

Create a Redshift Cluster

Overview of Redshift

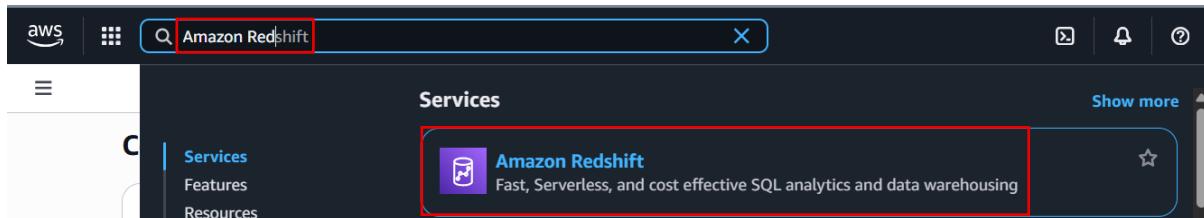
- **Amazon Redshift** is a fast, scalable, and cost-effective data warehousing service by AWS.
- Allows performing **complex queries** and **data analytics** on large datasets.
- Offers two deployment models:
 - **Redshift Serverless** – fully managed, pay-per-use model (no infrastructure management).
 - **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

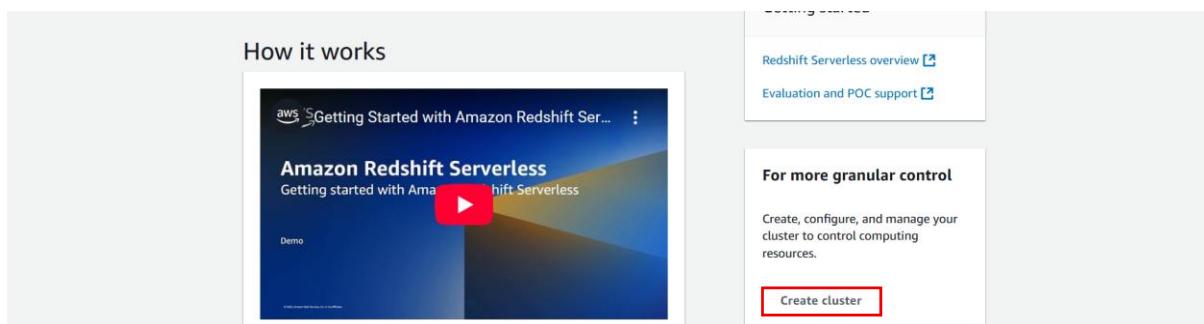
Summary of the Lab

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.

Cluster identifier
This is the unique key that identifies a cluster.
The identifier must be from 1-63 characters. Valid characters are a-z (lowercase only) and - (hyphen).

Choose the size of the cluster
 I'll choose
 Help me choose

Node type [Info](#)
Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.

AZ configuration [Info](#)
Choose if you want to deploy the Redshift cluster in another Availability Zone.

Single-AZ
Compute resources are deployed in a single Availability Zone. The cluster is default to use the **Current track**

Multi-AZ - new
Compute resources are deployed in two Availability Zones. The cluster is default to use the **Trailing track**

Number of nodes
Enter the number of nodes that you need.
Range (1-16)

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

Sample data [Info](#)

Load sample data
Load sample data to your Redshift cluster to start using the query editor to query data.

Ticket (28 MB)
Ticket is the sample data set that uses a sample database called TICKIT. Ticket contains individual sample data files: two fact tables and five dimensions.

- **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.

Database configurations

Admin user name
Enter a login ID for the admin user of your DB instance.
 awsuser

The name must be 1-128 alphanumeric characters, and it can't be a reserved word [\[?\]](#).

Admin password
Select an option to manage your admin password.

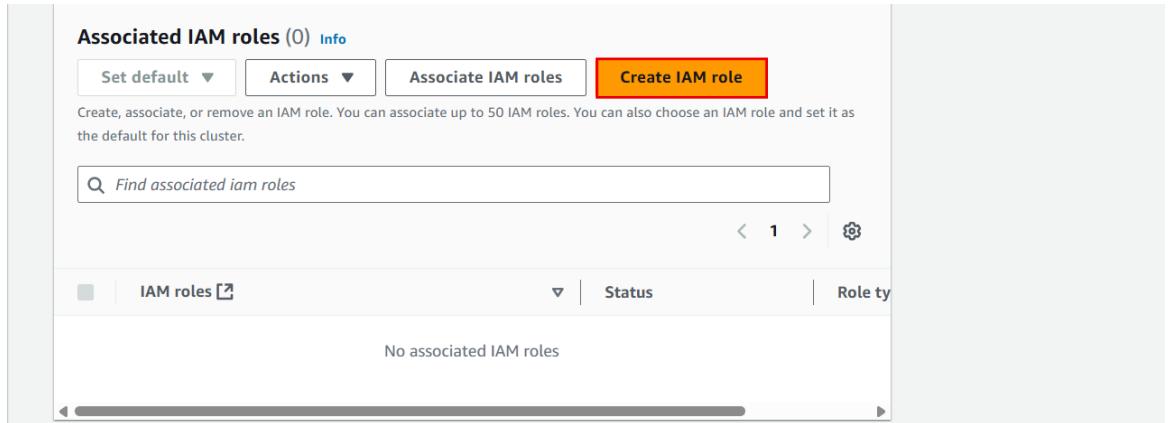
- Manage admin credentials in AWS Secrets Manager [Info](#)
AWS manages a KMS key that encrypts your data.
- Generate a password
Amazon Redshift generates an admin password.
- Manually add the admin password
Manually enter the admin password.

Admin user password

Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", "", or "@".

Show password

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



Associated IAM roles (0) [Info](#)

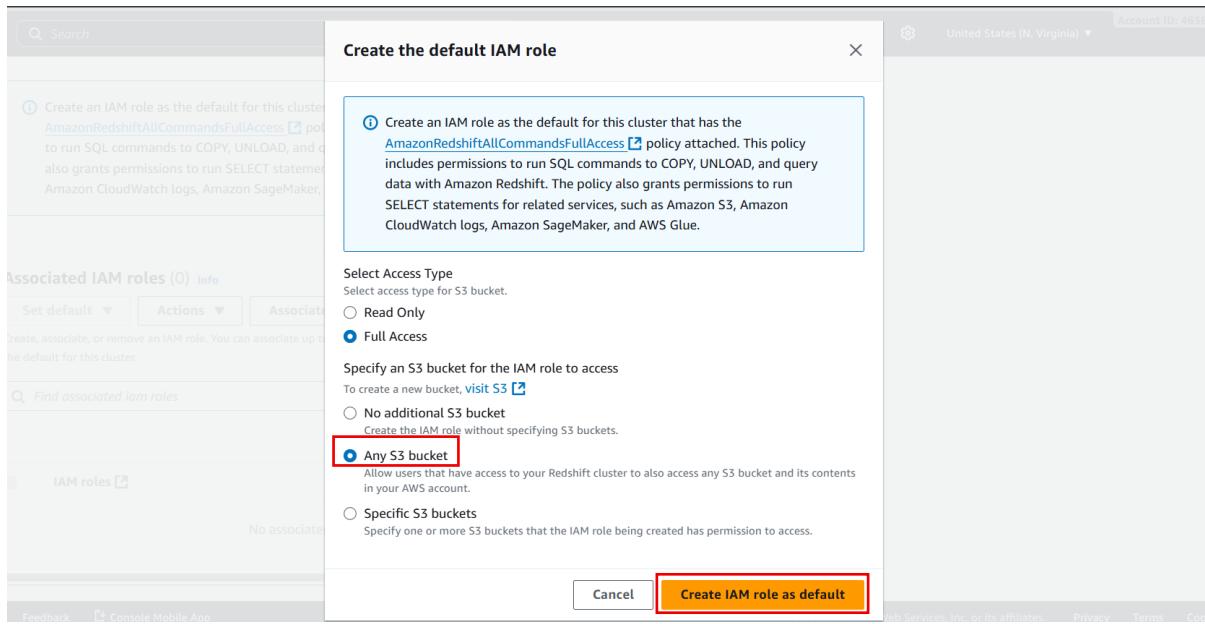
Create IAM role Create IAM role

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default for this cluster.

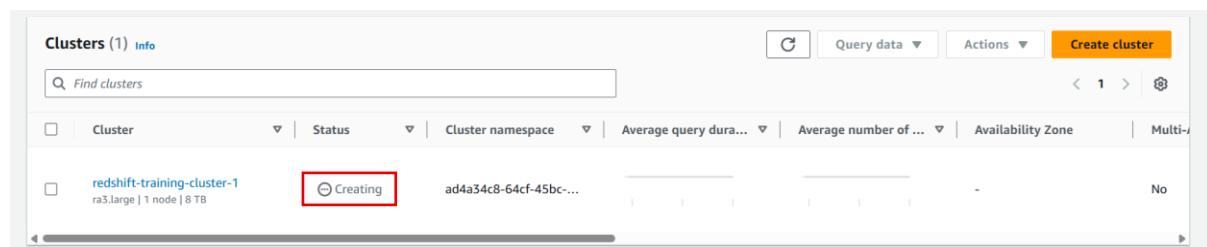
Find associated iam roles

| IAM roles | Status | Role type |
|-------------------------|--------|-----------|
| No associated IAM roles | | |

- 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**.

