Create a Redshift Cluster

Overview of Redshift

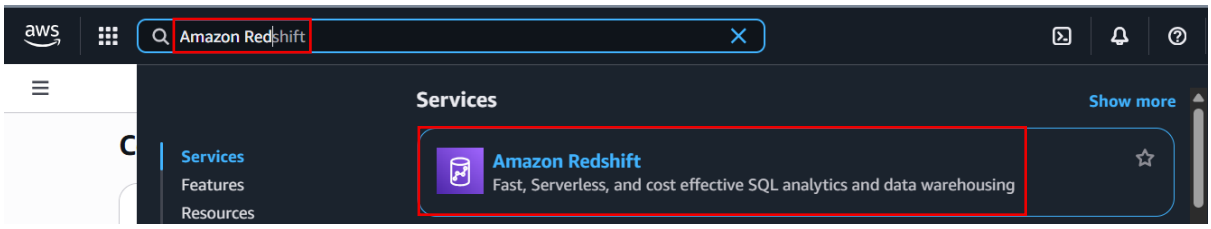
- **Amazon Redshift** is a **fast, scalable, and cost-effective data warehousing service** by AWS.
- Allows performing **complex queries** and **data analytics** on large datasets.
- Offers two deployment models:
    - **Redshift Serverless** – fully managed, pay-per-use model (no infrastructure management).
    - **Provisioned Cluster** – user manages nodes, performance, and scaling manually.
- This lab focuses on the **provisioned cluster** setup for better control and understanding.

To Begin with the Lab

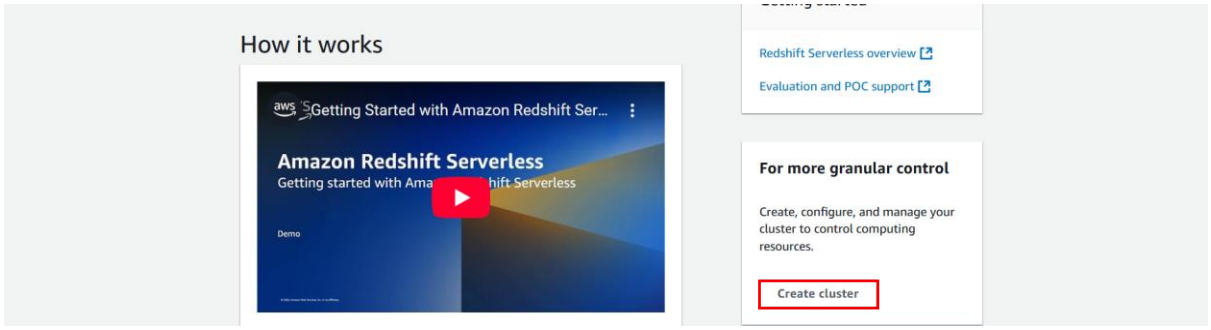
Summary of the Lab

Amazon Redshift is a fully managed, scalable, and cost-efficient data warehousing service for analytics on large datasets. In this lab, a **provisioned Redshift cluster** is created for greater control. The setup involves choosing a small, low-cost node type, configuring one node, and optionally loading sample data. Users set a database username and password, create an IAM role for S3 access, and accept default VPC, security, and snapshot settings. Automated backups and monitoring are enabled, and the cluster can be queried using Query Editor v2. Always delete the cluster afterward to avoid costs.

- Navigate to the Service, Open **AWS Management Console** → Search and select **Amazon Redshift**.

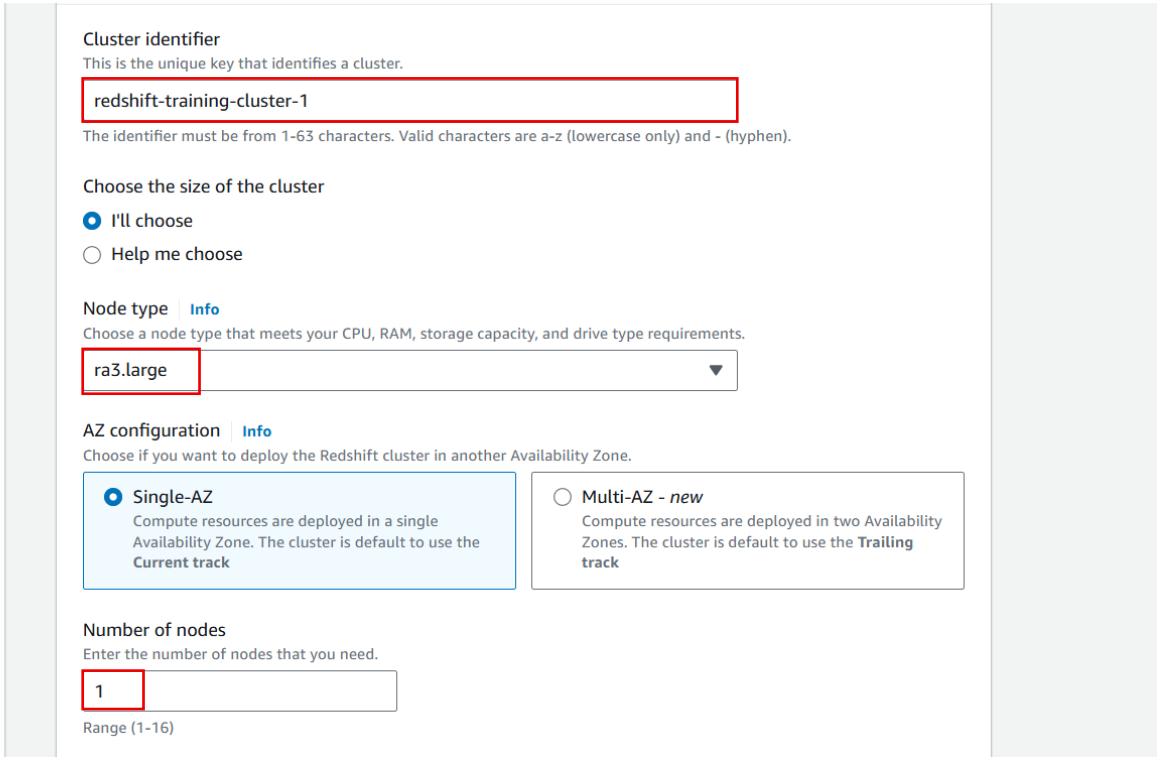


- Click Create cluster

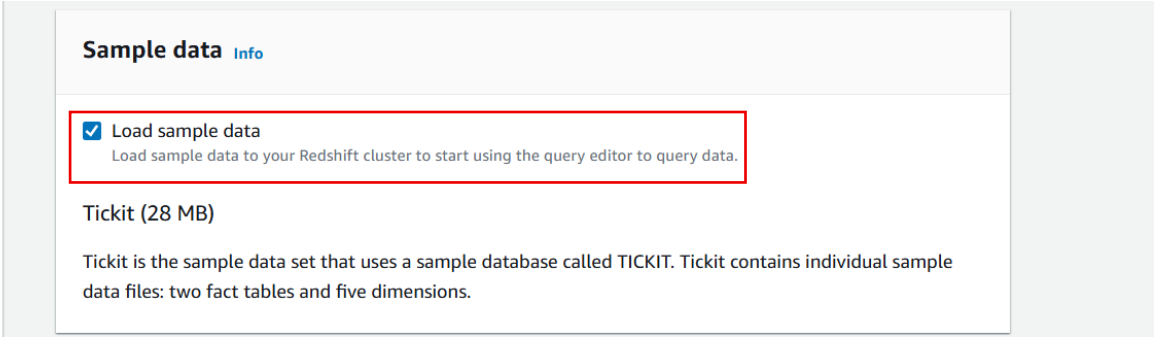


- **Cluster Identifier:** e.g., redshift-training-cluster-1.
- **Node Type:**

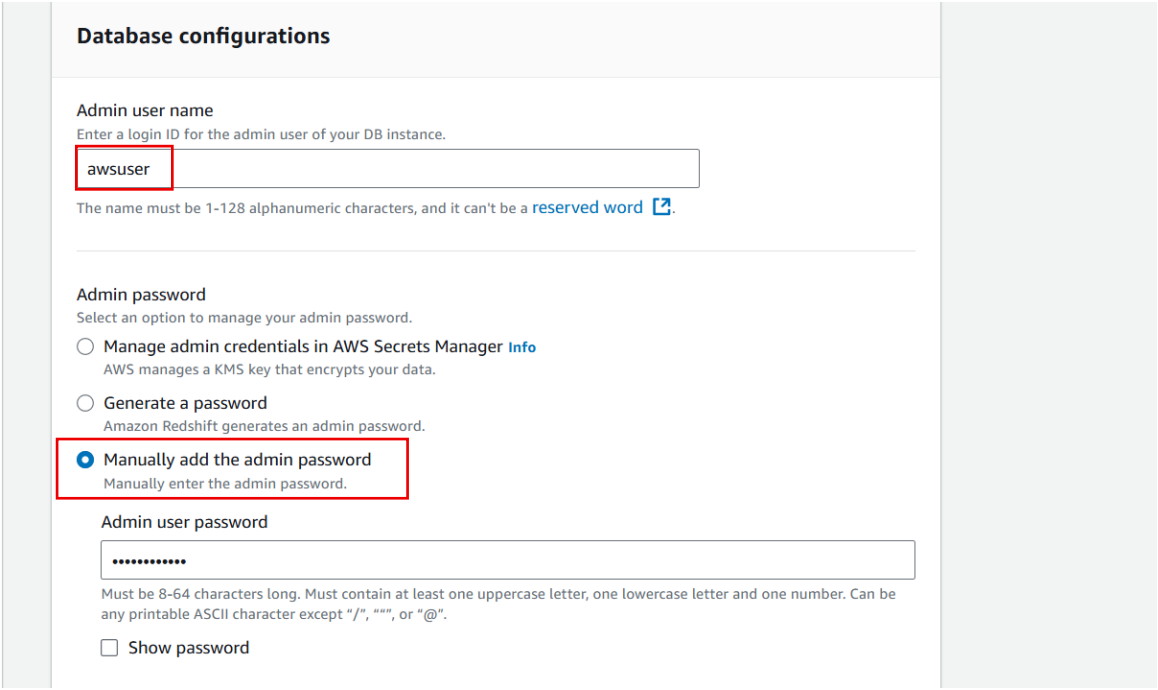
- Options: RA3, etc.
- **Choose:** ra3.large
- **Number of Nodes:**
- For demo purposes, use **1 node** to minimize cost.



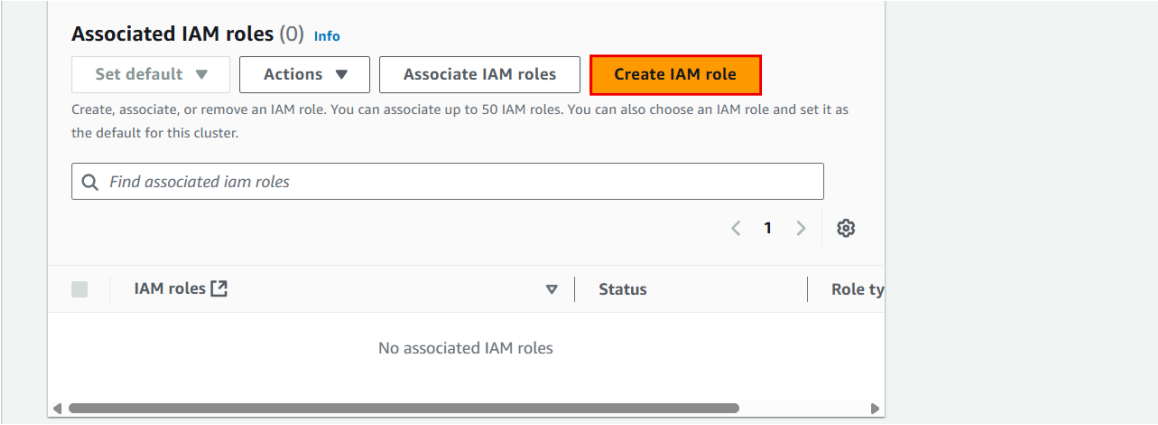
- Enable **"Load sample data"** option for quick testing.
- This automatically creates a test dataset for practice.



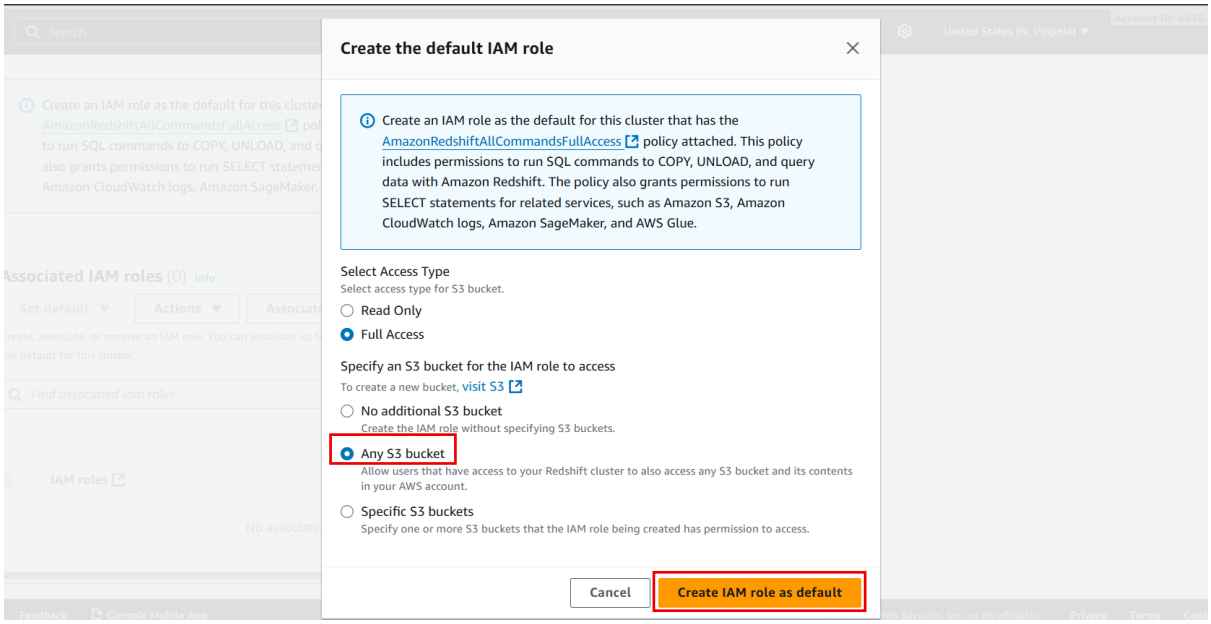
- **Database Configuration**
- **Admin Username:** Keep the default (awsuser).
- Select Manually add the admin password.
- **Password:** Enter and remember it (needed for database login).
- The database will be created within the Redshift cluster.



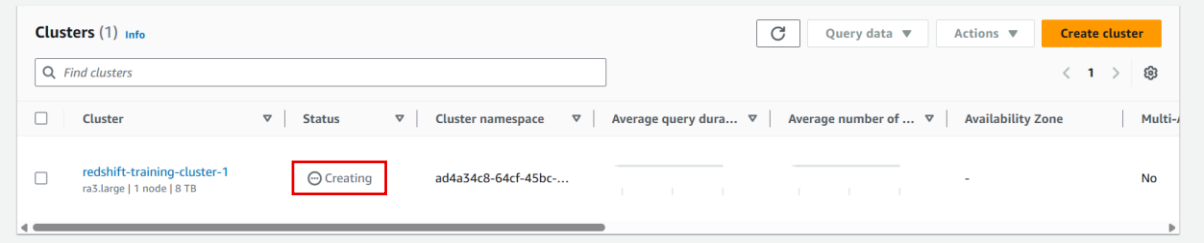
- Redshift requires an **IAM role** for permissions (e.g., to access S3).
- **Steps:**
- Click **Manage IAM Roles → Create IAM Role**.



- Add permission to access **S3 buckets** (for loading/unloading data).



- Set this role as **default** for the cluster.
- Additional roles can be added if needed.
- Leave all other setting on default.
- Now click create cluster.



- The creation process takes **1–2 minutes**.
- Refresh the page if the cluster status doesn't immediately show as **Available**.