

Creation of a Redshift Cluster

Screenshots of the configuration of the Redshift cluster that you have created:

Amazon Redshift > Clusters > redshift-cluster-1

redshift-cluster-1 ▾

Actions ▾

Edit

Add partner integration

Query data ▾

General information

Cluster identifier

redshift-cluster-1

Status

Available

Node type

dc2.large

Endpoint

redshift-cluster-1.ccyi511pvevc.us-east-...

Cluster namespace

5e51129c-869d-4e59-bcbd-4946ef3e8d1c

Date created

February 01, 2022, 13:39 (UTC+05:30)

Number of nodes

2

JDBC URL

jdbc:redshift://redshift-cluster-1.ccyi511...

Storage used

0.17% (0.54 of 320 GB used)

AQUA

Not available

ODBC URL

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

Database configurations

Change admin user password

Rotate encryption keys

Edit ▾

Database name

dev

Parameter group

Defines database parameter and query queues for all the databases.
default.redshift-1.0

Encryption

Disabled

Audit logging

Disabled

Port

5439

SSH ingestion setting (cluster public key)

ssh-rsa AAAAB3NzaC1yc2EAAAADAQ...

AWS KMS key ID

-

Admin user name

awsuser

Network and security settings

Edit

Virtual private cloud (VPC)

vpc-02991b629e47651b9

Availability Zone

us-east-1d

VPC security group

Specify which instances and devices can connect to the cluster.
sg-0a51938130894d373

Publicly accessible

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

Subnet

cluster-subnet-group-1

Enhanced VPC routing

Disabled

Endpoint URL

-

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

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

```
create schema atm_etl;
```

```
create table atm_etl.DIM_LOCATION(  
location_id int not null sortkey,  
location varchar(50),  
streetname varchar(255),  
street_number int,  
zipcode int,  
lat decimal(10,3),  
lon decimal(10,3)) diststyle all;
```

```
create table atm_etl.DIM_ATM(  
atm_id int not null sortkey,  
atm_number varchar(20) not null,  
atm_manufacturer varchar(50) not null,  
atm_location_id int ) diststyle all;
```

```
create table atm_etl.DIM_DATE(  
date_id int not null distkey sortkey,  
full_date_time timestamp,  
year int,  
month varchar(20),  
day int,  
hour int,  
weekday varchar(20));
```

```
create table atm_etl.DIM_CARD_TYPE(  
card_type_id int sortkey,  
card_type varchar(30)) diststyle all;
```

```
create table atm_etl.FACT_ATM_TRANS(  
trans_id bigint not null,  
atm_id int,  
weather_loc_id int distkey,  
date_id int,  
card_type_id int,  
atm_status varchar(20) sortkey,  
currency varchar(10),  
service varchar(20),  
transaction_amount int,  
message_code varchar(255),  
message_text varchar(255),  
rain_3h decimal(10,3),  
clouds_all int,  
weather_id int,  
weather_main varchar(50),  
weather_description varchar(255));
```

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 atm_etl.DIM_LOCATION from  
's3://upgrad-etl/atm_data/dim_location1.csv'  
iam_role 'arn:aws:iam::205639085974:role/upgrad-redshift-1'  
delimiter ',' region 'us-east-1';
```

```
copy atm_etl.DIM_ATM from  
's3://upgrad-etl/atm_data/dim_atm.csv'  
iam_role 'arn:aws:iam::205639085974:role/upgrad-redshift-1'  
delimiter ',' region 'us-east-1';
```

```
copy atm_etl.DIM_DATE from  
's3://upgrad-etl/atm_data/dim_date1.csv'  
iam_role 'arn:aws:iam::205639085974:role/upgrad-redshift-1'  
delimiter ',' region 'us-east-1'  
dateformat 'auto'  
timeformat 'auto'  
CSV;
```

```
copy atm_etl.DIM_CARD_TYPE from  
's3://upgrad-etl/atm_data/dim_card_type1.csv'  
iam_role 'arn:aws:iam::205639085974:role/upgrad-redshift-1'  
delimiter ',' region 'us-east-1';
```

```
copy atm_etl.FACT_ATM_TRANS from  
's3://upgrad-etl/atm_data/fact_atm_trans1.csv'  
iam_role 'arn:aws:iam::205639085974:role/upgrad-redshift-1'  
delimiter ',' region 'us-east-1'  
null 'NA'  
removequotes;
```