

## Creation of a Redshift Cluster

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

Amazon Redshift provisioned clusters

Redshift serverless New

Provisioned clusters dashboard

Clusters

Reserved nodes

Snapshots

Query editor

Query editor v2

Queries and loads

Datashares

Configurations

Workload management

Subnet groups

HSM

Manage Tags

Amazon Redshift > Clusters > redshift-cluster-2

redshift-cluster-2

Actions

Edit

Add partner integration

Query data

General information

Cluster identifier redshift-cluster-2	Status Available	Node type dc2.large	Endpoint redshift-cluster-2.cmbnqccs8kei.u...
Cluster namespace 9dd9e7c7-4aad-4e36-8c8a-61025e6d030c	Date created February 02, 2022, 22:56 (UTC+05:30)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-2.c...
	Storage used -	AQUA Not available	ODBC URL Driver={Amazon Redshift (x64)}; S...

Cluster performance

Query monitoring

Schedules

Maintenance

Properties

Recommendations (0)

To improve performance and decrease operating costs, the Amazon Redshift Advisor provides recommendations.

Alarms (0)

CloudWatch alarms are triggered when a metric threshold is met.

Events (1)

Amazon Redshift tracks events that occur on your cluster.

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 AWS KMS key ID -	Audit logging Disabled
Port 5550	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAA...		
Admin user name awsuser1			

Network and security settings

Edit

Virtual private cloud (VPC) vpc-00eff56a5b9c86fae	Availability Zone us-east-1a	VPC security group Specify which instances and devices can connect to the cluster. sg-045b52a18cb46c176	Publicly accessible Allow instances and devices outside the VPC to connect to your database through the cluster endpoint. Enabled
Subnet cluster-subnet-group-new	Enhanced VPC routing Disabled		
Endpoint URL -			

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 AWS KMS key ID -	Audit logging Disabled
Port 5550	SSH ingestion setting (cluster public key) ssh-rsa AAAAB3NzaC1yc2EAAA...		
Admin user name awsuser1			

Network and security settings

Edit

Virtual private cloud (VPC) vpc-00eff56a5b9c86fae	Availability Zone us-east-1a	VPC security group Specify which instances and devices can connect to the cluster. sg-045b52a18cb46c176	Publicly accessible Allow instances and devices outside the VPC to connect to your database through the cluster endpoint. Enabled
Subnet cluster-subnet-group-new	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:**

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**

create schema etl

1. Dim table - Location

```
create table etl.DIM_LOCATION(  
    location_id varchar distkey sortkey primary key,  
    location VARCHAR(50),  
    streetname VARCHAR(255),  
    street_number INT,  
    zipcode INT,  
    lat DECIMAL(10,3),  
    lon DECIMAL(10,3)  
)
```

--copying from S3 to location table

```
copy etl.DIM_LOCATION from  
's3://etlspark001/location/'  
iam_role 'arn:aws:iam::146883206406:role/NewRole_S3Access'  
delimiter ',' region 'us-east-1' removequotes;
```

2. Dim table - atm

```
create table etl.DIM_ATM(  
    atm_id INT distkey sortkey primary key,  
    atm_number varchar(20),  
    atm_manufacturer VARCHAR(50),  
    atm_location_id INT,  
    foreign key(atm_location_id) references etl.DIM_LOCATION(location_id)
```

)

-- copying from S3 S3 to atm table

```
copy etl.DIM_ATM from  
's3://etlspark001/atm/'  
iam_role 'arn:aws:iam::146883206406:role/NewRole_S3Access'  
delimiter ',' region 'us-east-1' removequotes;
```

### 3. Dim table - date

```
create table e_schema.DIM_DATE(  
    date_id INT distkey sortkey primary key,  
    full_date_time TIMESTAMP,  
    year INT,  
    month VARCHAR(20),  
    day INT,  
    hour INT,  
    weekday VARCHAR(20)  
)  
-- copying from S3 S3 to date table  
copy etl.DIM_DATE from  
's3://etlspark001/date/'  
iam_role 'arn:aws:iam::146883206406:role/NewRole_S3Access'  
delimiter ',' region 'us-east-1' removequotes;
```

### 4. Dim table – cardtype

```
create table etl.DIM_CARD_TYPE(  
    card_type_id INT distkey sortkey primary key,  
    card_type VARCHAR(30)  
)  
-- copying from S3 to cardtype table  
copy etl.DIM_CARD_TYPE from  
's3://etlspark001/card_type/'  
iam_role 'arn:aws:iam::146883206406:role/NewRole_S3Access'  
delimiter ',' region 'us-east-1' removequotes;
```

### 5. Fact table - FACT\_ATM\_TRANS

```
create table etl.FACT_ATM_TRANS(  
    trans_id BIGINT distkey sortkey primary key,  
    atm_id INT ,  
    weather_loc_id INT ,  
    date_id INT ,  
    card_type_id INT ,  
    atm_status VARCHAR(20) ,  
    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) ,  
foreign key(weather_loc_id) references etl.DIM_LOCATION(location_id),  
foreign key(atm_id) references etl.DIM_ATM(atm_id),  
foreign key(date_id) references etl.DIM_DATE(date_id),  
foreign key(card_type_id) references etl.DIM_CARD_TYPE(card_type_id)
```

```
)  
-- data importing from S3 to fact table  
copy etl.FACT_ATM_TRANS from  
s3://etlspark001/fact/  
iam_role 'arn:aws:iam::146883206406:role/NewRole_S3Access'  
delimiter ',' region 'us-east-1' removequotes;
```