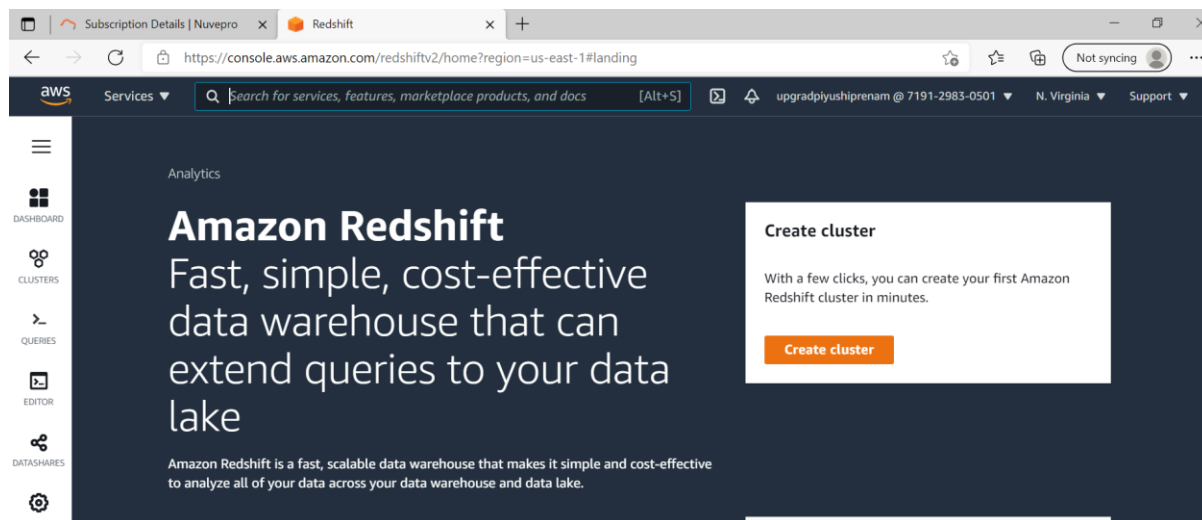


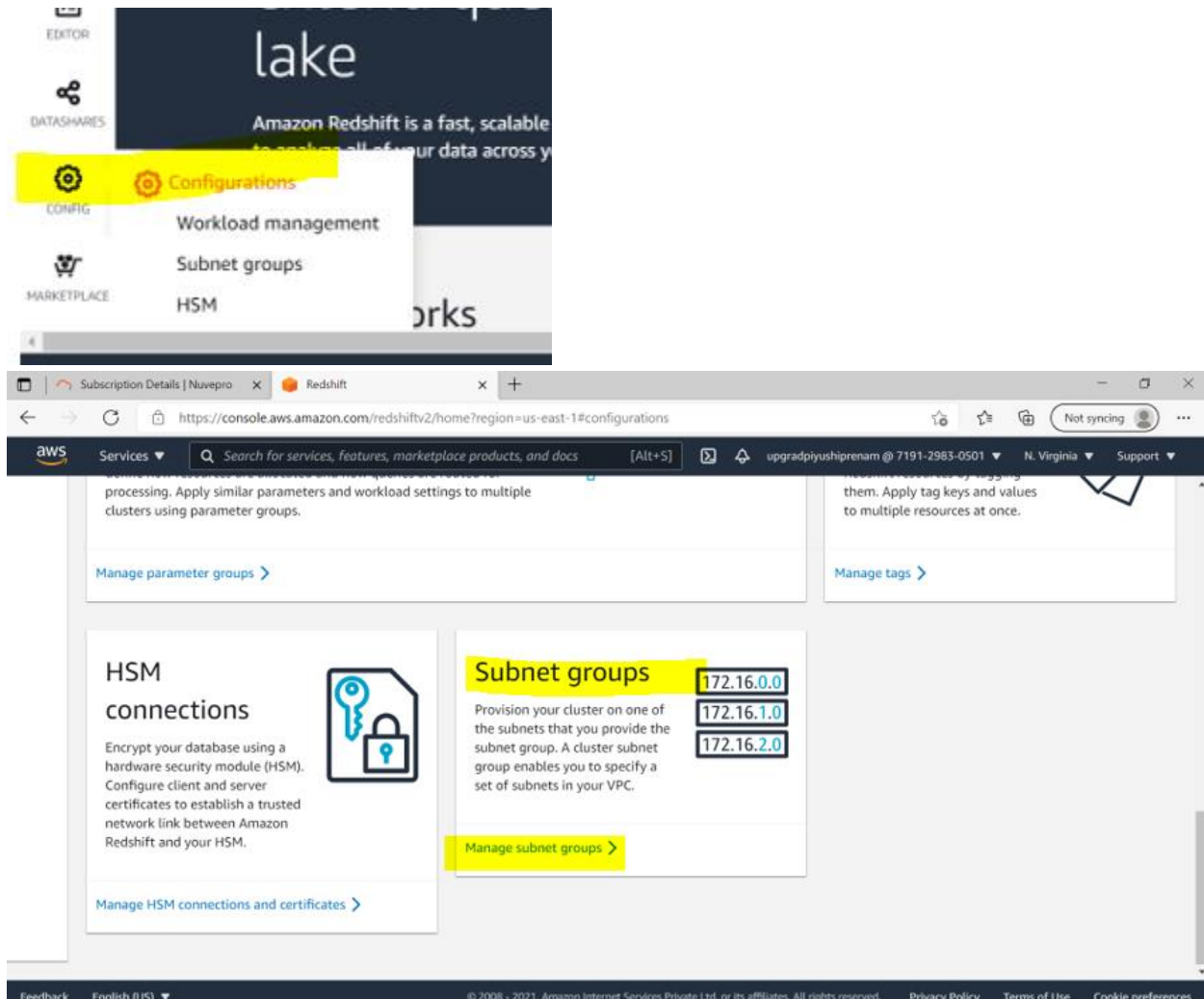
## Creation of a RedShift Cluster

**Screenshots of the configuration of the RedShift cluster that you have created:**

Navigate to aws console and search for Redshift



Click on Configurations



lake

Amazon Redshift is a fast, scalable

to analyze all of your data across y

Configurations

Workload management

Subnet groups

HSM

works

Subscription Details | Nuvepro x Redshift x +

https://console.aws.amazon.com/redshiftv2/home?region=us-east-1#configurations

aws Services Search for services, features, marketplace products, and docs [Alt+S] upgradpiyushiprenam @ 7191-2983-0501 N. Virginia Support

processing. Apply similar parameters and workload settings to multiple clusters using parameter groups.

Manage parameter groups >

them. Apply tag keys and values to multiple resources at once.

Manage tags >

HSM connections

Encrypt your database using a hardware security module (HSM). Configure client and server certificates to establish a trusted network link between Amazon Redshift and your HSM.

Manage HSM connections and certificates >

Subnet groups

Provision your cluster on one of the subnets that you provide the subnet group. A cluster subnet group enables you to specify a set of subnets in your VPC.

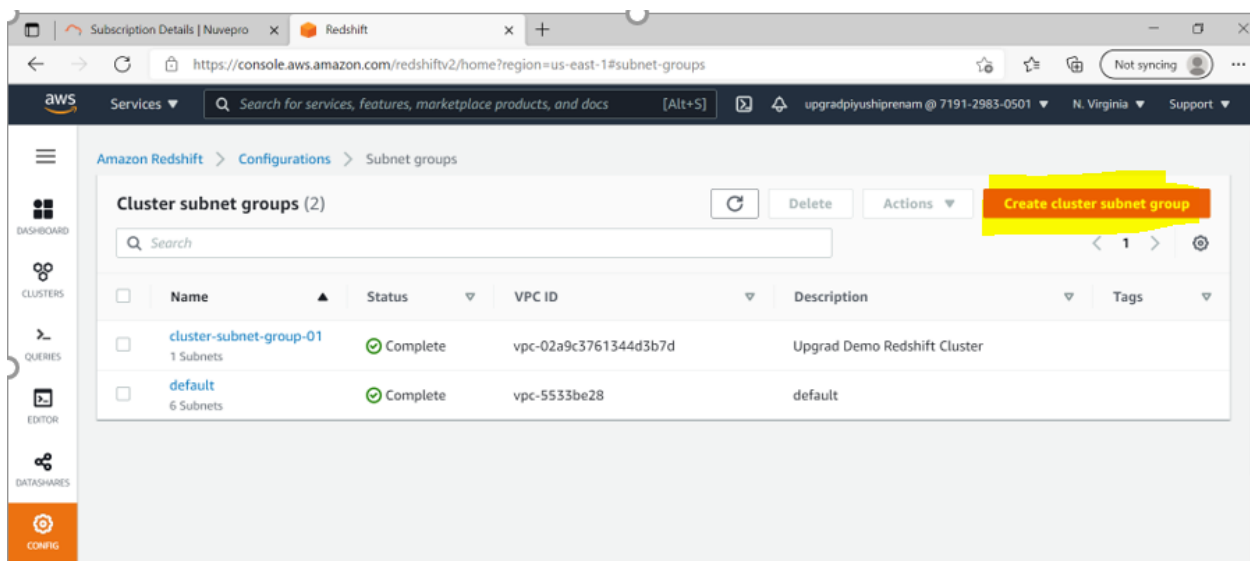
172.16.0.0

172.16.1.0

172.16.2.0

Manage subnet groups >

Feedback English (US) © 2008 - 2021, Amazon Internet Services Private Ltd. or its affiliates. All rights reserved. Privacy Policy Terms of Use Cookie preferences



Subscription Details | Nuvepro x Redshift x +

https://console.aws.amazon.com/redshiftv2/home?region=us-east-1#subnet-groups

aws Services Search for services, features, marketplace products, and docs [Alt+S] upgradpiyushiprenam @ 7191-2983-0501 N. Virginia Support

Amazon Redshift > Configurations > Subnet groups

Cluster subnet groups (2)

Search

Create cluster subnet group

<input type="checkbox"/>	Name	Status	VPC ID	Description	Tags
<input type="checkbox"/>	cluster-subnet-group-01 1 Subnets	Complete	vpc-02a9c3761344d3b7d	Upgrad Demo Redshift Cluster	
<input type="checkbox"/>	default 6 Subnets	Complete	vpc-5533be28	default	

DASHBOARD

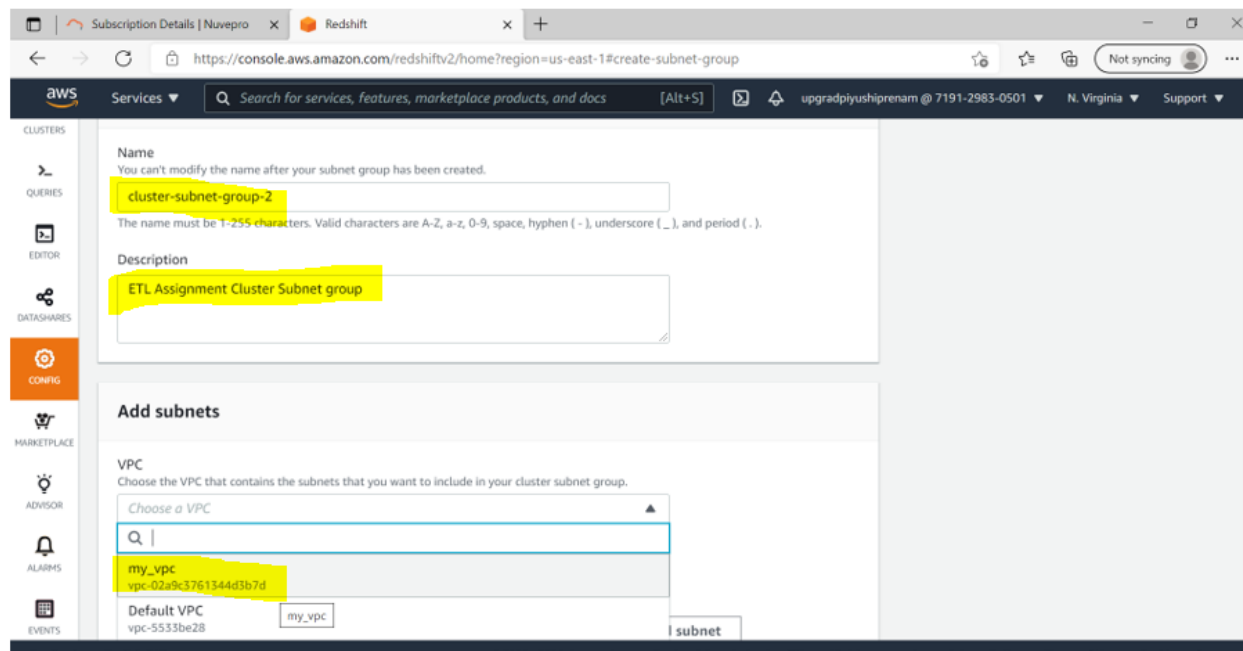
CLUSTERS

QUERIES

EDITOR

DATASHARES

CONFIG



**Name**  
You can't modify the name after your subnet group has been created.  
cluster-subnet-group-2  
The name must be 1-255 characters. Valid characters are A-Z, a-z, 0-9, space, hyphen ( - ), underscore ( \_ ), and period ( . ).

**Description**  
ETL Assignment Cluster Subnet group

**Add subnets**

**VPC**  
Choose the VPC that contains the subnets that you want to include in your cluster subnet group.  
Choose a VPC  
my\_vpc  
vpc-02a9c3761344d3b7d  
Default VPC  
vpc-5533be28

**Add subnets**

**VPC**  
Choose the VPC that contains the subnets that you want to include in your cluster subnet group.  
my\_vpc  
vpc-02a9c3761344d3b7d

**Add all the subnets for this VPC**

**Availability Zone**  
us-east-1a (1)

**Subnet**  
subnet-001e6189e79377d92

**Add subnet**

**Subnets in this cluster subnet group (0)** **Remove all**

Availability Zone	Subnet ID	CIDR block	Action
-------------------	-----------	------------	--------

VPC

Choose the VPC that contains the subnets that you want to include in your cluster subnet group.

my\_vpc  
vpc-02a9c3761344d3b7d

Add all the subnets for this VPC

Availability Zone: us-east-1a (1)

Subnet: subnet-001e6189e79377d92

Add subnet

Subnets in this cluster subnet group (1)

Remove all

Availability Zone	Subnet ID	CIDR block	Action
us-east-1a	subnet-001e6189e79377d92	10.0.0.0/24	Remove

Cancel Create cluster subnet group

Cluster subnet group cluster-subnet-group-2 was create successfully

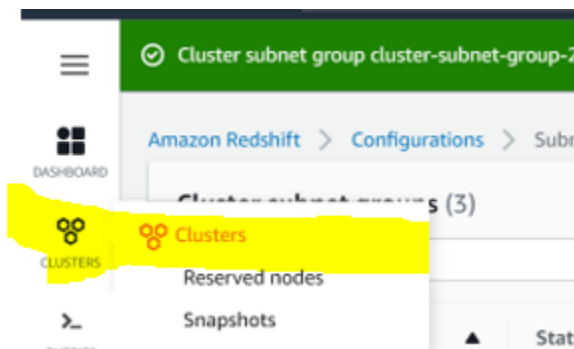
Amazon Redshift > Configurations > Subnet groups

Cluster subnet groups (3)

Search

	Name	Status	VPC ID	Description	Tags
<input type="checkbox"/>	cluster-subnet-group-01 1 Subnets	Complete	vpc-02a9c3761344d3b7d	Upgrad Demo Redshift Cluster	
<input type="checkbox"/>	cluster-subnet-group-2 1 Subnets	Complete	vpc-02a9c3761344d3b7d	ETL Assignment Cluster Subnet group	
<input type="checkbox"/>	default 6 Subnets	Complete	vpc-5533be28	default	

For Cluster creation



Subscription Details | Nuvepro x Redshift

https://console.aws.amazon.com/redshiftv2/home?region=us-east-1#create-cluster

Services Search for services, features, marketplace products, and docs [Alt+S] upgradpiyushiprenam @ 7191-2983-0501 N. Virginia Support

**Cluster identifier**  
This is the unique key that identifies a cluster.  
piyushietl  
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.

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

**Node type**  
Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.  
dc2.large

**Nodes**  
Enter the number of nodes that you need.  
2  
Range (1-32)

Services Search for services, features, marketplace products, and docs [Alt+S] upgradpiyushiprenam @ 7191-2983-0501 N. Virginia Support

☒ **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.

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

**Node type**  
Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.

dc2.large  
Storage: 160 GB/node  
\$0.25/node/hour  
2 vCPU (gen 2)

dc2.8xlarge  
Storage: 2.6 TB/node  
\$4.80/node/hour  
32 vCPU (gen 2)

DS2 (legacy)  
Large workloads with fixed local HDD storage

## 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](#).

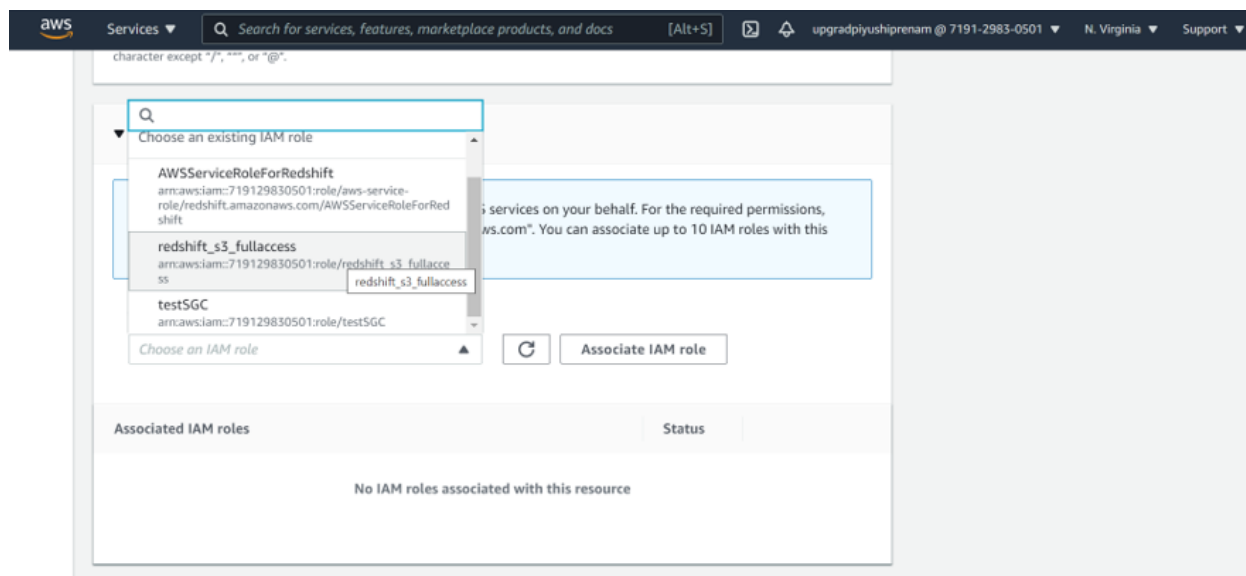
### ☐ Auto generate password

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 `@`.



The screenshot shows the AWS IAM console interface. At the top, there's a navigation bar with the AWS logo, 'Services' dropdown, a search bar, and user information. The main content area shows a step in the IAM role association process. A dropdown menu titled 'Choose an existing IAM role' is open, displaying a list of roles: 'AWSServiceRoleForRedshift', 'redshift\_s3\_fullaccess', and 'testSGC'. The 'redshift\_s3\_fullaccess' role is currently selected. To the right of the dropdown, there's a blue informational box. Below the dropdown, there's a 'Choose an IAM role' button and an 'Associate IAM role' button. At the bottom, there's a table with the header 'Associated IAM roles' and 'Status', which currently shows 'No IAM roles associated with this resource'.

## Cluster permissions

**i** Your cluster needs permissions to access other AWS services on your behalf. For the required permissions, add IAM roles with the principal "redshift.amazonaws.com". You can associate up to 10 IAM roles with this cluster. [Learn more](#)

### Available IAM roles

Choose an IAM role



Associate IAM role

### Associated IAM roles


### Status

redshift\_s3\_fullaccess

arn:aws:iam::719129830501:role/redshift\_s3\_fullaccess

Not applied

Remove


Services

Search for services, features, marketplace products, and docs

[Alt+S]

upgradpiyushiprenam @ 7191-2983-0501

N. Virginia

Support

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

my\_vpc  
vpc-02a9c3761344d3b7d

*i* You can't change the VPC associated with this cluster after the cluster has been created. [Learn more](#)

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

cloudera  
sg-0c511500d17f2bc6c

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

cluster-subnet-group-2

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

us-east-1a

Services
Search for services, features, marketplace products, and docs
[Alt+S]
upgradpiyushiprenam @ 7191-2983-0501
N. Virginia
Support

Cluster subnet group cluster-subnet-group-2 was create successfully

piyushietl is being created.

Amazon Redshift > Clusters

In my account
From other accounts

▼ Connect to Redshift clusters

**Query data using Redshift query editor**  
Use the query editor to run queries in your Redshift cluster.  
Query data

**Work with your client tools**  
You can connect to Amazon Redshift from your client tools, such as SQL clients, business intelligence (BI) tools, and extract, transform, load (ETL) tools, using JDBC or ODBC drivers.  
Cluster

**Choose your JDBC or ODBC driver**  
Use JDBC or ODBC drivers to connect to Amazon Redshift from your client tools, such as SQL clients, BI tools, and ETL tools. We recommend using the new Amazon Redshift-specific drivers for better performance and scalability.  
Driver

Clusters (1) Info
Refresh
Query cluster
Actions
Create cluster

Filter clusters by property or value
1

Cluster	Cluster namespace	Status	Storage capacity us...	CPU utilization	Snapsh...
<input type="checkbox"/> piyushietl dc2.large   2 nodes   320 GB	83312ddd-19ec-467c-...	Modifying Creating			

Clusters (1) Info
Refresh
Query cluster
Actions
Create cluster

Filter clusters by property or value
1

Cluster	Cluster namespace	Status	Storage capacity us...	CPU utilization	Snapsh...
<input type="checkbox"/> piyushietl dc2.large   2 nodes   320 GB	83312ddd-19ec-467c-...	Available			



→ ↺ https://console.aws.amazon.com/redshiftv2/home?region=us-east-1#cluster-details?cluster=piyushietl

WS Services 🔍 Search for services, features, marketplace products, and docs [Alt+S] 📄 🔔 upgradpiyushiprenam @ 7191-2983-0501 ▼ N. Virginia ▼ Support ▼

🟢 Cluster subnet group cluster-subnet-group-2 was create successfully

Amazon Redshift > Clusters > piyushietl

piyushietl

Actions ▼ Edit Add partner integration Query cluster

General information

Cluster identifier piyushietl	Status 🟢 Available	Node type dc2.large	Endpoint 📄 piyushietl.c7recyb5vuic.us-east-1.r...
Cluster namespace 83312ddd-19ec-467c-afbf-0514835cebda	Date created June 19, 2021, 01:09(UTC+05:30)	Number of nodes 2	JDBC URL 📄 jdbc:redshift://piyushietl.c7recyb5v...
	Storage used -	AQUA Not available	ODBC URL 📄 Driver={Amazon Redshift (x64)}; Se...

Amazon Redshift > Clusters > piyushietl

piyushietl

Actions

Edit

Add partner integration

Query cluster

General information

Cluster identifier

piyushietl

Cluster namespace

83312ddd-19ec-467c-afb-0514835cebdda

Status

Available

Date created

June 19, 2021, 01:09(UTC+05:30)

Storage used

-

Node type

dc2.large

Number of nodes

2

AQUA

Not available

Endpoint

piyushietl.c7recyb5vuic.us-east-1.r...

JDBC URL

jdbc:redshift://piyushietl.c7recyb5v...

ODBC URL

Driver={Amazon Redshift (x64)}; Se...

Database name

dev

Port

5439

Admin user name

awsuser

Parameter group

Defines database parameter and query queues for all the databases.

default.redshift-1.0

SSH ingestion setting (cluster public key)

ssh-rsa AAAAB3NzaC1yc2EAAA...

Encryption

Disabled

AWS KMS key ID

-

Audit logging

Disabled

Network and security settings

Edit publicly accessible

Edit

Virtual private cloud (VPC)

vpc-02a9c3761344d3b7d

Subnet

cluster-subnet-group-2

Endpoint URL

Availability Zone

us-east-1a

Enhanced VPC routing

Disabled

VPC security group

Specify which instances and devices can connect to the cluster.

sg-0c511500d17f2bc6c

Publicly accessible

Allow instances and devices outside the VPC to connect to your database through the cluster endpoint.

Disabled

Cluster permissions (1)

Your cluster needs permissions to access other AWS services on your behalf. For the required permissions, add IAM roles with the principal "redshift.amazonaws.com". You can associate up to 10 IAM roles with this cluster. [Learn more](#)

Manage IAM roles

Associated IAM roles

redshift\_s3\_fullaccess

Status

in-sync

Amazon Resource Name (ARN)

arn:aws:iam::719129830501:role/redshift\_s3\_fullaccess

## Setting up a database in the RedShift cluster and running queries to create the dimension and fact tables

[https://s3.console.aws.amazon.com/s3/bucket/create?region=us-east-1](#)

aws

Services ▾

Search for services, features, marketplace products, and docs

[Alt+S]

upgradpiyushiprenam @ 7191-2983-050

We're continuing to improve the S3 console to make it faster and easier to use. If you have feedback on the updated experience, choose **Provide feedback**.

Amazon S3 > Create bucket

Create bucket

Buckets are containers for data stored in S3. [Learn more](#)

General configuration

Bucket name

redshiftpiyushi

Bucket name must be unique and must not contain spaces or uppercase letters. [See rules for bucket naming](#)

AWS Region

US East (N. Virginia) us-east-1

Copy settings from existing bucket - optional

Only the bucket settings in the following configuration are copied.

redshiftpiyushi

US East (N. Virginia) us-east-1

Objects can be public

June 19, 2021, 13:18:00 (UTC+05:30)

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Copy S3 URI

Copy URL

Download

Open

Delete

Actions ▾

Create folder

Upload

Find objects by prefix

< 1 >

Name

Type

Last modified

Size

Storage class

ETL/

Folder

-

-

-

© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

Amazon S3
>
redshiftpiyushi
>
ETL/

ETL/

Copy S3 URI

Objects

Properties

Objects (5)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh

Copy S3 URI

Copy URL

Download

Open

Delete

Actions

Create folder

Upload

Find objects by prefix

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	DIM_ATM/	Folder	-	-	-
<input type="checkbox"/>	DIM_CARD/	Folder	-	-	-

Objects (5)

Objects are the fundamental entities stored in Amazon S3. You can use [Amazon S3 inventory](#) to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. [Learn more](#)

Refresh

Copy S3 URI

Copy URL

Download

Open

Delete

Actions

Create folder

Upload

Find objects by prefix

< 1 >

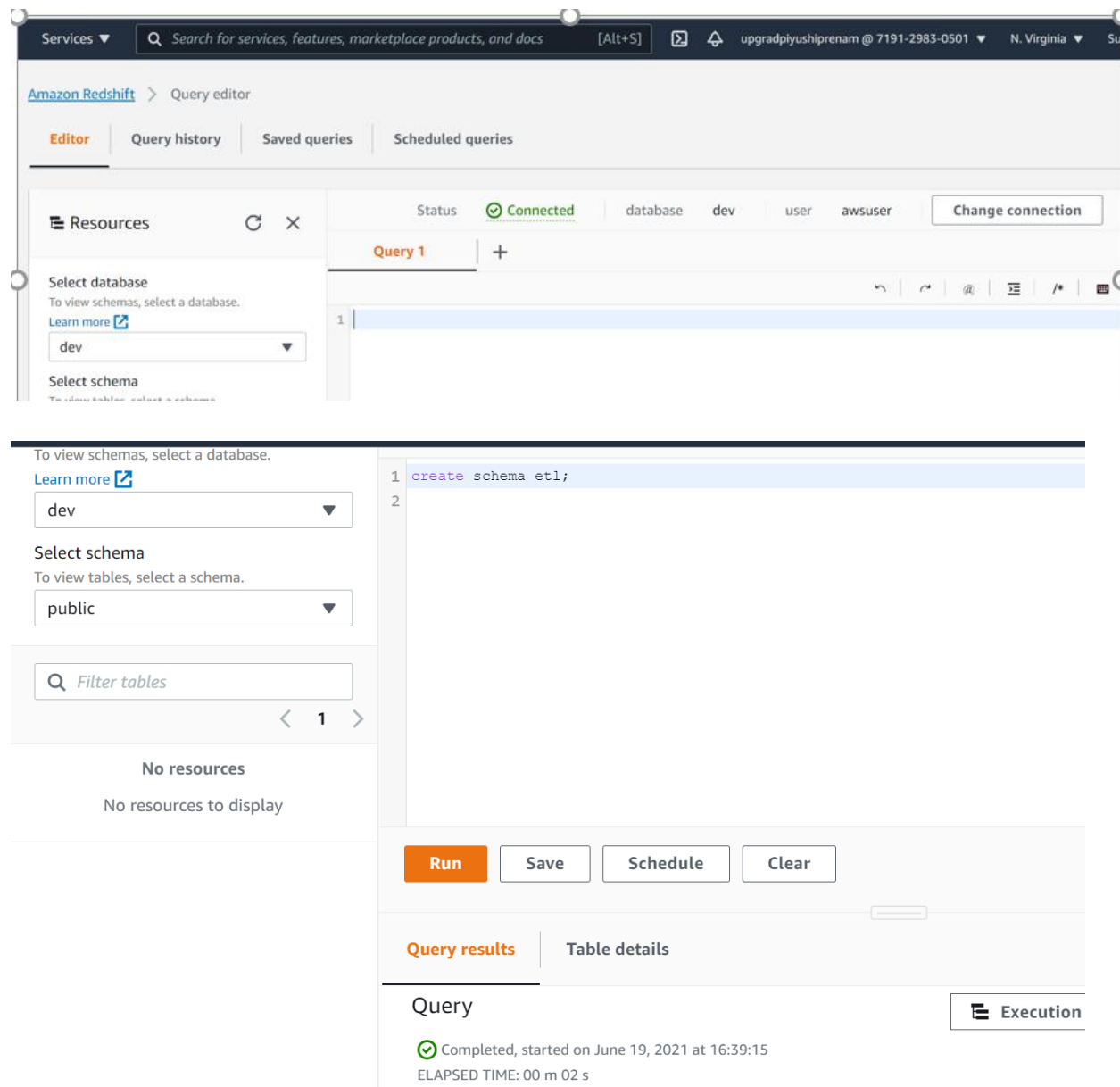
<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	DIM_ATM/	Folder	-	-	-
<input type="checkbox"/>	DIM_CARD/	Folder	-	-	-
<input type="checkbox"/>	DIM_DATE/	Folder	-	-	-
<input type="checkbox"/>	DIM_LOC/	Folder	-	-	-
<input type="checkbox"/>	FACT_ATM_TRANS/	Folder	-	-	-

Find objects by prefix

< 1 >

<input type="checkbox"/>	Name	Type	Last modified	Size	Storage class
<input type="checkbox"/>	_SUCCESS	-	June 19, 2021, 14:32:26 (UTC+05:30)	0 B	Standard
<input type="checkbox"/>	part-00000-3c54cfc2-4caa-4ae2-a1e7-51b05fae8995-c000.csv	csv	June 19, 2021, 14:32:25 (UTC+05:30)	2.5 KB	Standard

**Queries to create the various dimension and fact tables with appropriate primary and foreign keys:**



The screenshot displays the Amazon Redshift Query Editor interface. At the top, there's a search bar and user information. The main area is divided into a left sidebar and a central query editor. The sidebar contains a 'Resources' section with dropdowns for 'Select database' (set to 'dev') and 'Select schema' (set to 'public'). Below these are filters and a 'No resources' message. The central query editor shows a single query: `create schema etl;`. At the bottom, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Query results' tab is active, showing a successful execution status: 'Completed, started on June 19, 2021 at 16:39:15' with an elapsed time of '00 m 02 s'.

Services

Search for services, features, marketplace products, and docs [Alt+S]

upgradpiyushiprenam @ 7191-2983-0501 N. Virginia Support

Select schema

To view tables, select a schema.

etl

Filter tables

1

atm\_pkey

card\_pkey

date\_pkey

fact\_atm\_trans\_pkey

loc\_pkey

atm

card

date

fact\_atm\_trans

```

49 clouds_all integer,
50 weather_id integer,
51 weather_main varchar(50),
52 weather_description varchar(255),
53 primary key(trans_id),
54 foreign key(atm_prim_id) references etl.atm(atm_prim_id),
55 foreign key(location_id) references etl.loc(location_id),
56 foreign key(date_id) references etl.date(date_id),
57 foreign key(card_type_id) references etl.card(card_type_id)
58 );

```

Run

Save

Schedule

Clear

Send feedback

Query results

Table details

Query

Execution

Data

Visualize

Completed, started on June 19, 2021 at 16:42:25

ELAPSED TIME: 00 m 02 s

**create schema etl;**

**create table etl.loc(  
location\_id integer not null,  
atm\_location varchar(50),  
atm\_streetname varchar(255),  
atm\_street\_number integer,  
atm\_zipcode integer,  
atm\_lat NUMERIC(10,3),  
atm\_lon NUMERIC(10,3),  
primary key(location\_id));**

**create table etl.atm(  
atm\_prim\_id integer not null,  
atm\_id varchar(20),  
atm\_manufacturer varchar(50),  
location\_id integer,  
primary key(atm\_prim\_id),  
foreign key(location\_id) references etl.loc(location\_id));**

**create table etl.date(  
year integer,  
month varchar(20),  
day integer,  
hour integer,  
weekday varchar(20),**

```
full_date_time timestamp,  
date_id integer,  
primary key(date_id));
```

```
create table etl.card(  
card_type varchar(23),  
card_type_id integer,  
primary key(card_type_id));
```

```
create table etl.FACT_ATM_TRANS(  
trans_id BIGINT not null,  
atm_prim_id integer,  
location_id integer,  
date_id integer,  
card_type_id integer,  
atm_status varchar(20),  
currency varchar(10),  
service varchar(20),  
transaction_amount integer,  
message_code varchar(255),  
message_text varchar(255),  
rain_3h NUMERIC(10,3),  
clouds_all integer,  
weather_id integer,  
weather_main varchar(50),  
weather_description varchar(255),  
primary key(trans_id),  
foreign key(atm_prim_id) references etl.atm(atm_prim_id),  
foreign key(location_id) references etl.loc(location_id),  
foreign key(date_id) references etl.date(date_id),  
foreign key(card_type_id) references etl.card(card_type_id)  
);
```

Loading data into a RedShift cluster from Amazon S3 bucket

Queries to copy the data from S3 buckets to the RedShift cluster in the appropriate tables

```
copy etl.atm from
's3://redshiftpiyushi/ETL/DIM_ATM/part-00000-3c54cfc2-4caa-4ae2-a1e7-51b05fae8995-
c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```

```
copy etl.loc from
's3://redshiftpiyushi/ETL/DIM_LOC/part-00000-d7f4f92b-ca8f-4101-be12-9e1f16e7d5ae-
c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```

```
copy etl.date from
's3://redshiftpiyushi/ETL/DIM_DATE/part-00000-9e8fdccc-8934-4538-b283-939e185b68aa-
c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
timeformat 'auto'
region 'us-east-1';
```

```
copy etl.card from
's3://redshiftpiyushi/ETL/DIM_CARD/part-00000-57508156-8187-4127-8c41-f8f550a90b82-
c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```

```
copy etl.FACT_ATM_TRANS from
's3://redshiftpiyushi/ETL/FACT_ATM_TRANS/part-00000-98475c40-758c-455f-ae3f-
09300371c62e-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1'
TRUNCATECOLUMNS
CSV
```