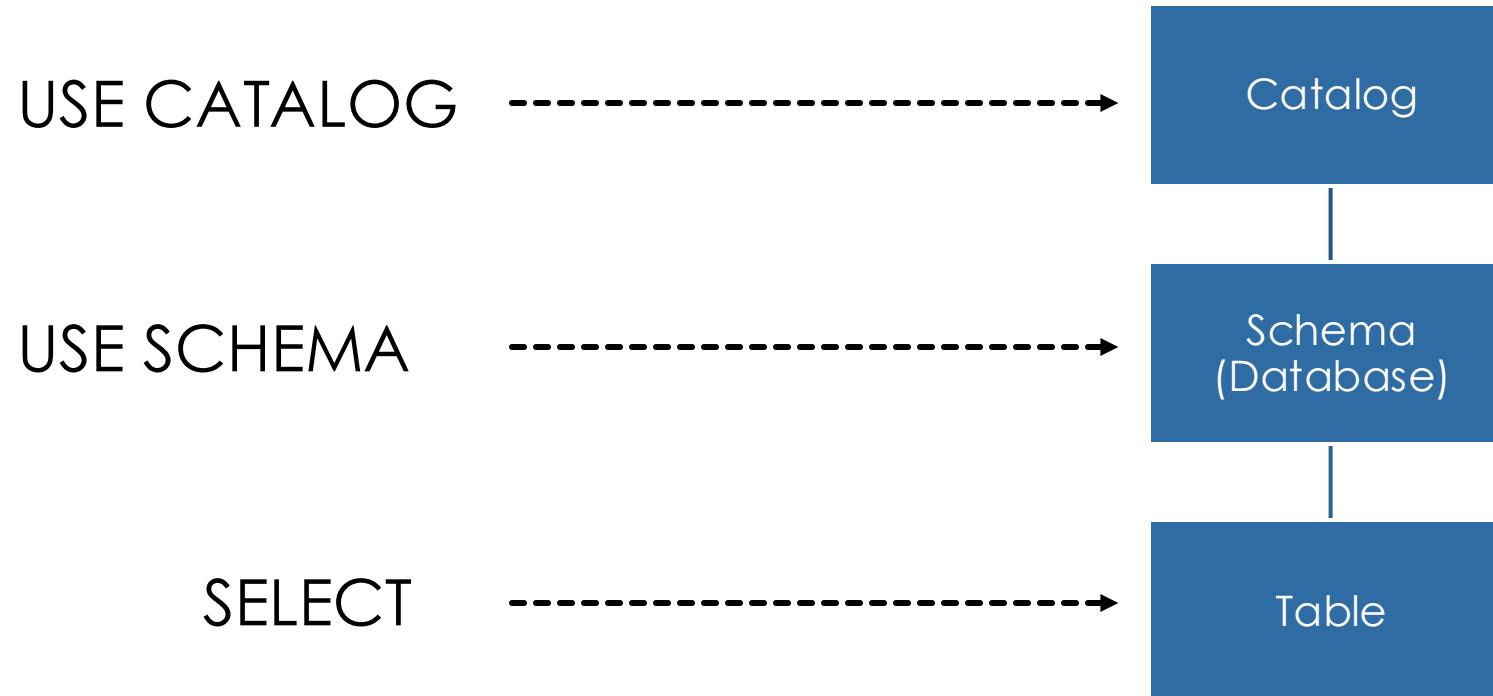
# Identity Federation



# Privileges

- ▶ CREATE
- ▶ SELECT
- ▶ MODIFY
- ▶ READ FILES
- ▶ WRITE FILES
- ▶ EXECUTE

# Usage Privileges



# Security model

GRANT **Privilege** ON **Securable\_Object** TO **Principal**

## Privileges

CREATE

SELECT

MODIFY

READ FILES

WRITE FILES

EXECUTE

USE CATALOG

USE SCHEMA

## Securable objects

Storage  
Credential

External  
Location

Catalog

Share

Recipient

Schema  
(Database)

Table

View

Function

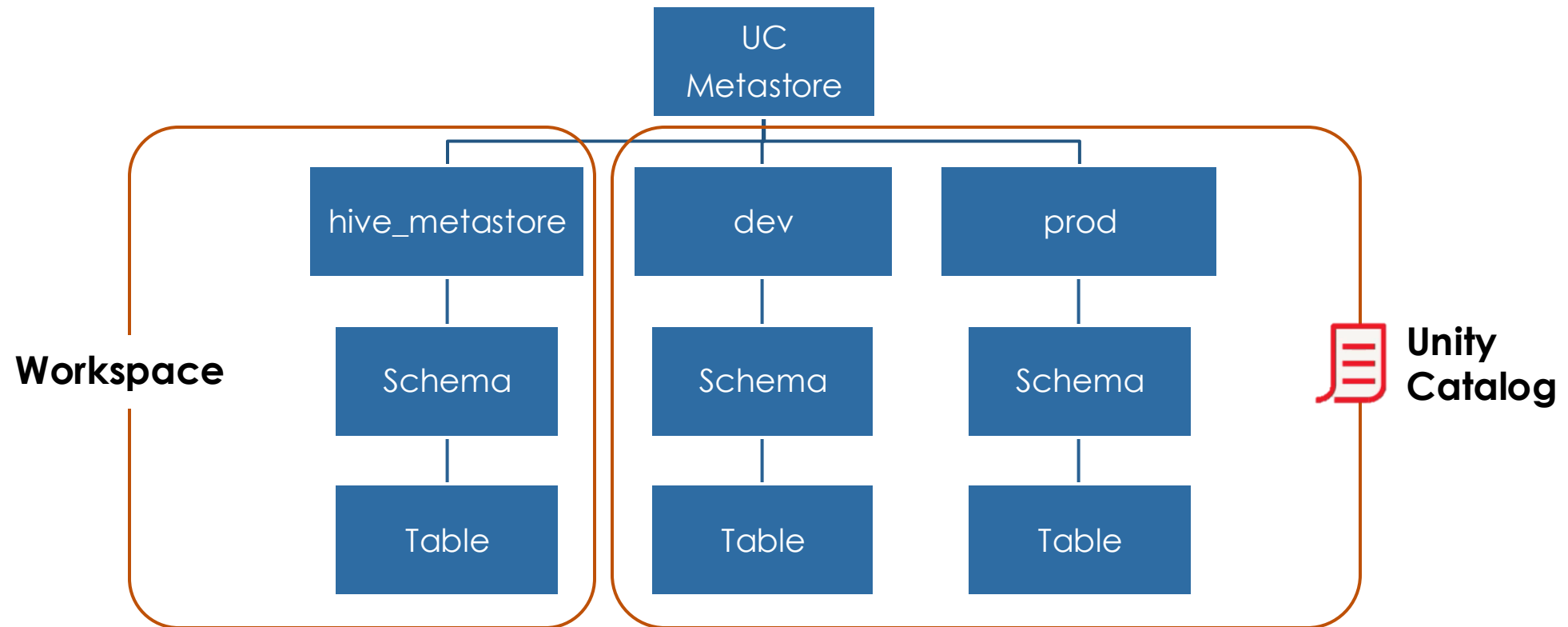
## Principles

User

Service principals

group

# Accessing legacy Hive metastore



# Features

- ▶ Centralized governance for data and AI
- ▶ Built-in data search and audit logging
- ▶ Automated lineage
- ▶ No hard migration required

# Account Console

- ▶ Log in as account administrator
- ▶ AWS: <https://accounts.cloud.databricks.com>
- ▶ Azure: <https://accounts.azure.databricks.net/>
- ▶ GCP: <https://accounts.gcp.databricks.com/>