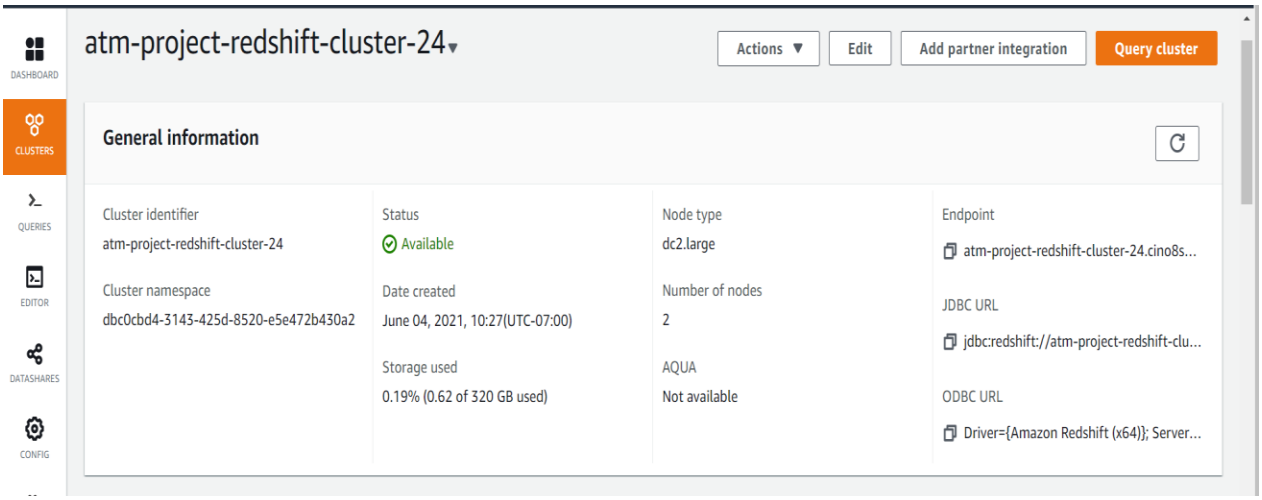
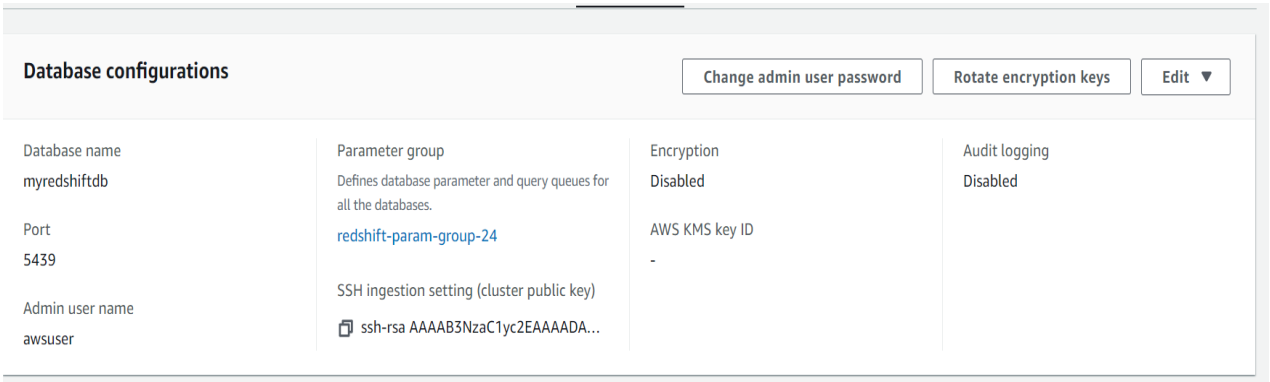
Creation of a RedShift Cluster

Screenshots of the configuration of the RedShift cluster created:

<Screenshot of the type of machine used along with number of nodes>





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 foriegn keys:

<Queries>

```
create schema transactions;
```

<Creating table DIM_LOCATION>

```
create table transactions.DIM_LOCATION(
    location_id int NOT NULL,
    location VARCHAR(50),
    streetname VARCHAR(255),
    street_number int,
    zipcode int,
    lat DECIMAL(10,
    3),
    lon DECIMAL(10,
    3),
    PRIMARY KEY (location_id)
);
```

<Creating table DIM_ATM>

```
create table transactions.DIM_ATM(
  atm_id int NOT NULL,
  atm_number VARCHAR(20),
  atm_manufacturer VARCHAR(50),
  atm_location_id INT NOT NULL,

  PRIMARY KEY (atm_id),
  foreign key(atm_location_id) references transactions.DIM_LOCATION (location_id)
  );
```

<Creating table DIM_DATE>

```
  create table transactions.DIM_DATE(
    date_id INT NOT NULL,
    full_date_time TIMESTAMP,
    year INT,
    month VARCHAR(20),
```

```
  day INT,
  hour INT,
  weekday VARCHAR(20),
  PRIMARY KEY (date_id)
  );
```

<Creating table DIM_CARD_TYPE>

```
create table transactions.DIM_CARD_TYPE(
  card_type_id INT NOT NULL,
  card_type VARCHAR(50),
  PRIMARY KEY (card_type_id)
  );
```

<Creating table FACT_ATM_TRANS >

```
create table transactions.FACT_ATM_TRANS(
  trans_id BIGINT NOT NULL,
  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),

  PRIMARY KEY (trans_id),
  FOREIGN KEY (weather_loc_id) REFERENCES transactions.DIM_LOCATION (location_id),
  FOREIGN KEY (atm_id) REFERENCES transactions.DIM_ATM (atm_id),
  FOREIGN KEY (date_id) REFERENCES transactions.DIM_DATE (date_id),
  FOREIGN KEY (card_type_id) REFERENCES transactions.DIM_CARD_TYPE (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

<Queries>

<Copying dim_card_type data from S3 to RedShift>

```
copy transactions.dim_card_type (card_type, card_type_id)
from 's3://my-redshift-bucket-24/card_type'
iam_role 'arn:aws:iam::164541431167:role/my-redshift-role-permission-24'
csv
```

<Copying dim_location data from S3 to RedShift>

```
copy transactions.dim_location (location, streetname, street_number, zipcode, lat, lon,
location_id)
from 's3://my-redshift-bucket-24/location'
iam_role 'arn:aws:iam::164541431167:role/my-redshift-role-permission-24'
csv
```

<Copying dim_atm data from S3 to RedShift>

```
copy transactions.dim_atm (atm_number, atm_manufacturer, atm_location_id, atm_id)
from 's3://my-redshift-bucket-24/atm'
iam_role 'arn:aws:iam::164541431167:role/my-redshift-role-permission-24'
csv
```

<Copying dim_date data from S3 to RedShift>

```
copy transactions.dim_date (year,month,day,hour,weekday,date_id)
from 's3://my-redshift-bucket-24/date'
iam_role 'arn:aws:iam::164541431167:role/my-redshift-role-permission-24'
csv
```

<Copying fact_atm_trans data from S3 to RedShift>

```
copy transactions.fact_atm_trans (atm_status, currency, service, transaction_amount,
message_code, message_text, rain_3h, clouds_all, weather_id, weather_main,
weather_description, weather_loc_id, date_id, card_type_id, atm_id, trans_id)
from 's3://my-redshift-bucket-24/trans'
iam_role 'arn:aws:iam::164541431167:role/my-redshift-role-permission-24'
emptyasnull
csv
```