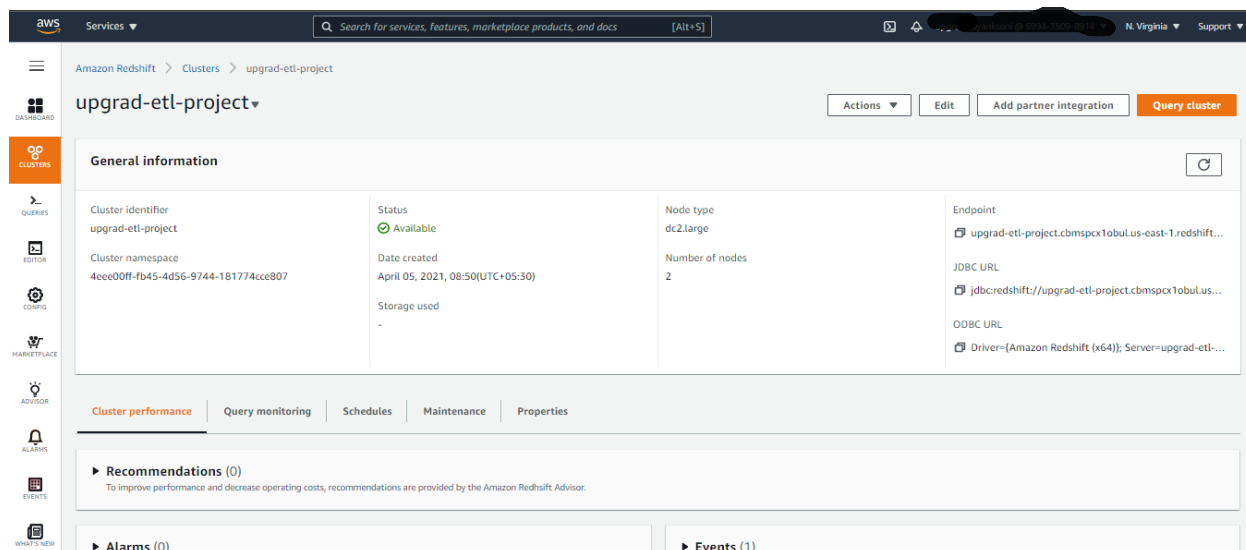


Creation of a RedShift Cluster

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



The screenshot displays the AWS Management Console interface for an Amazon RedShift cluster. The breadcrumb navigation shows the path: Amazon Redshift > Clusters > upgrad-etl-project. The cluster name 'upgrad-etl-project' is prominently displayed at the top, along with action buttons: Actions, Edit, Add partner integration, and Query cluster.

The 'General information' tab is active, showing the following details:

Cluster identifier	Status	Node type	Endpoint
upgrad-etl-project	Available	dc2.large	upgrad-etl-project.cbmspcx1obul.us-east-1.redshift...
Cluster namespace	Date created	Number of nodes	JDBC URL
4ee00ff-fb45-4d56-9744-181774cce807	April 05, 2021, 08:50(UTC+05:30)	2	jdbc:redshift://upgrad-etl-project.cbmspcx1obul.us...
	Storage used		ODBC URL
	-		Driver=[Amazon Redshift (x64)]; Server=upgrad-etl-...

Below the general information, there are tabs for Cluster performance, Query monitoring, Schedules, Maintenance, and Properties. The 'Recommendations' section shows 0 recommendations, and the 'Alarms' section shows 0 alarms. The 'Events' section shows 1 event.

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:

--Creating Schema

```
create schema etl_schema;
```

--Creating Tables

```
create table etl_schema.dim_location(  
    location_id integer not null distkey sortkey,  
    location varchar(50),  
    streetname varchar(255),  
    street_number integer,  
    zipcode integer,  
    lat decimal(10,3),  
    lon decimal(10,3),  
    primary key(location_id)  
);
```

```
create table etl_schema.dim_atm(  
    atm_id integer not null distkey sortkey,  
    atm_number varchar(20),  
    atm_manufacturer varchar(50),  
    atm_location_id integer,  
    primary key (atm_id),  
    foreign key (atm_location_id) references etl_schema.dim_location(location_id)  
);
```

```
create table etl_schema.dim_date(  
    date_id integer not null distkey sortkey,  
    full_date_time timestamp,  
    year integer,  
    month varchar(20),  
    day integer,  
    hour integer,  
    weekday varchar(20),  
    primary key(date_id)  
);
```

```
create table etl_schema.dim_card_type(  
    card_type_id integer not null distkey sortkey,
```

```
card_type varchar(25),
primary key (card_type_id)
);

create table etl_schema.fact_atm_trans(
trans_id bigint not null distkey,
atm_id integer not null,
weather_loc_id integer not null,
date_id integer not null,
card_type_id integer not null,
atm_status varchar(20),
currency varchar(10),
service varchar(20),
transaction_amount integer,
message_code varchar(255),
message_text varchar(255),
rain_3h decimal(10,3),
clouds_all integer,
weather_id integer,
weather_main varchar(50),
weather_description varchar(255),
primary key(trans_id),
foreign key(weather_loc_id) references etl_schema.dim_location (location_id),
foreign key(atm_id) references etl_schema.dim_atm (atm_id),
foreign key(date_id) references etl_schema.dim_date (date_id)
) sortkey(trans_id, atm_id, weather_loc_id, date_id, 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

```
--load dim_location
copy etl_schema.dim_location
from 's3://etl-project/csv/dim_location.csv'
iam_role 'arn:aws:iam::XXXXXXXXXX:role/XXXXXXXXXX'
delimiter ',' region 'us-east-1'
CSV
IGNOREHEADER 1;

--load dim_atm
copy etl_schema.dim_atm
```

```
from 's3://etl-project/csv/dim_atm.csv'
iam_role 'arn:aws:iam::XXXXXXXXXX:role/XXXXXXXXXX'
delimiter ',' region 'us-east-1'
CSV
IGNOREHEADER 1;

--load dim_date
copy etl_schema.dim_date
from 's3://etl-project/csv/dim_date.csv'
iam_role 'arn:aws:iam::XXXXXXXXXX:role/XXXXXXXXXX'
delimiter ',' region 'us-east-1'
CSV
IGNOREHEADER 1;

--load dim_card_type
copy etl_schema.dim_card_type
from 's3://etl-project/csv/dim_card_type.csv'
iam_role 'arn:aws:iam::XXXXXXXXXX:role/XXXXXXXXXX'
delimiter ',' region 'us-east-1'
CSV
IGNOREHEADER 1;

--load fact_atm_trans
copy etl_schema.fact_atm_trans
from 's3://etl-project/csv/fact_atm_trans.csv'
iam_role 'arn:aws:iam::XXXXXXXXXX:role/XXXXXXXXXX'
delimiter ',' region 'us-east-1'
CSV
IGNOREHEADER 1;
```