



Search for anything

CATALOG

MY PROGRAMS



CONCEPTS



- ✓ 30. Create an AWS Redshift Cluster
- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES



LAUNCH CLOUD GATEWAY



- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES



LAUNCH CLOUD GATEWAY



- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES



LAUNCH CLOUD GATEWAY



- 31. Demo: Exercise 6: Building the S3 to Redshift DAG

SEND PAGE FEEDBACK



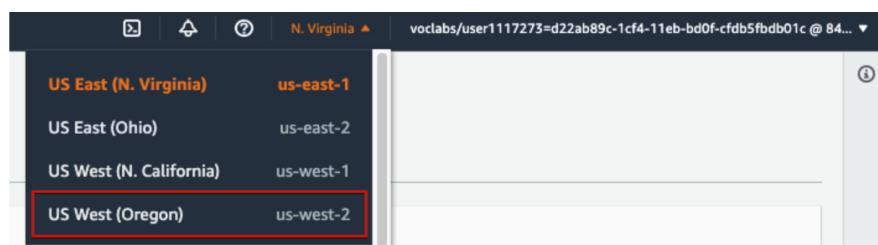
DATA ENGINEERING > AUTOMATE DATA PIPELINES

Data Pipelines

17 minutes remaining

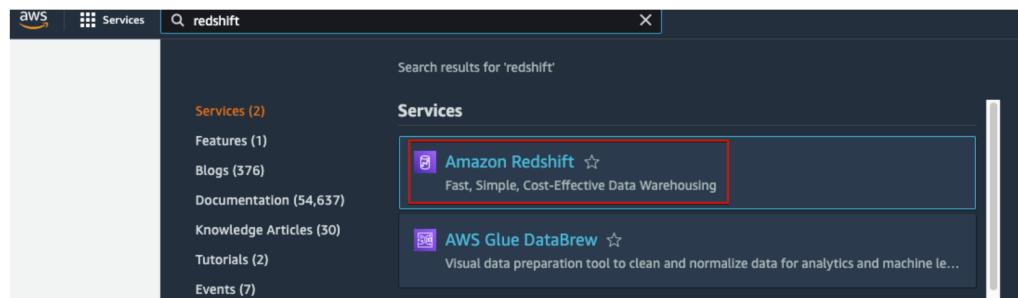
Create an AWS Redshift Cluster

- Open the AWS console by clicking on the **Launch Cloud Gateway** button followed the **Open Cloud Console** button in the classroom.
- Select US West (Oregon) **us-west-2** region.



Set region to us-west-2

- Search **Redshift** in the search bar, and then click on **Amazon Redshift**.



Select Amazon Redshift

- Click on **Create cluster**.

Analytics

Amazon Redshift

Accelerate your time to value with fast, easy, and secure analytics at scale.

Amazon Redshift makes it easier for you to run and scale analytics without having to manage their data warehouse. Get insights running real-time and predictive analytics on all your data across your operational databases, data lake, data warehouse, and thousands of third-party datasets.

Get to powerful insights fast

The Amazon Redshift serverless experience makes it easy for customers to run and scale analytics without having to provision and manage their data warehouse. Simply load and query data.

Try Amazon Redshift Serverless (Preview)

How it works

Provision and manage clusters

With a few clicks, you can create your first Amazon Redshift provisioned cluster in minutes.

Create cluster

- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES ▲

LAUNCH CLOUD GATEWAY 

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES ▲

LAUNCH CLOUD GATEWAY 

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES ▲

LAUNCH CLOUD GATEWAY 

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG

- Input **Cluster Identifier** and select **Free trial**.

Create cluster 

Cluster configuration

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).

What are you planning to use this cluster for?

 Production

Configure for fast and consistent performance at the best price.

 Free trial

Configure for learning about Amazon Redshift. This configuration is free for a limited time if your organization has never created an Amazon Redshift cluster.

 When the free trial ends, delete your cluster to avoid incurring charges at [on-demand rate](#) for compute and storage. If you want to take a final snapshot of your cluster and store the snapshot on an S3, our on-demand rate applies.

Cluster Configuration

- Input **Admin user name**. This can be same as the IAM user you created earlier.
- Input **Admin user password**. Keep the username and password saved locally, as they will be needed in Airflow.
- Click on **Create cluster**.

Database configurations

Admin user name

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

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

 Auto generate password

Amazon Redshift can generate a password for you, or you can specify your own password.

Admin user password

 Show password

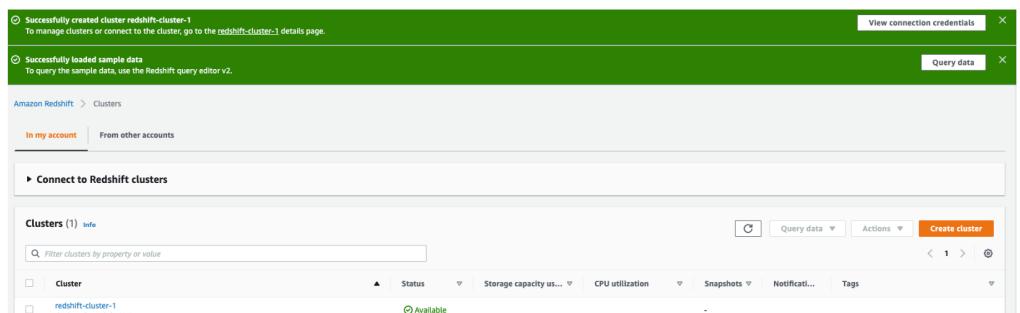
Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", "", or "@".

Cancel

Create cluster

Set username and password

- On successful creation of the cluster, you will see Status Available, as shown below:



The screenshot shows the 'Clusters' section of the Amazon Redshift console. A green success message at the top states: "Successfully created cluster redshift-cluster-1. To manage clusters or connect to the cluster, go to the redshift-cluster-1 details page." Below this, another message says: "Successfully loaded sample data. To query the sample data, use the Redshift query editor v2." At the bottom right of the message area, there are 'View connection credentials' and 'Query data' buttons. The main table lists the cluster 'redshift-cluster-1' with its status as 'Available'. Other columns include 'Cluster', 'Status', 'Storage capacity us...', 'CPU utilization', 'Schemas', 'Notifications', and 'Tags'.

34. Solution 6: Build the S3 to Redshift DAG

35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

Redshift cluster created successfully

- Go to **redshift-cluster-1**. Next, we are going to make this cluster publicly accessible as we would like to connect to this cluster via Airflow.
- Click on **Actions** and select **Modify publicly accessible setting**.

Modify public accessibility

- Click on **Enable** and **Save changes**.

Enable public accessibility

- Enable VPC Routing by going to the **Properties** tab and clicking on **Edit** button in the **Network and security settings** section.

Edit network and security settings

- Enable **Enhanced VPC Routing**.

Edit network and security

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

LAUNCH CLOUD GATEWAY

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

▼ Network and security

Virtual private cloud (VPC)
This VPC defines the virtual networking environment for this cluster.

vpc-0d5f291374f076211

VPC security groups
This VPC security group defines which subnets and IP ranges the cluster can use in the VPC.

Choose one or more security groups

default sg-00c3e40f066d2b48c

Cluster subnet group
Choose the Amazon Redshift subnet group to launch the cluster in.

default

Availability Zone
Specify the Availability Zone that you want the cluster to be created in. Otherwise, Amazon Redshift chooses an Availability Zone for you.

No preference

Enhanced VPC routing
Enabling this option forces network traffic between your cluster and data repositories through a VPC, instead of the internet. [Learn more](#)

Disabled
 Enabled

Enable Enhanced VPC routing

Cancel **Save changes**

Network and security settings

Virtual private cloud (VPC) vpc-0d5f291374f076211	Availability Zone us-east-1f	VPC security group Specify which instances and devices can connect to the cluster. sg-00c3e40f066d2b48c
Subnet default	Enhanced VPC routing Enabled	
Endpoint URL - -		

Open VPC security group

Security Groups (1/1) Info

Name	Security group ID	Security group name	VPC ID	Description	Owner	Inbound rules count	Outbound rules count
-	sg-00c3e40f066d2b48c	default	vpc-0d5f291374f076211	default VPC security gr...	84586960839	1 Permission entry	1 Permission entry

sg-00c3e40f066d2b48c - default

Inbound rules (1/1)

Name	Security group rule...	IP version	Type	Protocol	Port range	Source	Description
-	sg-025a8750d835cec90	-	All traffic	All	All	sg-00c3e40f066d2b48...	-

Edit inbound rules

EC2 > Security Groups > sg-00c3e40f066d2b48c - default > Edit inbound rules

Edit inbound rules

- Add an inbound rule, as shown in the image below.
 - Type = Custom TCP
 - Port range = 0 - 5500
 - Source = Anywhere-iPv4

- 30. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

- 31. Demo: Exercise 6: Building the S3 to Redshift DAG
- 32. Add Airflow Connections to AWS Redshift
- 33. Exercise 6: Build the S3 to Redshift DAG
- 34. Solution 6: Build the S3 to Redshift DAG
- 35. Conclusion

RESOURCES

LAUNCH CLOUD GATEWAY



Inbound rules control the incoming traffic that's allowed to reach the instance.

Inbound rules

Security group rule ID	Type	Protocol	Port range	Source	Description - optional	Actions
sg-025ab750d835ec90	All traffic	All	All	Custom	sg-003e40f0f66d2b48: 0.0.0.0/0	<input type="button" value="Delete"/>
-	Custom TCP	TCP	0 - 5500	Anywhere-IPv4	0.0.0.0/0	<input type="button" value="Delete"/>

Add rule

Save rules

Add inbound rule

- Now our Redshift cluster should be accessible from Airflow.
- Go back to the Redshift cluster and copy the endpoint. Store this locally as we will need this while configuring Airflow.

General information

Cluster identifier	redshift-cluster-1	Status	Available	Node type	d2.2xlarge	Endpoint	redshift-cluster-1.cvevrggq9.us-east-1.redshift.amazonaws.com:54...
Cluster namespace	37155767-cd08-4bae-99a1-d3fc246dcce7	Date created	February 14, 2022, 15:42 (UTC+05:30)	Number of nodes	1	ODBC URL	jdbc:redshift://redshift-cluster-1.cvevrggq9.us-east-1.redshift.am...
		Storage used	0.21% (0.34 of 160 GB used)	Aqua	Not available	DB URL	Driver=(Amazon Redshift (x64)); Server=redshift-cluster-1.cvevrgg...

Copy the redshift cluster endpoint

← PREVIOUS

NEXT →