

DW WITH AWS REDSHIFT

MODULE : REDSHIFT SERVERLESS

MODULE OBJECTIVES

At the end of this module, you should be able to:

- Explain Overview of Redshift Serverless
- Understand the Compute Capacity of Redshift Serverless
- Overview of Amazon Redshift Serverless workgroups and namespaces
- Getting started with Amazon Redshift Serverless using the AWS Console
- Migrating a provisioned cluster to Amazon Redshift Serverless.
- Running complex Analytical queries on Redshift Serverless using Redshift Spectrum



WHAT IS REDSHIFT SERVERLESS ?

- Amazon Redshift Serverless makes it convenient for you to run and scale analytics without having to provision and manage data warehouses.
- With Amazon Redshift Serverless, data analysts, developers, and data scientists can now use Amazon Redshift to get insights from data in seconds by loading data into and querying records from the data warehouse.
- Amazon Redshift automatically provisions and intelligently scales data warehouse capacity to deliver fast performance for demanding and unpredictable workloads.
- You pay only for the capacity that you use. You can benefit from this simplicity without changing your existing analytics and business intelligence applications.
- Amazon Redshift Serverless adjusts capacity in seconds to deliver consistently high performance and simplified operations for even the most demanding and volatile workloads.

BENEFITS OF REDSHIFT SERVERLESS

- Access and analyze data without the need to set up, tune, and manage Amazon Redshift provisioned clusters.
- Use the superior Amazon Redshift SQL capabilities, industry-leading performance, and data-lake integration to seamlessly query across a data warehouse, a data lake, and operational data sources.
- Deliver consistently high performance and simplified operations for the most demanding and volatile workloads with intelligent and automatic scaling.
- Pay only when the data warehouse is in use.

COMPUTE CAPACITY FOR REDSHIFT SERVERLESS

RPUs

Amazon Redshift Serverless measures data warehouse capacity in Redshift Processing Units (RPUs). RPUs are resources used to handle workloads.

Base capacity

- This setting specifies the base data warehouse capacity Amazon Redshift uses to serve queries.
- Base capacity is specified in RPUs.
- Setting higher base capacity improves query performance, especially for data processing jobs that consume a lot of resources.
- The default base capacity for Amazon Redshift Serverless is 128 RPUs.
- You can adjust the Base capacity setting from 32 RPUs to 1024 RPUs in units of 8 (32,40,48... 1024), using the management console

OVERVIEW OF AMAZON REDSHIFT SERVERLESS NAMESPACES

Namespace :

- Is a collection of database objects and users.
- The storage-related namespace groups together schemas, tables, users, or AWS Key Management Service keys for encrypting data.
- Storage properties include the database name and password of the admin user, permissions, and encryption and security.
- You can configure these storage properties using the Amazon Redshift Serverless console, the AWS Command Line Interface, or the Amazon Redshift Serverless APIs for the specific resource.

OVERVIEW OF AMAZON REDSHIFT SERVERLESS WORKGROUP

Workgroup :

- Workgroup is a collection of compute resources.
- The compute-related workgroup groups together compute resources like RPU, VPC subnet groups, and security groups.
- Properties for the workgroup include network and security settings.
- Other resources that are grouped under workgroups include access and usage limits.
- You can configure these compute properties using the Amazon Redshift Serverless console, the AWS Command Line Interface, or the Amazon Redshift Serverless APIs.
- You can create one or more namespaces and workgroups.
- A namespace can exist without any workgroup associated with it.
- Each namespace can have only one workgroup associated with it. Conversely, each workgroup can be associated with only one namespace.

GETTING STARTED WITH AMAZON REDSHIFT SERVERLESS USING THE CONSOLE

Steps to setup Redshift serverless :

- Create a namespace and workgroup,
- Associate them with each other.
- To get started setting Amazon Redshift Serverless configuration using the Amazon Redshift Serverless console, we need to choose
 - Get started with Amazon Redshift Serverless to set up Amazon Redshift Serverless.
 - We can choose an environment with default settings, which makes for quicker setup, or
 - Explicitly configure the settings per your organization's requirements.
 - During this process, you specify settings for your workgroup and namespace.

GETTING STARTED WITH AMAZON REDSHIFT SERVERLESS USING THE AWS CONSOLE

Creating a workgroup with an accompanying namespace

Creating a workgroup :

1. Choose Create workgroup.
2. Enter the workgroup name.
3. Choose a Virtual private cloud (VPC) for Amazon Redshift Serverless. This assigns the workgroup to a specific virtual network in your AWS environment.
4. Choose one or more VPC security groups.
5. Under Subnet, specify one or more subnets to associate with your database. These subnets are contained in the VPC you chose previously and must be in three distinct Availability Zones. Choose Continue.

GETTING STARTED WITH AMAZON REDSHIFT SERVERLESS USING THE AWS CONSOLE

Choosing a namespace :

1. Choose either Create a new namespace, and enter the namespace name, or Add to an existing namespace, and select the namespace from the drop-down list.
2. For Database name and password, specify the name of the first database. You can also specify an admin other than your default console admin, by editing the Admin user credentials.
3. For Permissions, you choose Associate IAM role to associate specific IAM roles with your namespace and workgroup.

GETTING STARTED WITH AMAZON REDSHIFT SERVERLESS USING THE AWS CONSOLE

Review workgroup selections :

1. Review your settings under Review and create. It shows the settings you chose in the previous steps.
2. Choose Save.
3. After you create the workgroup, it's added to the Workgroups list.

MIGRATING A PROVISIONED CLUSTER TO AMAZON REDSHIFT SERVERLESS

Creating a snapshot of your provisioned cluster :

To transfer data from your provisioned cluster to Amazon Redshift Serverless, create a snapshot of your provisioned cluster, and then restore the snapshot in Amazon Redshift Serverless.

To create a snapshot of your provisioned cluster follow the below steps :

1. On the AWS Redshift Management console -> navigation menu, choose Clusters, Snapshots, and then choose Create snapshot.
2. Enter the properties of the snapshot definition, then choose Create snapshot. It might take some time for the snapshot to be available.

MIGRATING A PROVISIONED CLUSTER TO AMAZON REDSHIFT SERVERLESS

To restore a provisioned cluster snapshot to an Amazon Redshift Serverless instance follow the below steps:

1. Start on the Amazon Redshift provisioned cluster console and navigate to the Clusters, Snapshots page.
2. Choose a snapshot to use.
3. Choose Restore snapshot, Restore to serverless namespace.
4. Choose a namespace to restore your snapshot to.
5. Confirm you want to restore from your snapshot. This action replaces all the databases in your serverless endpoint with the data from your provisioned cluster. Choose Restore.

REDSHIFT SERVERLESS BILLING

Please refer the below link for Redshift Serverless Billing details :

<https://docs.aws.amazon.com/redshift/latest/mgmt/serverless-billing.html>

MODULE SUMMARY

Now, you should be able to:

- Explain Overview of Redshift Serverless
- Understand the Compute Capacity of Redshift Serverless
- What are Amazon Redshift Serverless workgroups and namespaces
- Provision Amazon Redshift Serverless using the AWS Console
- How to Migrating a provisioned cluster to Amazon Redshift Serverless
- Know how to run complex Analytical queries on Redshift Serverless using Redshift Spectrum



THANK YOU