



Data Administration in Databricks



Learning Objectives

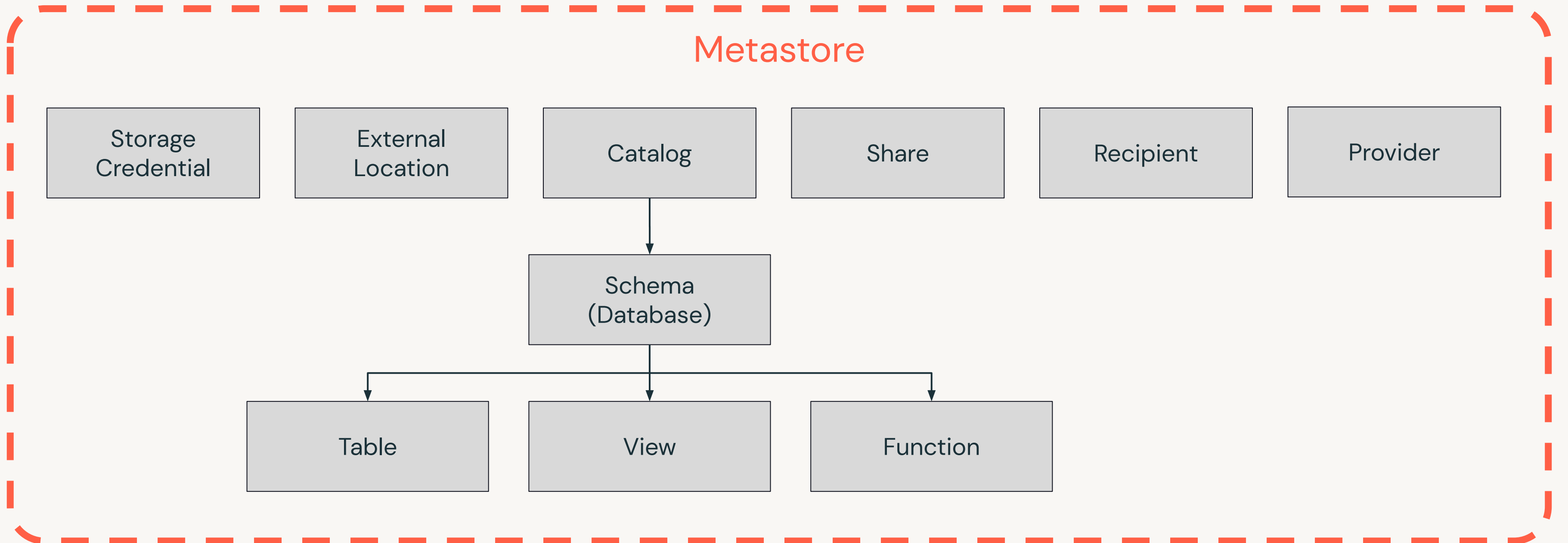
In this course, you will learn how to:

- Describe data access patterns in Databricks
- Define data access rules and manage data ownership
- Secure access to external storage
- Store tables, schemas and catalogs and external storage
- Upgrade legacy data assets to Unity Catalog

Key Concepts

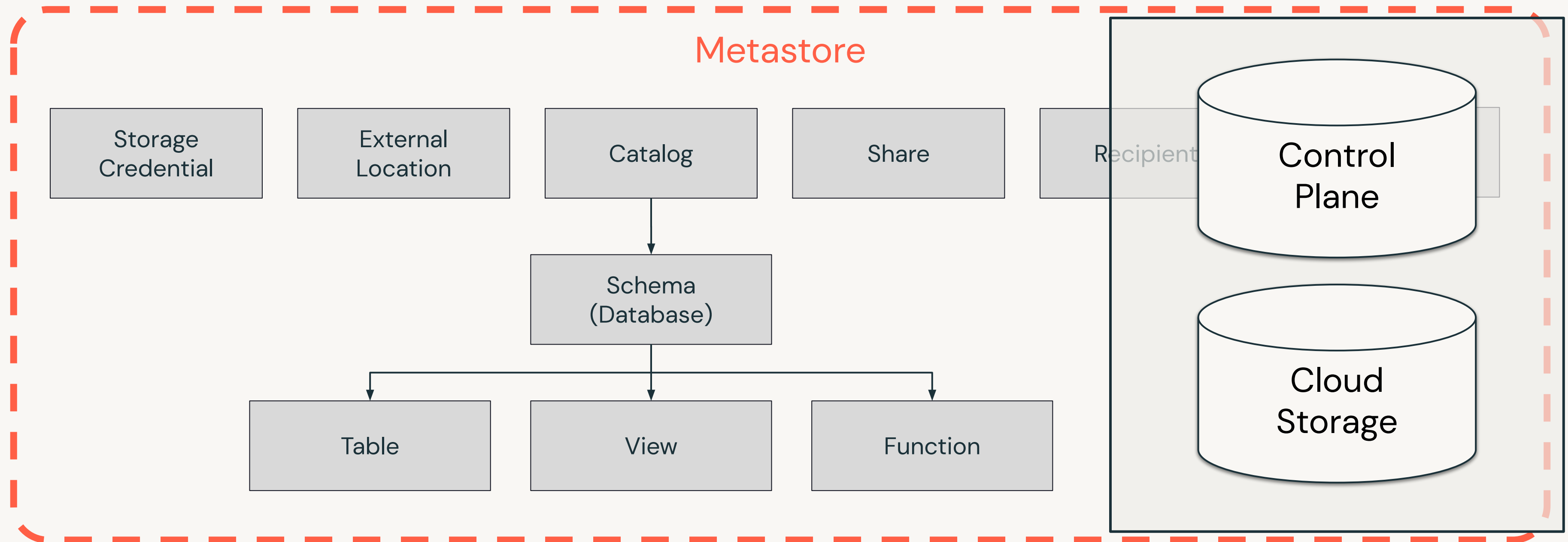
Key Concepts

Metastore



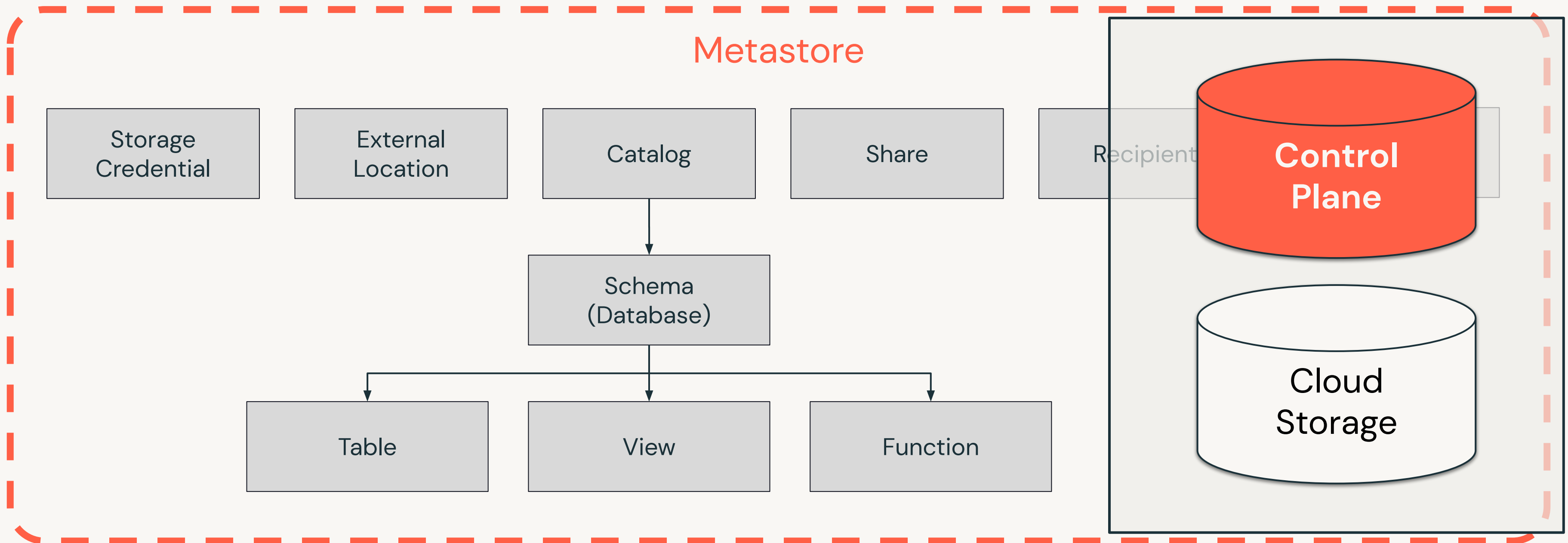
Key Concepts

Metastore storage



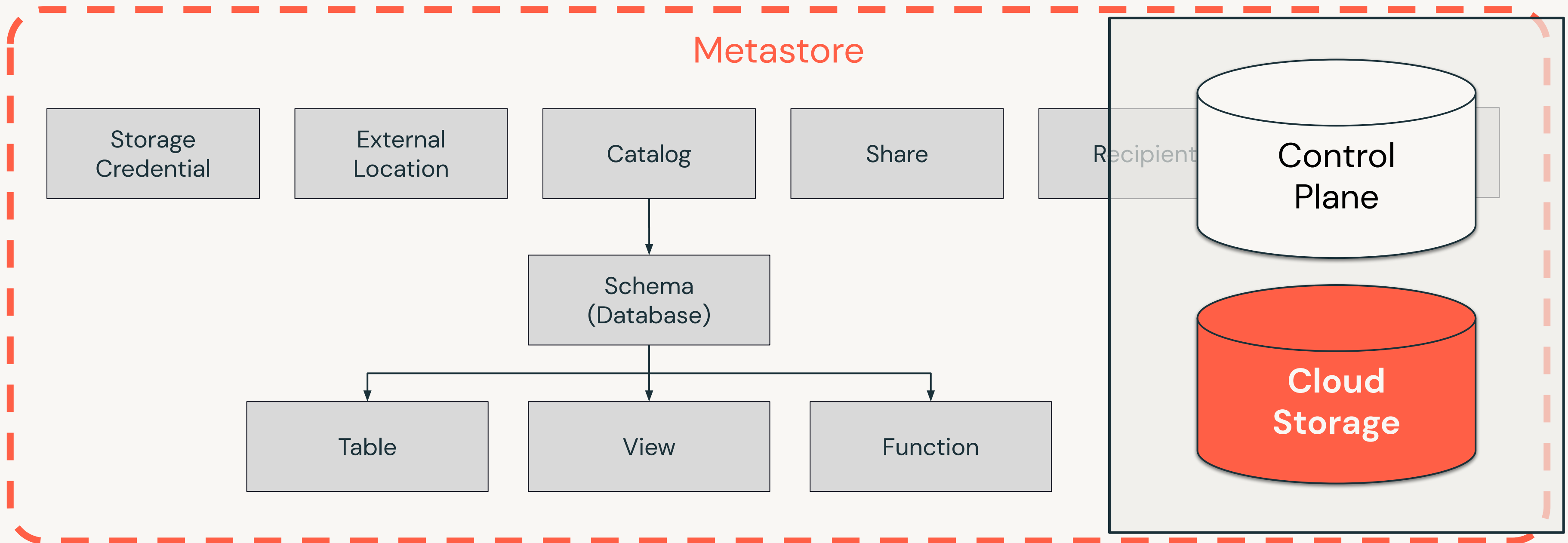
Key Concepts

Metastore storage



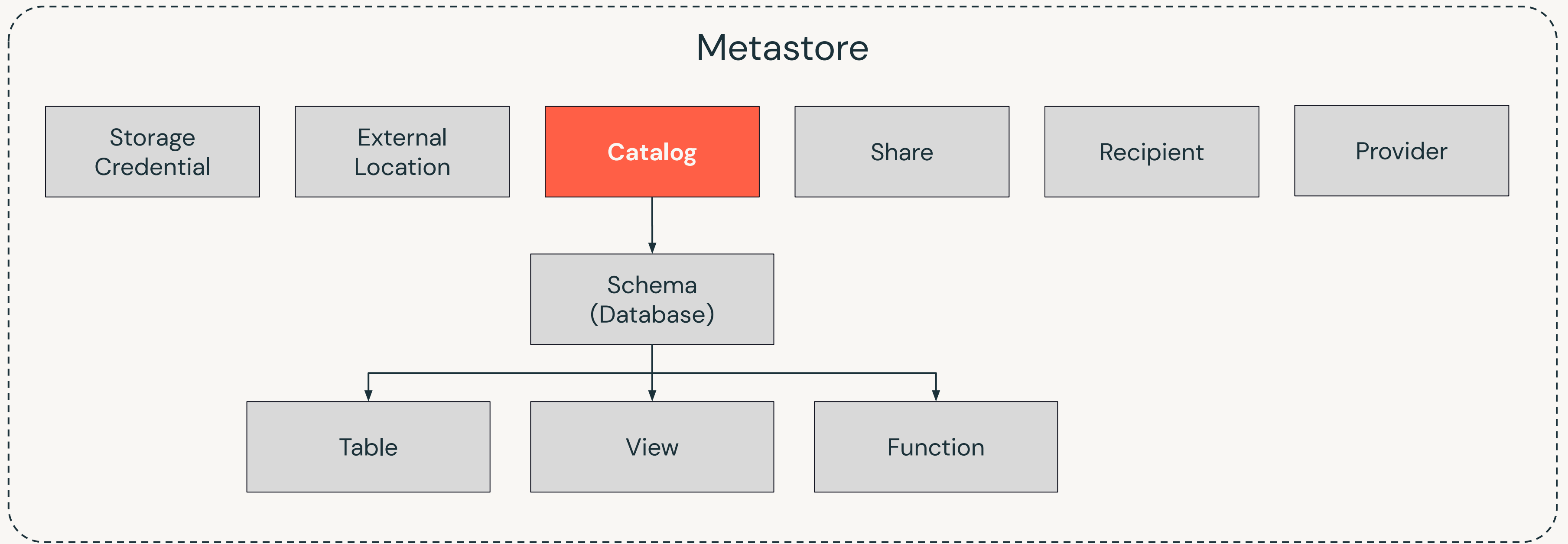
Key Concepts

Metastore storage



Key Concepts

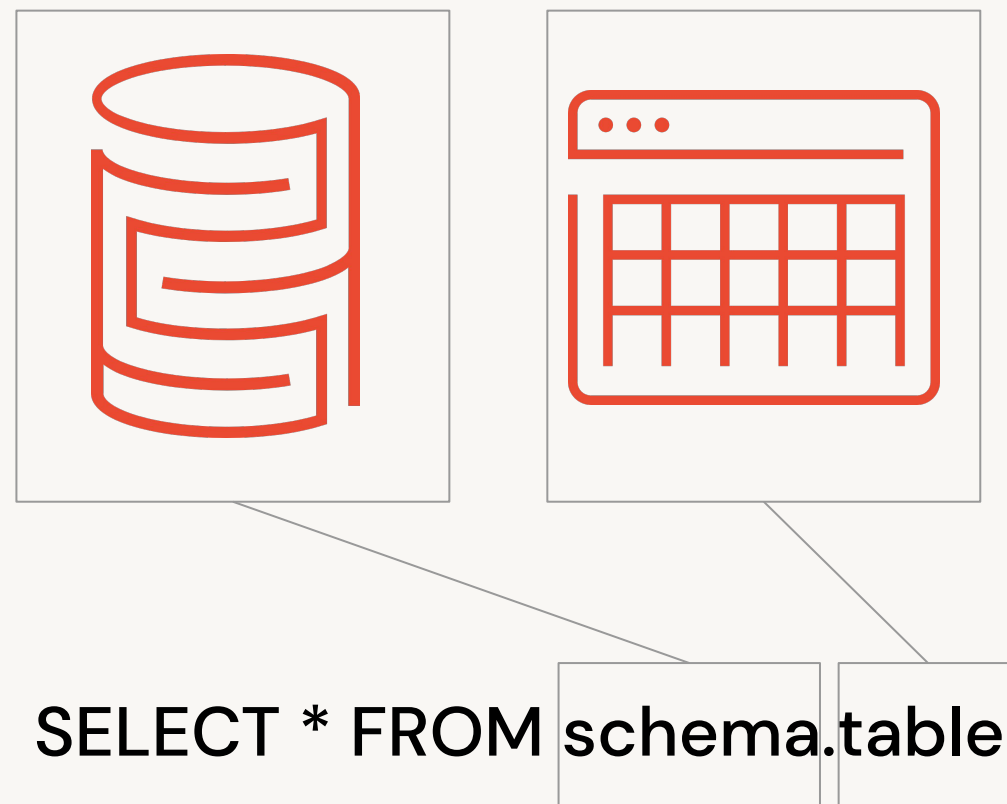
Catalog



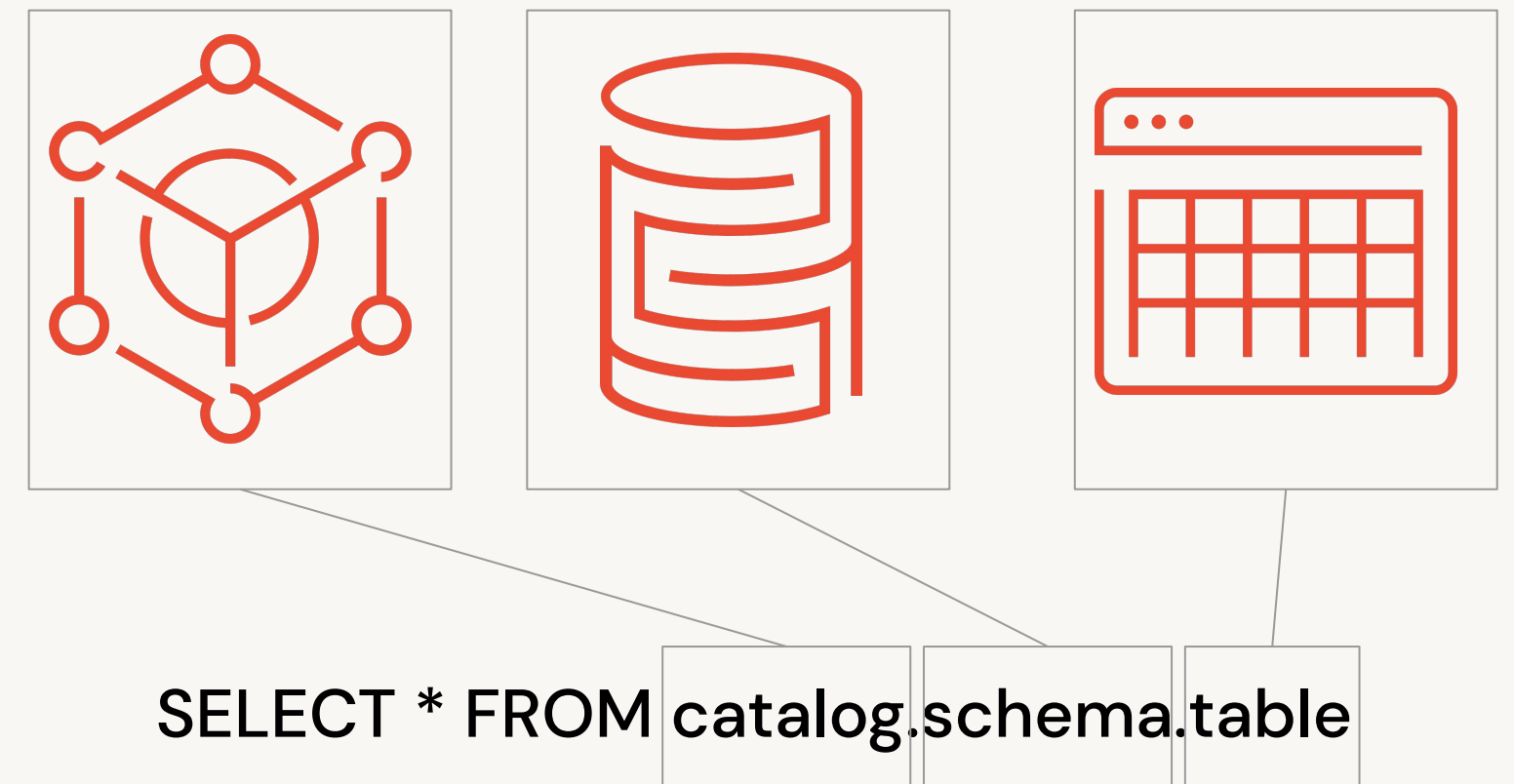
Key Concepts

Three-level namespace

Traditional SQL two-level namespace

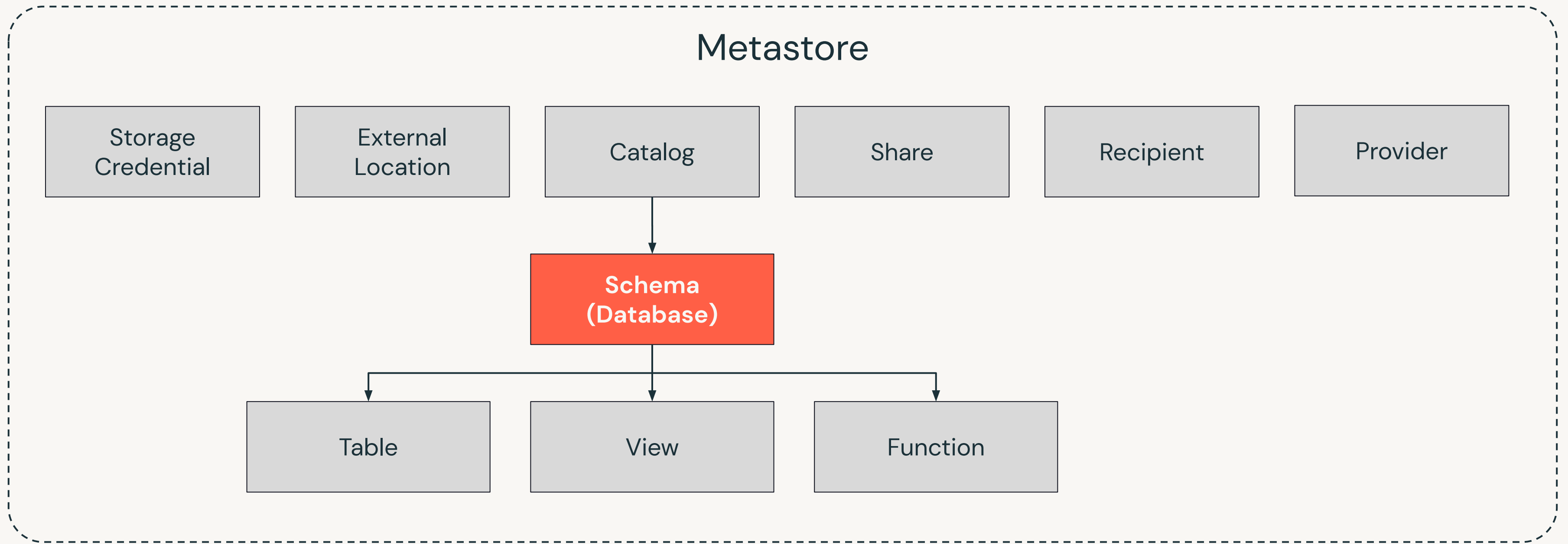


Unity Catalog three-level namespace



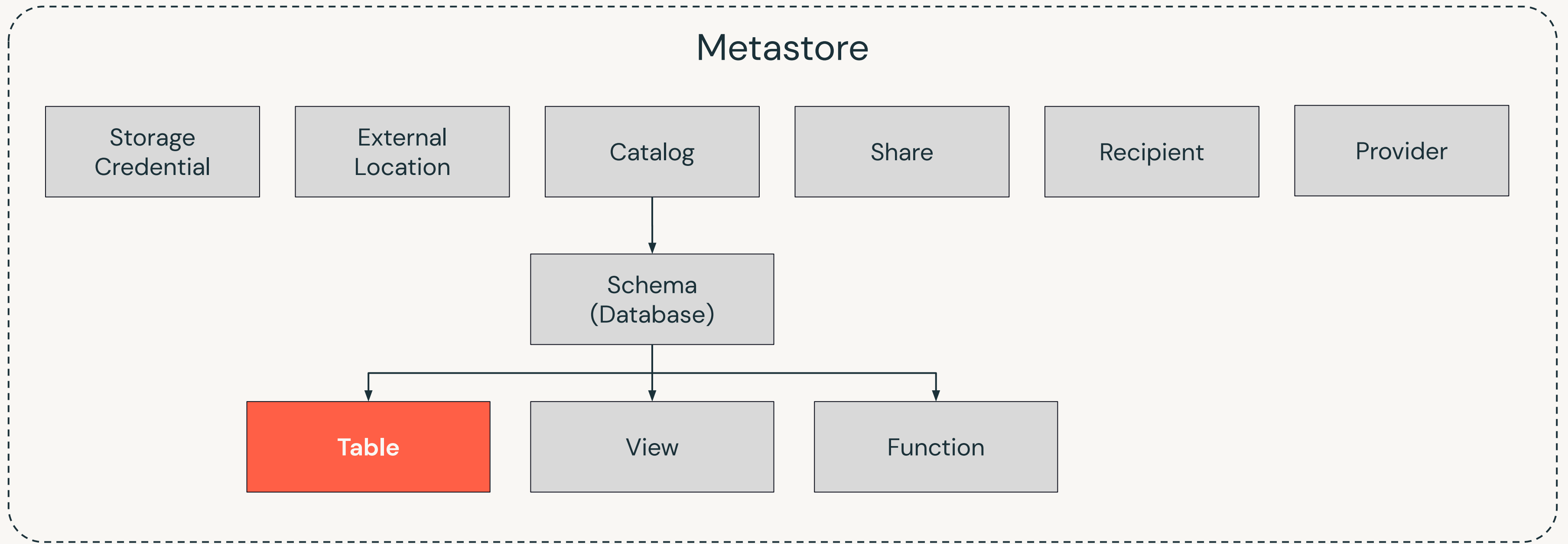
Key Concepts

Schema



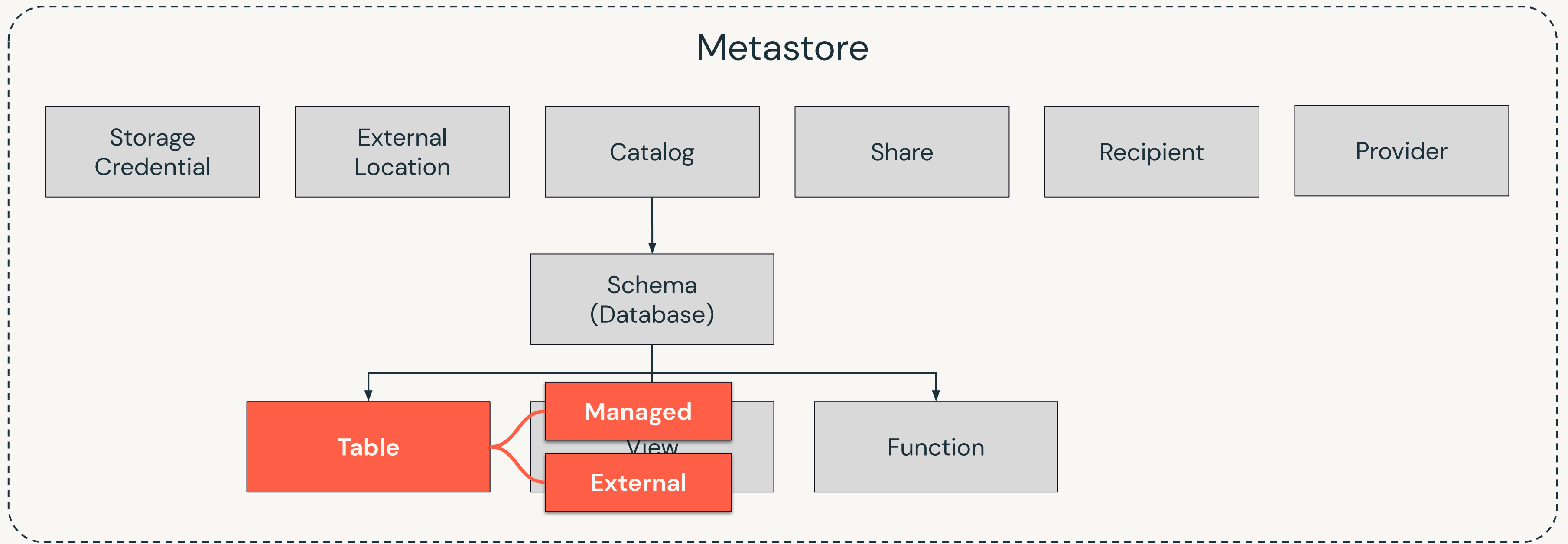
Key Concepts

Table



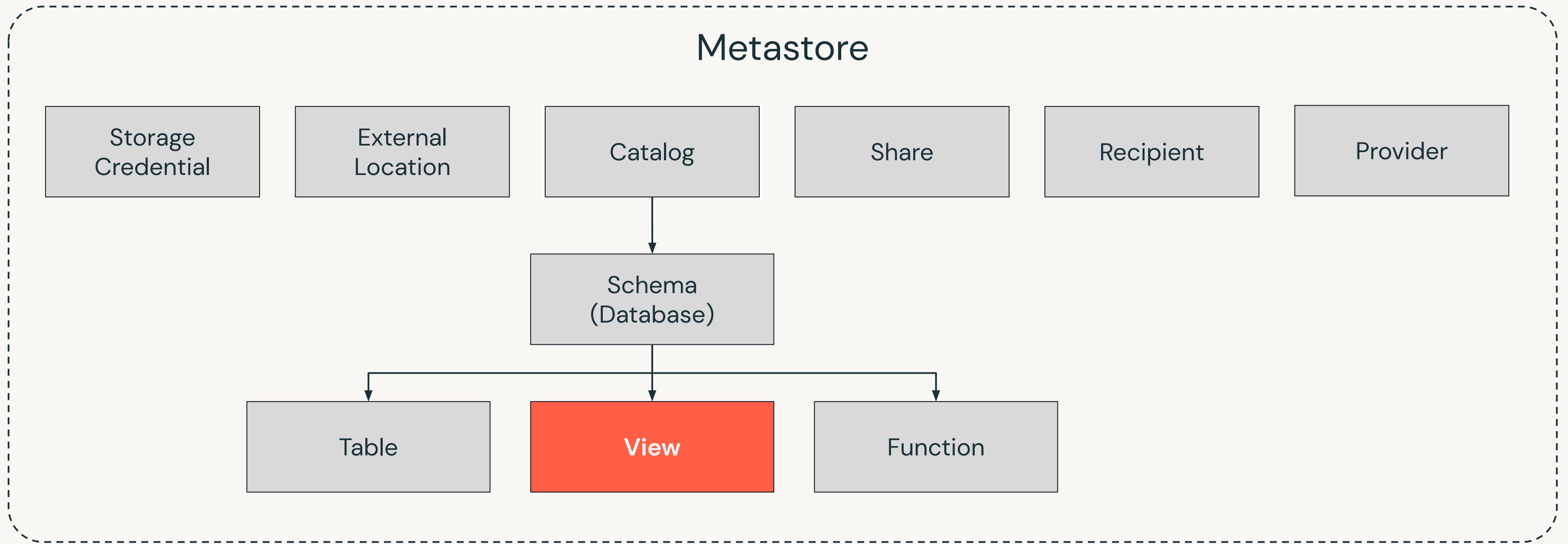
Key Concepts

Table



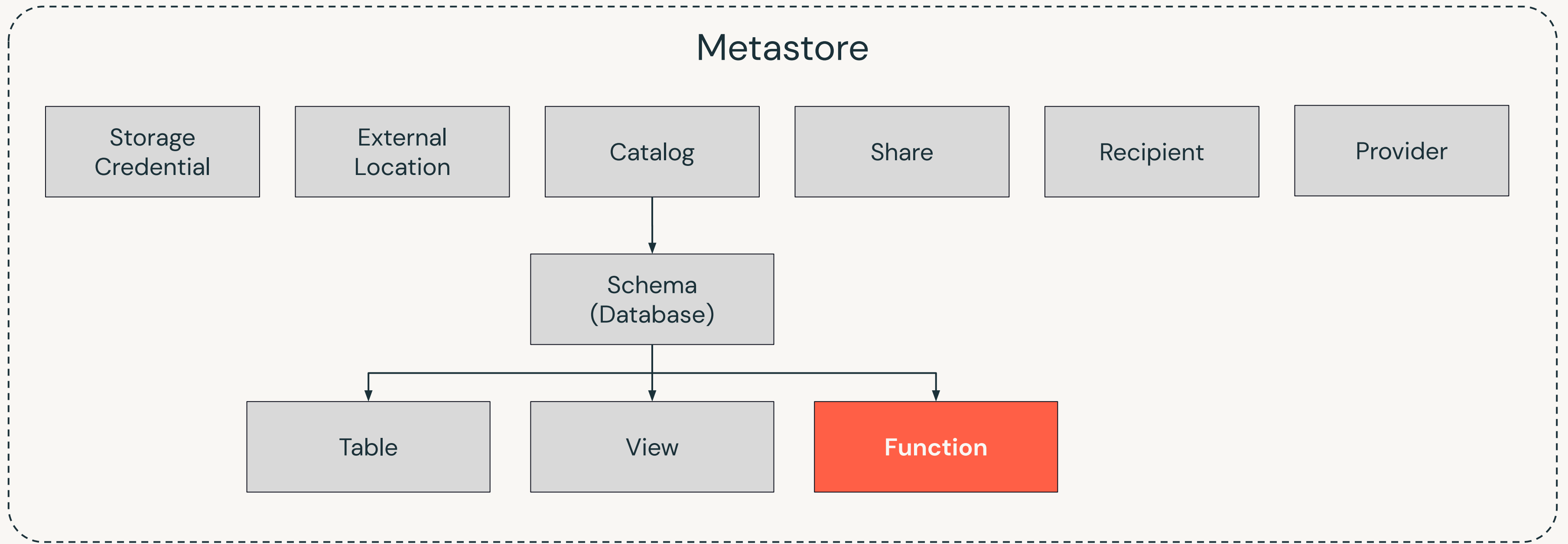
Key Concepts

View



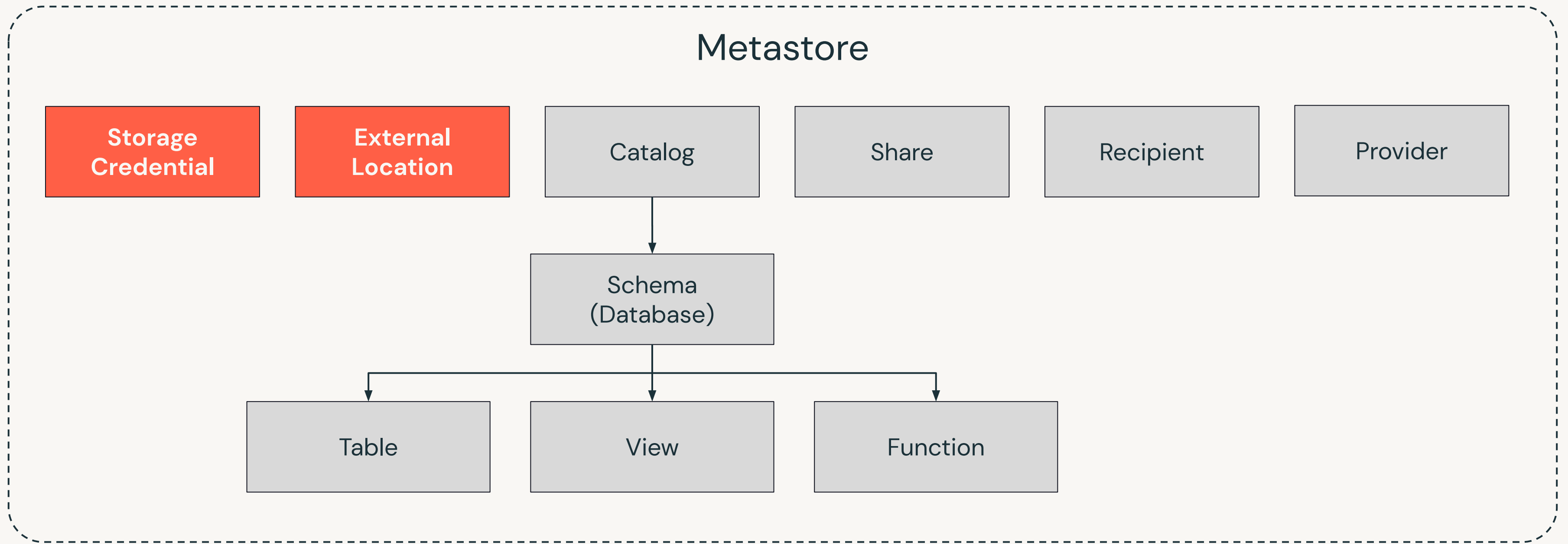
Key Concepts

User-defined function



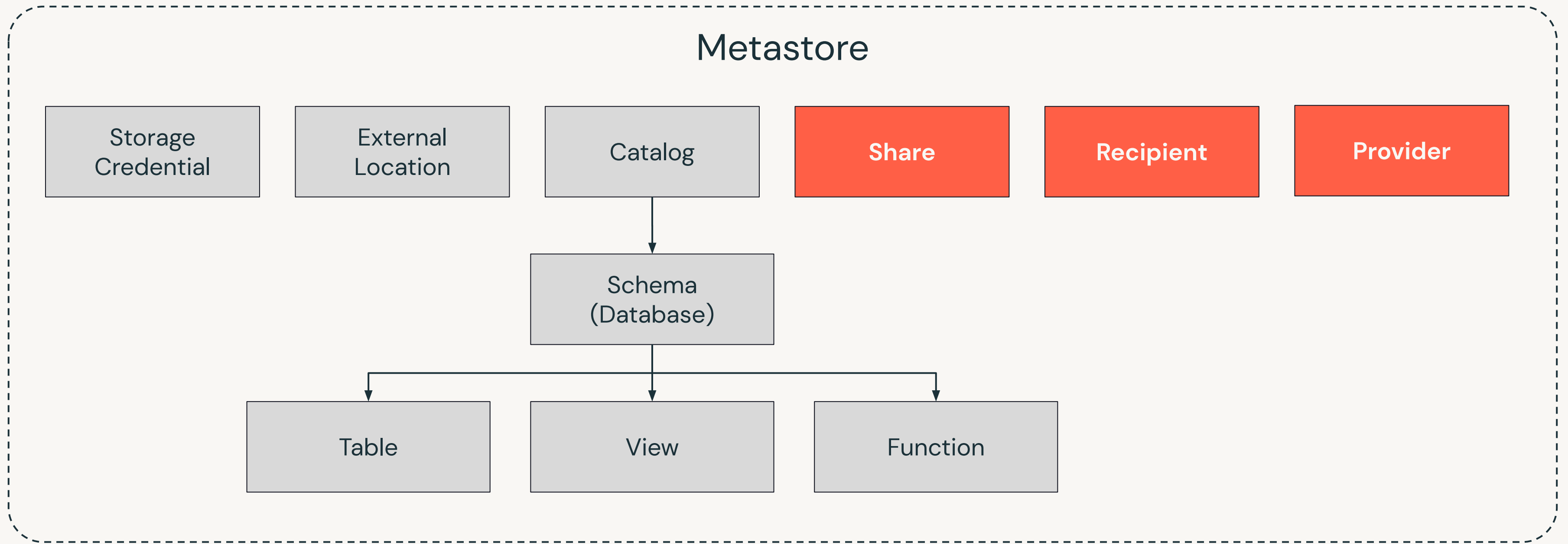
Key Concepts

Storage constructs



Key Concepts

Delta sharing



Lab: Populating the Metastore

Learning Objectives

In this lab, you will learn how to:

- Use the Unity Catalog three-level namespace
- Use SQL to create and explore data objects like catalogs, schemas and tables
- Use the Data explorer
- Access data lineage

Prerequisites

In order to do this lab, you will need:

- Metastore admin capability, or `CREATE CATALOG` privileges

Data Security Model

Data Security Model

Access control lists (ACLs)

Data Security Model

Access control lists (ACLs)

Table **t**

Data Security Model

Access control lists (ACLs)

Table `t`

Principal	Privilege
<code>analysts</code>	<code>SELECT</code>

Data Security Model

Access control lists (ACLs)

Table t

Principal	Privilege
analysts	SELECT
dbadmins	ALL PRIVILEGES

Data Security Model

Access control lists (ACLs)

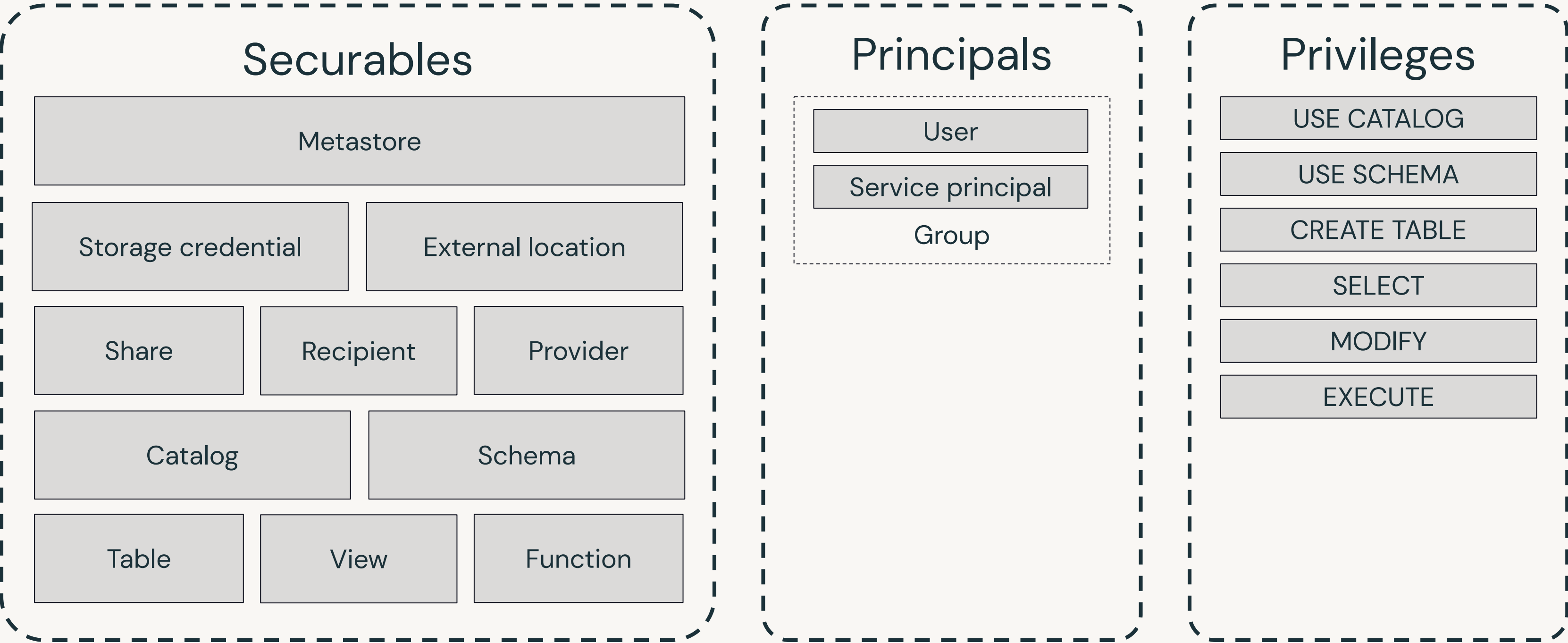
Table `t`

Owned by `jdope@company.com`

Principal	Privilege
analysts	SELECT
dbadmins	ALL PRIVILEGES

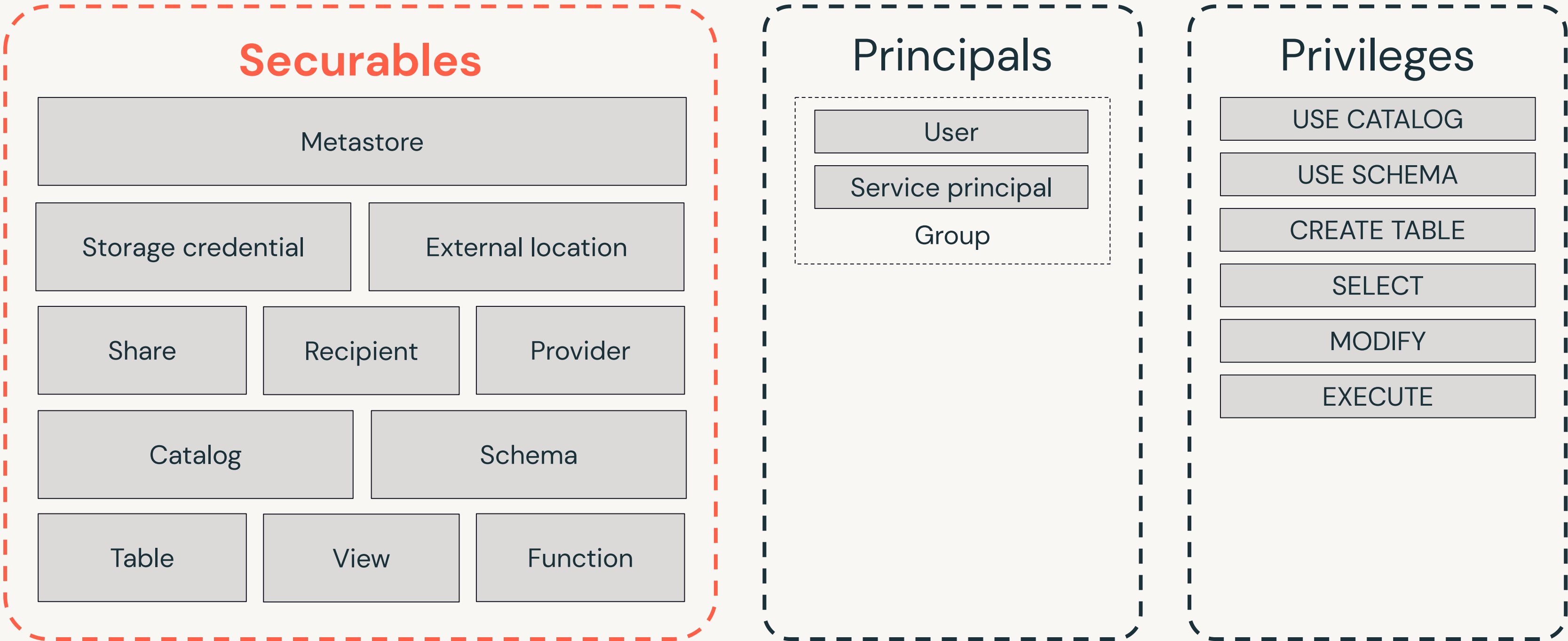
Data Security Model

Access control lists (ACLs)



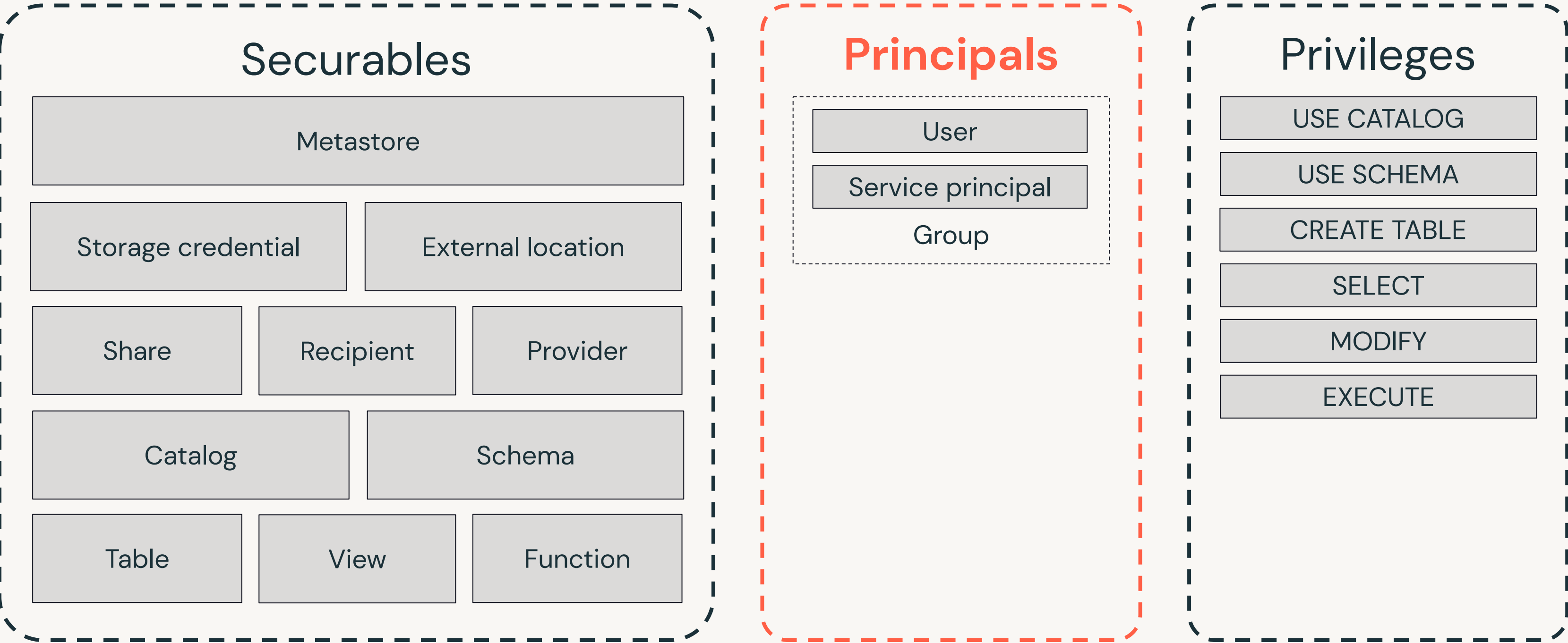
Data Security Model

Access control lists (ACLs)



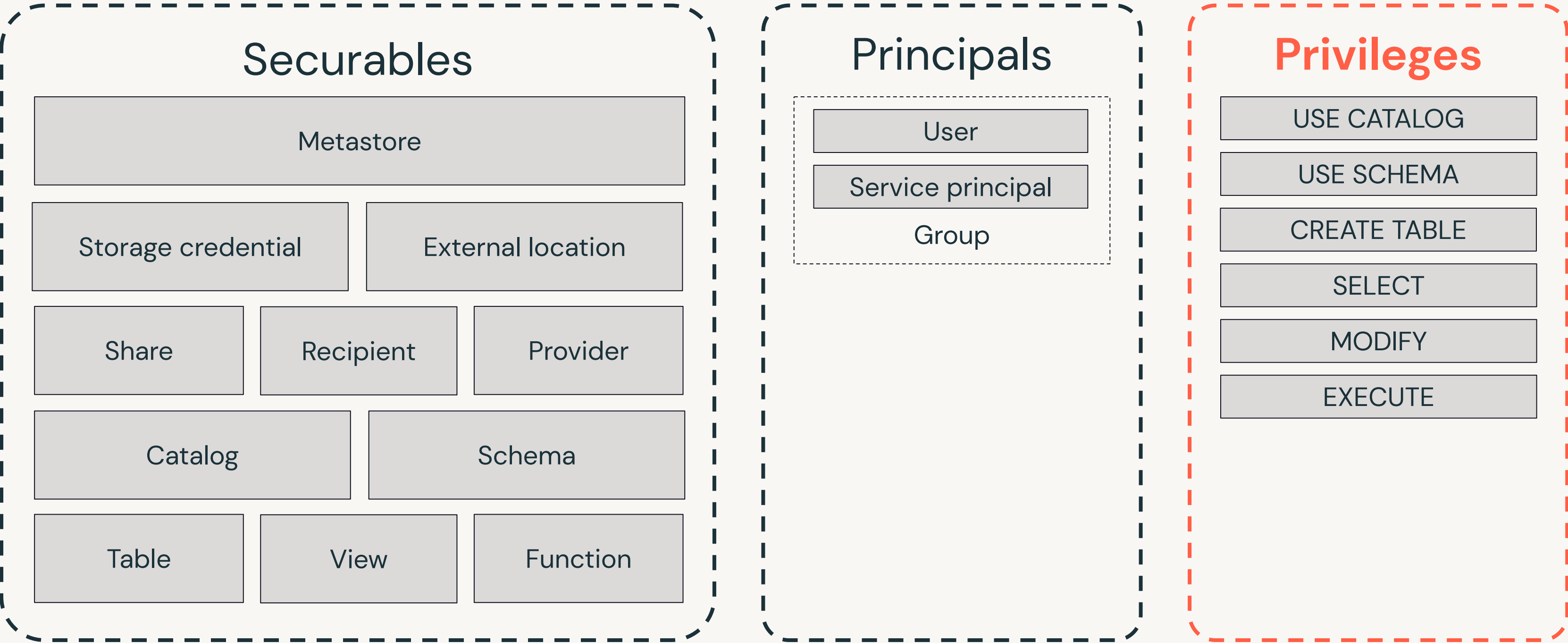
Data Security Model

Access control lists (ACLs)



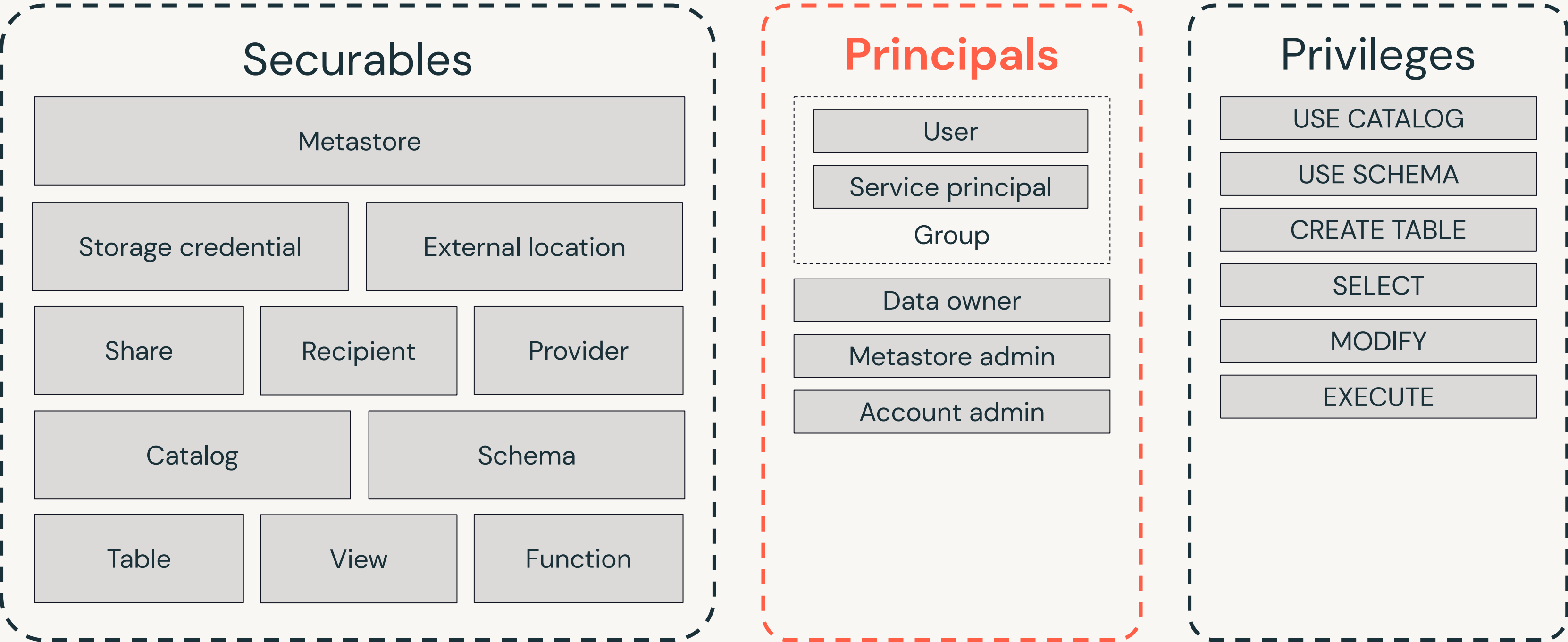
Data Security Model

Access control lists (ACLs)



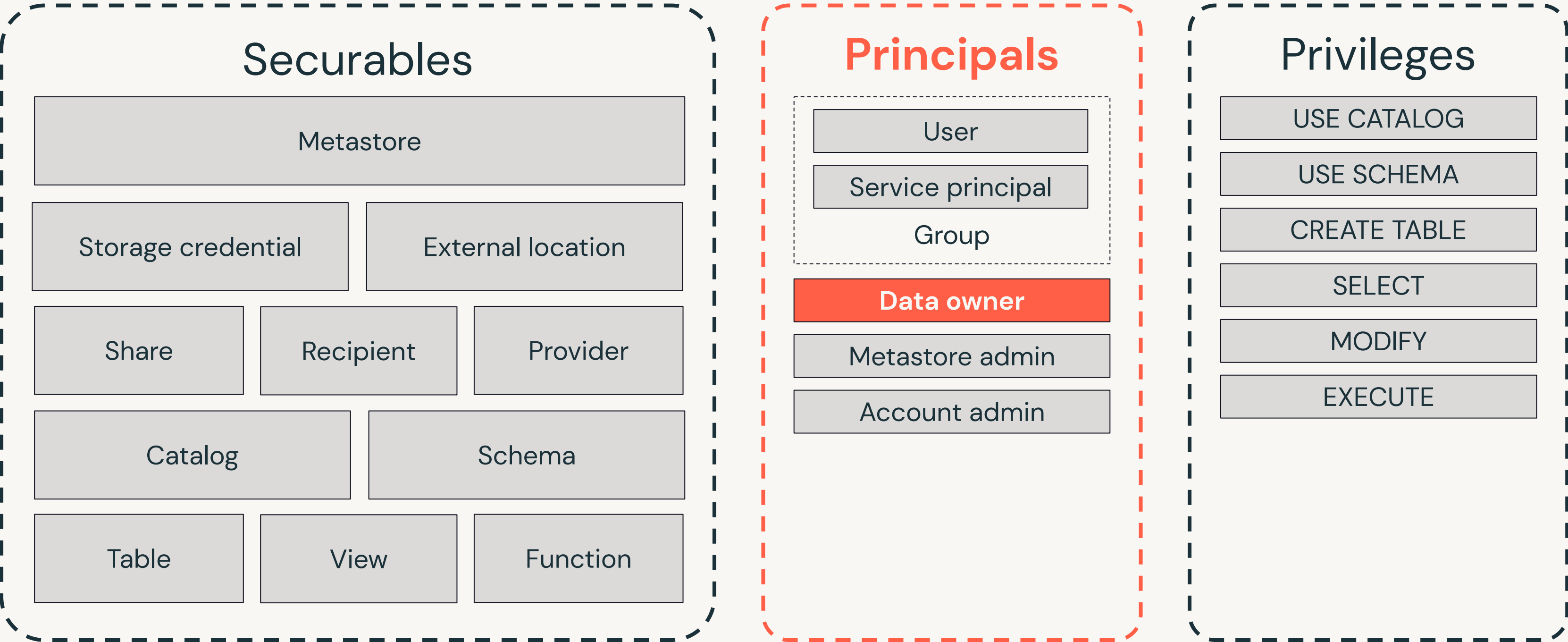
Data Security Model

Access control lists (ACLs)



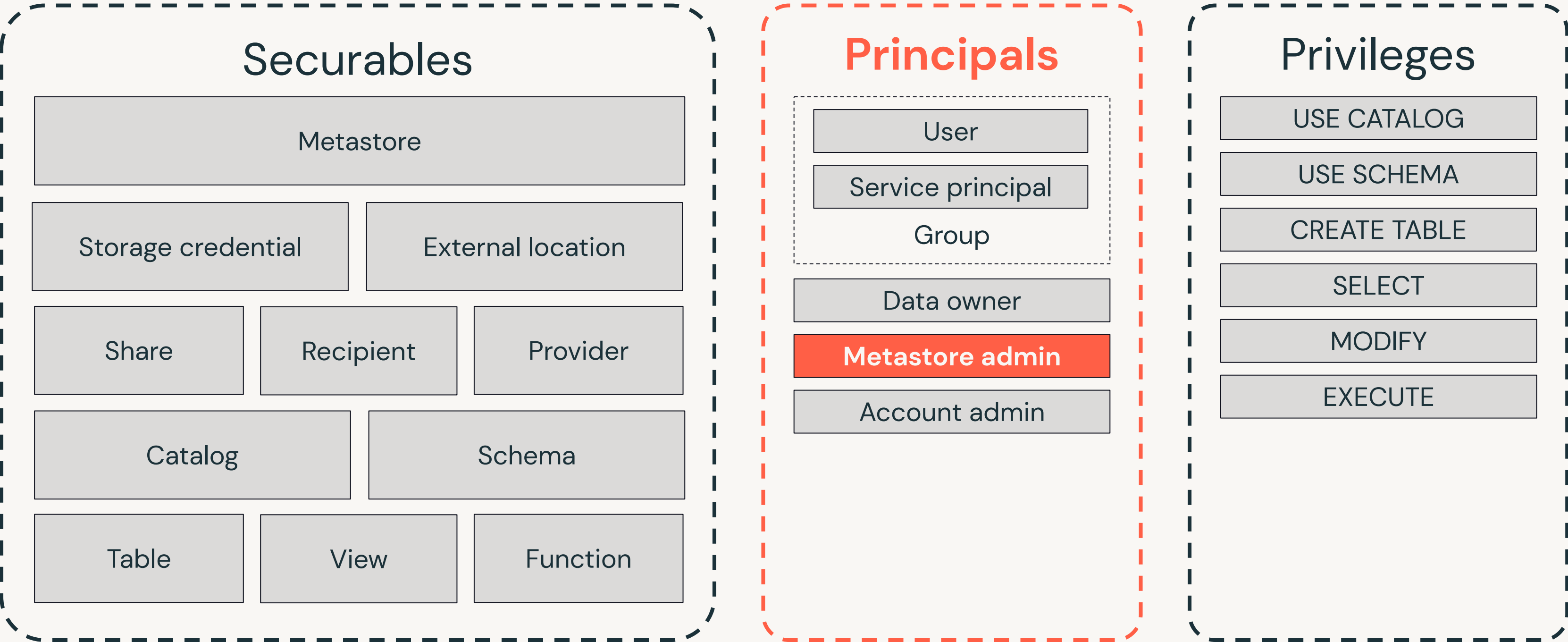
Data Security Model

Access control lists (ACLs)



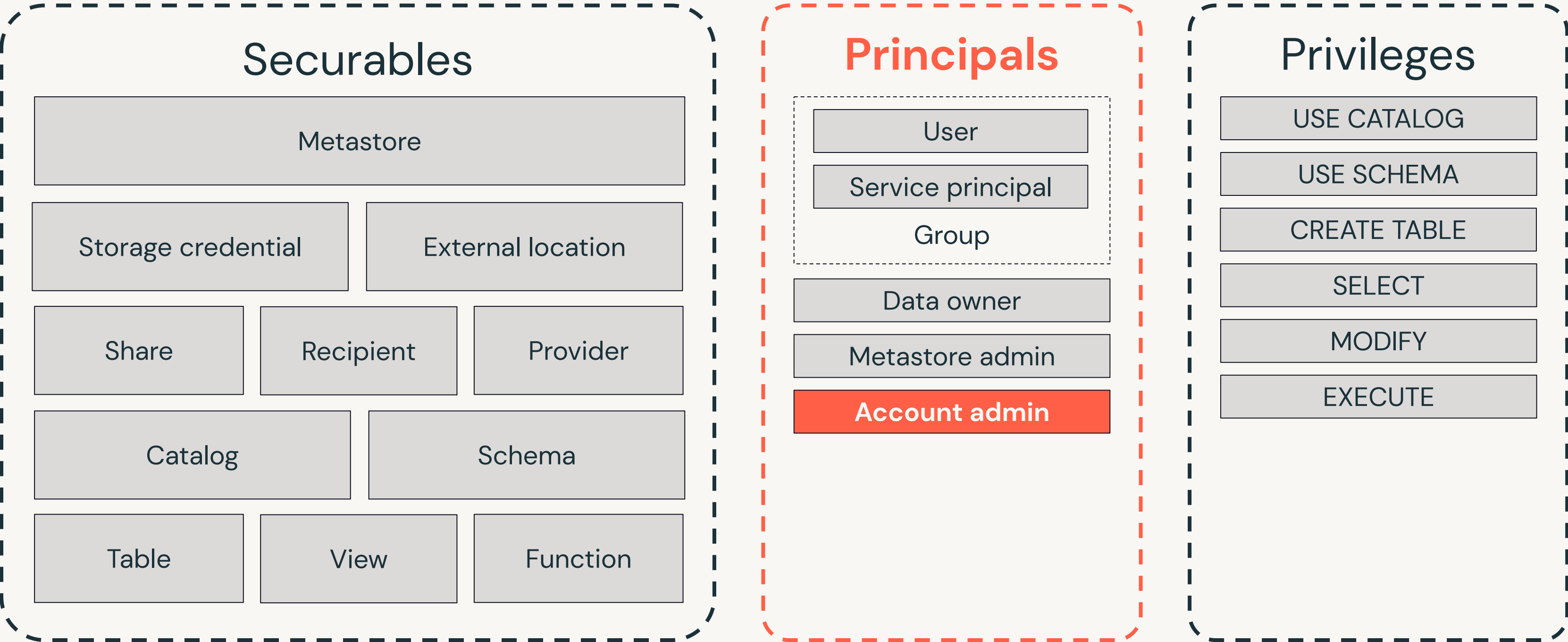
Data Security Model

Access control lists (ACLs)



Data Security Model

Access control lists (ACLs)



Data Security Model

Managing ACLs

Data Security Model

Managing ACLs

1. SQL

GRANT SELECT ON TABLE t TO analysts

REVOKE SELECT ON TABLE t FROM analysts

Data Security Model

Managing ACLs

1. SQL

GRANT SELECT ON TABLE t TO analysts
REVOKE SELECT ON TABLE t FROM analysts



Notebook

Databricks SQL

Data Security Model

Managing ACLs

1. SQL

GRANT SELECT ON TABLE t TO analysts

REVOKE SELECT ON TABLE t FROM analysts

2. Data explorer

Data Science & Engineering workspace or DBSQL

Data Security Model

Managing ACLs

1. SQL

GRANT SELECT ON TABLE t TO analysts

REVOKE SELECT ON TABLE t FROM analysts

2. Data explorer

Data Science & Engineering workspace or DBSQL

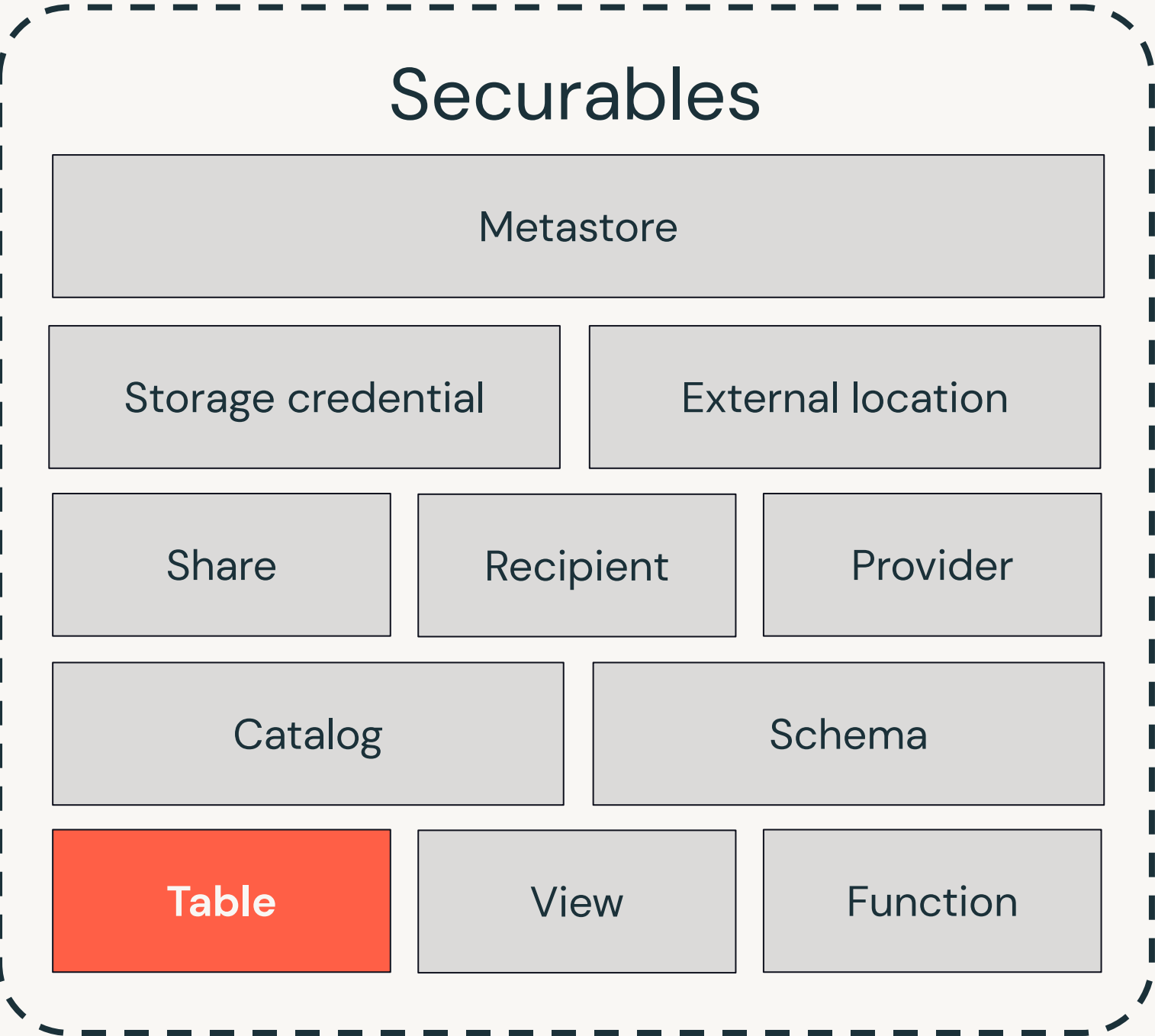
3. Programmatically

Databricks CLI, Terraform, and REST APIs

Privilege Types

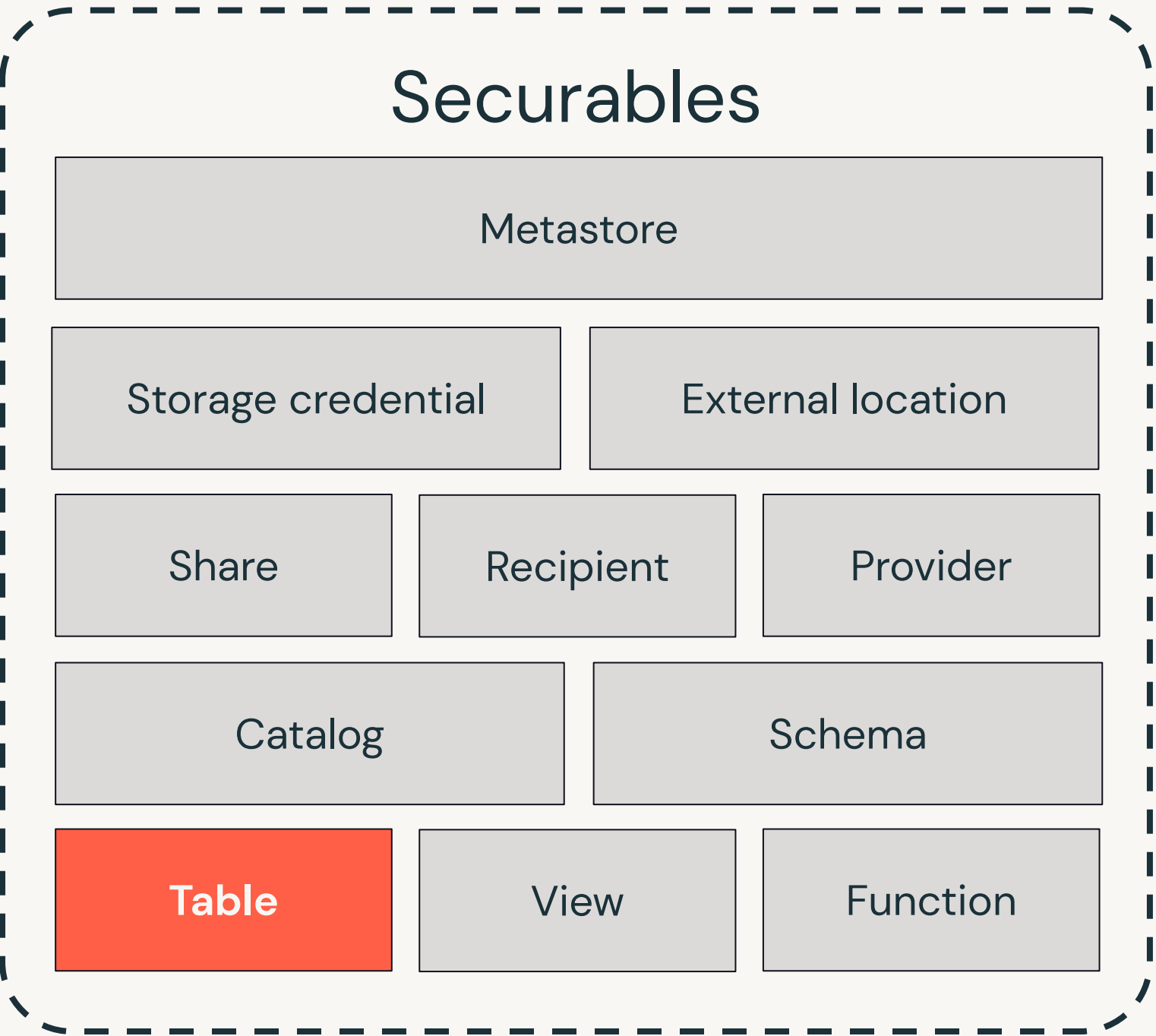
Privilege Types

Table



Privilege Types

Table

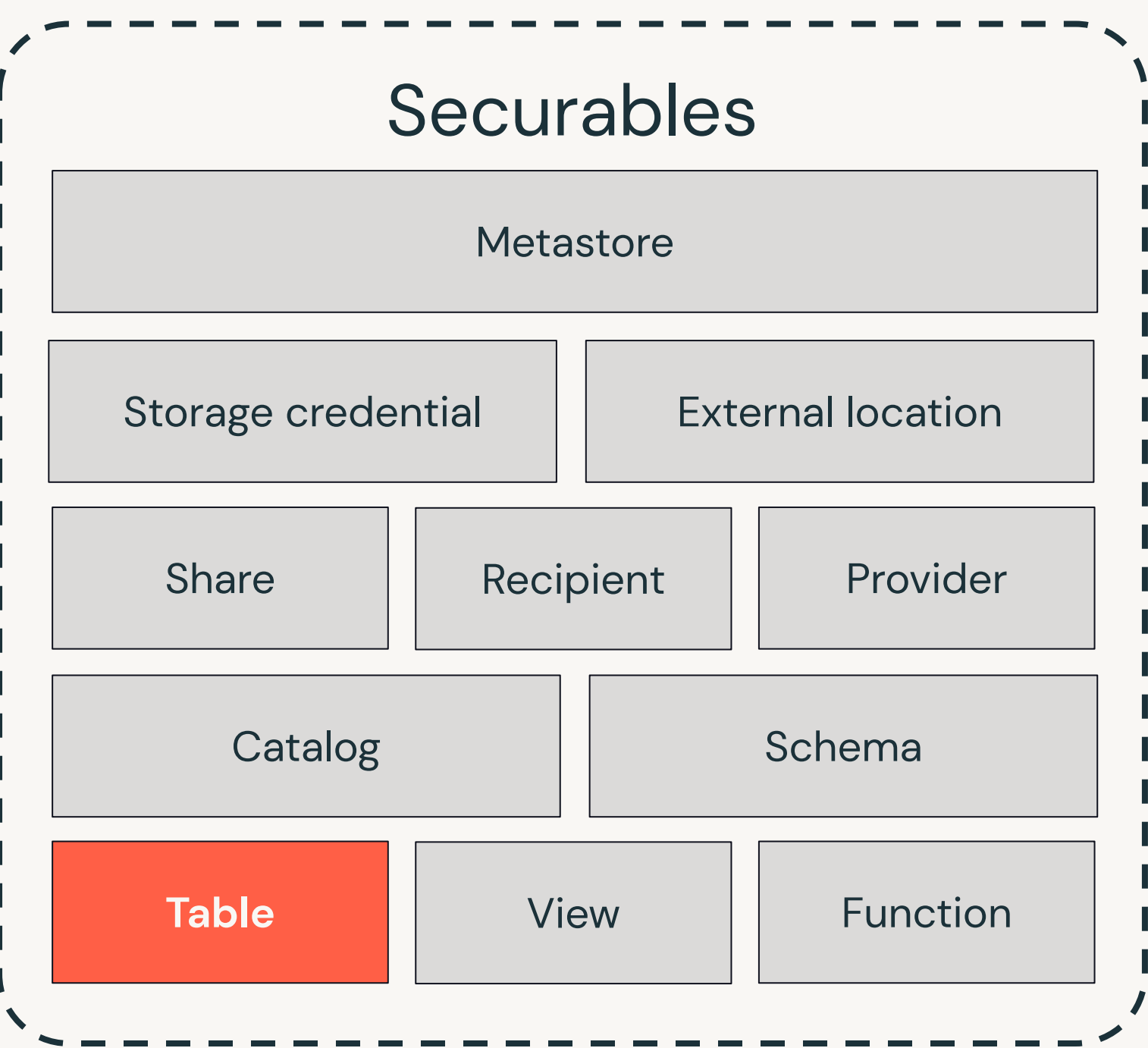


Privileges

SELECT
Read from a table (SELECT)

Privilege Types

Table



Privileges

SELECT

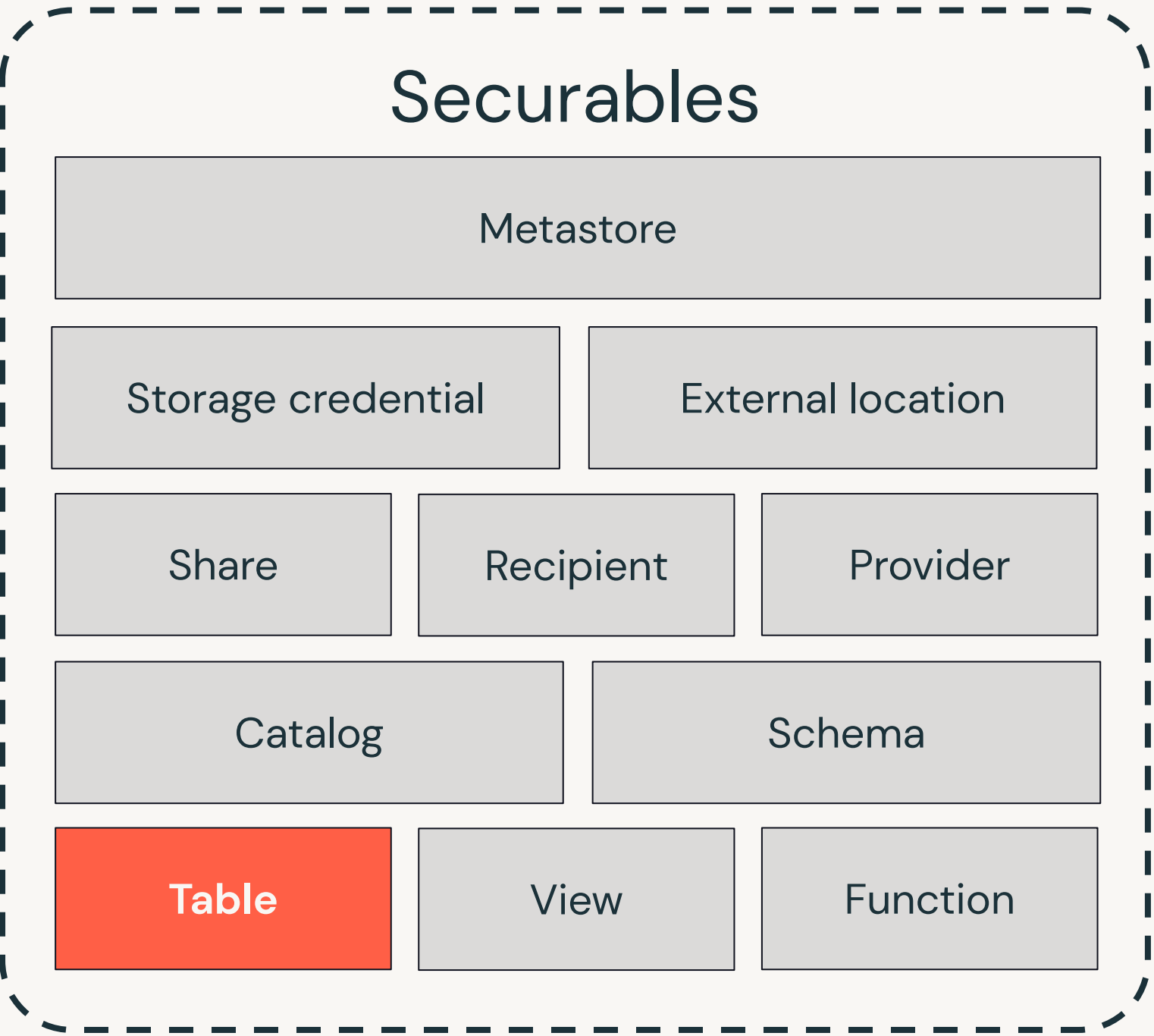
Read from a table (SELECT)

MODIFY

Modify table data (UPDATE) or metadata (ALTER)

Privilege Types

Table



Privileges

SELECT

Read from a table (SELECT)

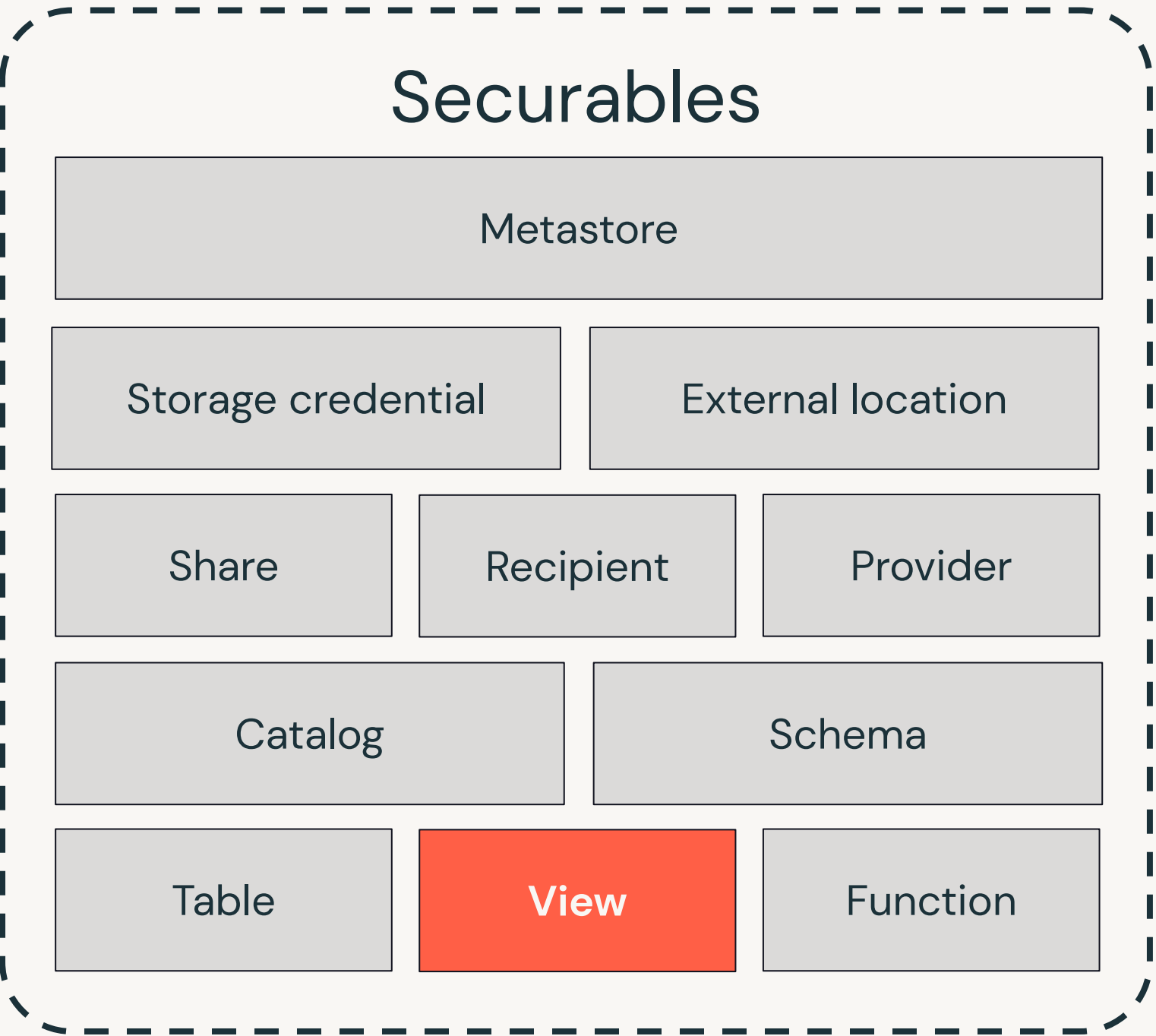
MODIFY

Modify table data (UPDATE) or metadata (ALTER)

ALL PRIVILEGES

Privilege Types

View



Privileges

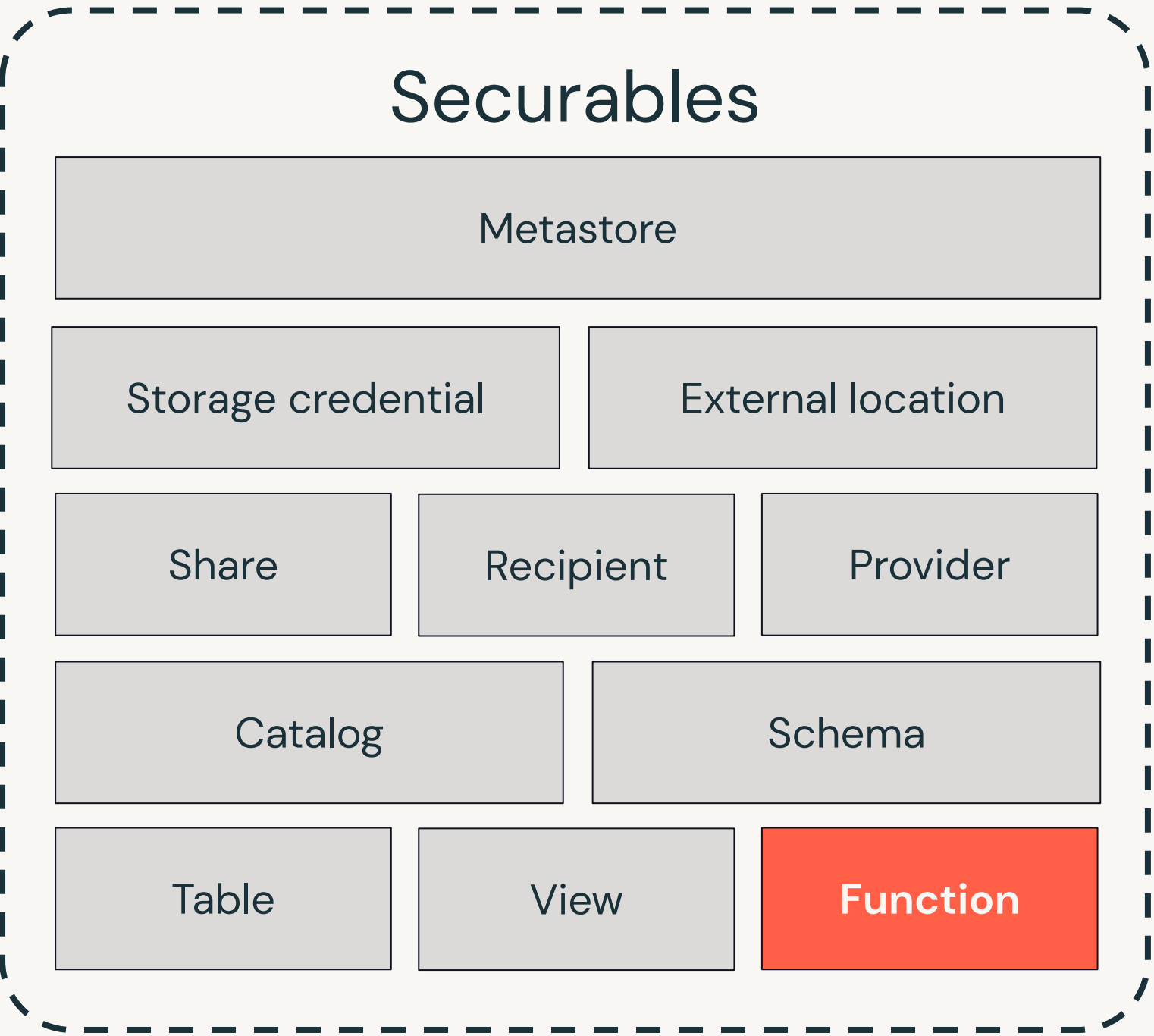
SELECT

Query view (SELECT)

ALL PRIVILEGES

Privilege Types

Function



Privileges

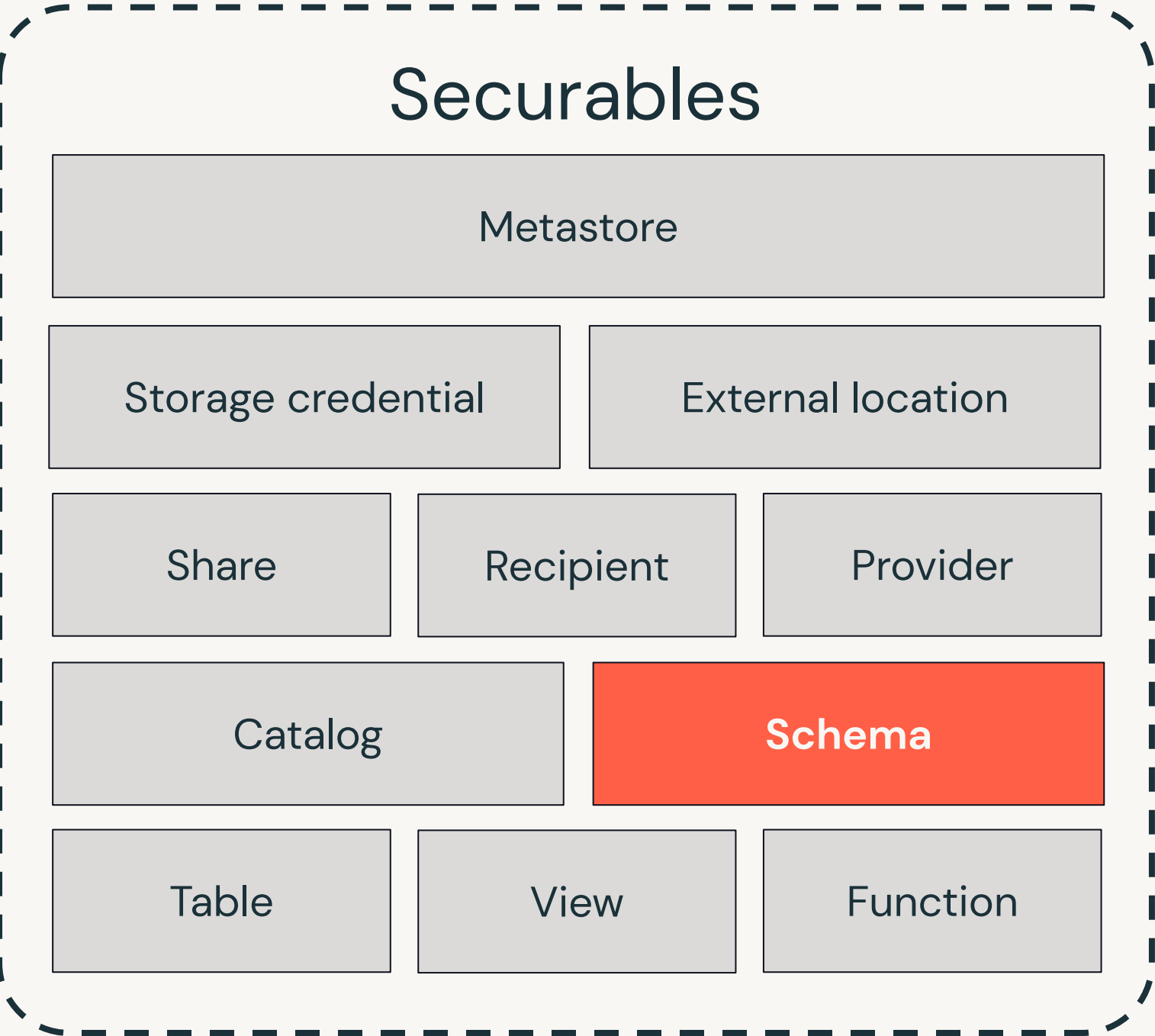
EXECUTE

Use function in a query

ALL PRIVILEGES

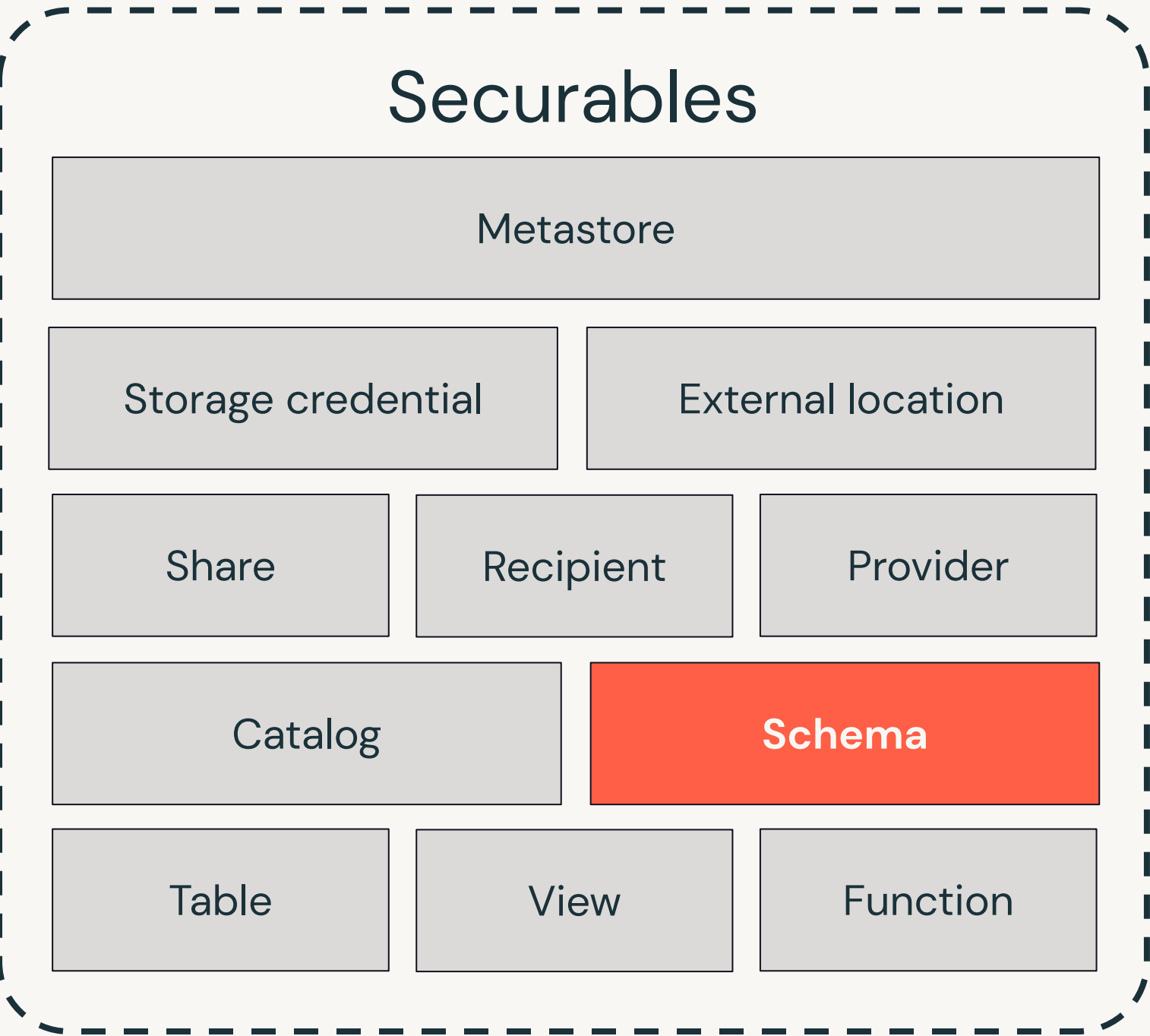
Privilege Types

Schema



Privilege Types

Schema

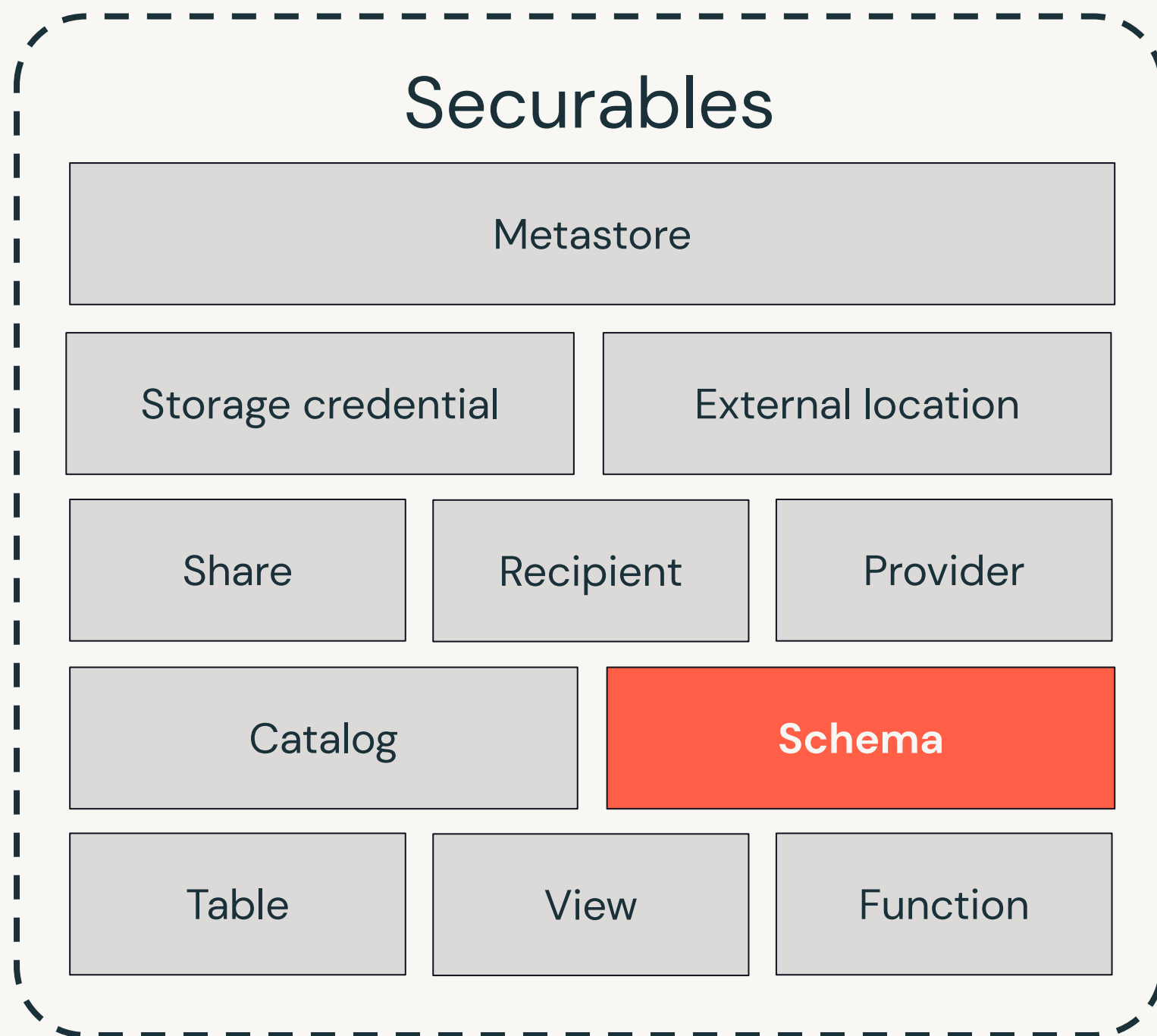


Privileges

USE SCHEMA
Access objects in schema

Privilege Types

Schema



Privileges

USE SCHEMA

Access objects in schema

CREATE TABLE

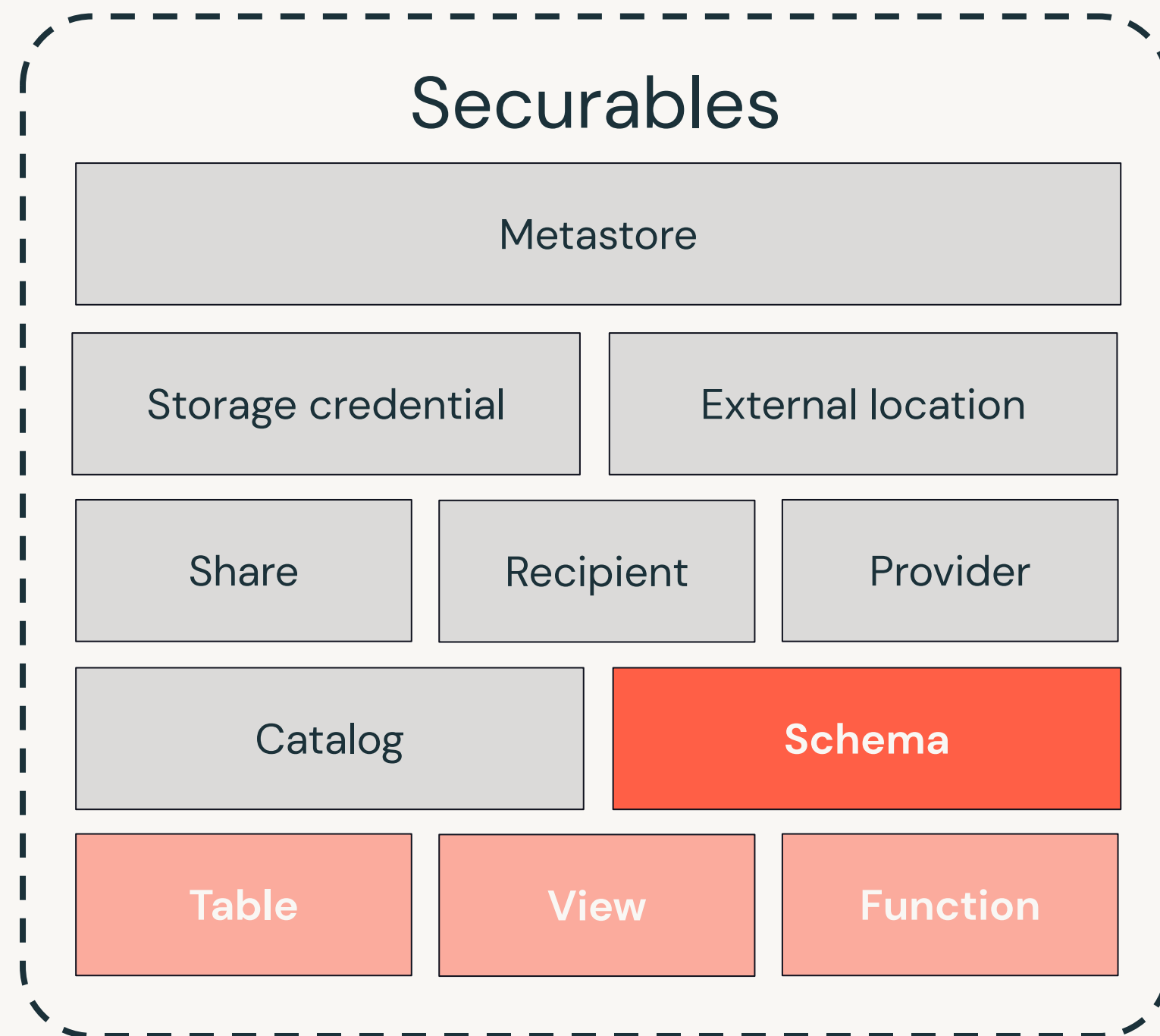
Create table or view in schema

CREATE FUNCTION

Create user-defined function in schema

Privilege Types

Schema privilege inheritance



Bequeathed privileges

SELECT

Read from any table or view in schema

MODIFY

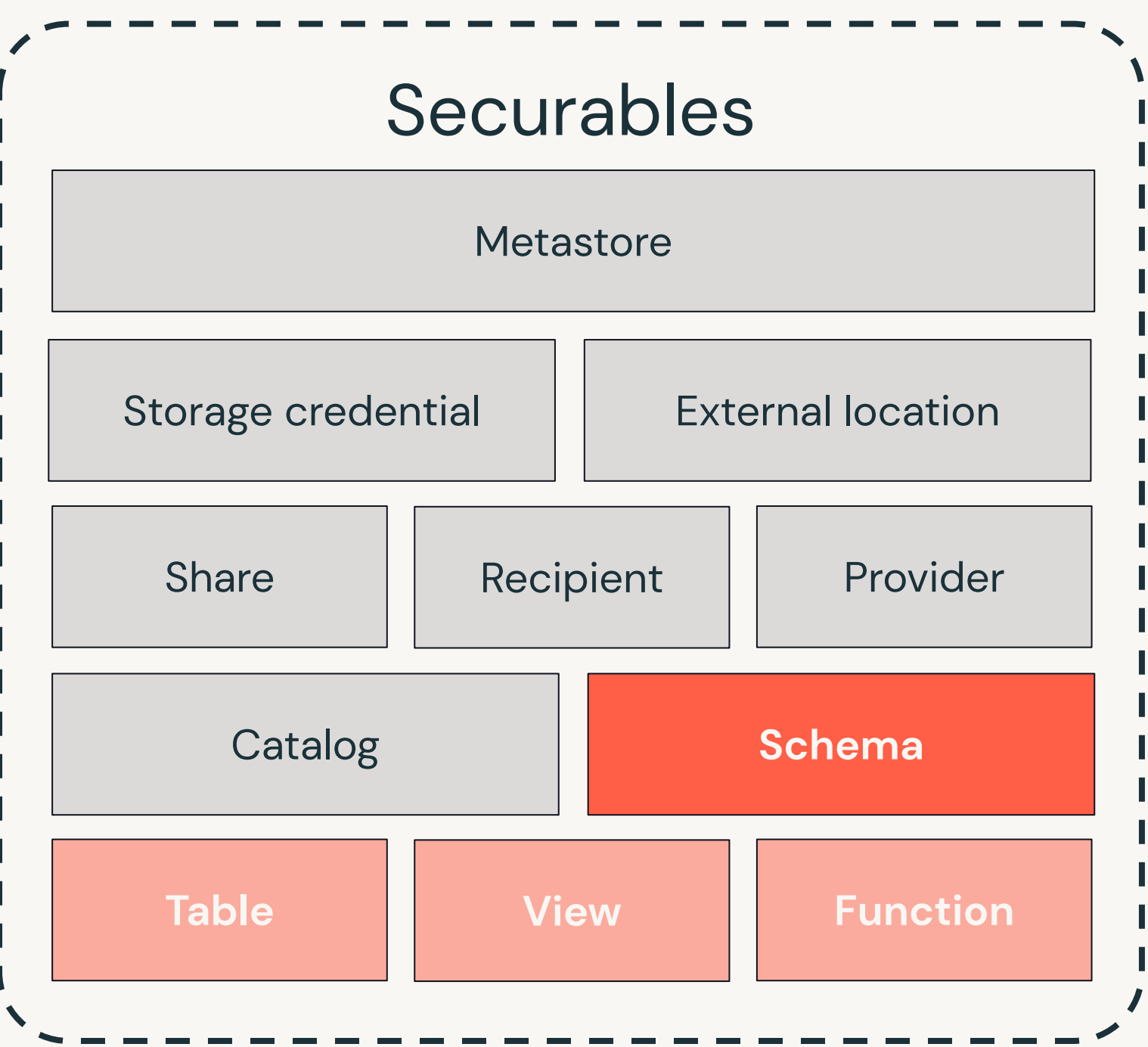
Modify data or metadata of any table in schema

EXECUTE

Use any function in schema

Privilege Types

Schema



ALL PRIVILEGES

USE SCHEMA

CREATE TABLE

CREATE FUNCTION

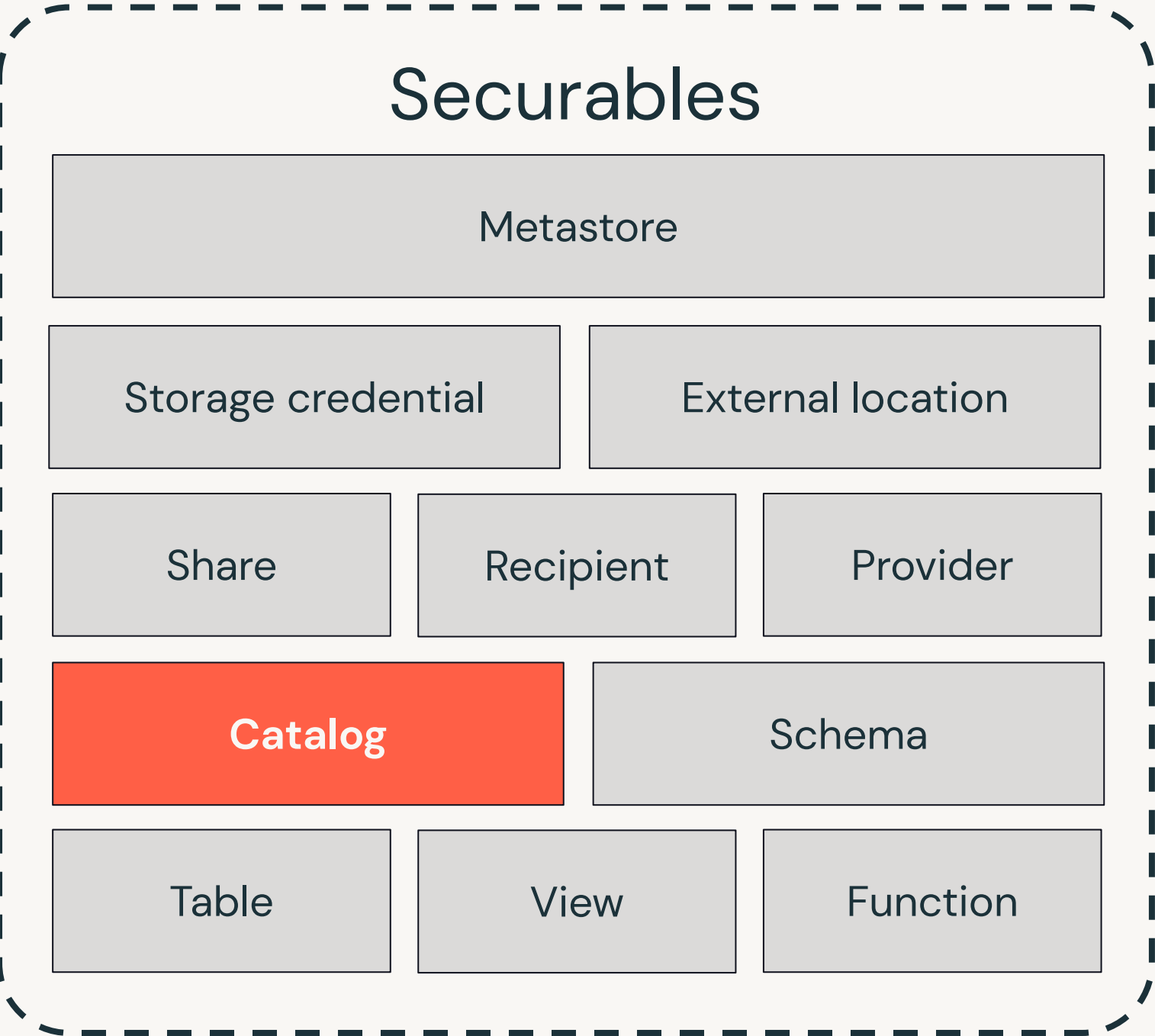
SELECT

MODIFY

EXECUTE

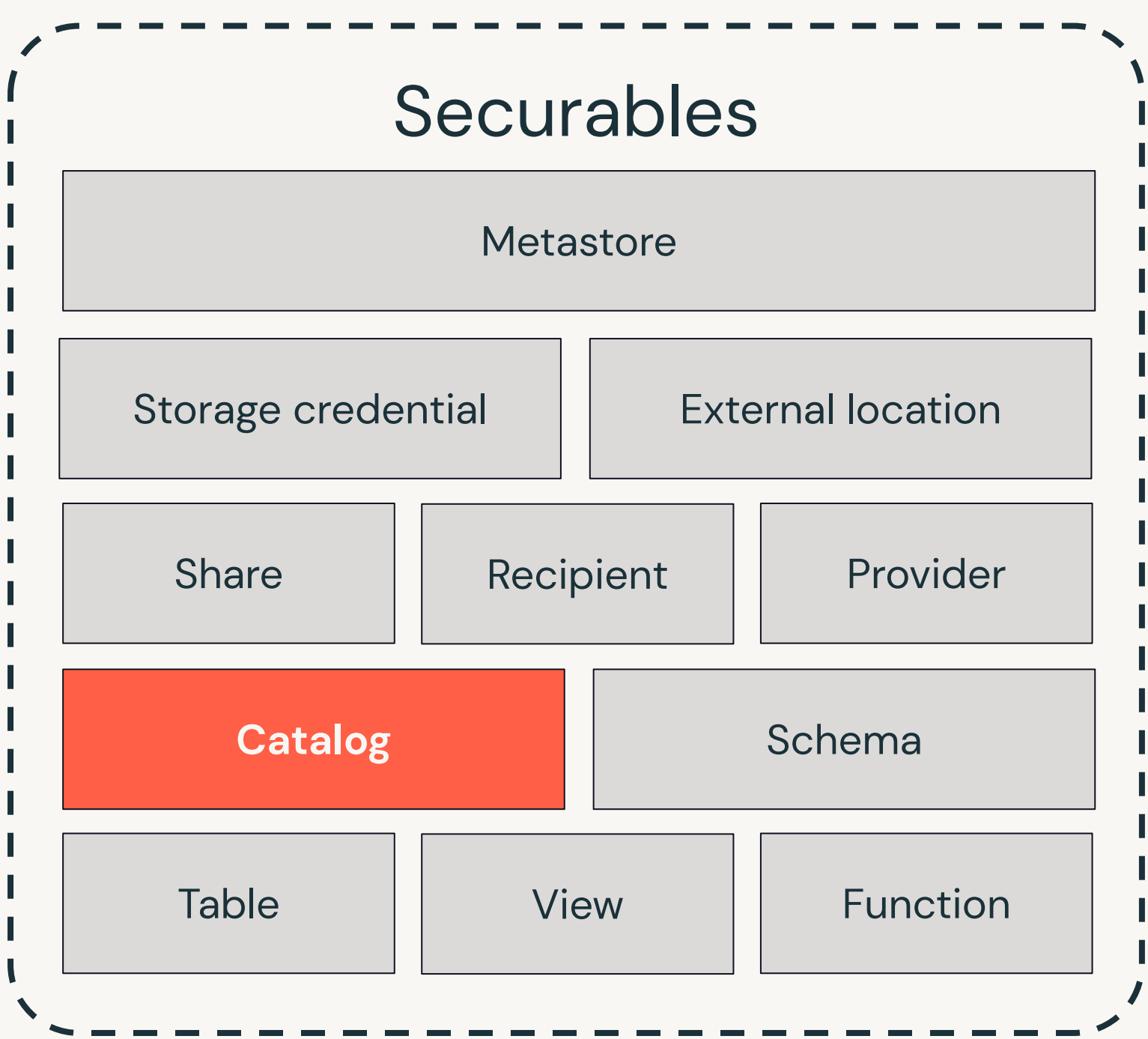
Privilege Types

Catalog



Privilege Types

Catalog

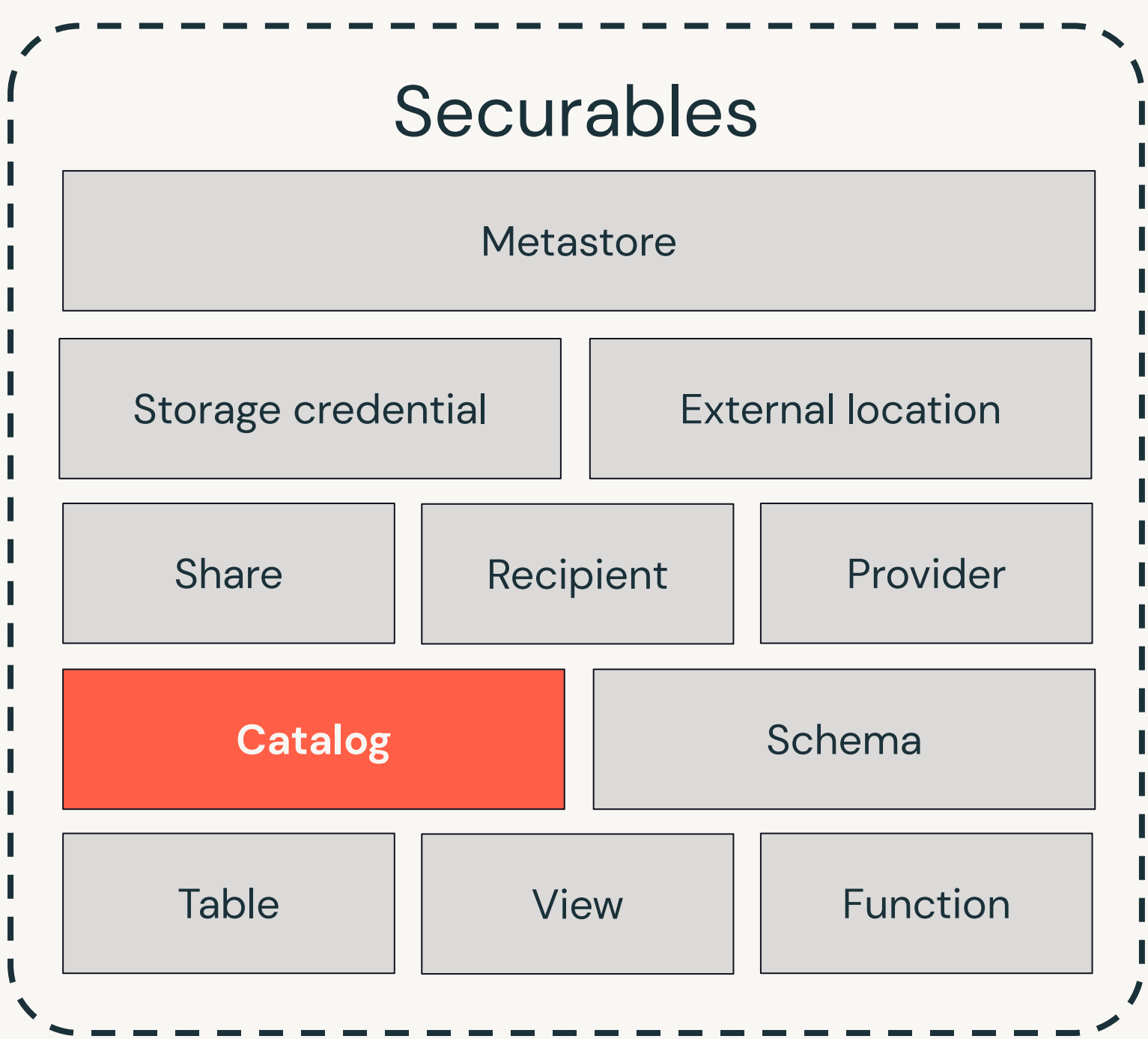


Privileges

USE CATALOG
Access schemas in catalog

Privilege Types

Catalog



Privileges

USE CATALOG

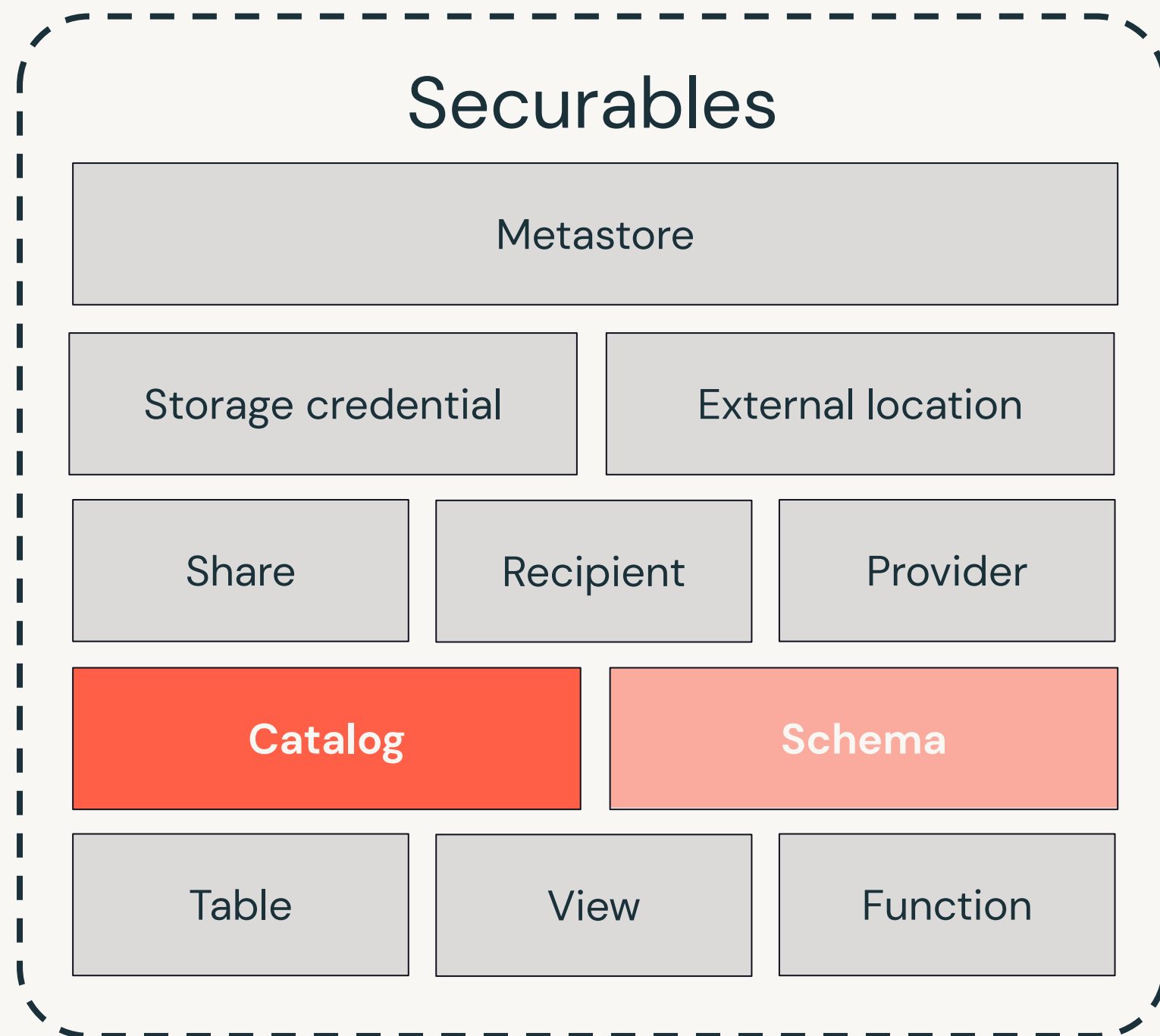
Access schemas in catalog

CREATE SCHEMA

Create schema in catalog

Privilege Types

Catalog privilege inheritance



Bequeathed privileges

CREATE TABLE

Create table or view in any schema

CREATE FUNCTION

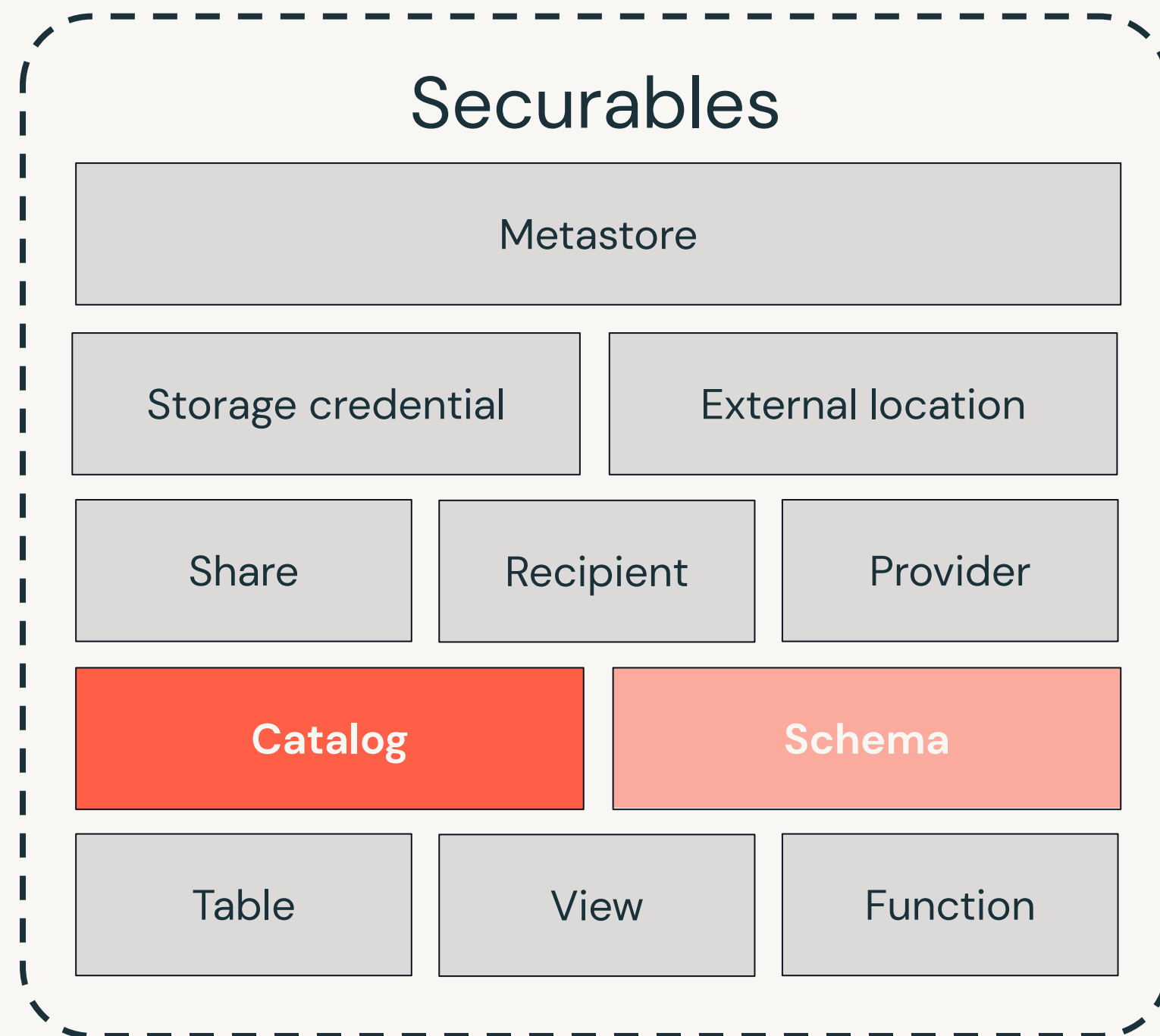
Create user-defined function in any schema

USE SCHEMA

Access objects in any schema

Privilege Types

Catalog privilege inheritance



Bequeathed privileges

CREATE TABLE

Create table or view in any schema

CREATE FUNCTION

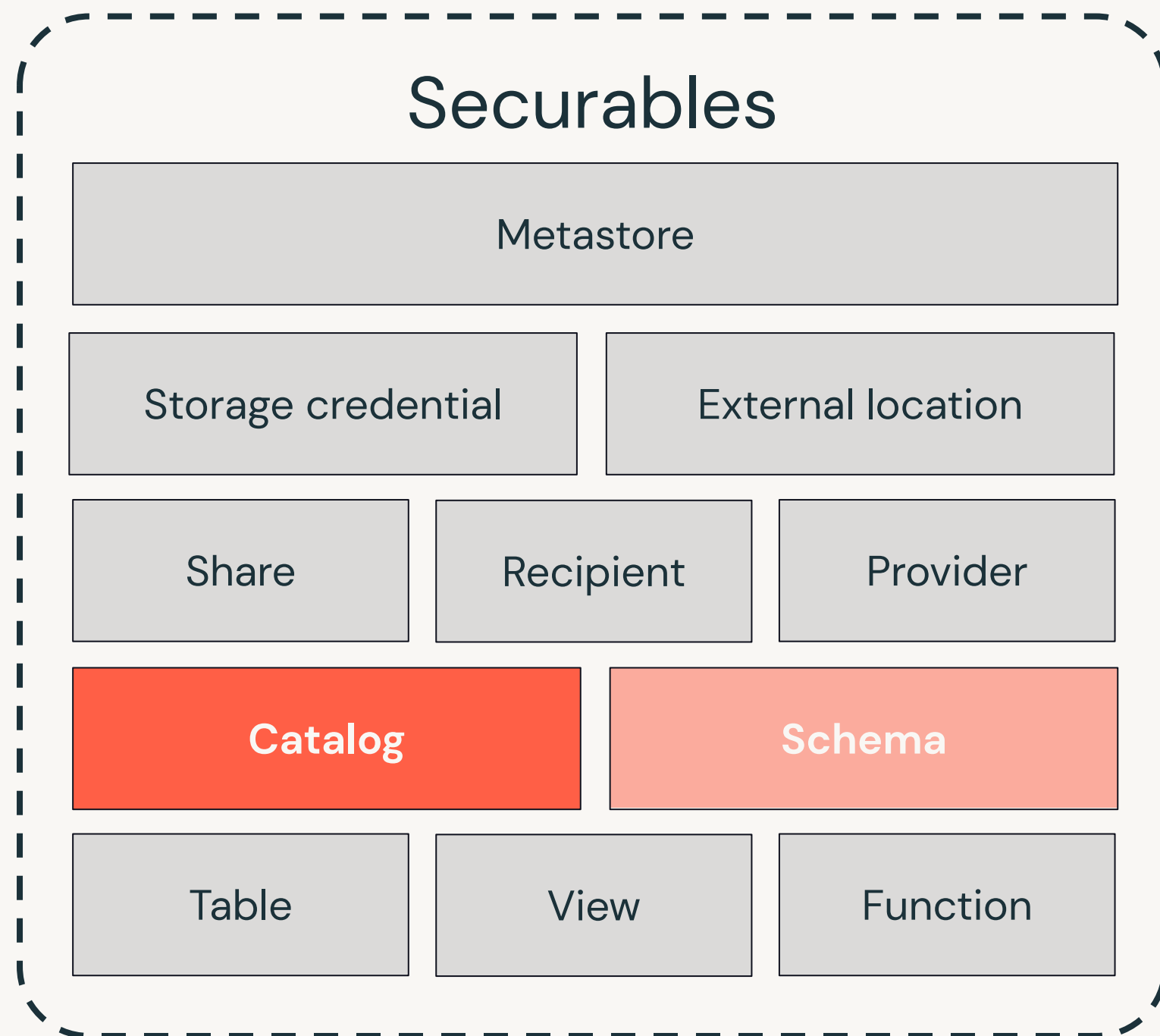
Create user-defined function in any schema

USE SCHEMA

Access objects in any schema

Privilege Types

Catalog privilege inheritance



Bequeathed privileges

CREATE TABLE

Create table or view in any schema

CREATE FUNCTION

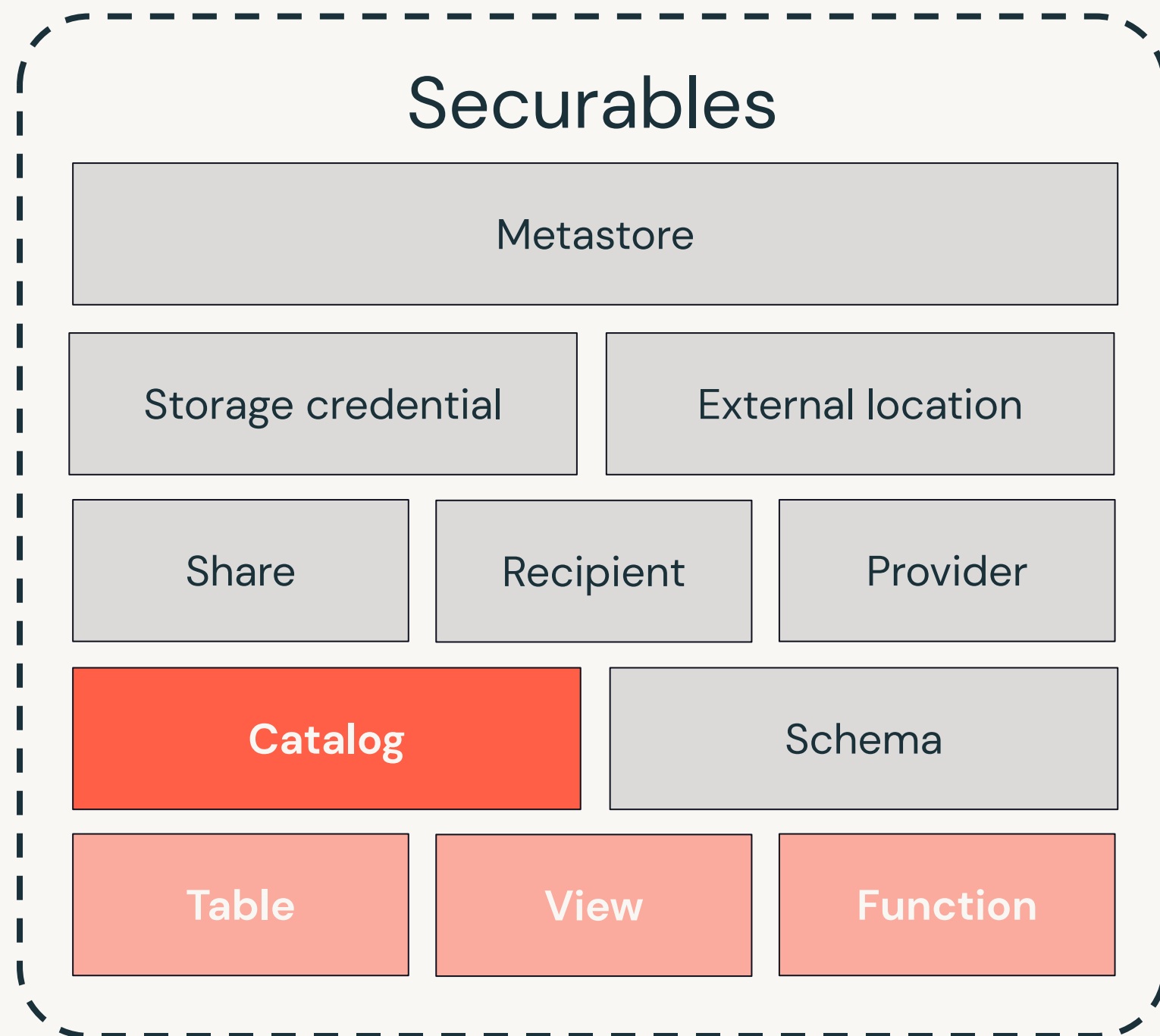
Create user-defined function in any schema

USE SCHEMA

Access objects in any schema

Privilege Types

Catalog privilege inheritance



Bequeathed privileges

SELECT

Read from any table or view from any schema

MODIFY

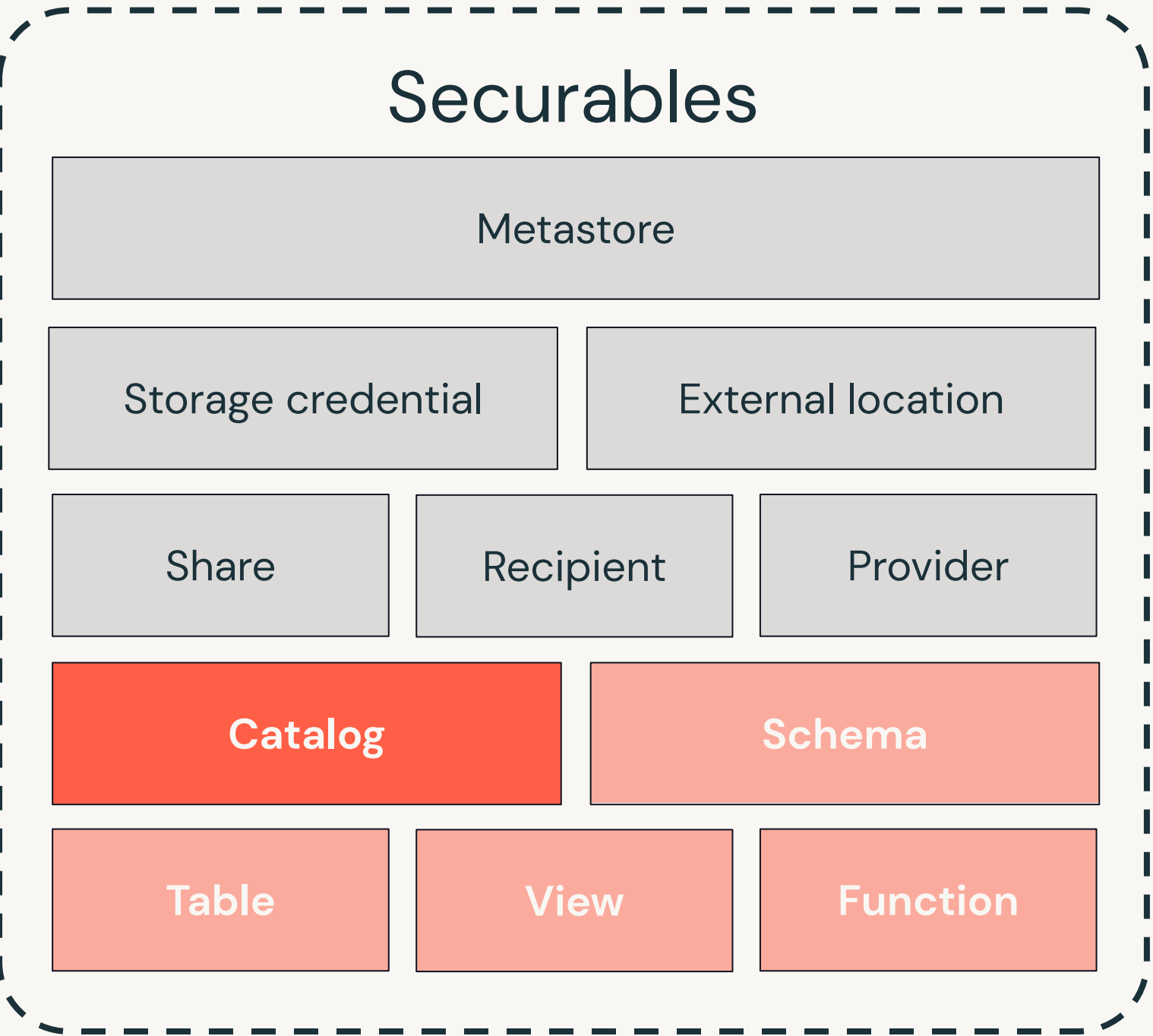
Modify data or metadata of any table in any schema

EXECUTE

Use any function from any schema

Privilege Types

Catalog

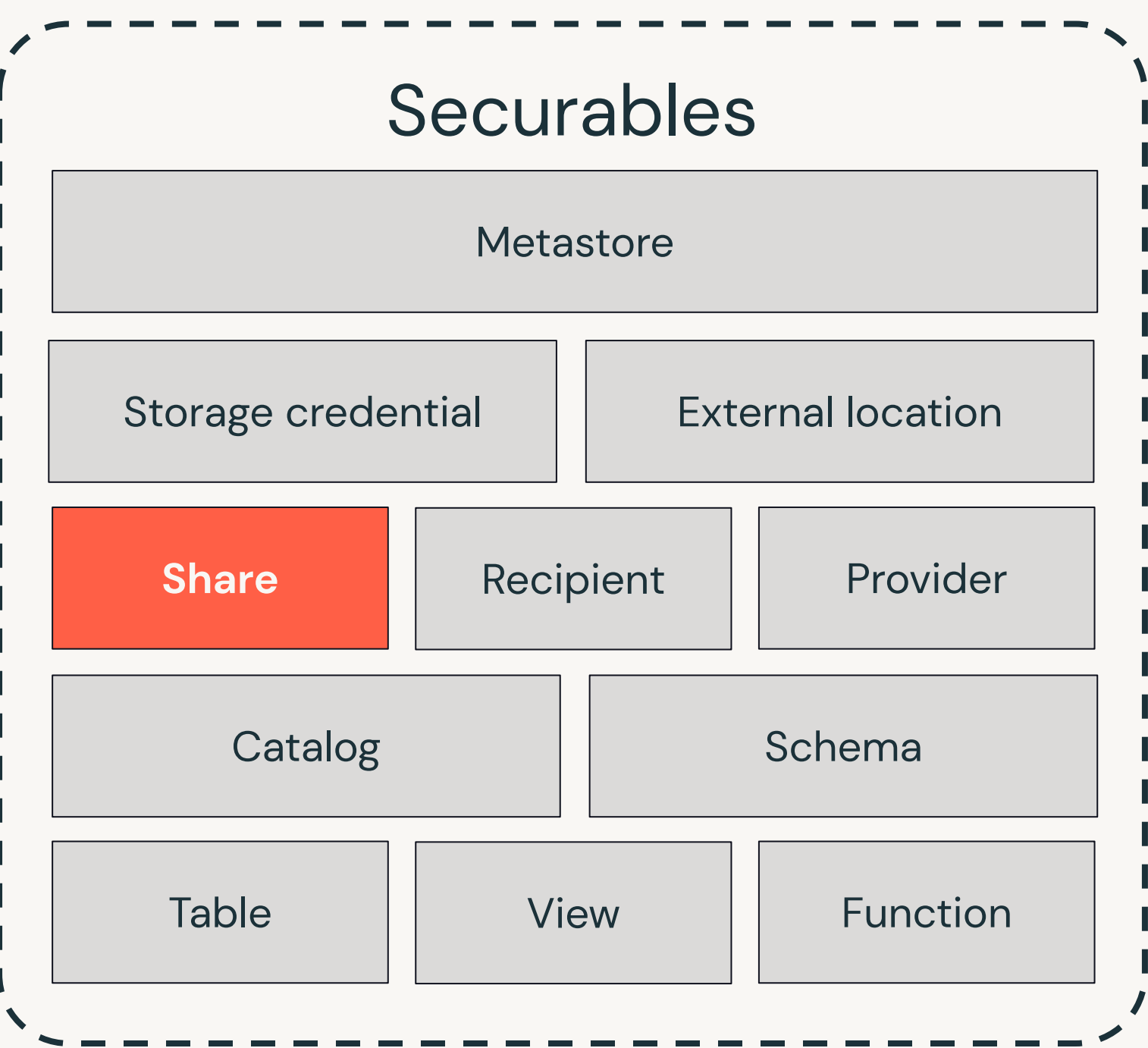


ALL PRIVILEGES

- | | |
|-----------------|------------|
| USE CATALOG | USE SCHEMA |
| CREATE SCHEMA | SELECT |
| CREATE TABLE | MODIFY |
| CREATE FUNCTION | EXECUTE |

Privilege Types

Delta sharing



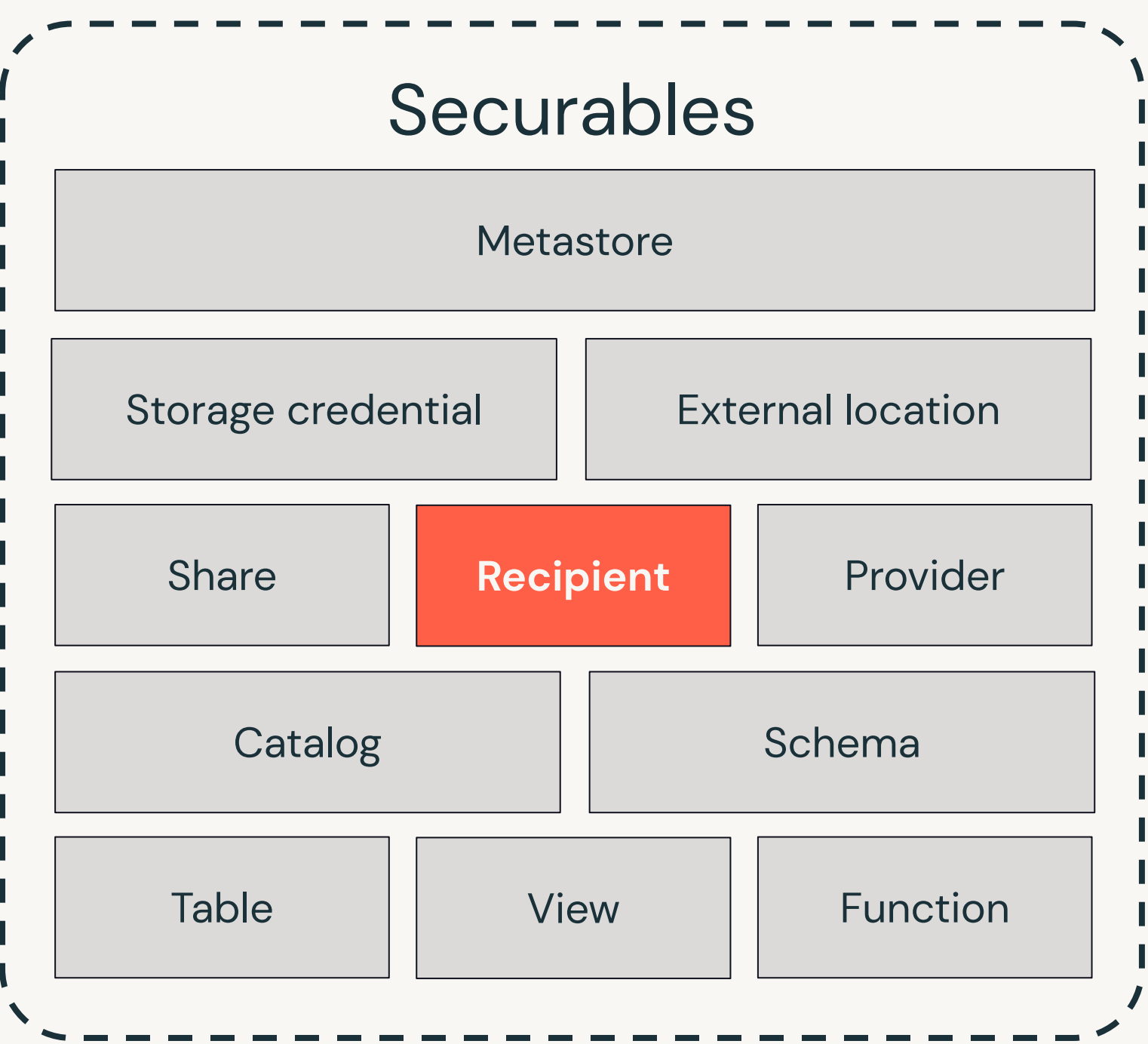
Privileges

SELECT

Provide read ability to a recipient

Privilege Types

Delta sharing

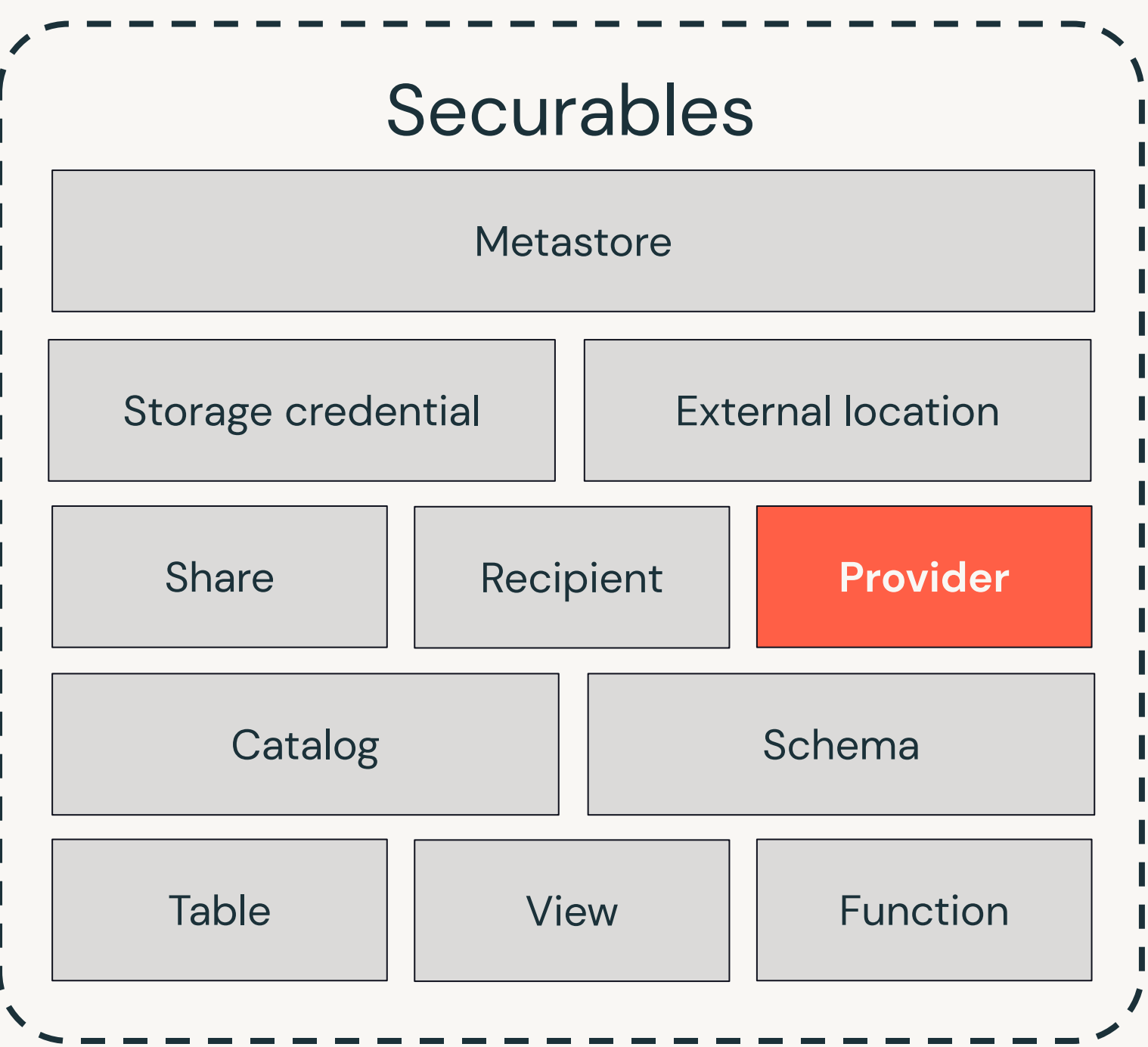


Privileges

No privileges

Privilege Types

Delta sharing

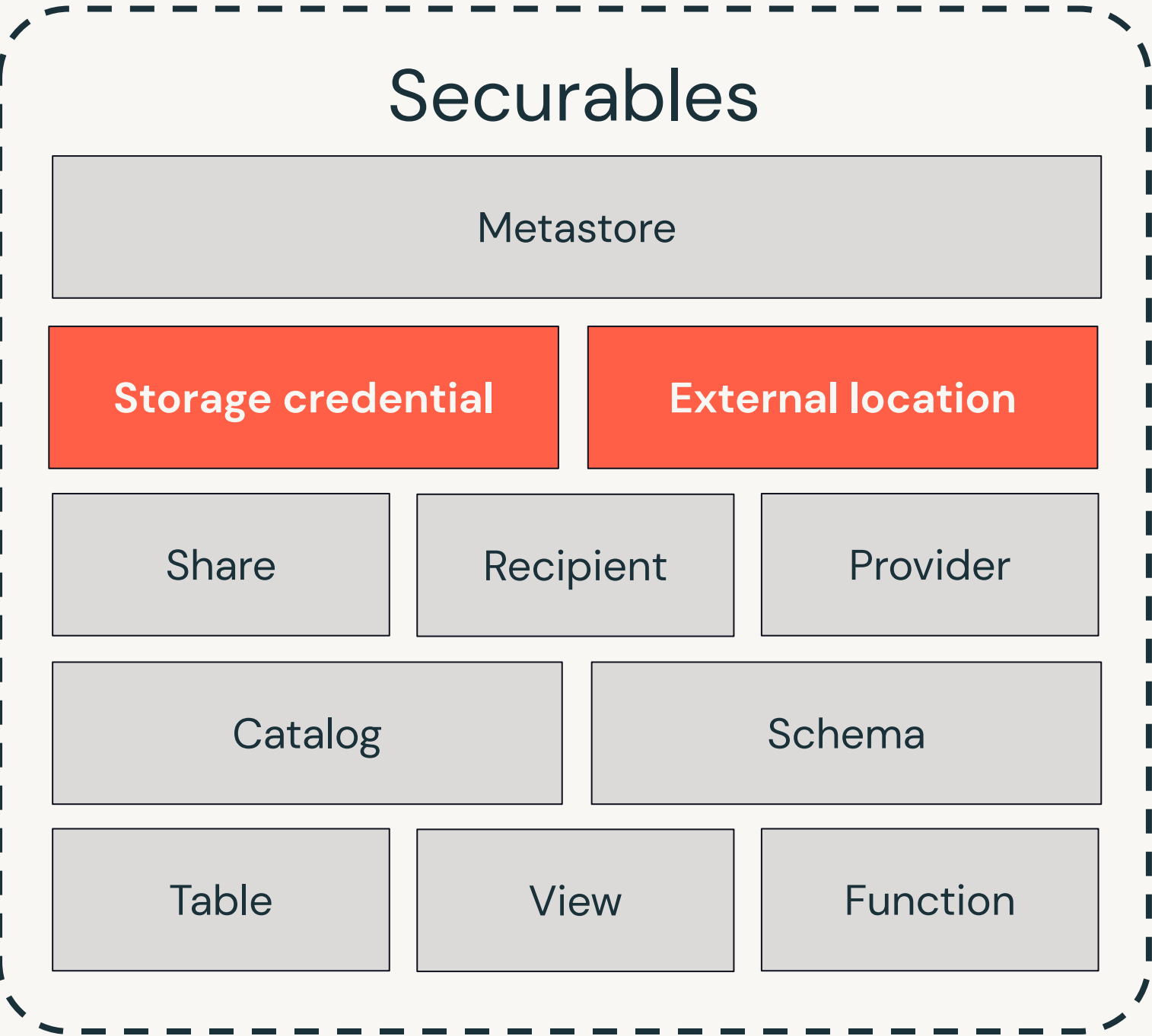


Privileges

No privileges

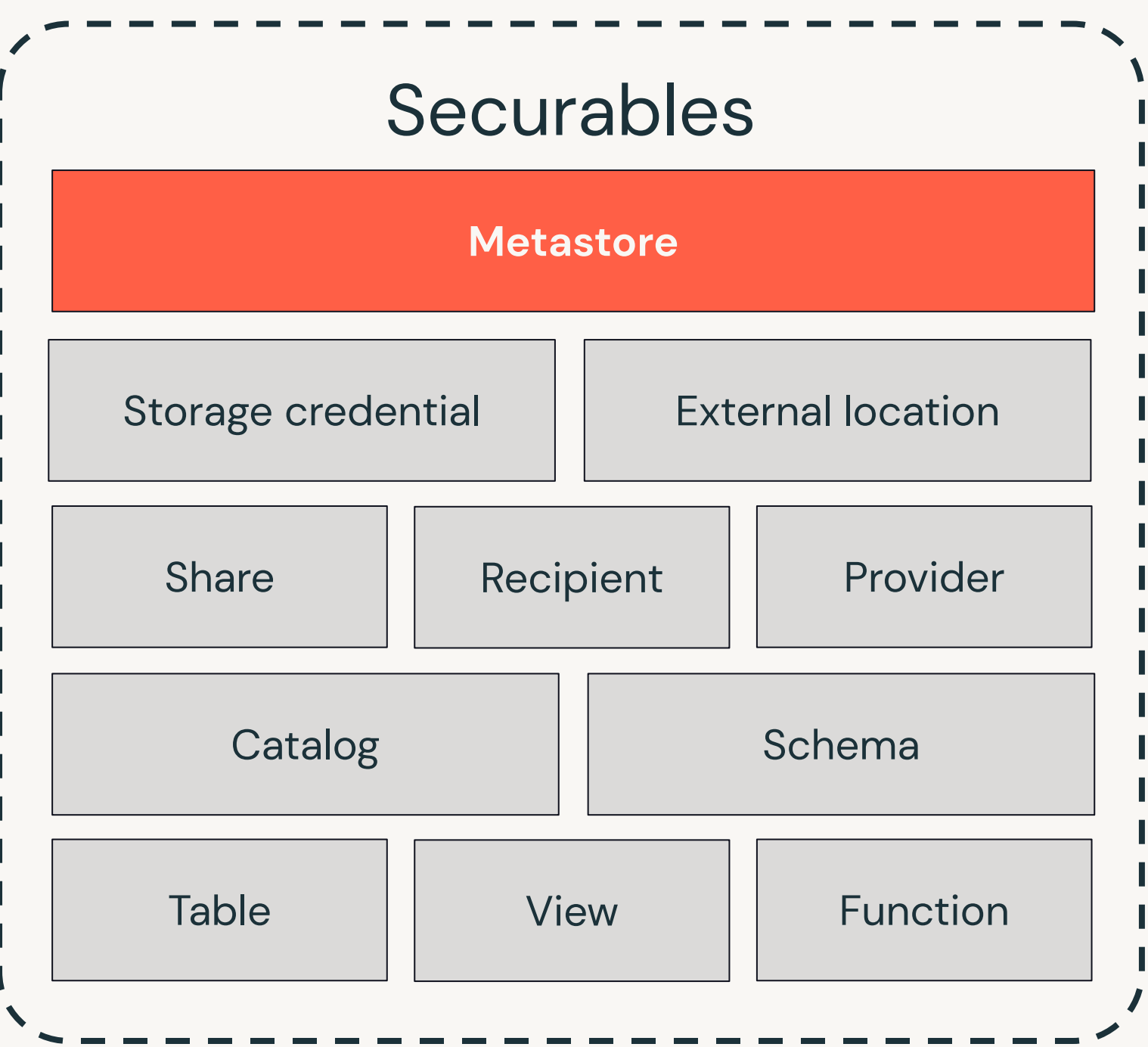
Privilege Types

External storage management



Privilege Types

Metastore

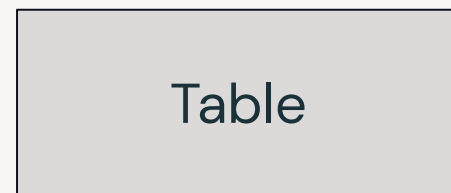


Privileges

- CREATE CATALOG
- CREATE EXTERNAL LOCATION
- CREATE RECIPIENT
- CREATE SHARE
- CREATE PROVIDER

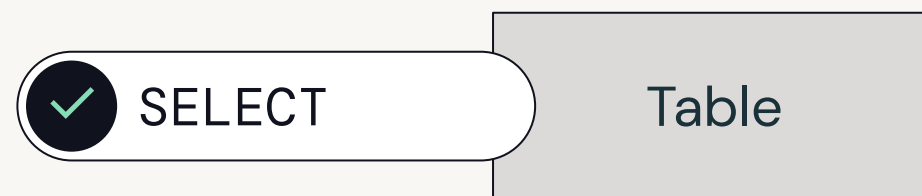
Privilege Scenarios

Querying a table



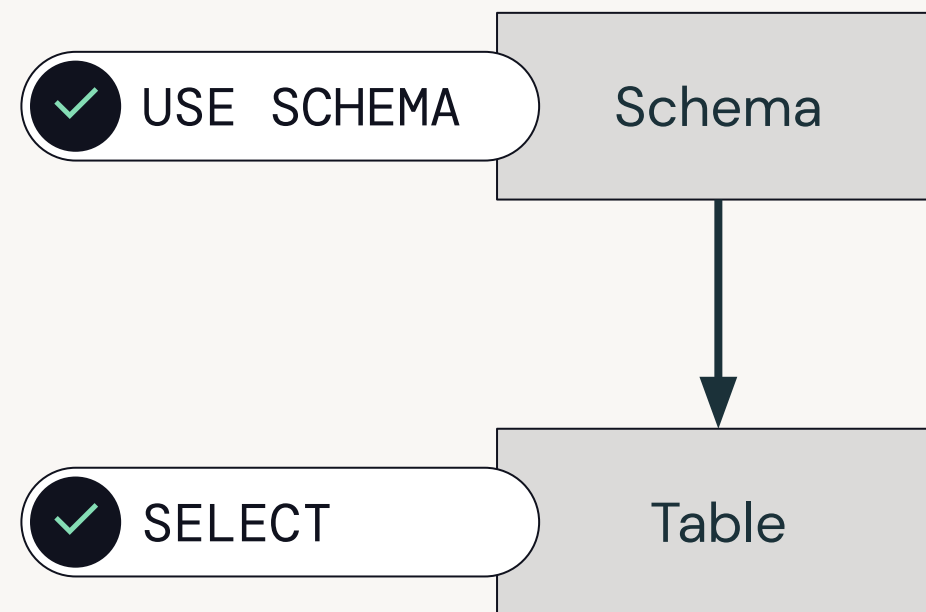
Privilege Scenarios

Querying a table



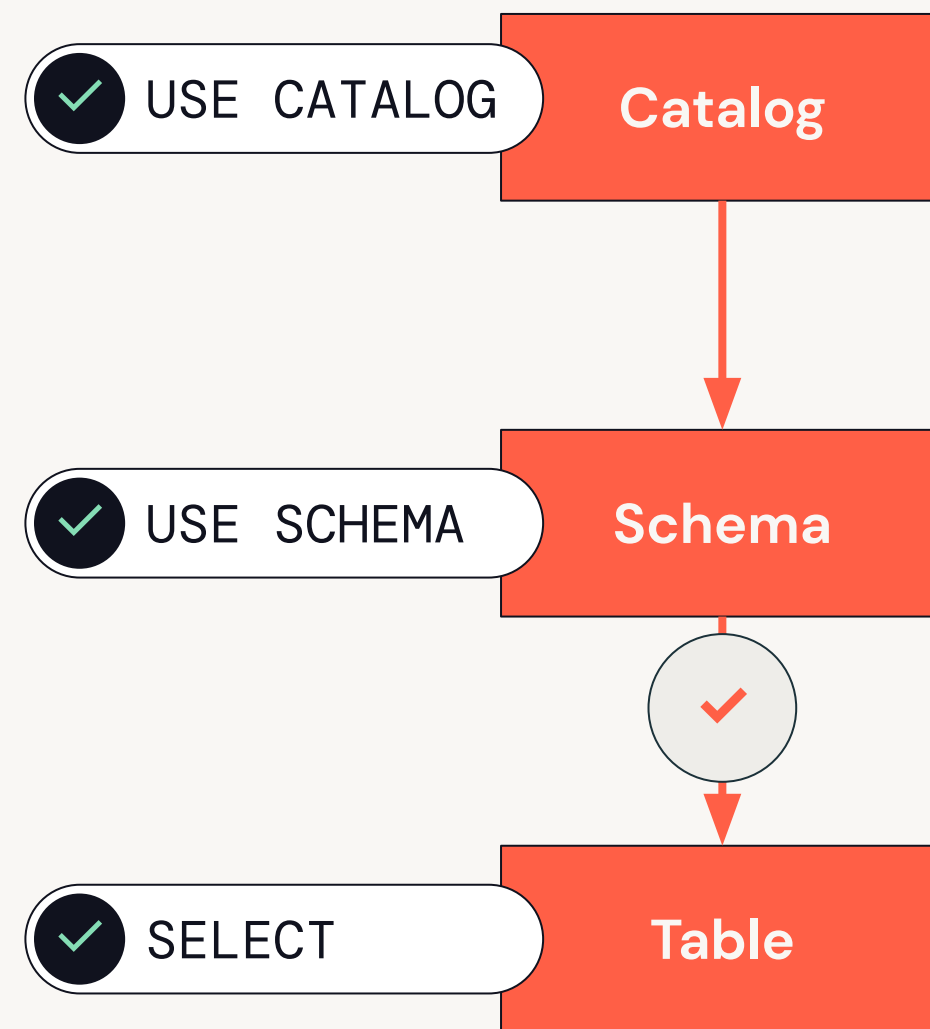
Privilege Scenarios

Querying a table



Privilege Scenarios

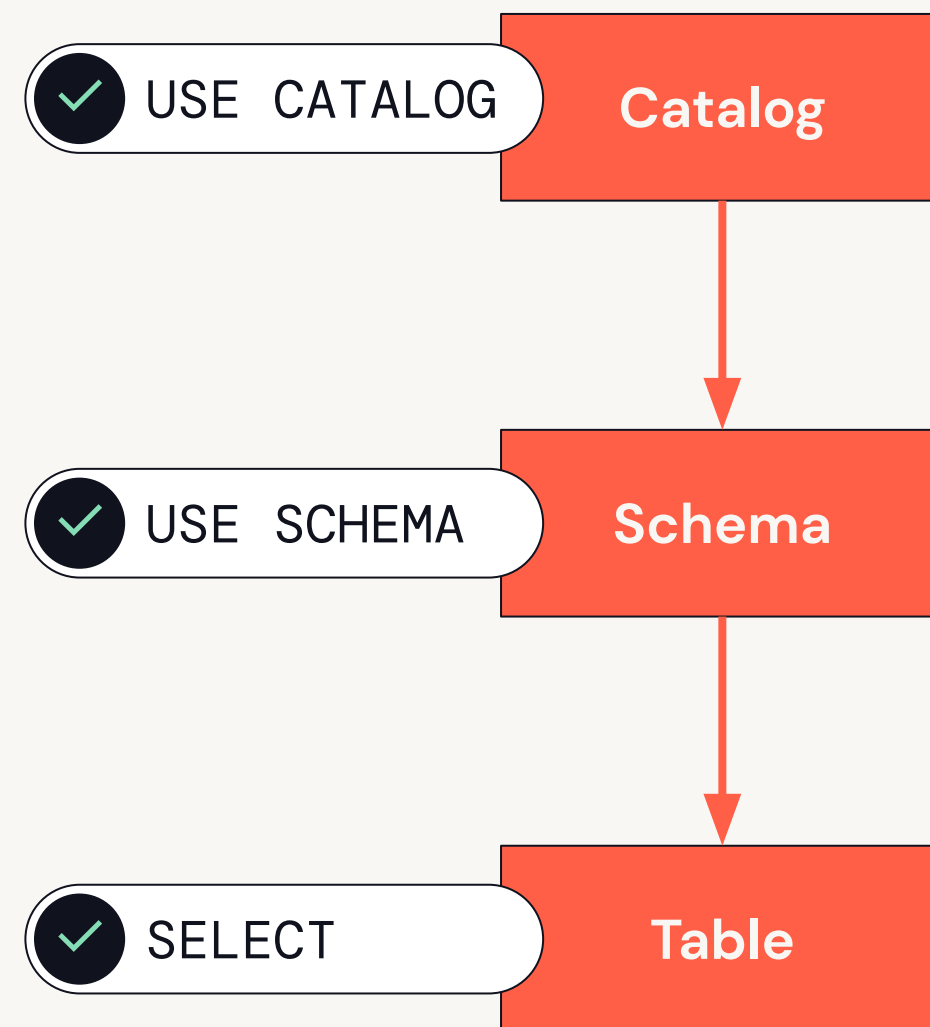
Querying a table



Privilege Scenarios

Querying a table

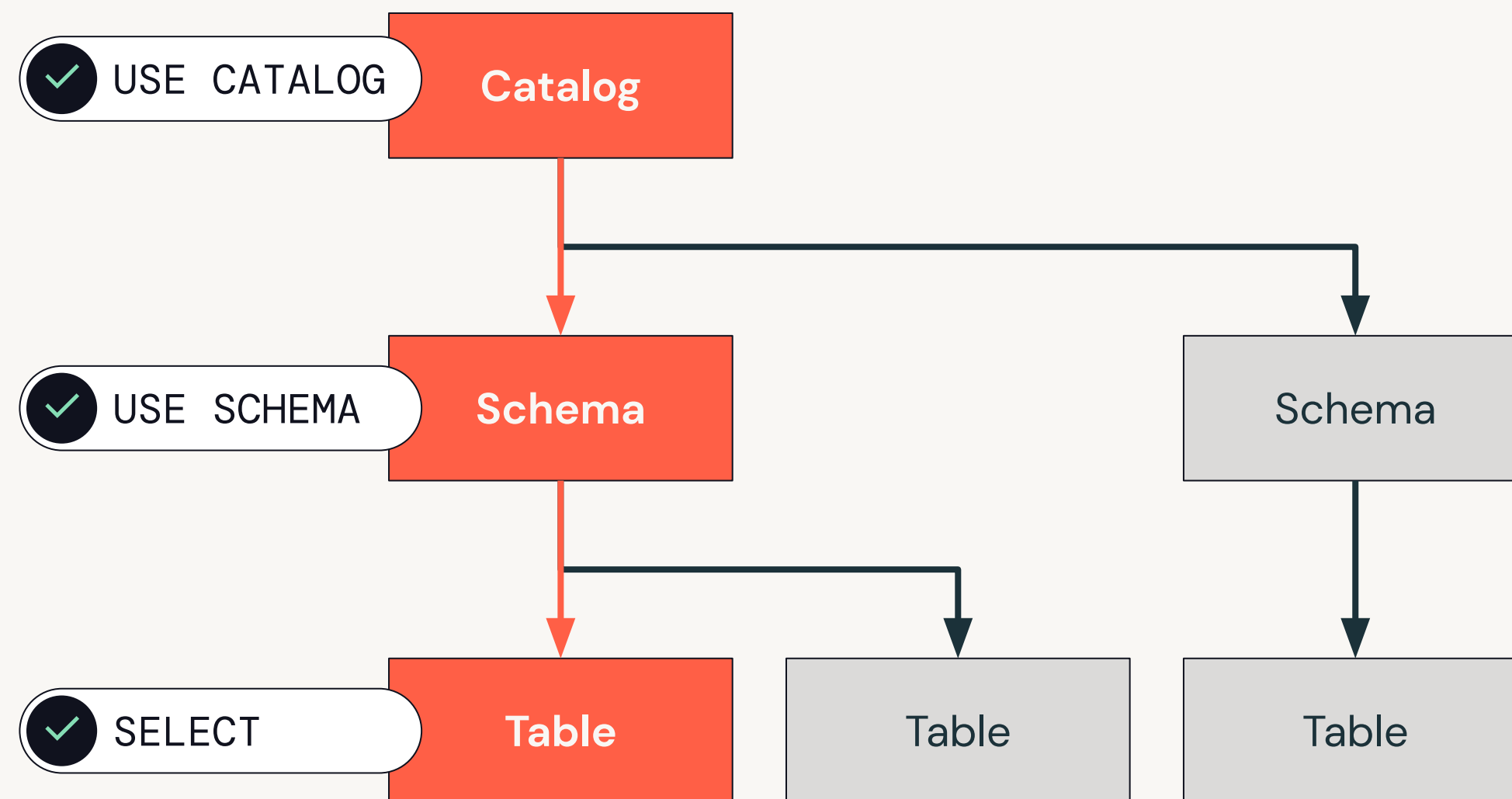
Explicit



Privilege Scenarios

Querying a table

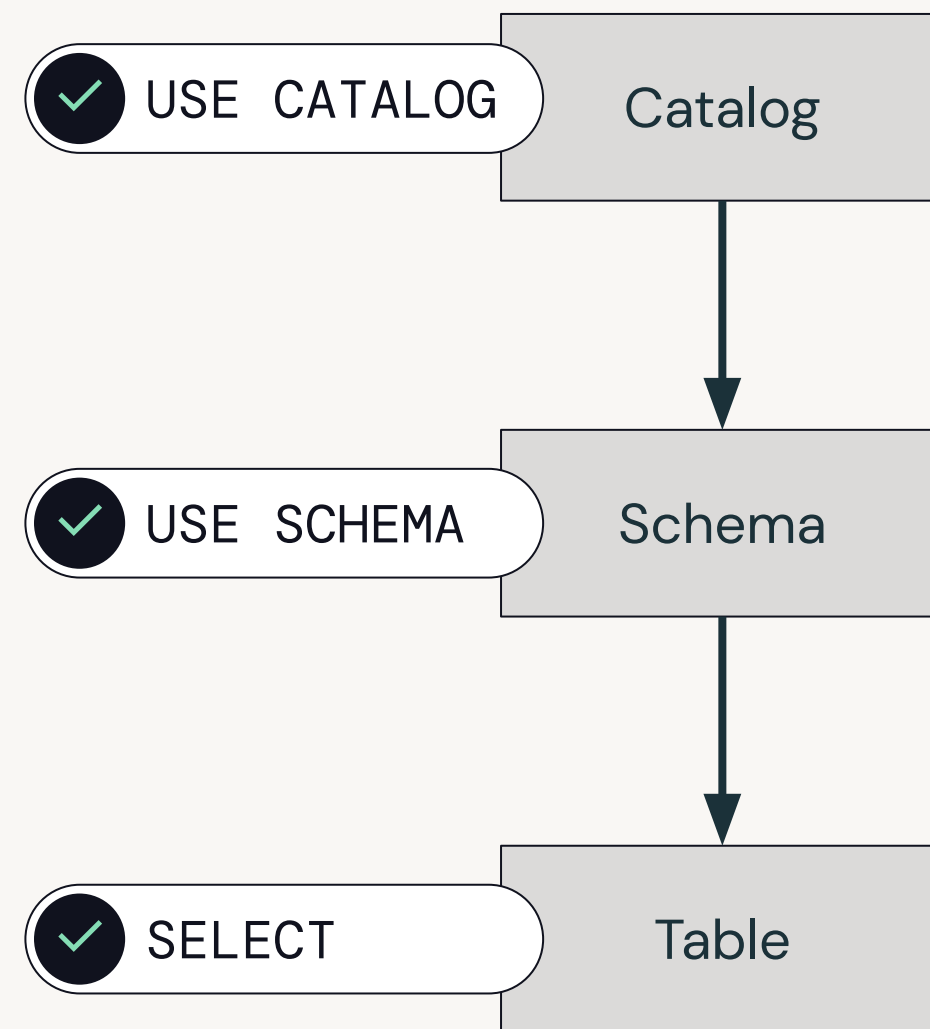
Explicit



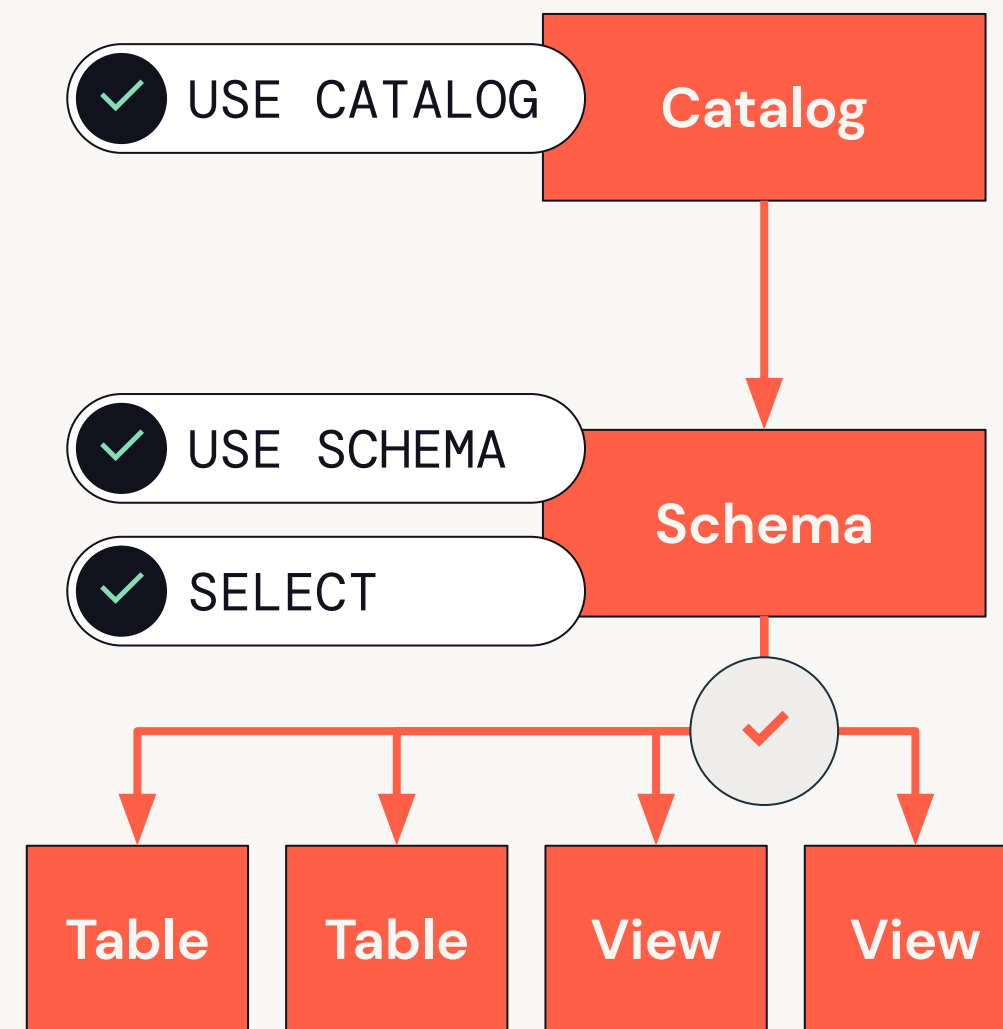
Privilege Scenarios

Querying a table

Explicit



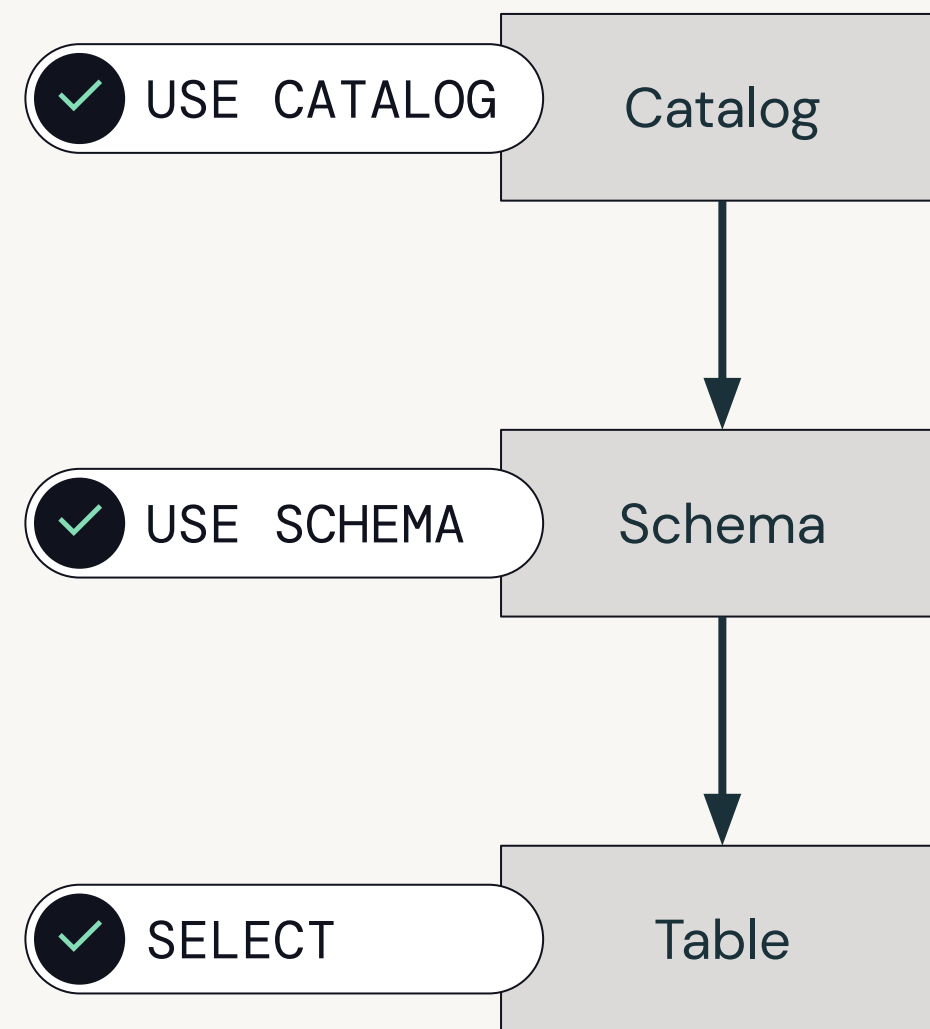
Inherited from schema



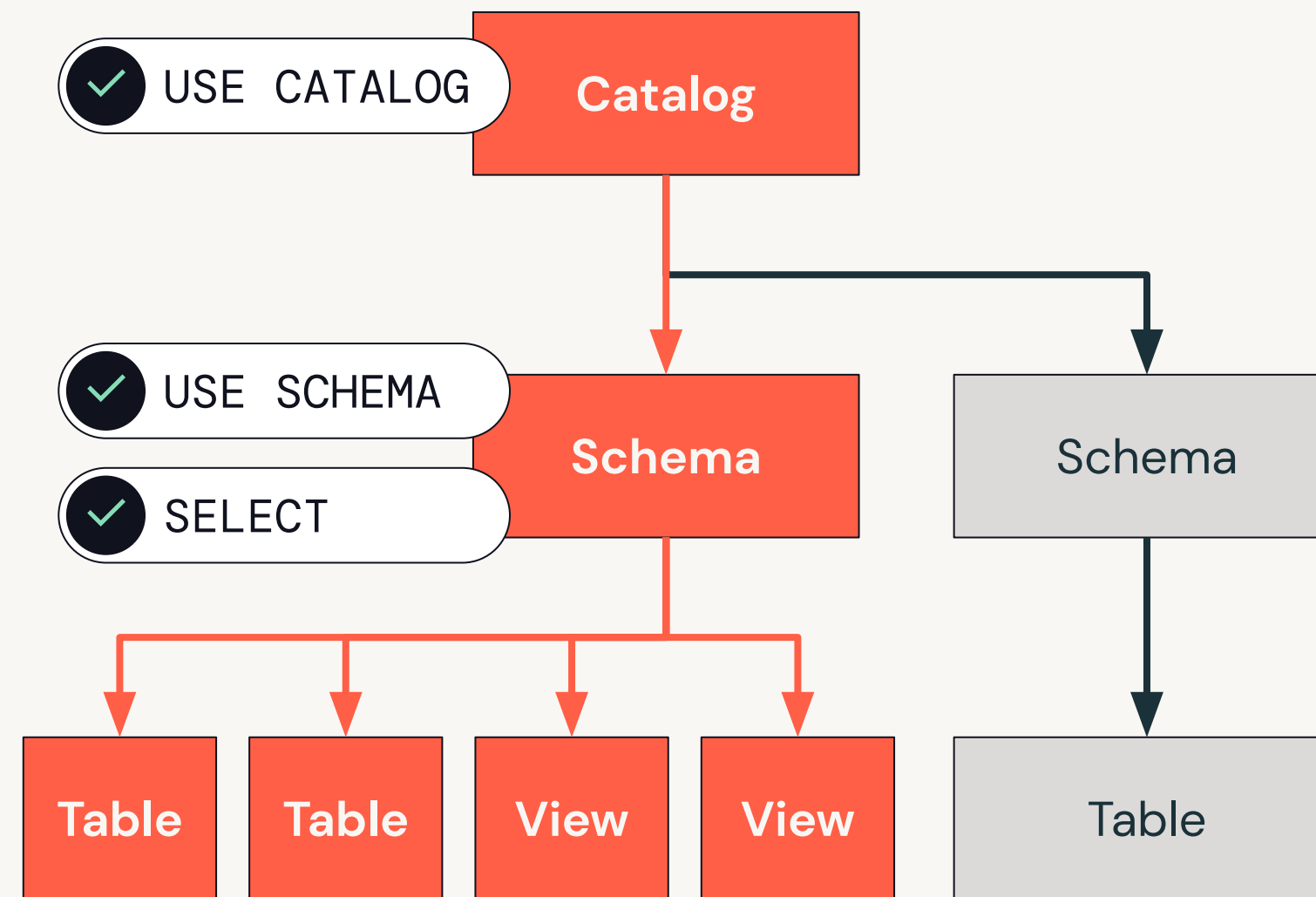
Privilege Scenarios

Querying a table

Explicit



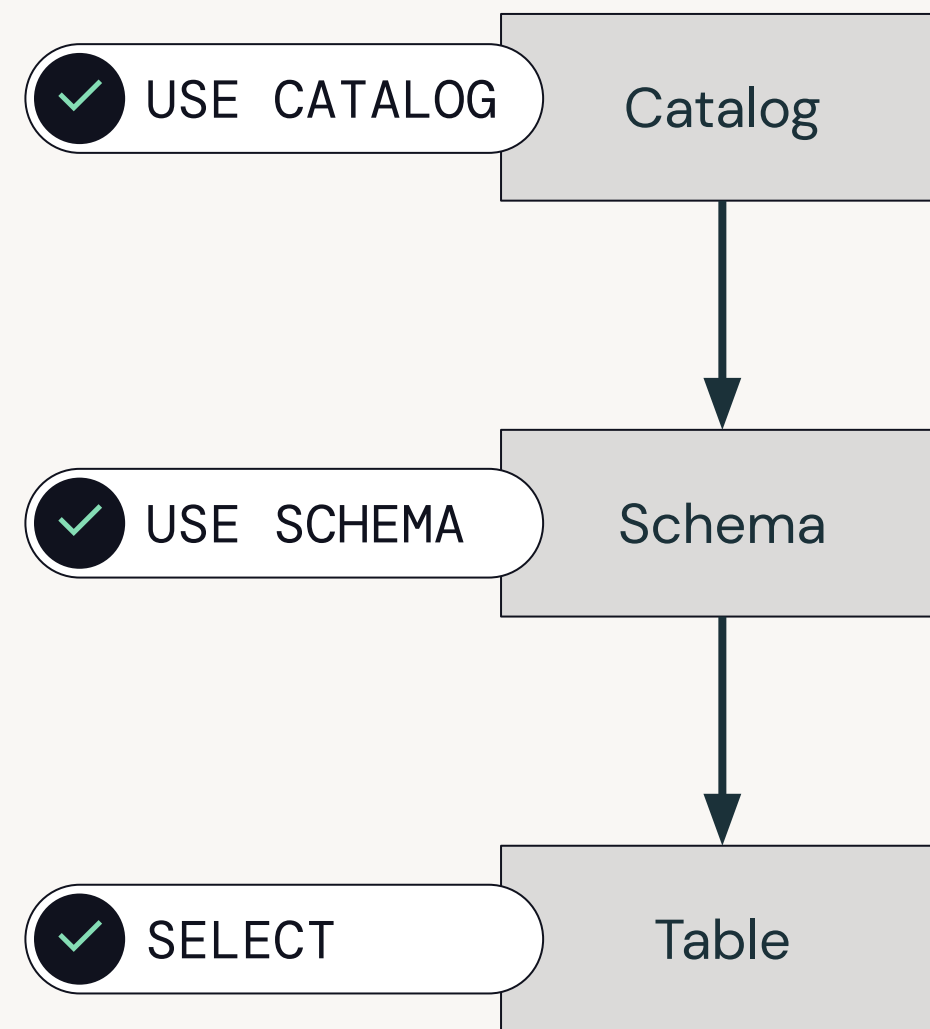
Inherited from schema



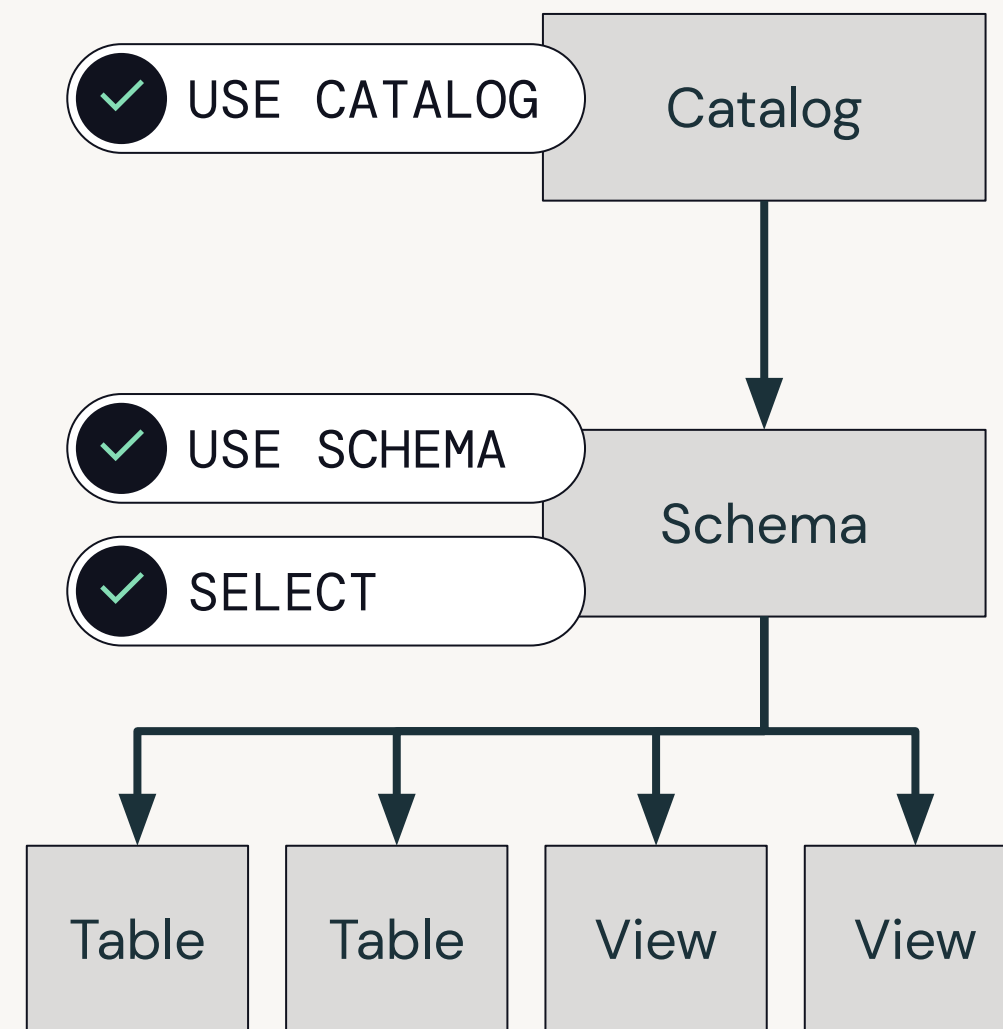
Privilege Scenarios

Querying a table

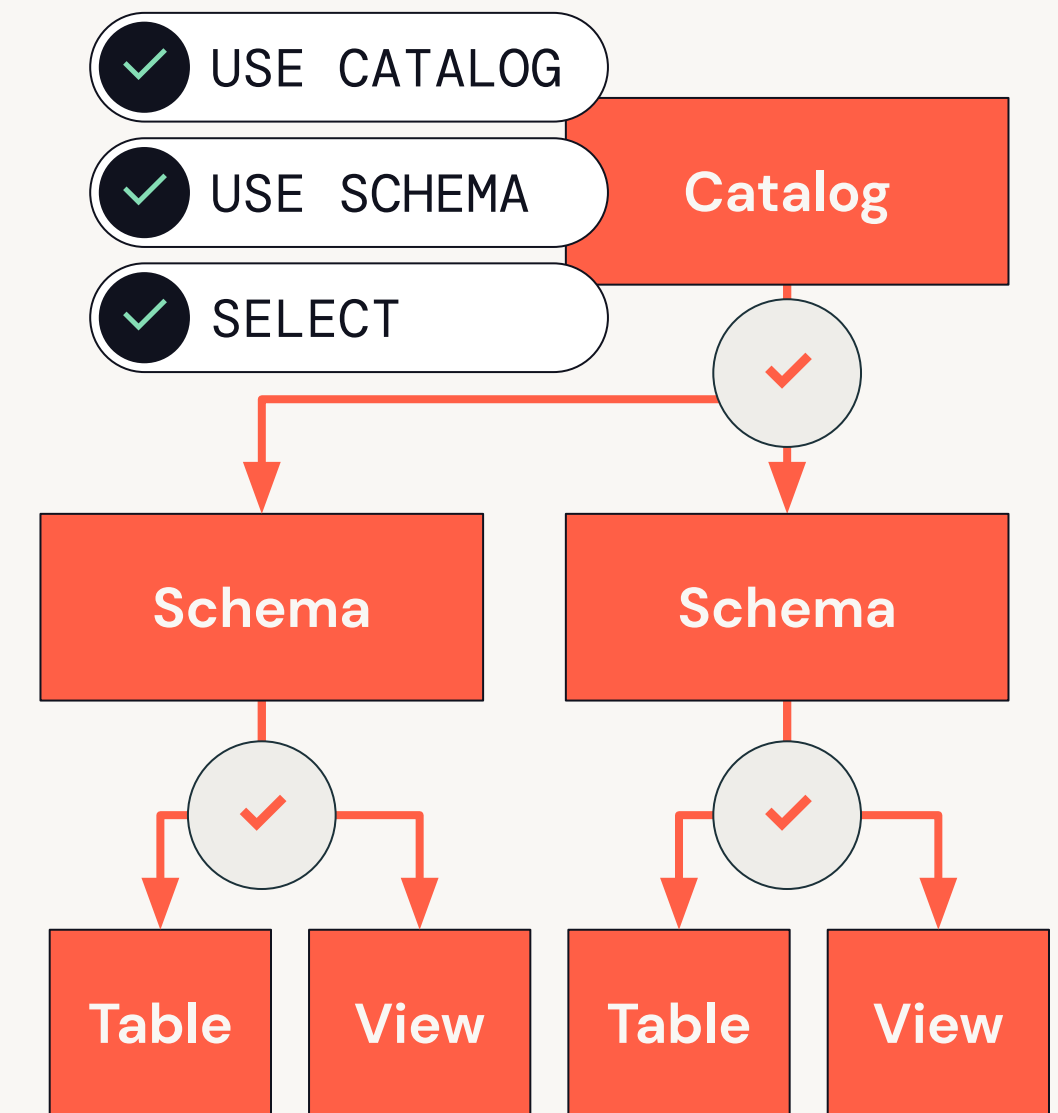
Explicit



Inherited from schema



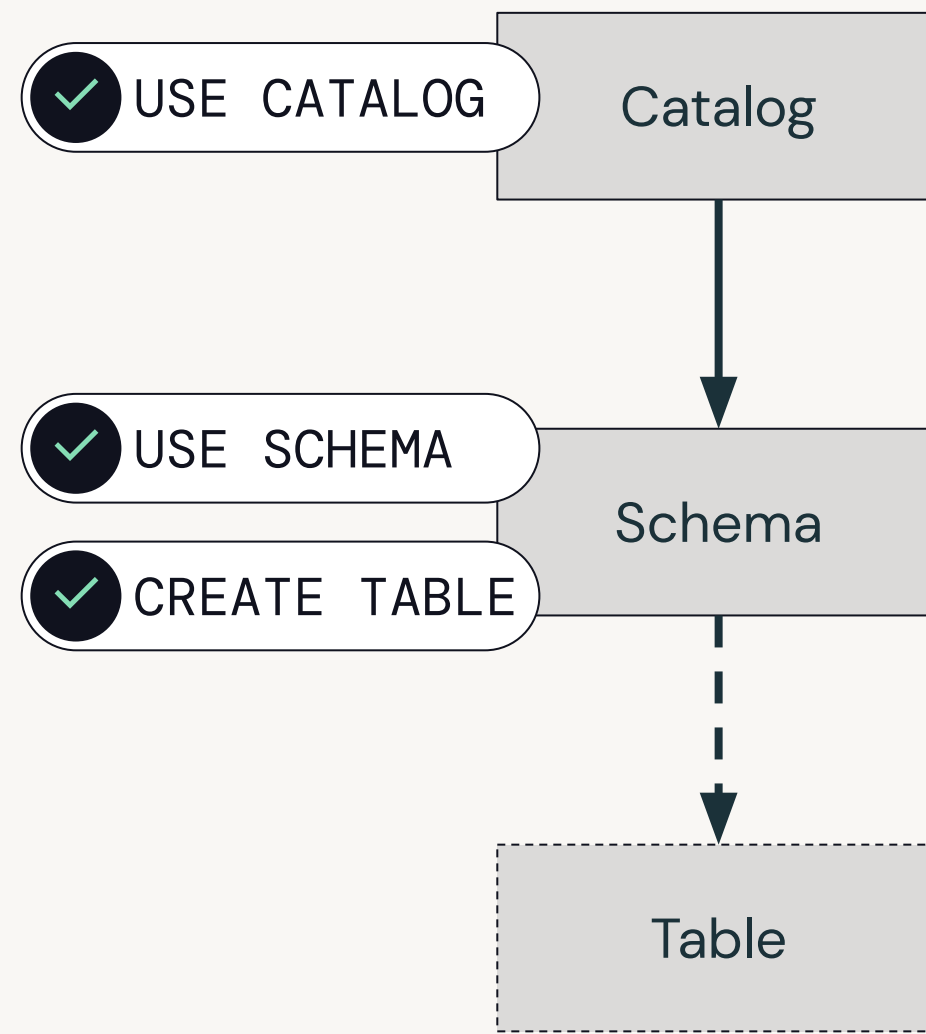
Inherited from catalog



Privilege Scenarios

Creating a table

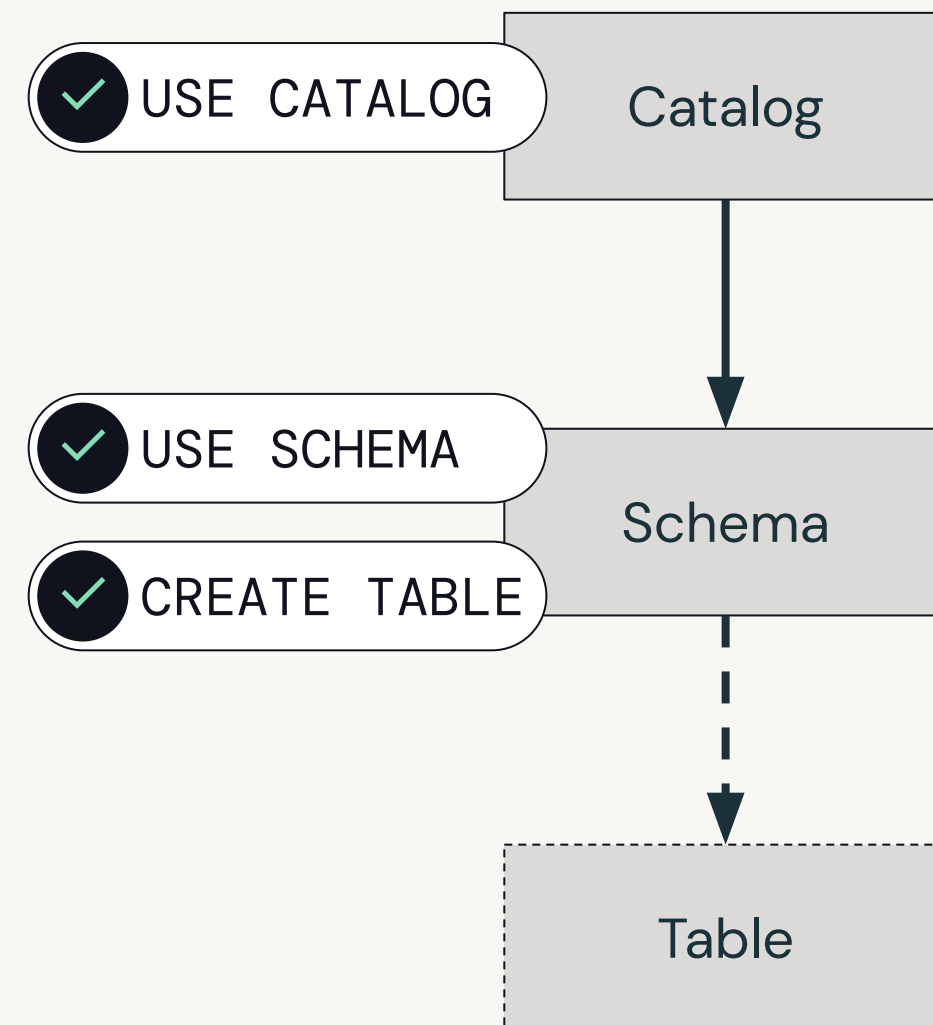
Explicit



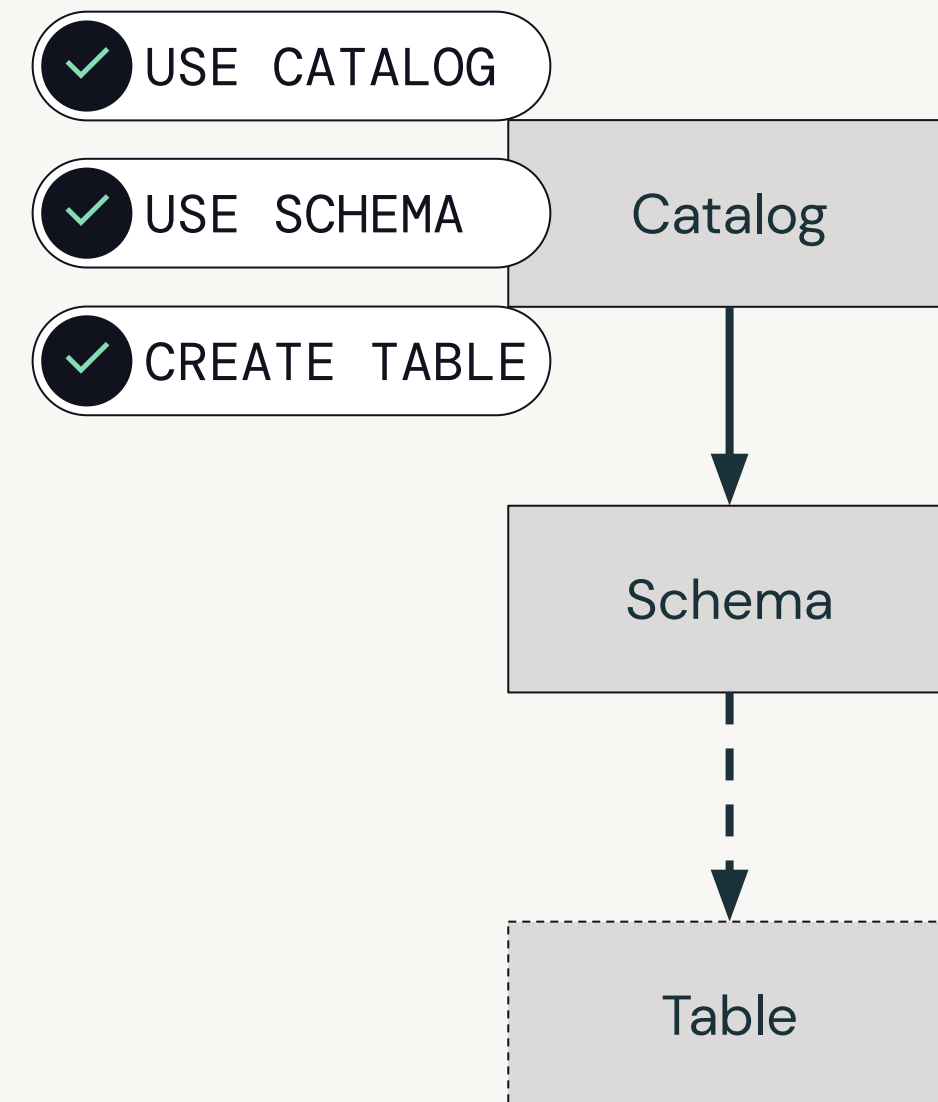
Privilege Scenarios

Creating a table

Explicit



Inherited from catalog



Lab: Controlling access to data

Learning Objectives

In this lab, you will learn how to:

- Grant access to data objects
- Revoke access to data objects
- Validate access
- Apply an explicit versus inherit grant model

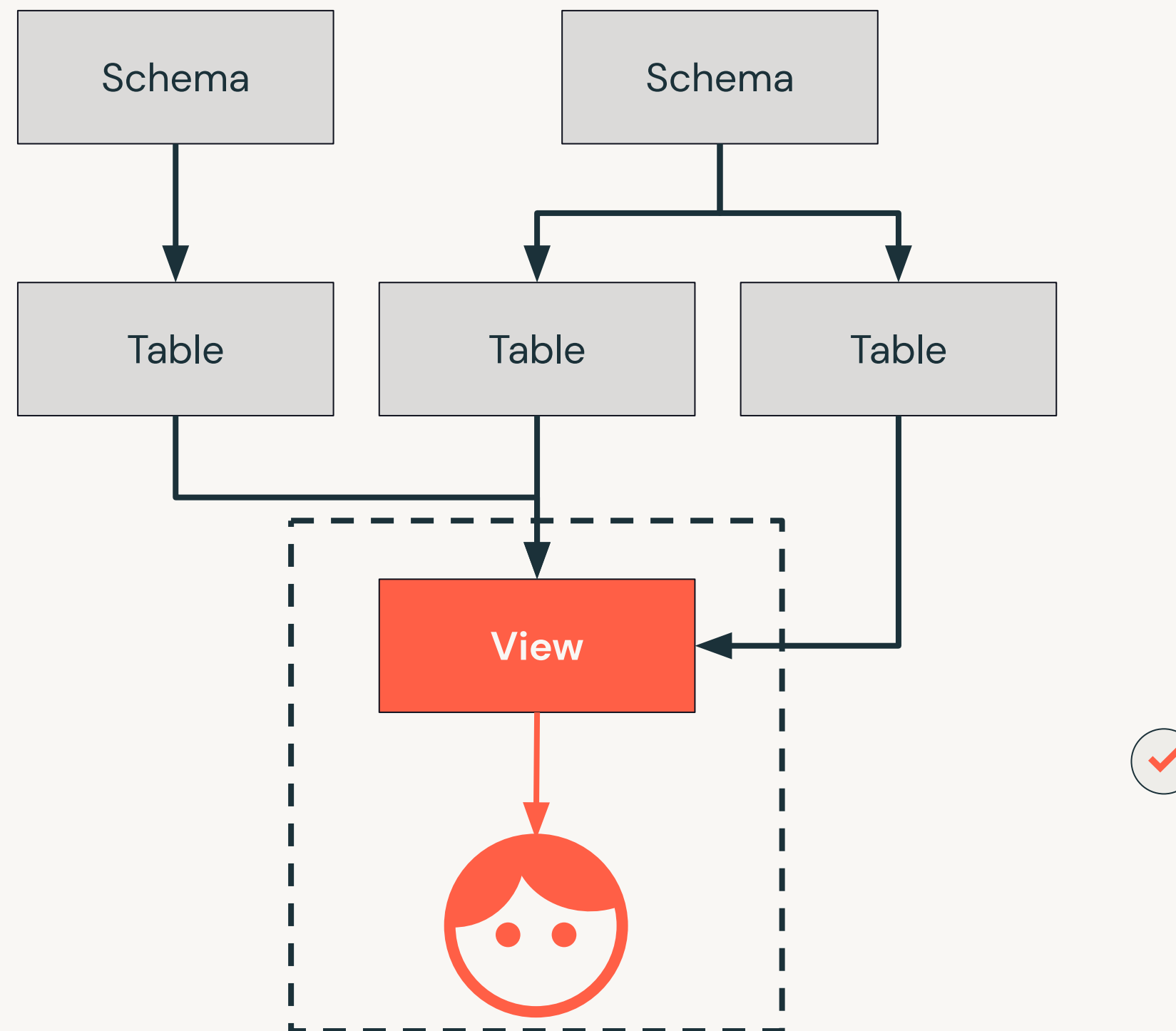
Prerequisites

In order to do this lab, you will need:

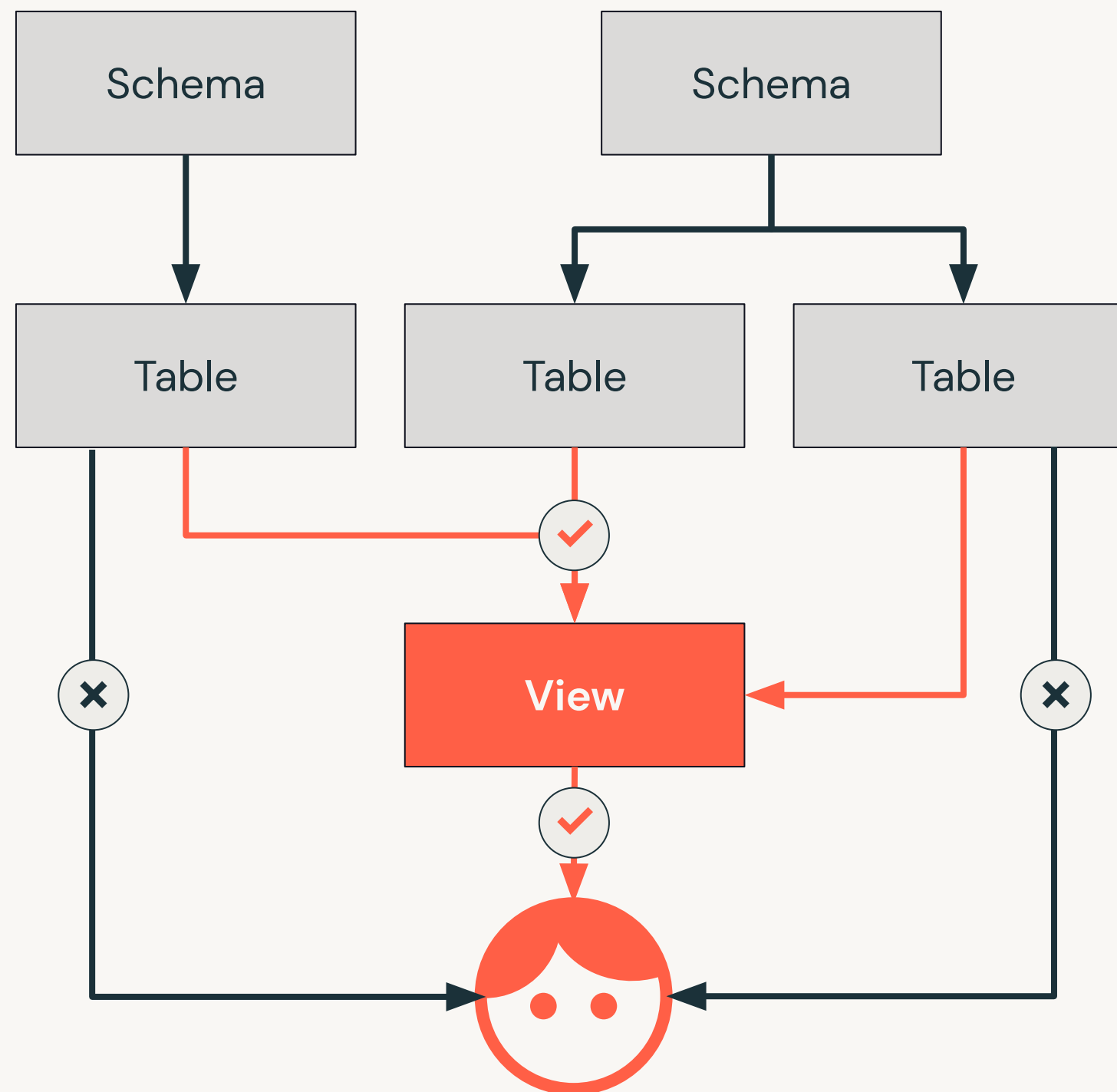
- Account admin capability

Dynamic Views

Dynamic Views



Dynamic Views



Dynamic Views

Three common use cases

Control access to columns

Dynamic Views

Three common use cases

Control access to
columns

Control access to rows

Dynamic Views

Three common use cases

Control access to columns

Control access to rows

Mask data

.....@databricks.com

Lab: Protecting columns and rows

Learning Objectives

In this lab, you will learn how to:

- Redact columns
- Filter rows
- Mask data



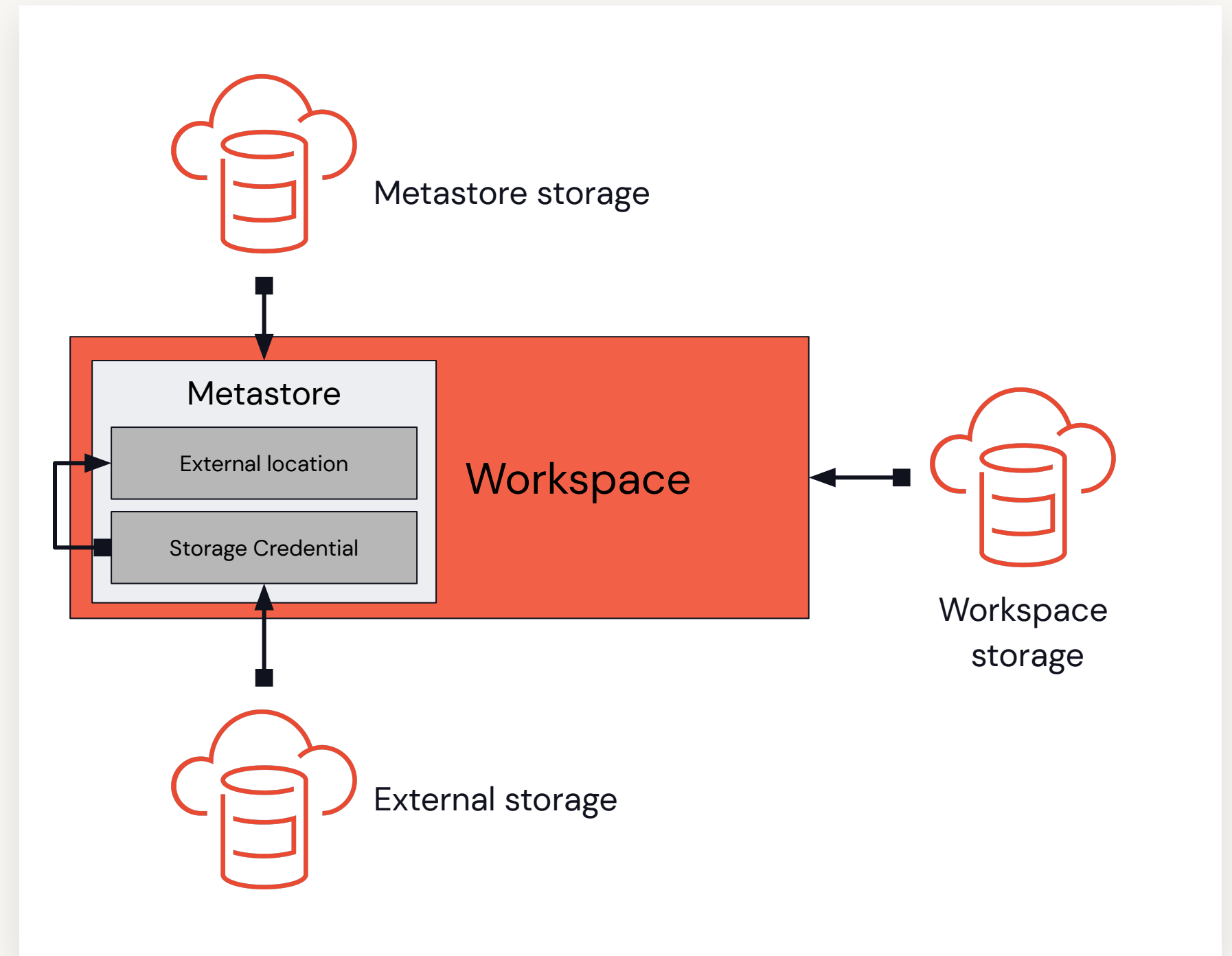
External Data Storage

External Data Storage

Integrates arbitrary cloud storage containers to address important use cases

- Data exploration
- Data ingestion
- External tables

Can be done any time

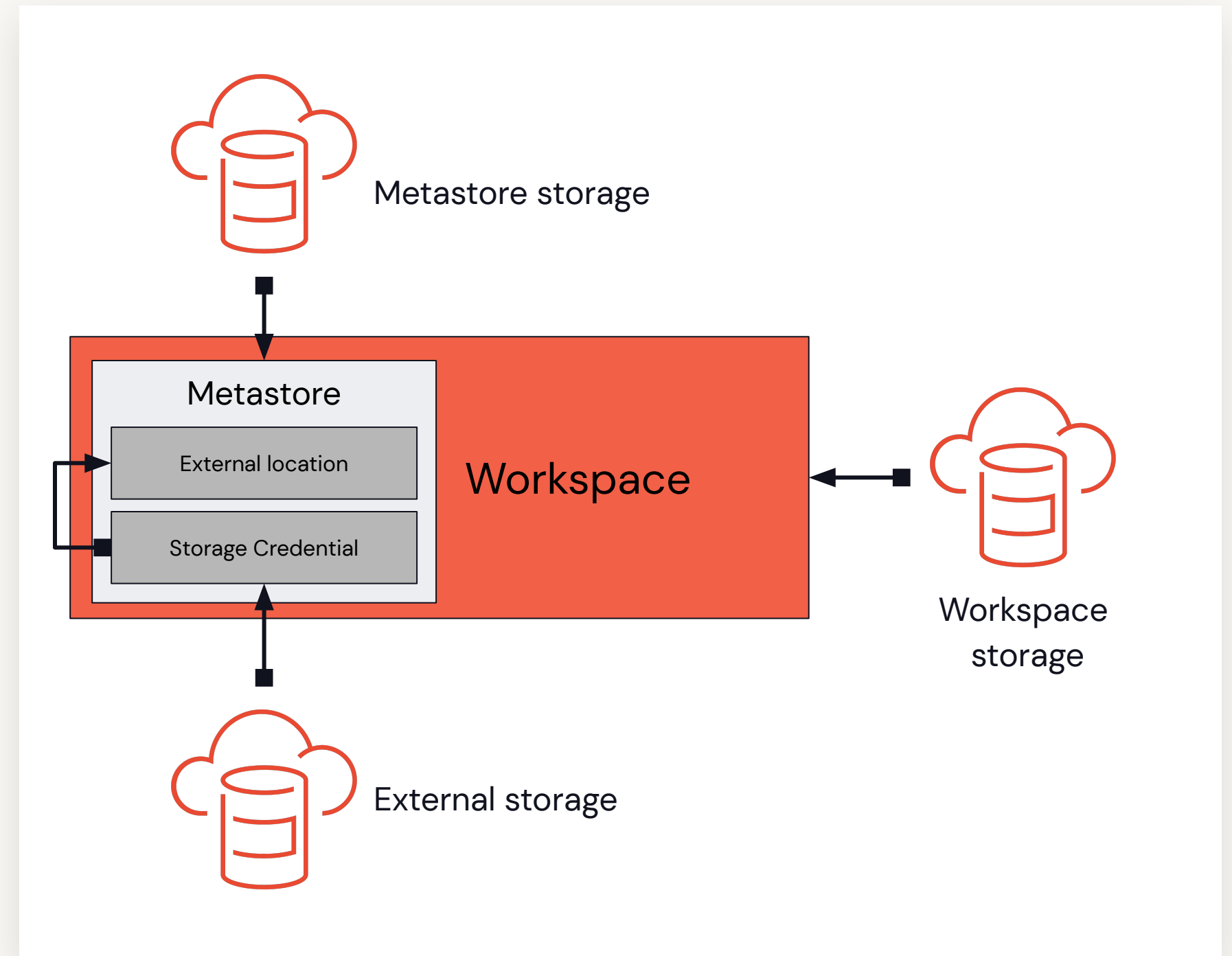


External Data Storage

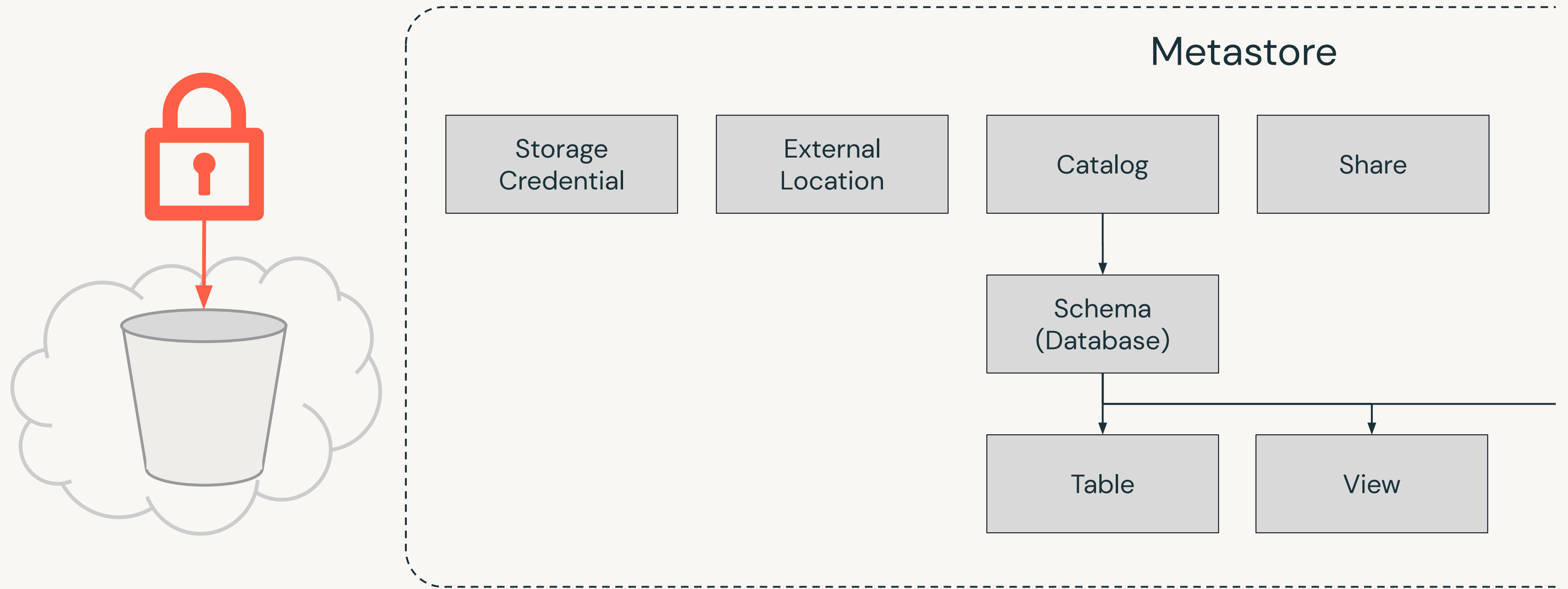
Integrates arbitrary cloud storage containers to address important use cases

- Data exploration
- Data ingestion
- External tables

Can be done any time

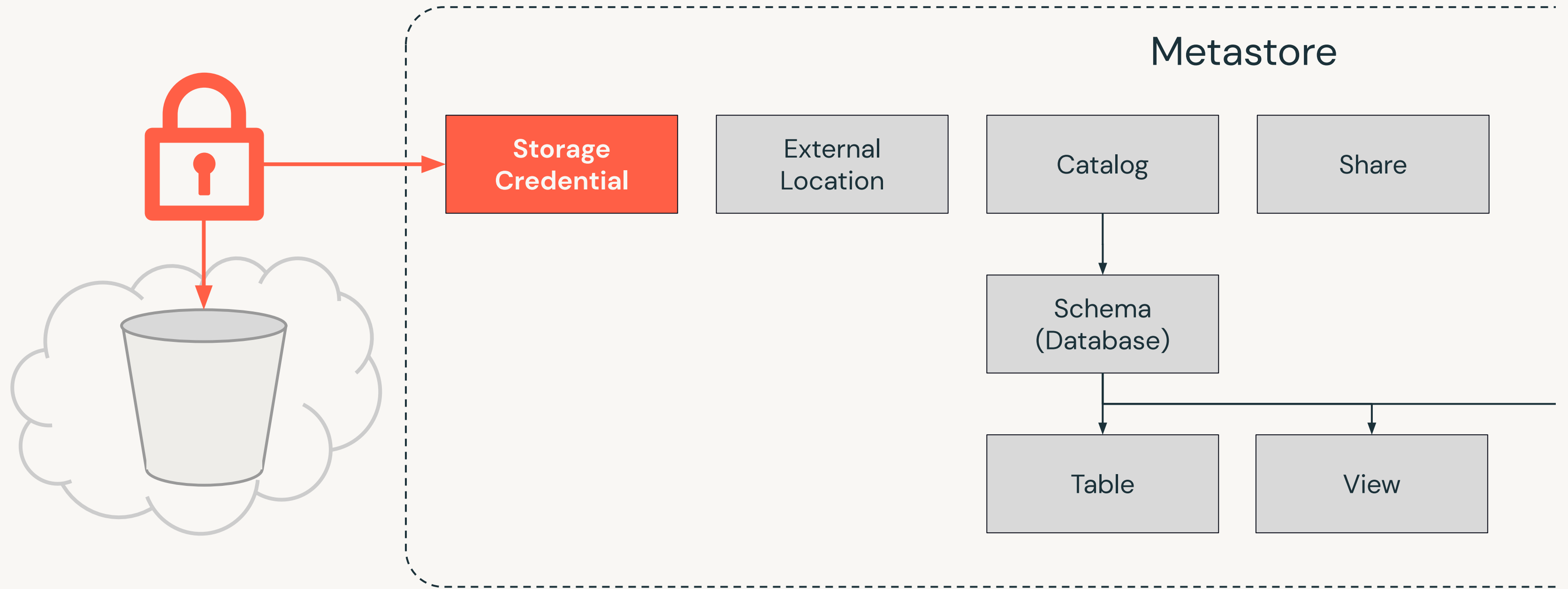


External Data Storage



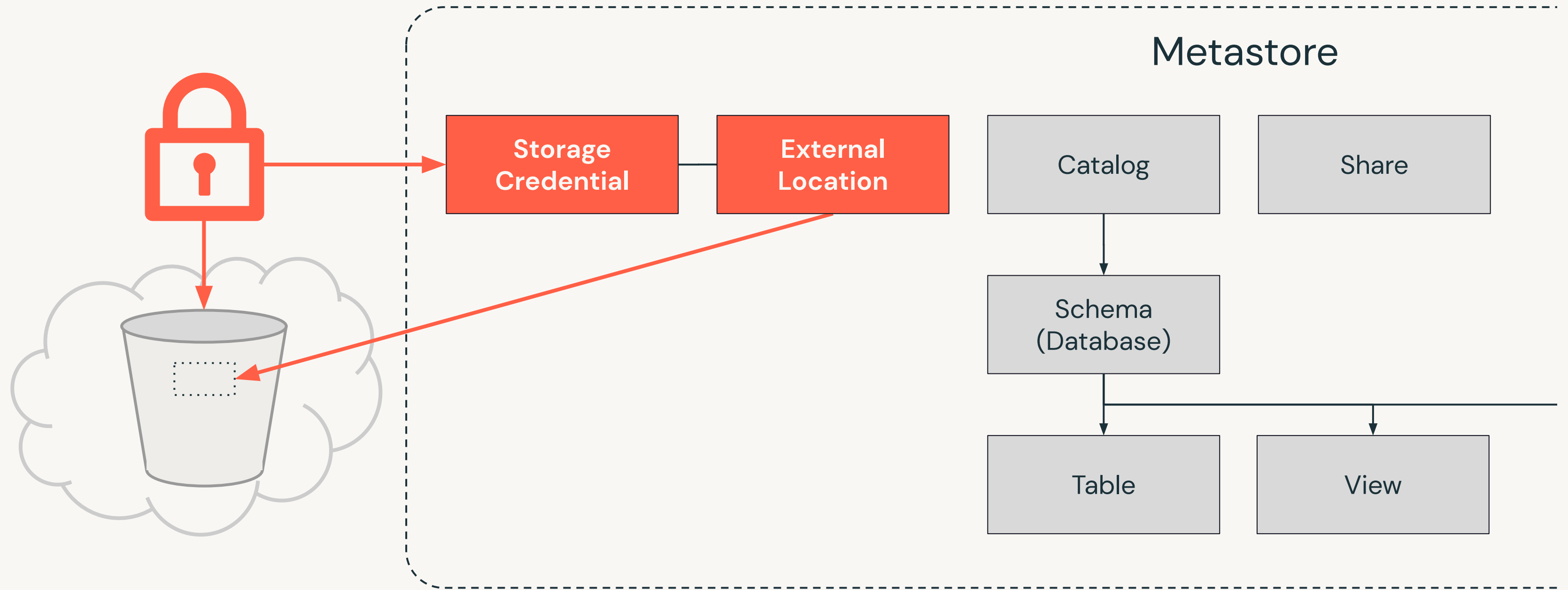
External Data Storage

Storage credential



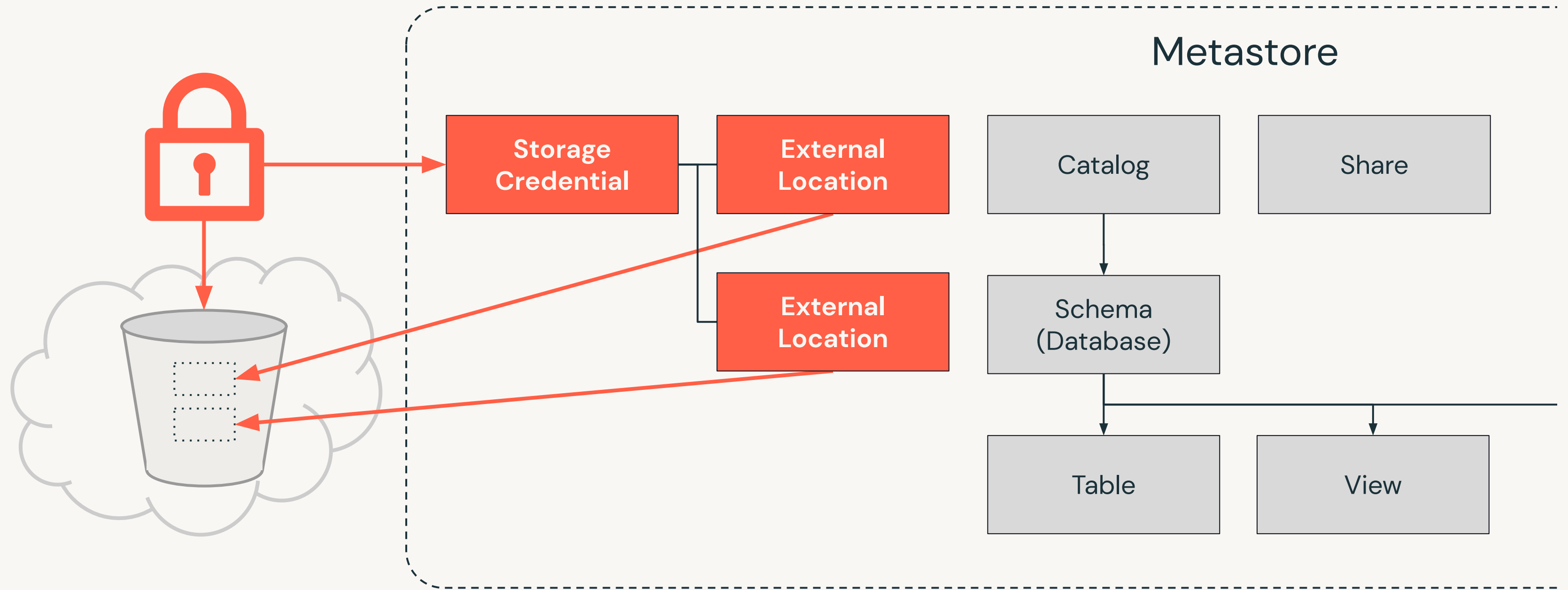
External Data Storage

External location



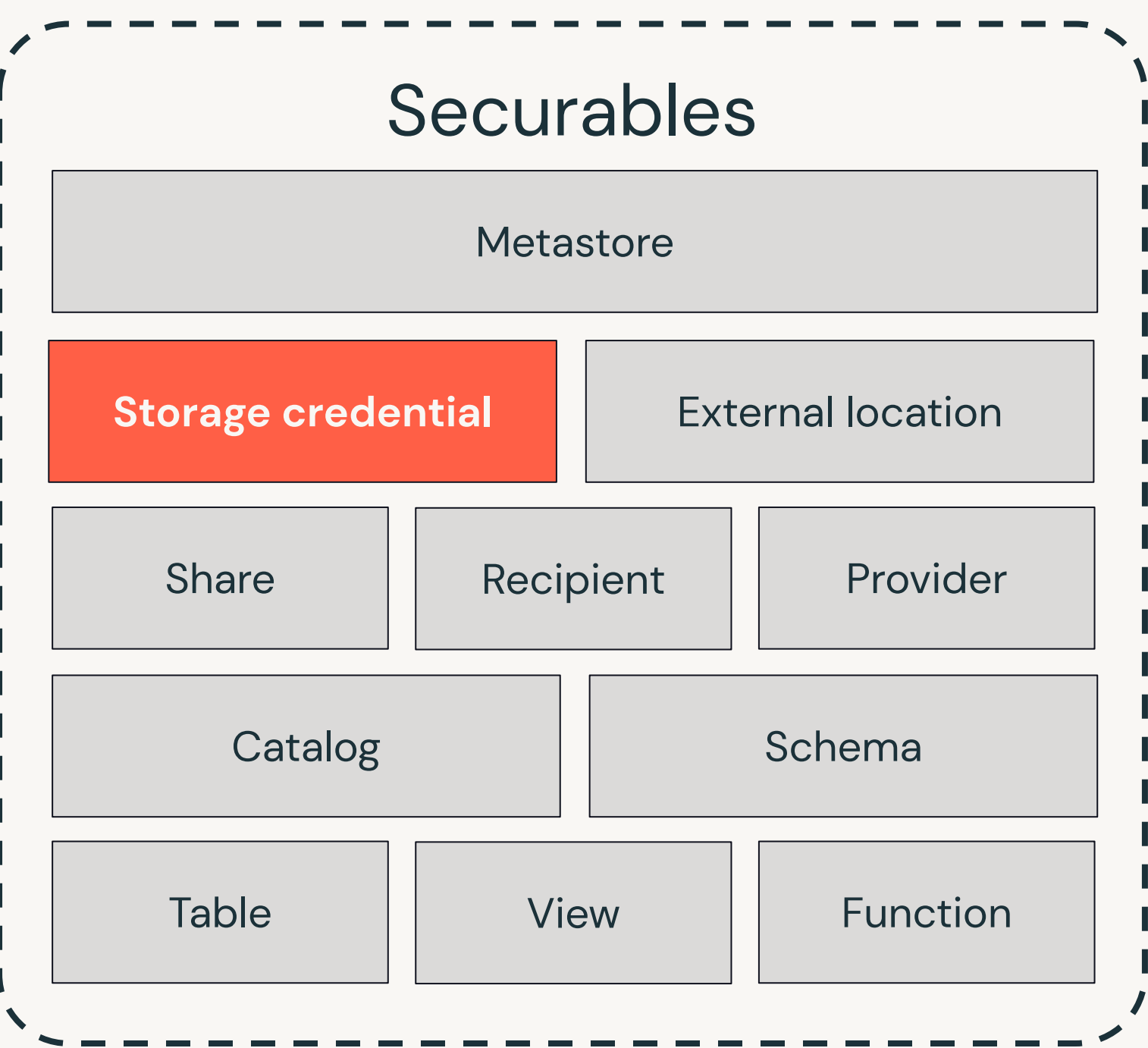
External Data Storage

External location



External Data Storage

Privilege types

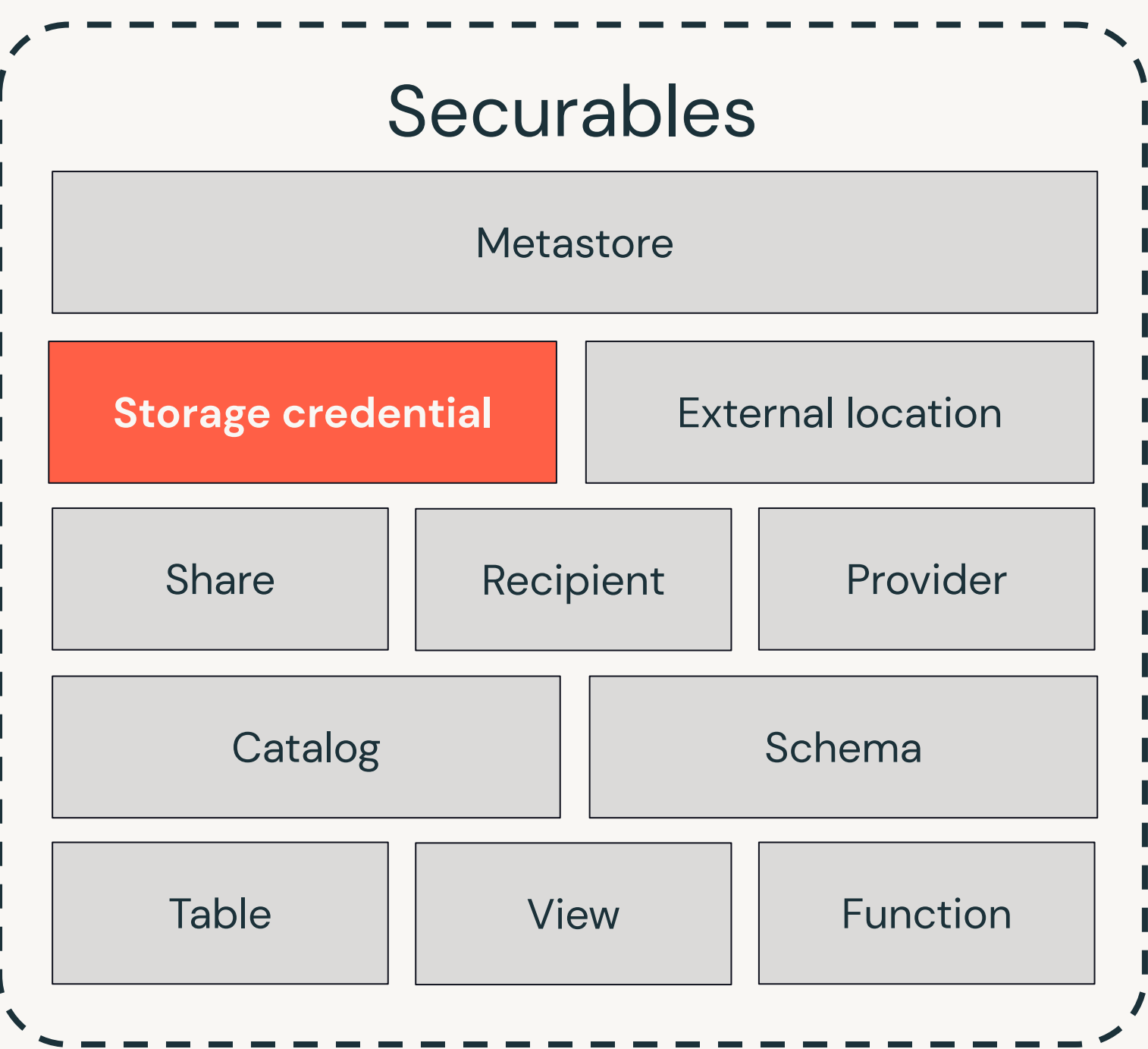


Privileges

CREATE EXTERNAL LOCATION
Create referencing external location

External Data Storage

Privilege types



Privileges

CREATE EXTERNAL LOCATION

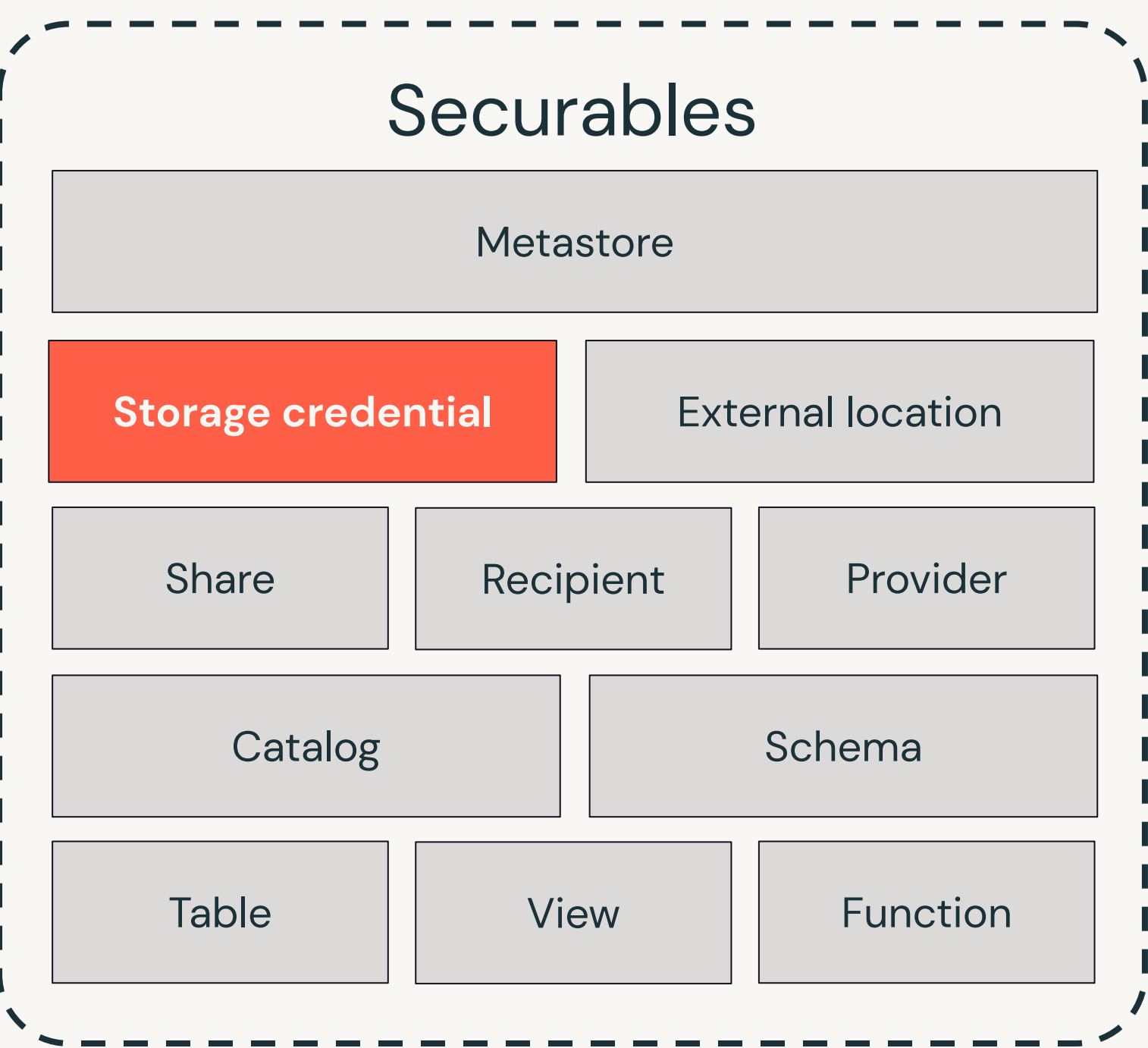
Create referencing external location

CREATE EXTERNAL TABLE

Create table based on contained files

External Data Storage

Privilege types



Privileges

CREATE EXTERNAL LOCATION

Create referencing external location

CREATE EXTERNAL TABLE

Create table based on contained files

READ FILES

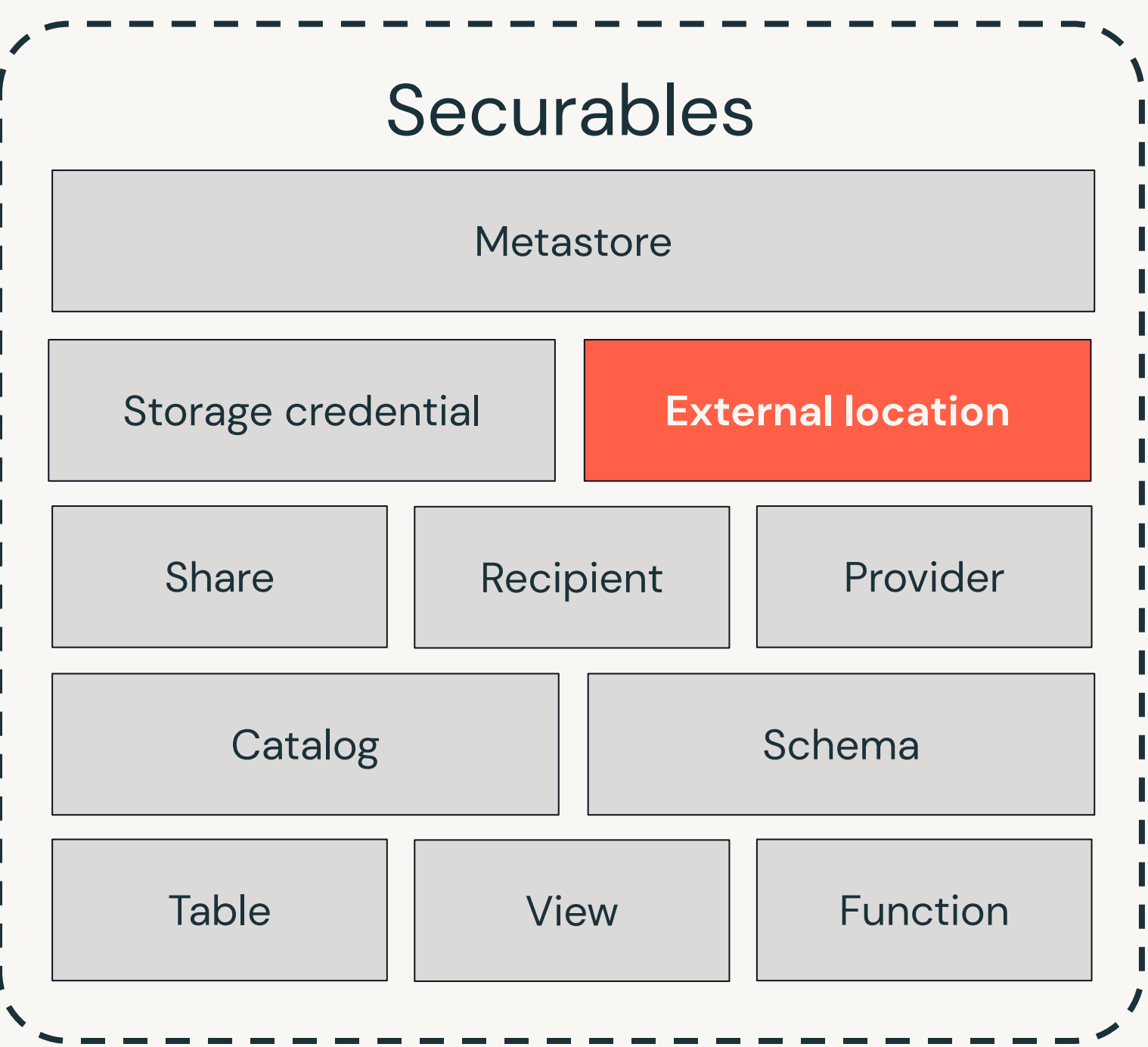
Read contained files

WRITE FILES

Modify contained files

External Data Storage

Privilege types



Privileges

CREATE EXTERNAL TABLE

Create table based on contained files

READ FILES

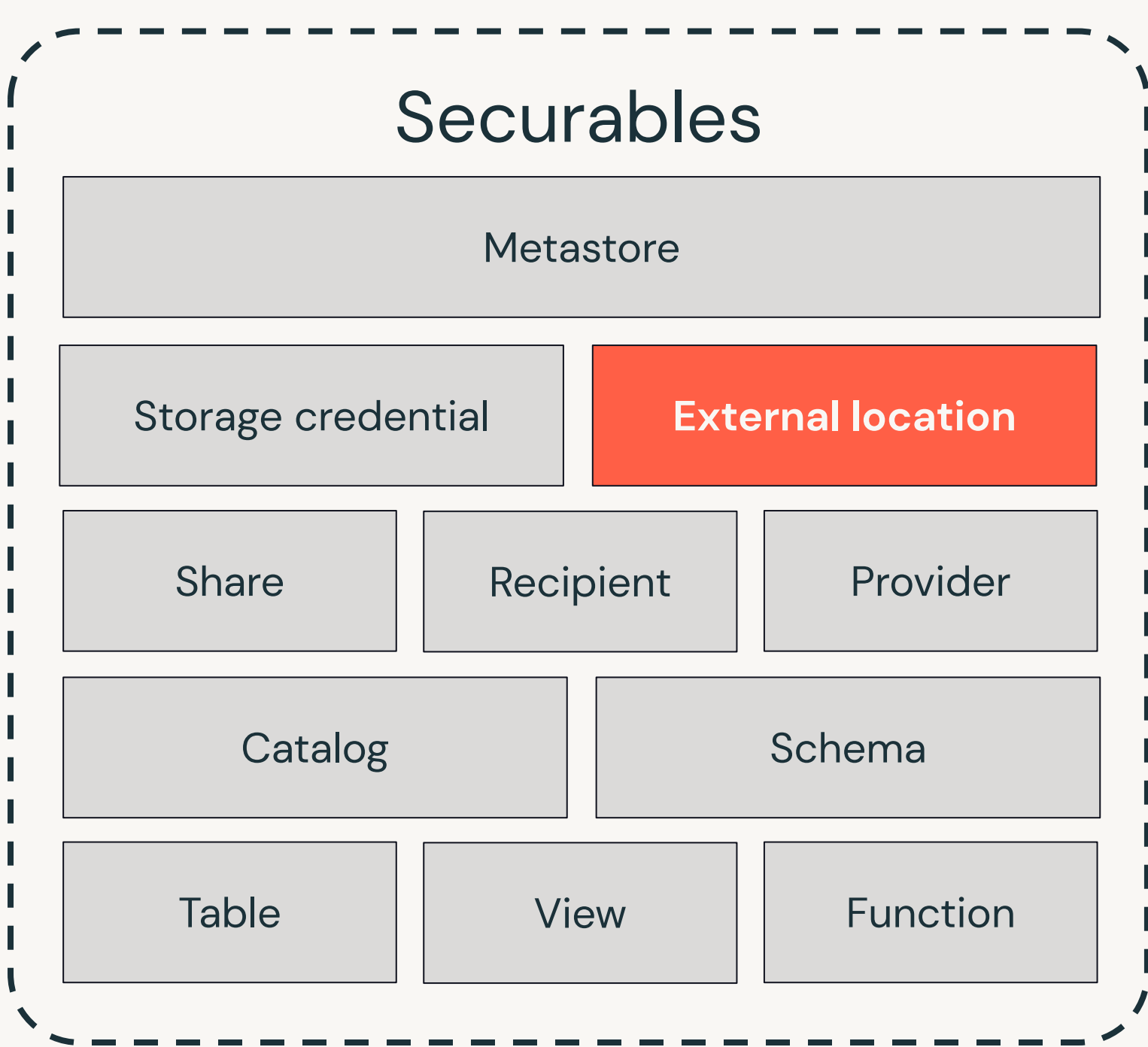
Read contained files

WRITE FILES

Modify contained files

External Data Storage

Privilege types



Privileges

CREATE EXTERNAL TABLE

Create table based on contained files

READ FILES

Read contained files

WRITE FILES

Modify contained files

CREATE MANAGED STORAGE

Create referencing external location

External Data Storage

Storage Credential

Enables Unity Catalog to access cloud data storage

Securable object

- ACL applies to entire storage container

External Location

Combines storage credential and cloud storage path

Securable object

- ACL applies to the portion of the storage container represented by the path

Lab: Managing external storage

Learning Objectives

In this lab, you will learn how to:

- Create and and manage a storage credential
- Create and and manage an external location
- Control access to files using an external location

Prerequisites

In order to do this lab, you will need:

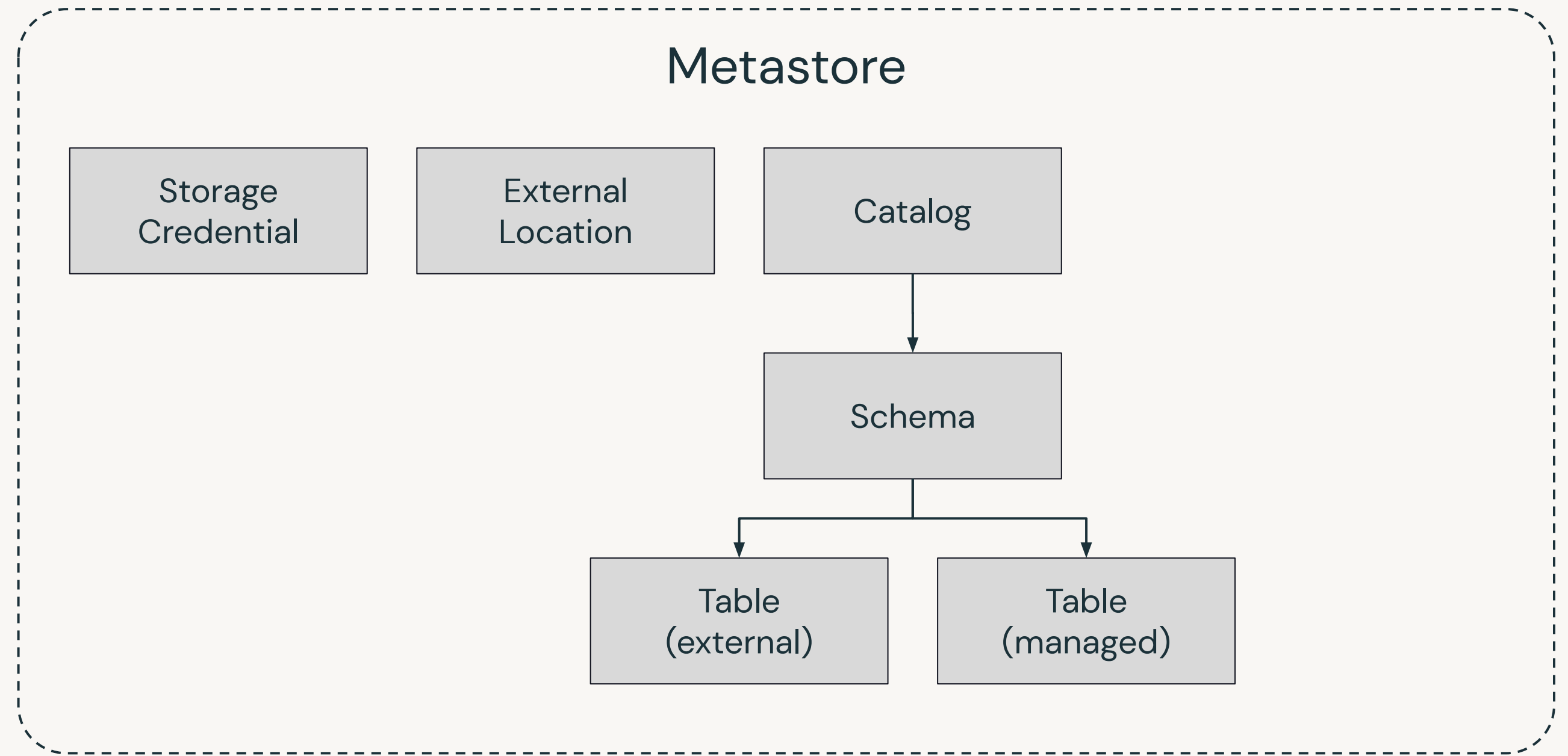
- Cloud resources to support an external storage container
- Account admin capability
- Workspace admin capability

The background is a dark blue-grey color. In the top right corner, there is a large orange circle. Below it, towards the center-right, is a smaller orange square. In the bottom right corner, there is a small orange circle. On the right side, there are several geometric shapes: a teal triangle pointing left, a red triangle pointing right, and a large orange triangle pointing left. The text "Managed and External Tables" is written in a large, white, sans-serif font on the left side of the image.

Managed and External Tables

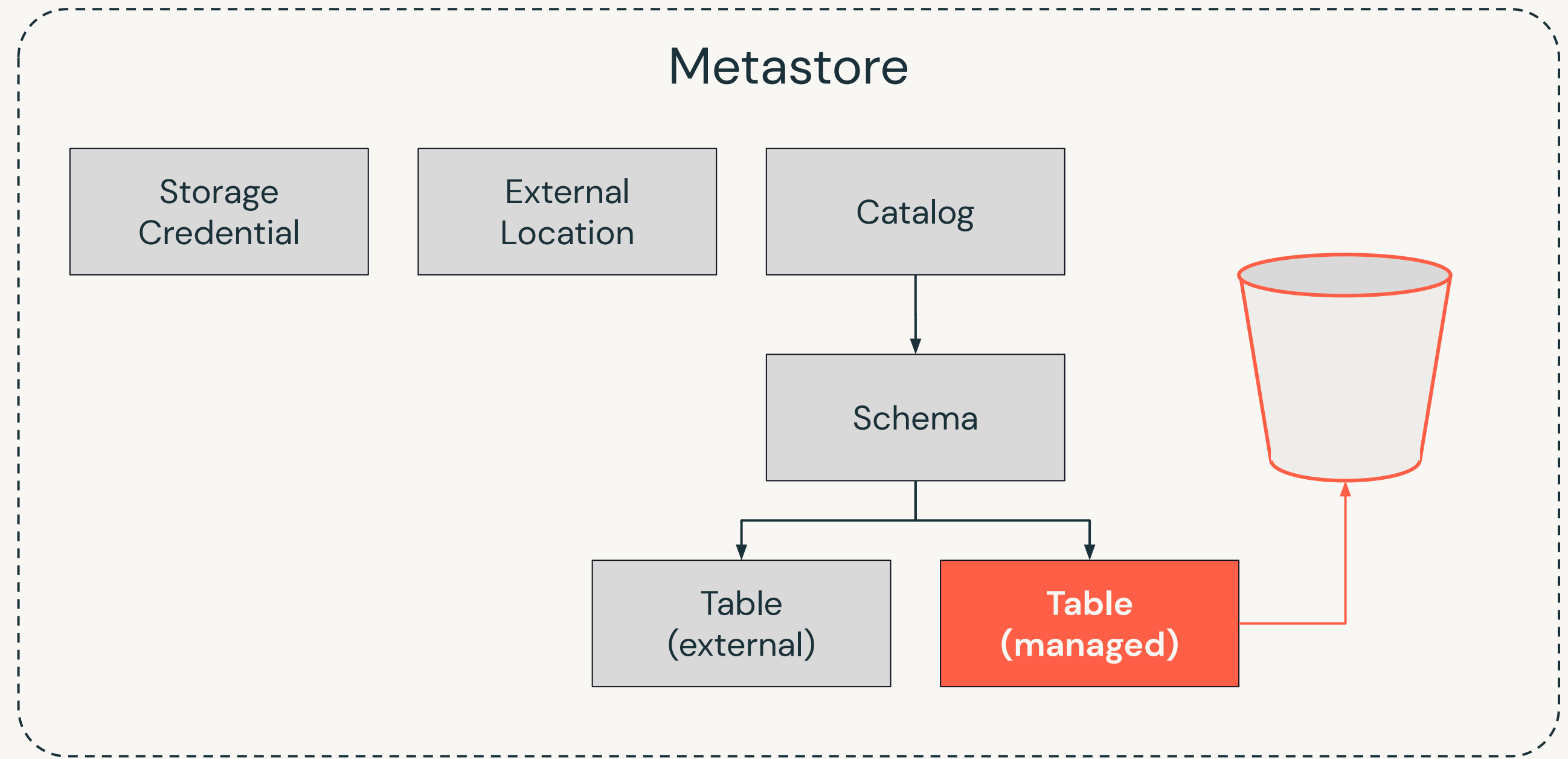
Managed and External Tables

Tables



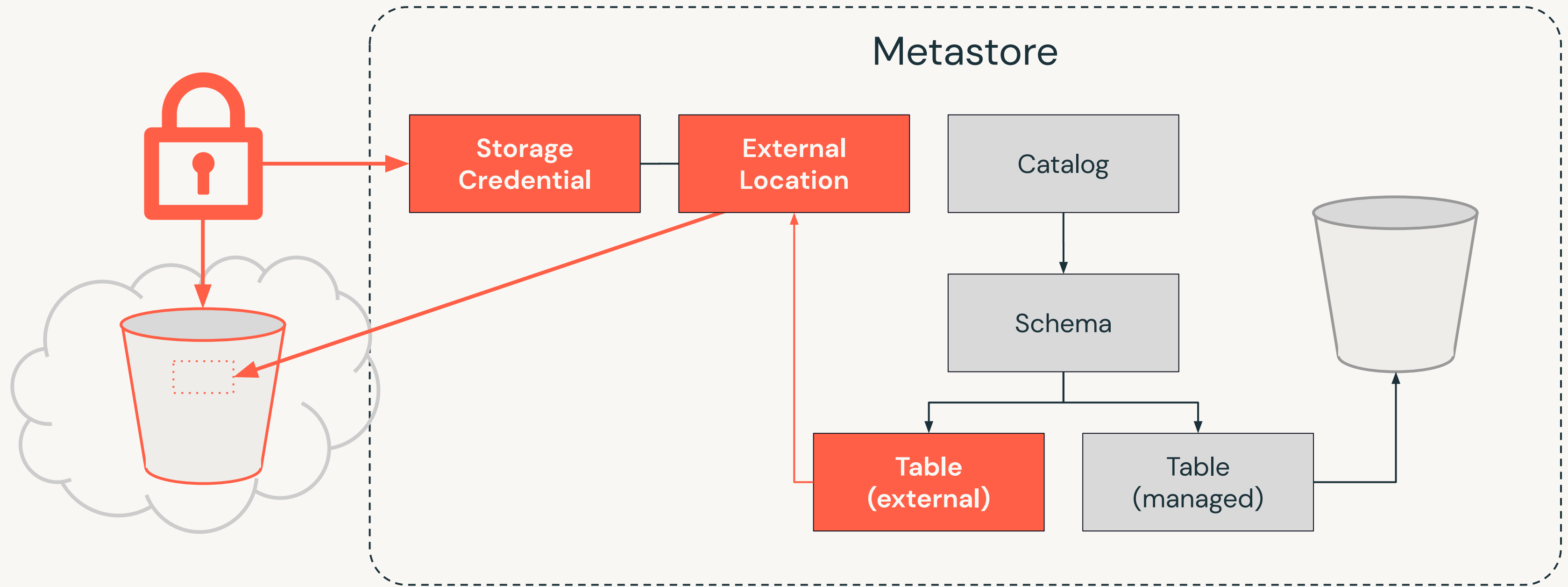
Managed and External Tables

Managed tables



Managed and External Tables

External tables



Managed and External Tables

Recap

Managed Tables

Metadata lives in control plane

Data lives in metastore managed storage location

DROP discards metadata and data

Delta format only

New product features supported first

External Tables

Metadata lives in control plane

Data lives in user-provided storage location

DROP discards metadata only

Several formats supported

Managed and External Tables

Recap

Managed Tables

Metadata lives in control plane

Data lives in metastore managed storage location

DROP discards metadata and data

Delta format only

New product features supported first

External Tables

Metadata lives in control plane

Data lives in user-provided storage location

DROP discards metadata only

Several formats supported

Managed and External Tables

When to use external tables

Quick and easy upgrade from external table in Hive metastore

Managed and External Tables

When to use external tables

Quick and easy upgrade from external table in Hive metastore

External writers

Managed and External Tables

When to use external tables

Quick and easy upgrade from external table in Hive metastore

External writers

Non-Delta support requirement

Managed and External Tables

When to use external tables

Quick and easy upgrade from external table in Hive metastore

External writers

Non-Delta support requirement

Requirement for specific storage naming or hierarchy

Infrastructure-level isolation requirements

Lab: Managing external tables

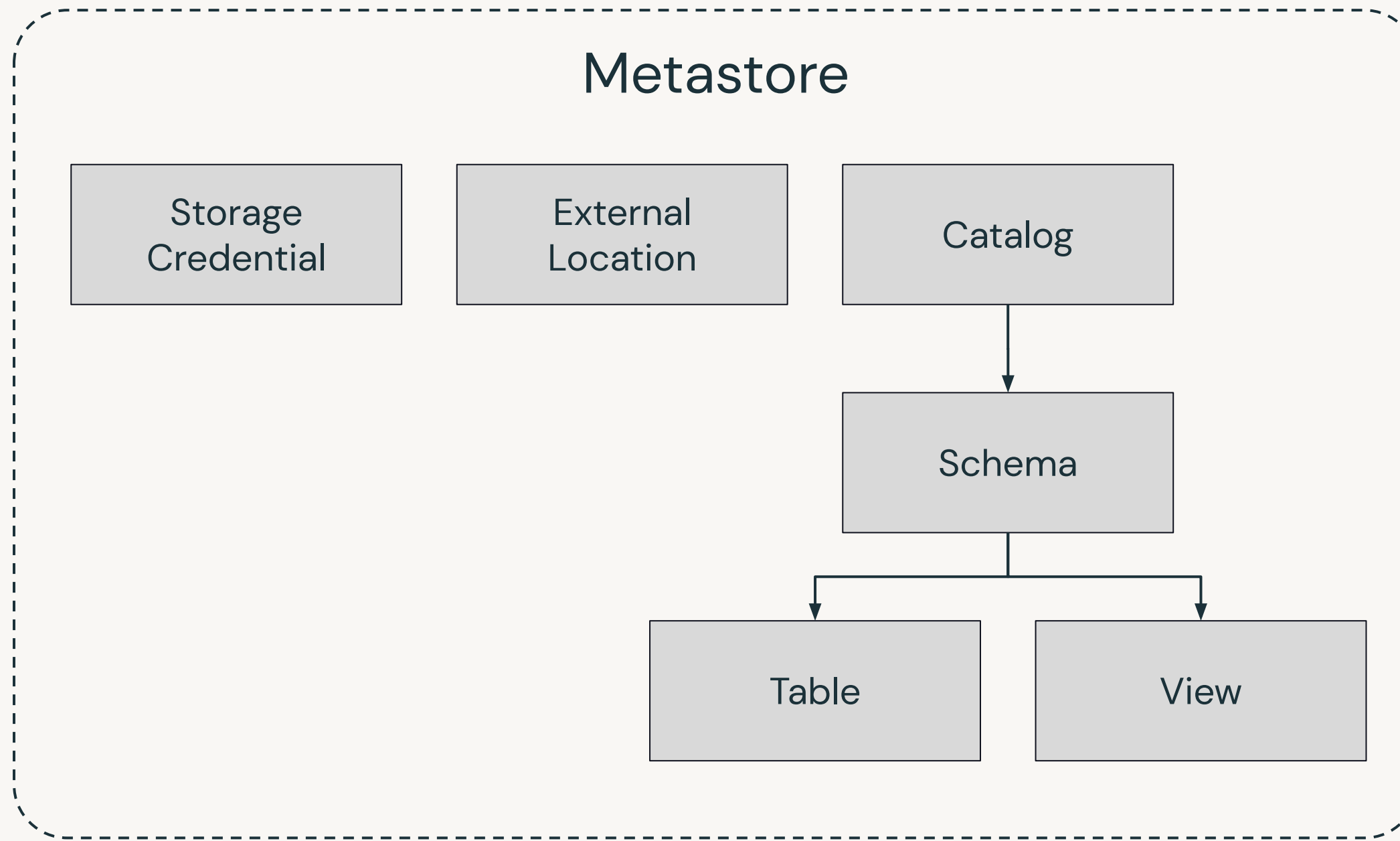
Learning Objectives

In this lab, you will learn how to:

- Create and and manage external tables
- Compare external and managed table behaviour

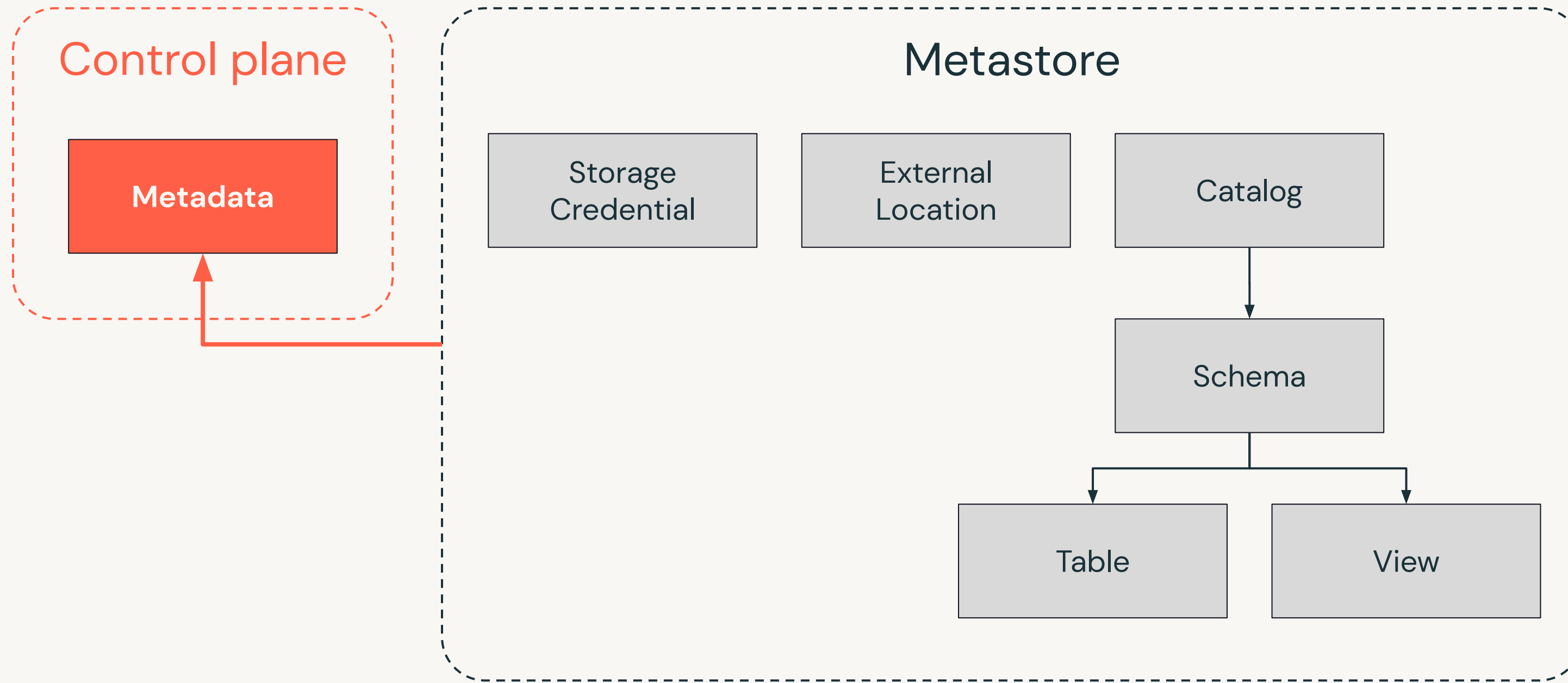
Data Segregation

Data Segregation



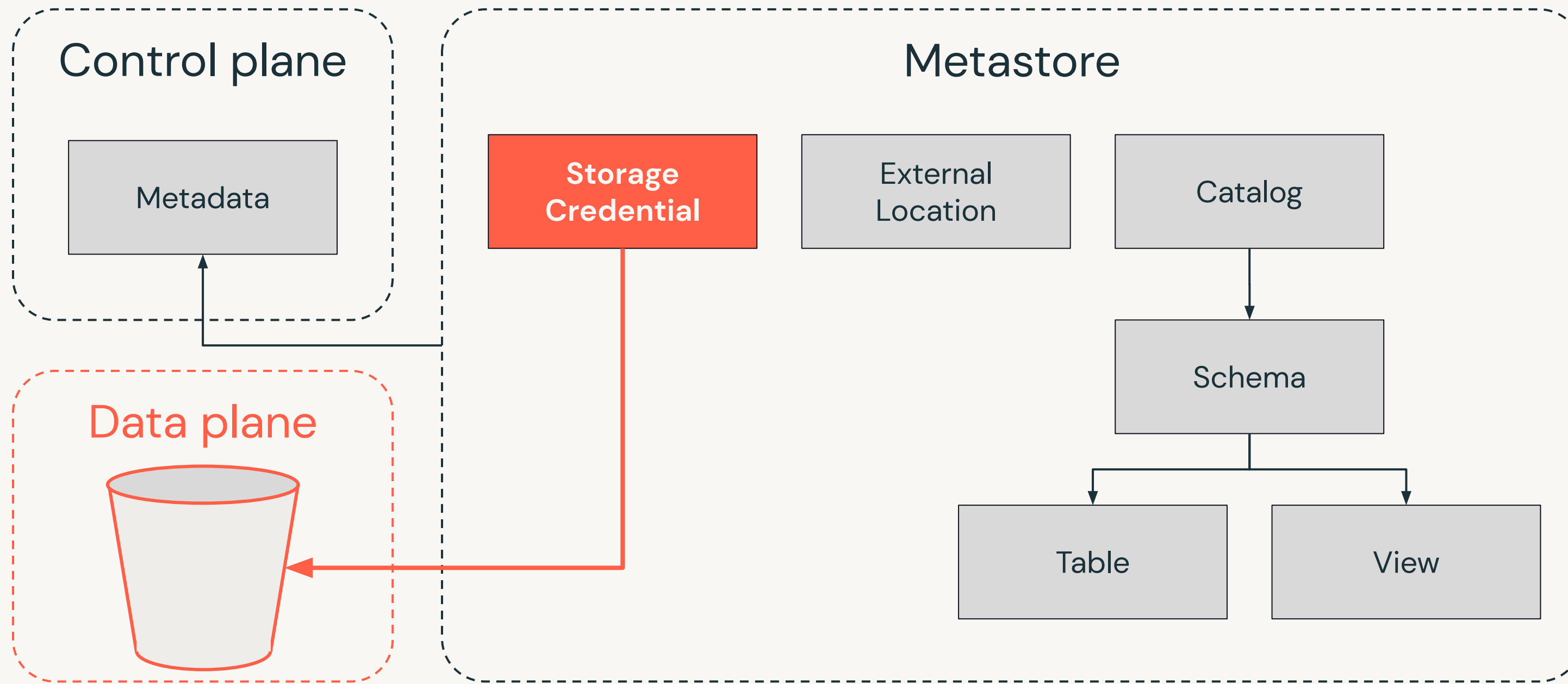
Data Segregation

Metastore requirements



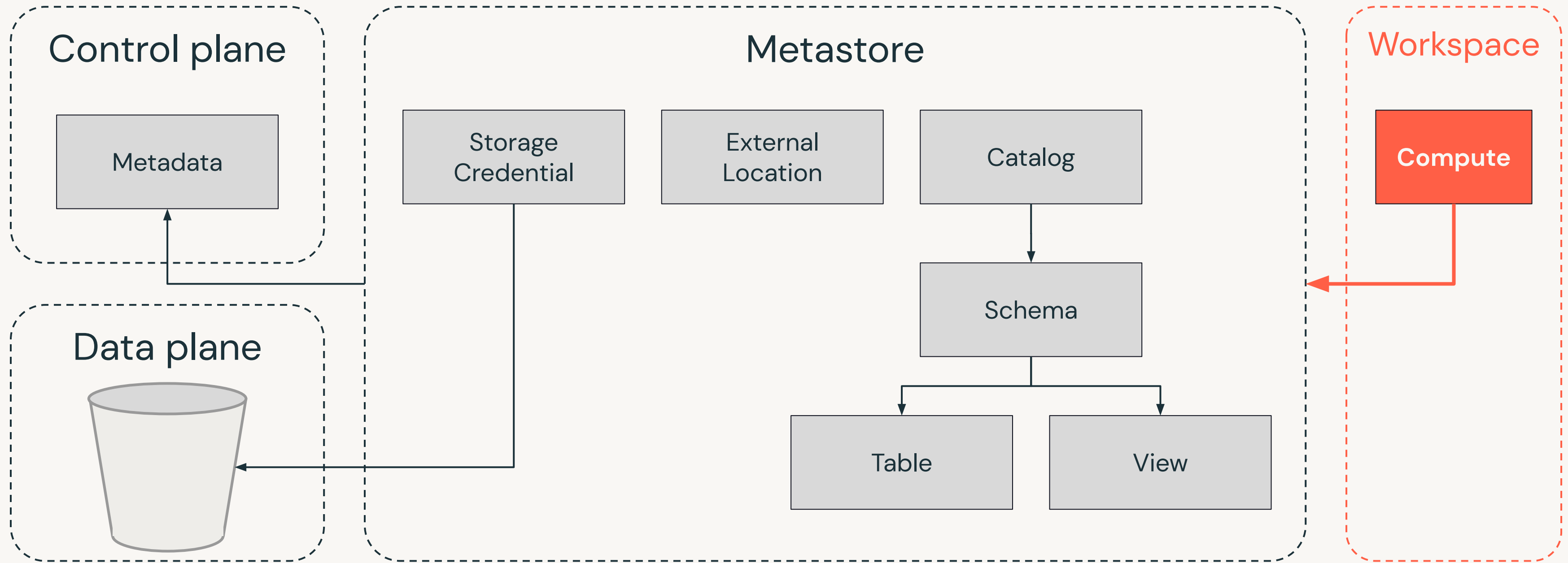
Data Segregation

Metastore requirements



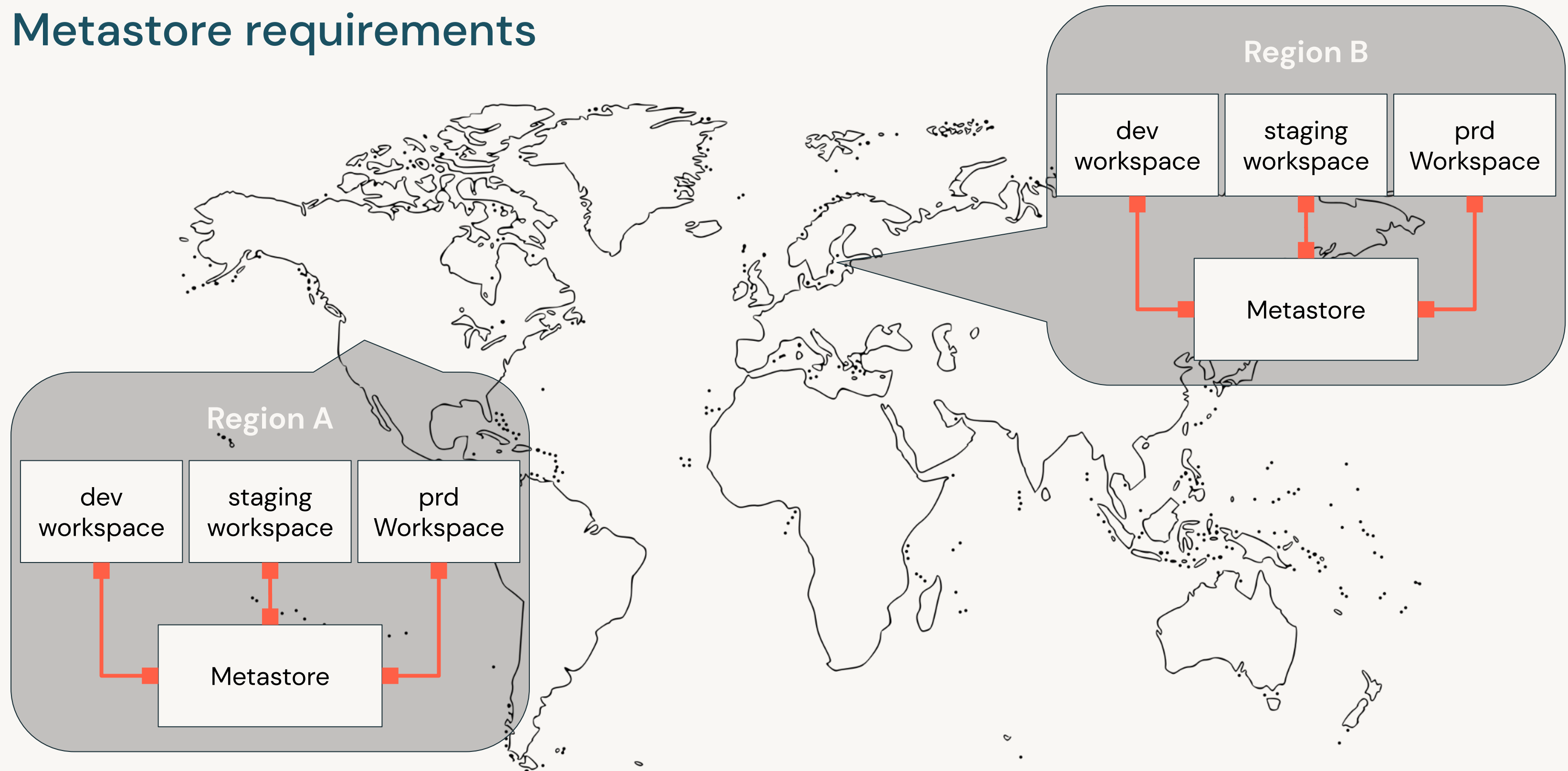
Data Segregation

Metastore requirements



Data Segregation

Metastore requirements



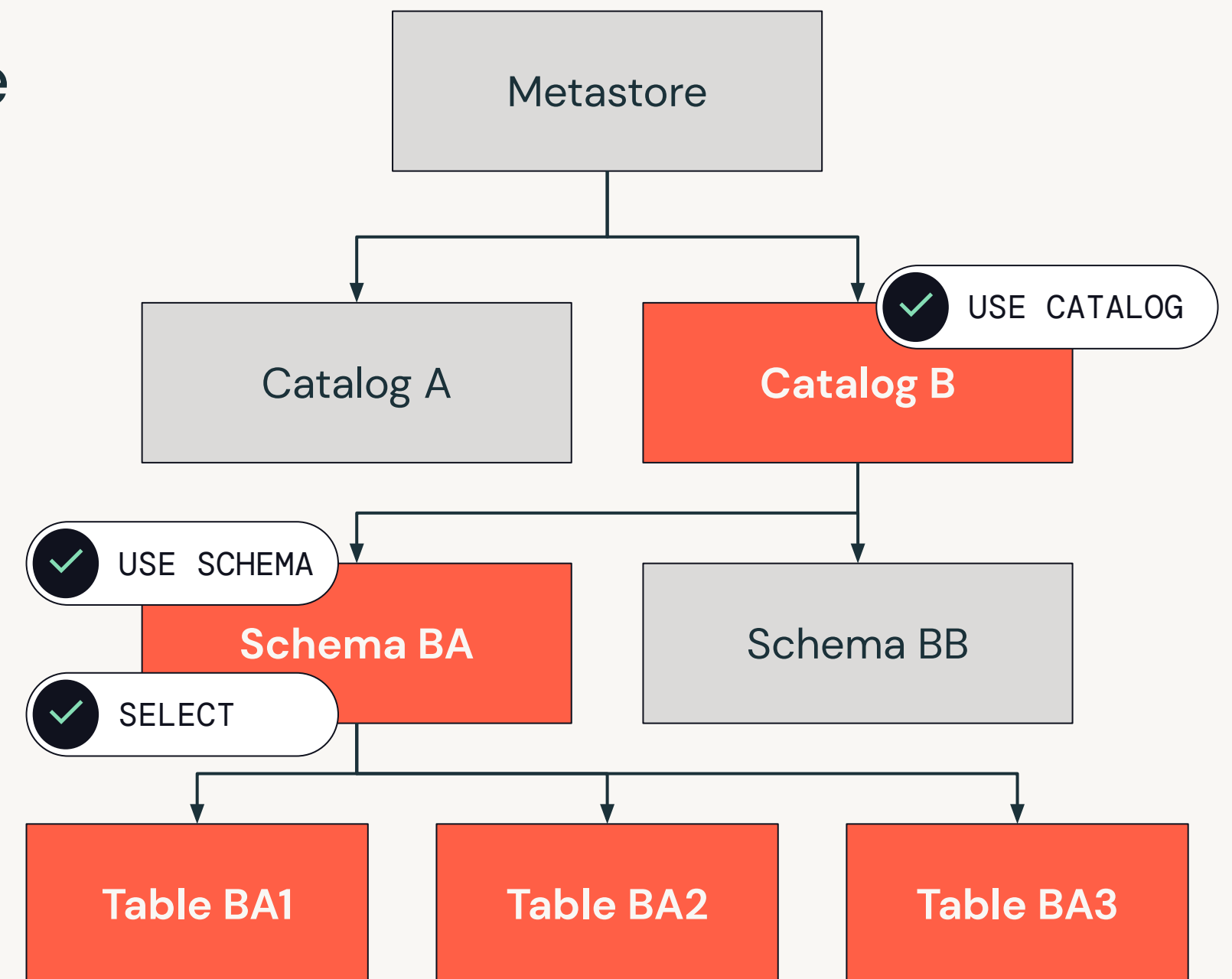
Data Segregation

Logical separation

Use catalogs and/or schemas to segregate data

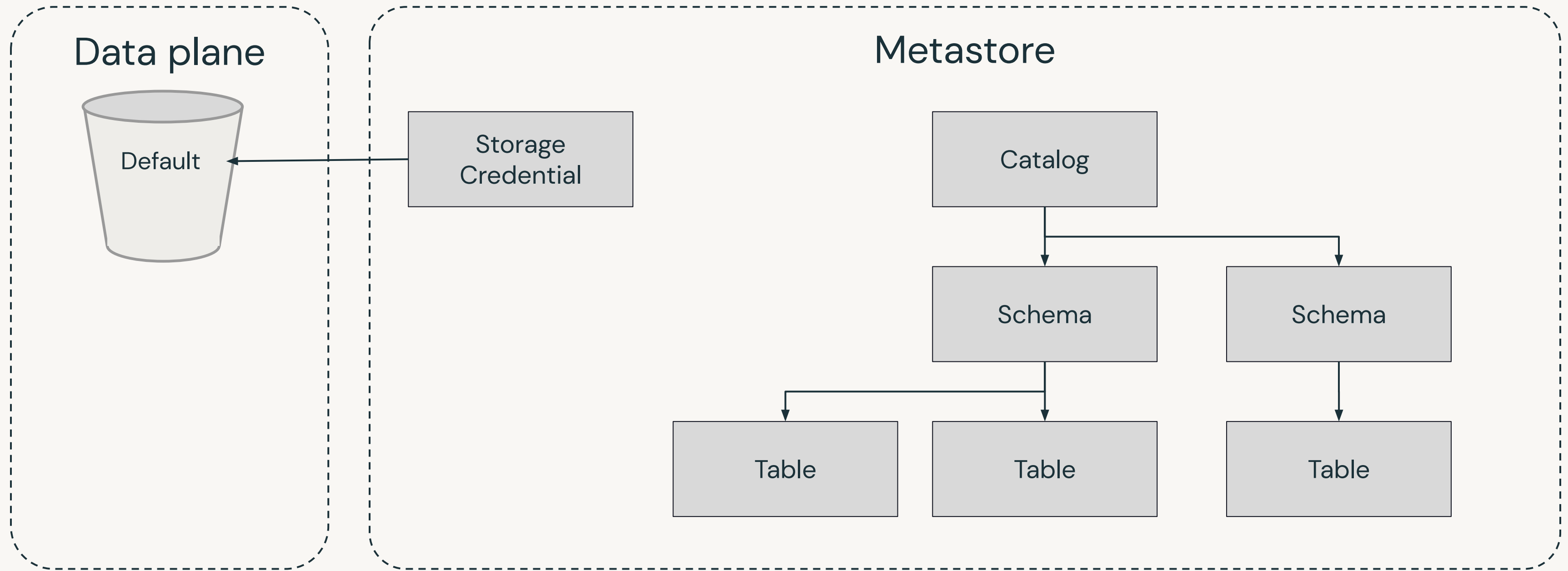
Apply permissions appropriately

- **USE CATALOG** on catalog B
- **USE SCHEMA** on schema BA
- **SELECT** on schema BA



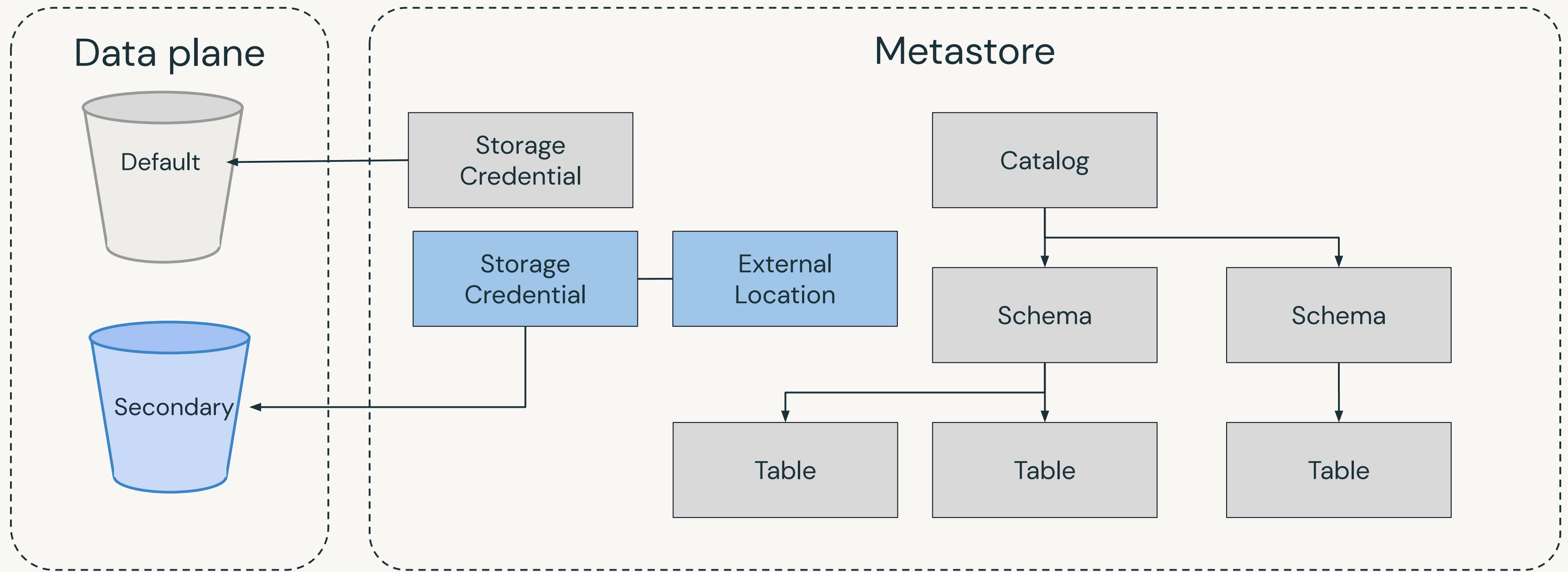
Data Segregation

Physical separation



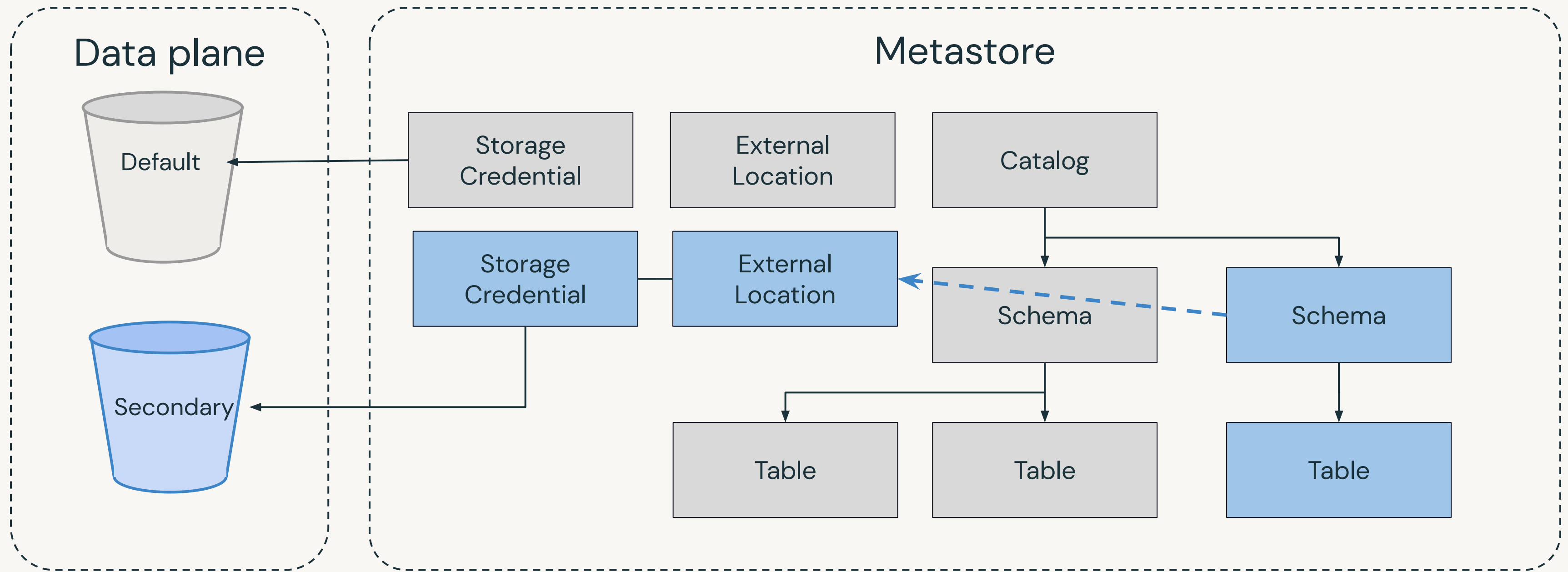
Data Segregation

Physical separation



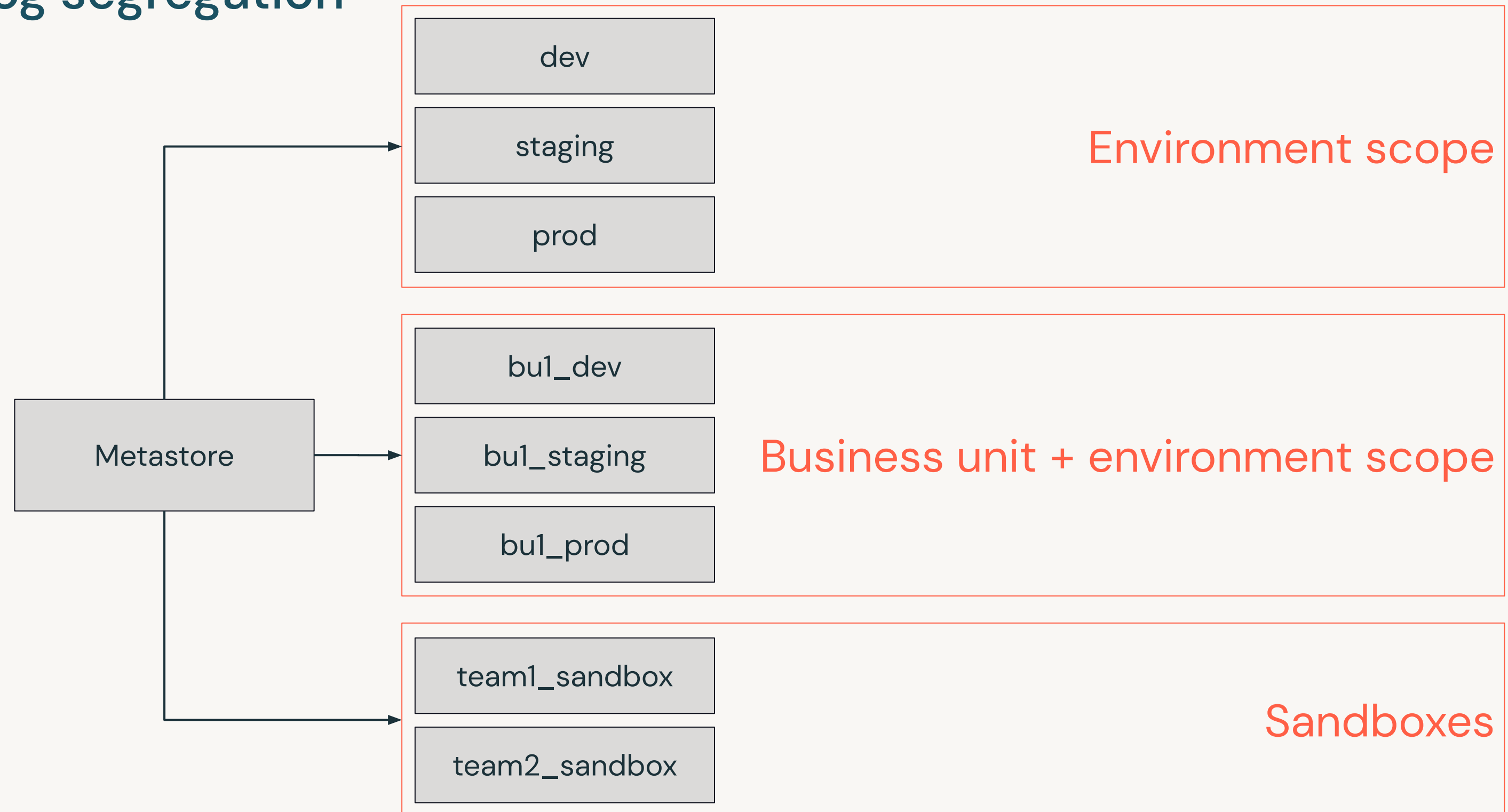
Data Segregation

Physical separation



Data Segregation

Catalog segregation



Lab: Segregating data

Learning Objectives

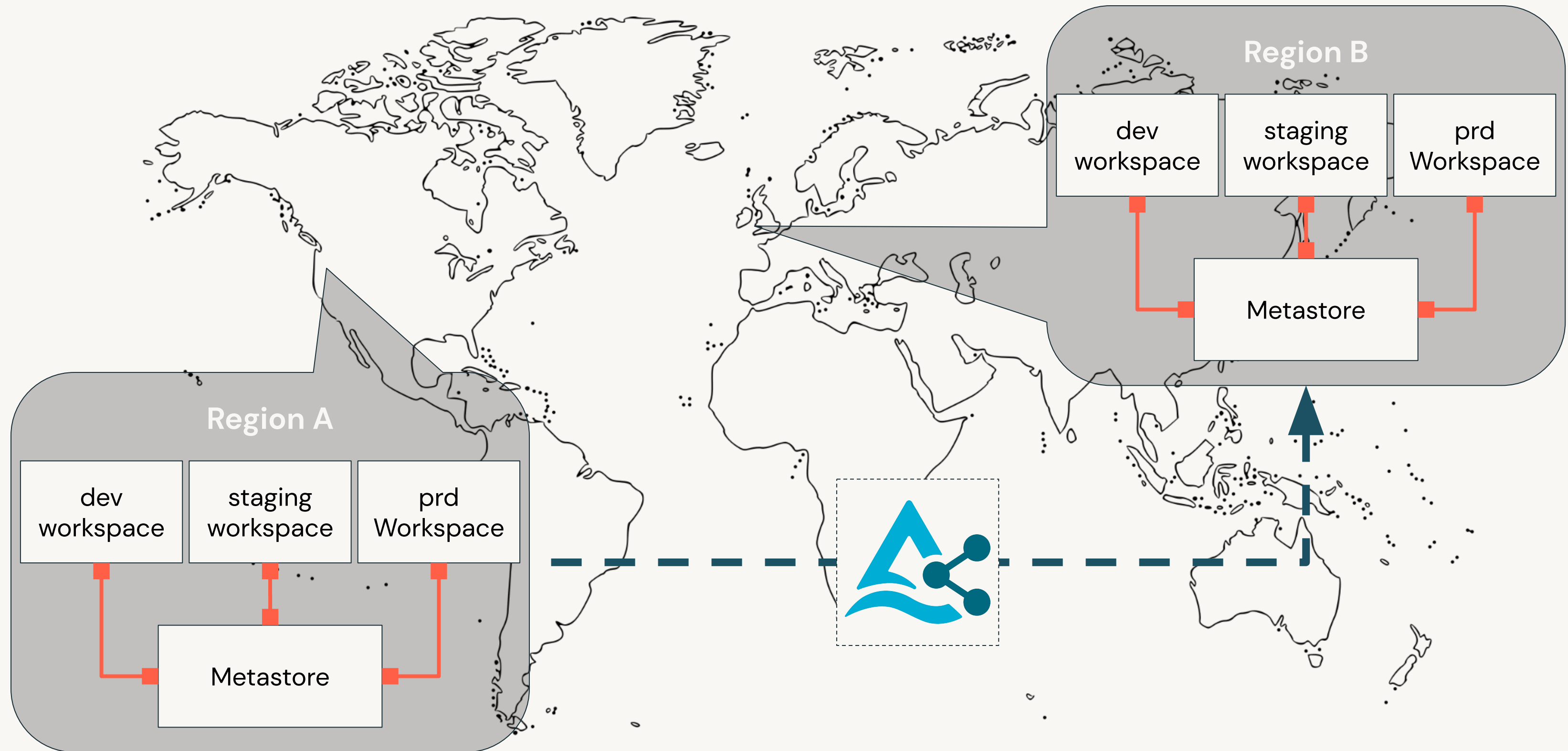
In this lab, you will learn how to:

- Redirect schemas and catalogs to external storage

Lab: Upgrading tables to Unity Catalog

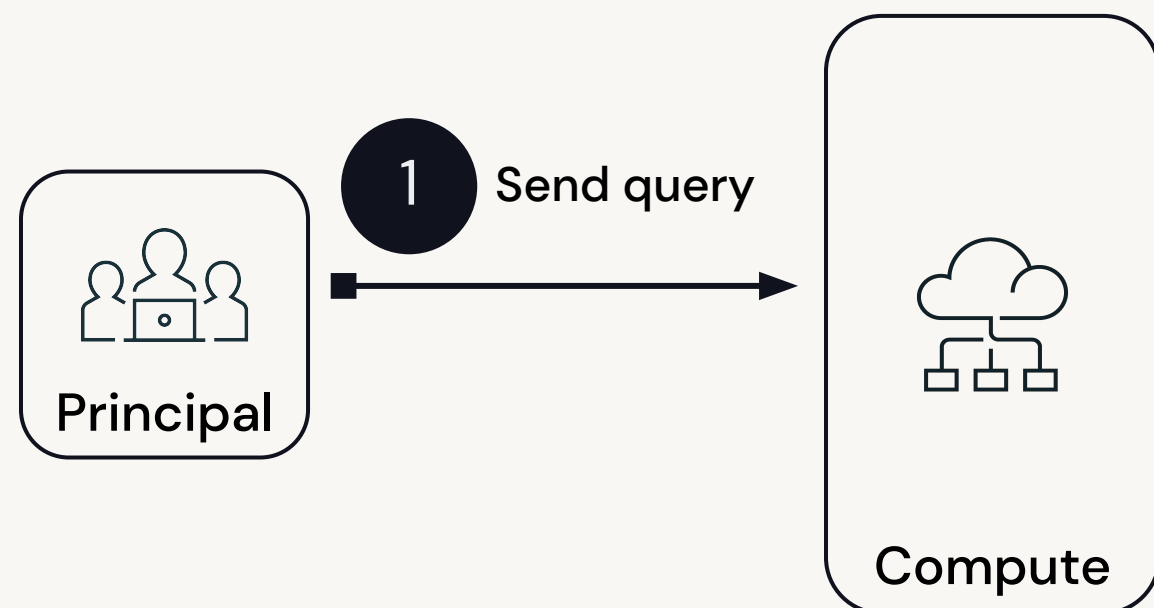


Metastore Location



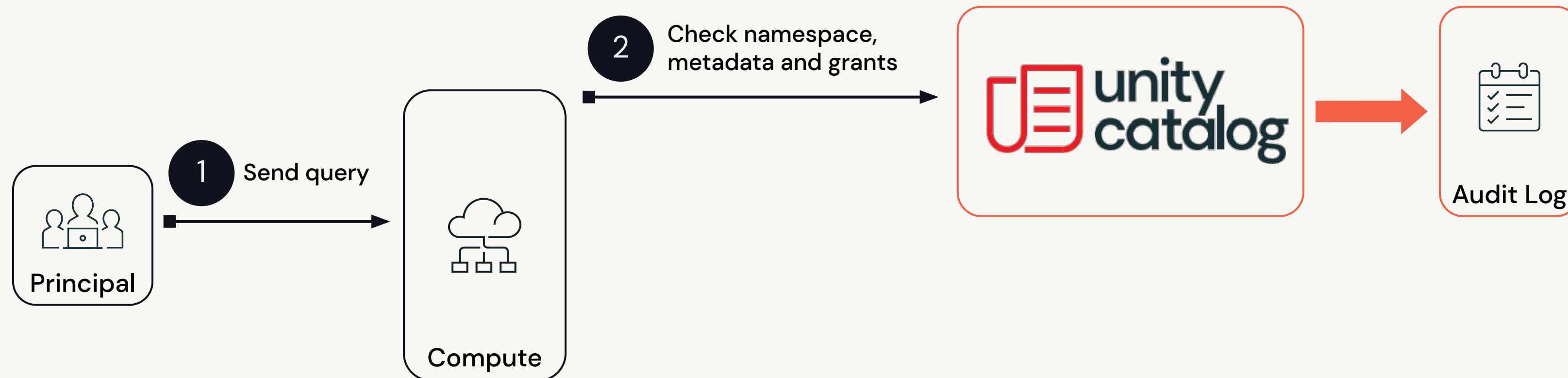
Data Security Model

Query Lifecycle



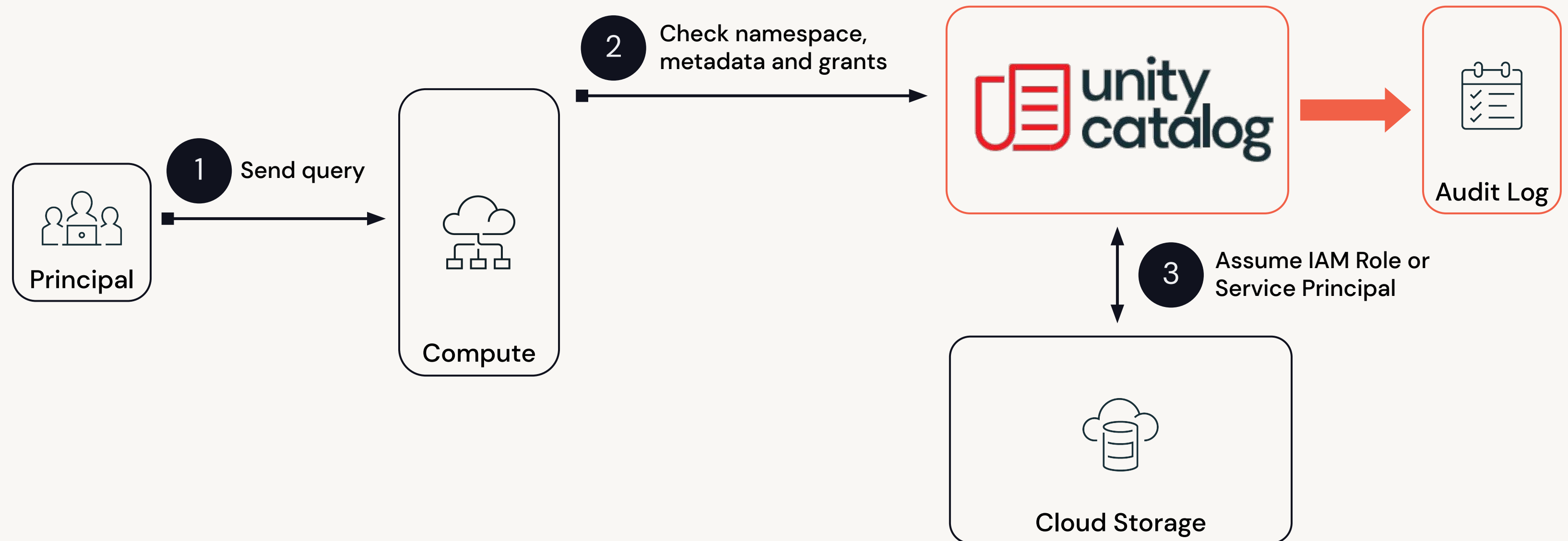
Data Security Model

Query Lifecycle



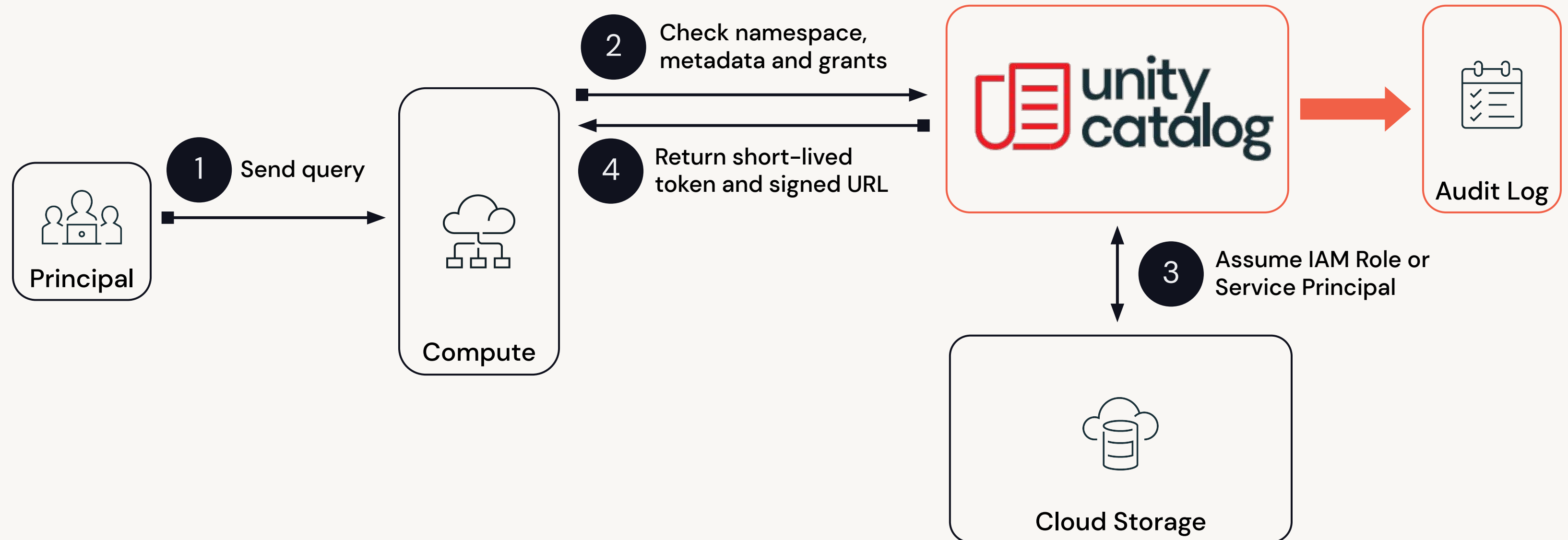
Data Security Model

Query Lifecycle



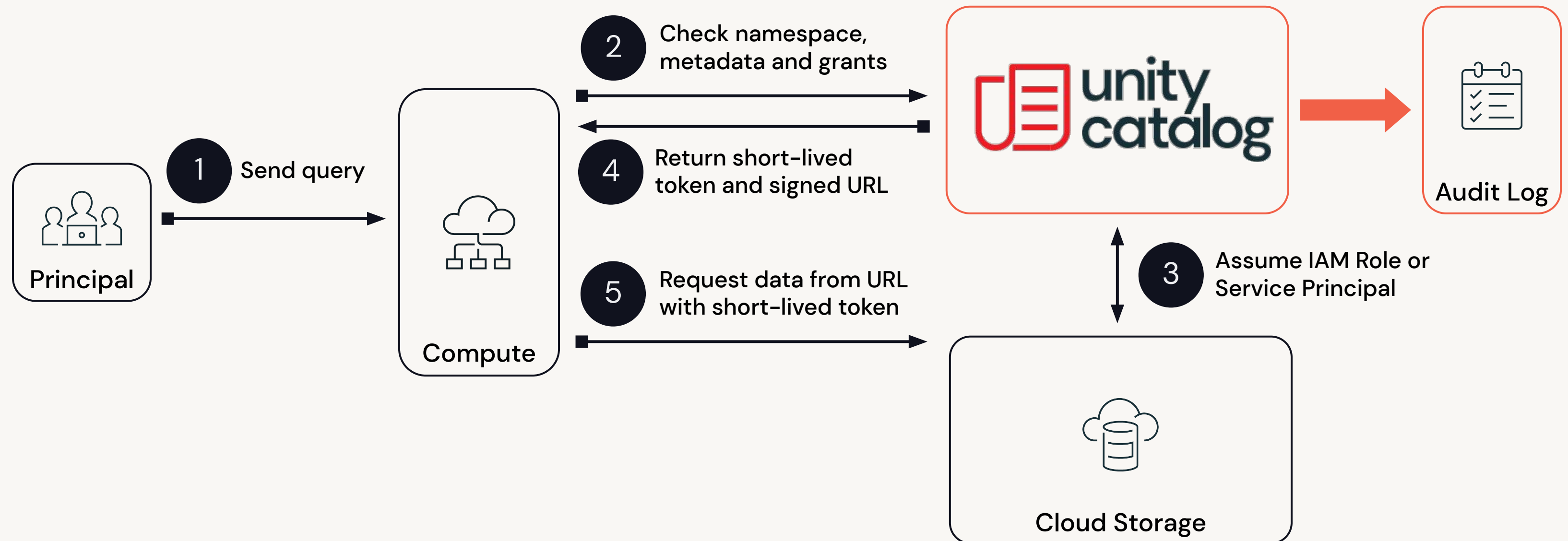
Data Security Model

Query Lifecycle



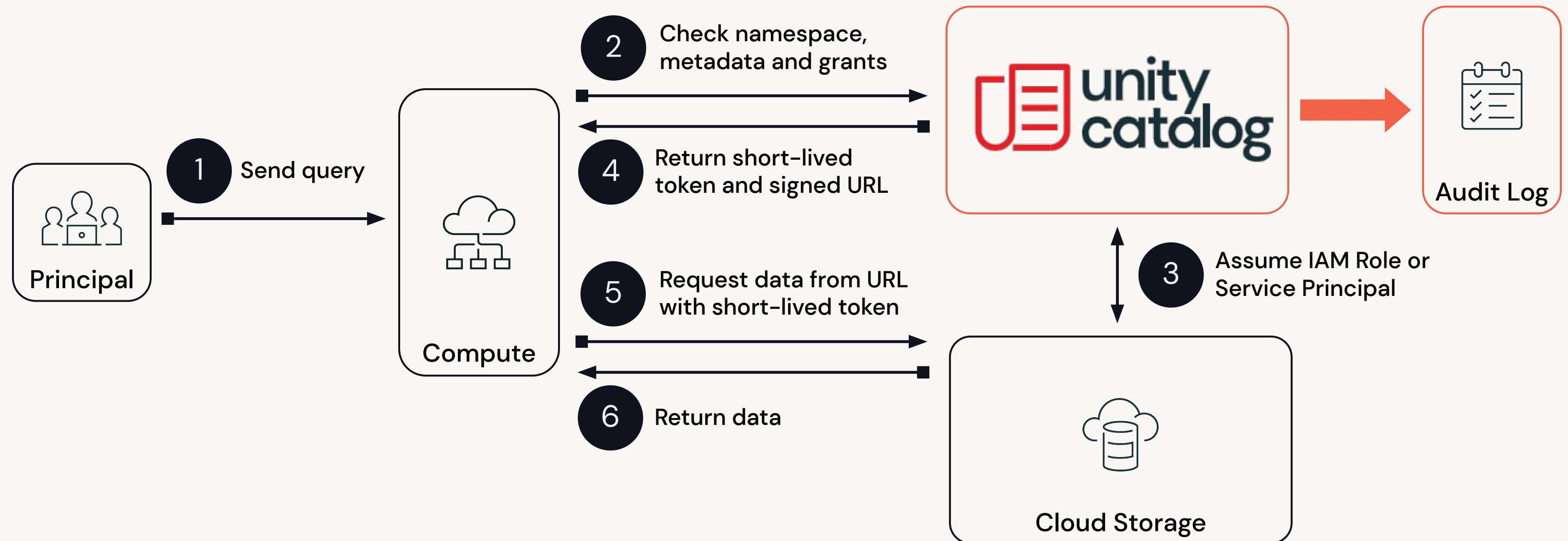
Data Security Model

Query Lifecycle



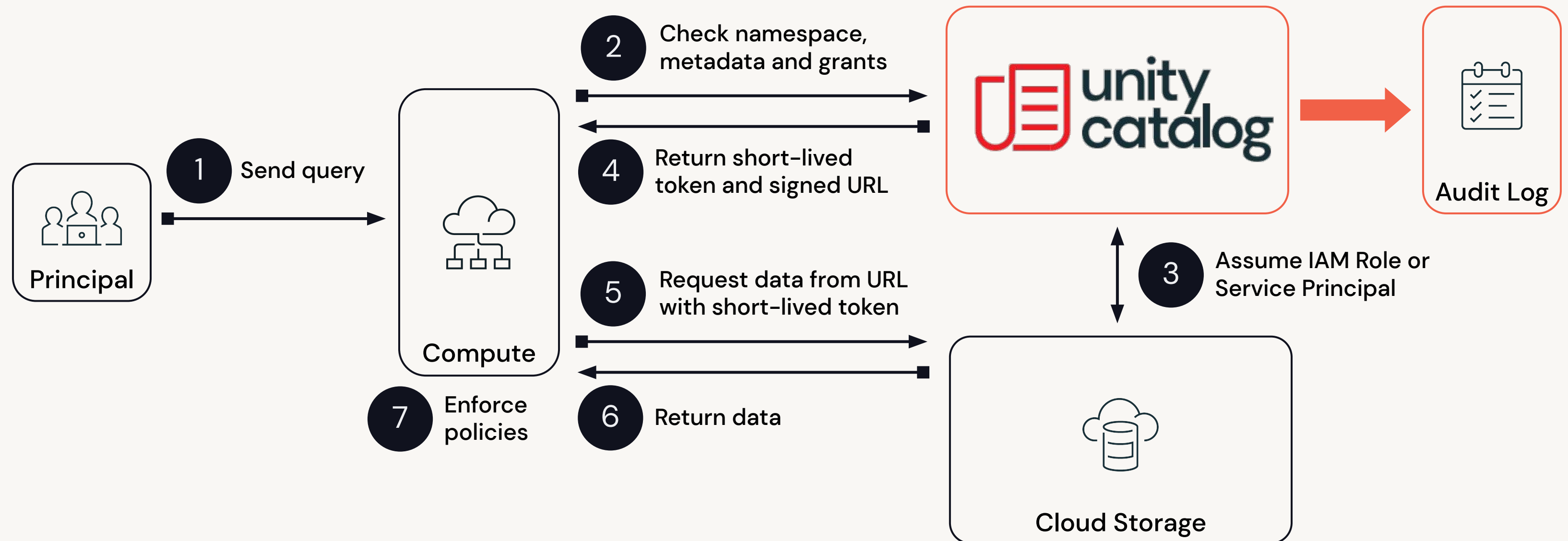
Data Security Model

Query Lifecycle



Data Security Model

Query Lifecycle



Data Security Model

Query Lifecycle

