

# Aws lake formation

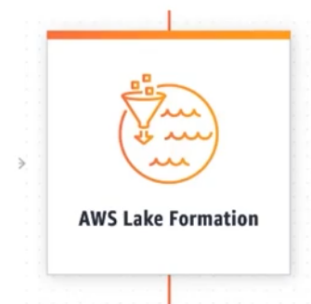
---

aws lake formation you can secure data lake in a days

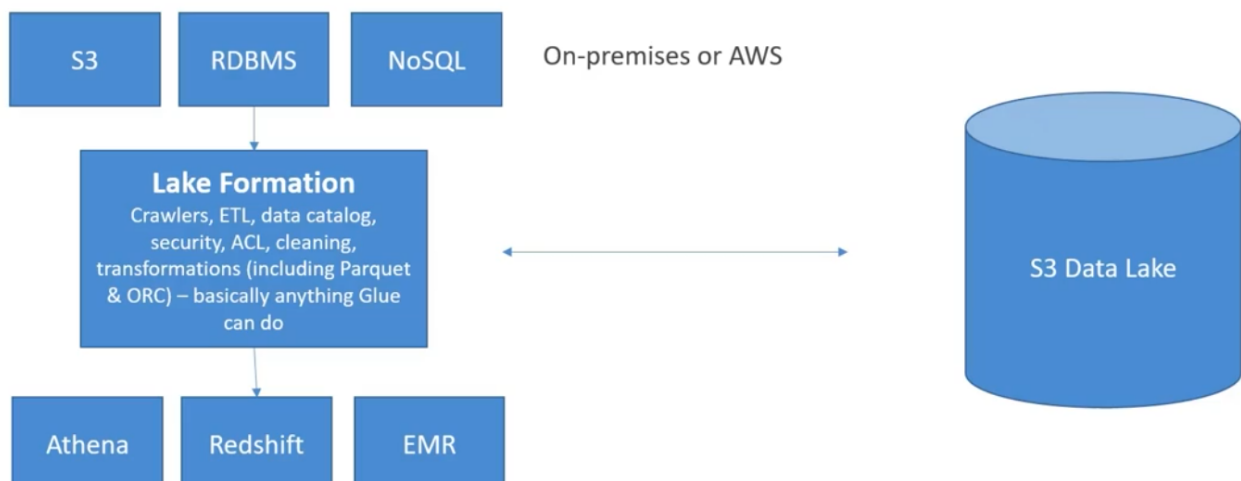
you cannot create a data lake in day it takes to more than week are more than few days

## AWS Lake Formation

- “Makes it easy to set up a secure data lake in days”
- Loading data & monitoring data flows
- Setting up partitions
- Encryption & managing keys
- Defining transformation jobs & monitoring them
- Access control
- Auditing
- Built on top of Glue



## AWS Lake Formation



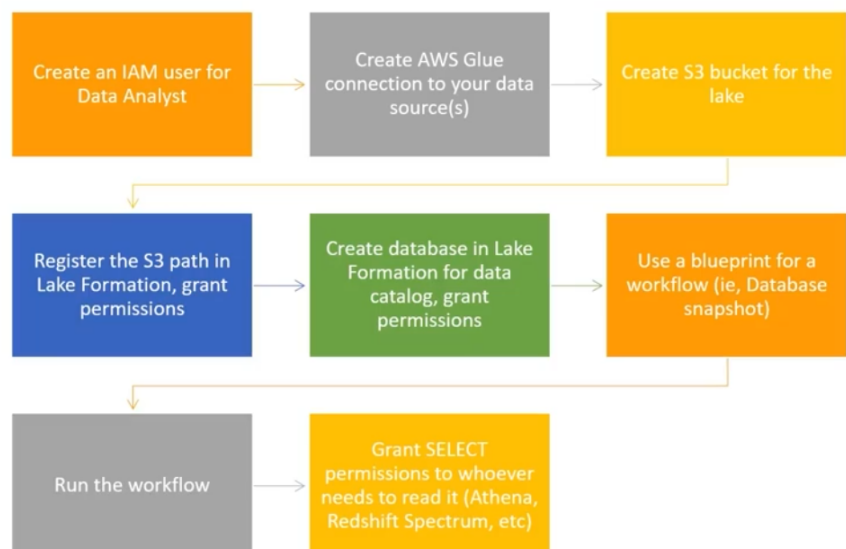
after create lake formation you can access are query the data using redshift athe are emr (emr also access the data lake to query using hive ,spark , etc..)

# AWS Lake Formation: Pricing

- No cost for Lake Formation itself
- But underlying services incur charges
  - Glue
  - S3
  - EMR
  - Athena
  - Redshift



## AWS Lake Formation: Building a Data Lake



## AWS Lake Formation: The Finer Points

- Cross-account Lake Formation permission
  - Recipient must be set up as a data lake administrator
  - Can use AWS Resource Access Manager for accounts external to your organization
  - IAM permissions for cross-account access
- Lake Formation does not support manifests in Athena or Redshift queries
- IAM permissions on the KMS encryption key are needed for encrypted data catalogs in Lake Formation
- IAM permissions needed to create blueprints and workflows

# AWS Lake Formation: Governed Tables and Security

- Now supports “Governed Tables” that support ACID transactions across multiple tables
  - New type of S3 table
  - Can't change choice of governed afterwards
  - Works with streaming data too (Kinesis)
  - Can query with Athena
- Storage Optimization with Automatic Compaction
- Granular Access Control with Row and Cell-Level Security
  - Both for governed and S3 tables
- Above features incur additional charges based on usage

## Data Permissions in Lake Formation

- Can tie to IAM users/roles, SAML, or external AWS accounts

The screenshot shows the 'Grant Permissions' console in AWS Lake Formation. The interface is divided into several sections:

- Principals:** A section for selecting the principal to grant permissions to. It includes three radio buttons: 'IAM users and roles' (selected), 'IAM users and groups' (disabled), and 'External accounts' (disabled). Below these is a dropdown menu for 'Choose IAM principal to add' and a 'Require it admin' button.
- Policy tags or catalog resources:** A section for selecting the policy tags or catalog resources. It includes two radio buttons: 'Resources matched by policy tags' (selected) and 'Normal data catalog resources' (disabled). Below these is a dropdown menu for 'Choose database' and a 'Require it admin' button.
- Permissions:** A section for selecting the permissions to grant. It includes two radio buttons: 'Table permissions' (selected) and 'Column-based permissions' (disabled). Below these is a dropdown menu for 'Choose database' and a 'Require it admin' button.
- Table permissions:** A section for selecting the table permissions to grant. It includes a dropdown menu for 'Choose database' and a 'Require it admin' button.
- Grantable permissions:** A section for selecting the grantable permissions to grant. It includes a dropdown menu for 'Choose database' and a 'Require it admin' button.

At the bottom of the console, there are 'Cancel' and 'Grant' buttons.