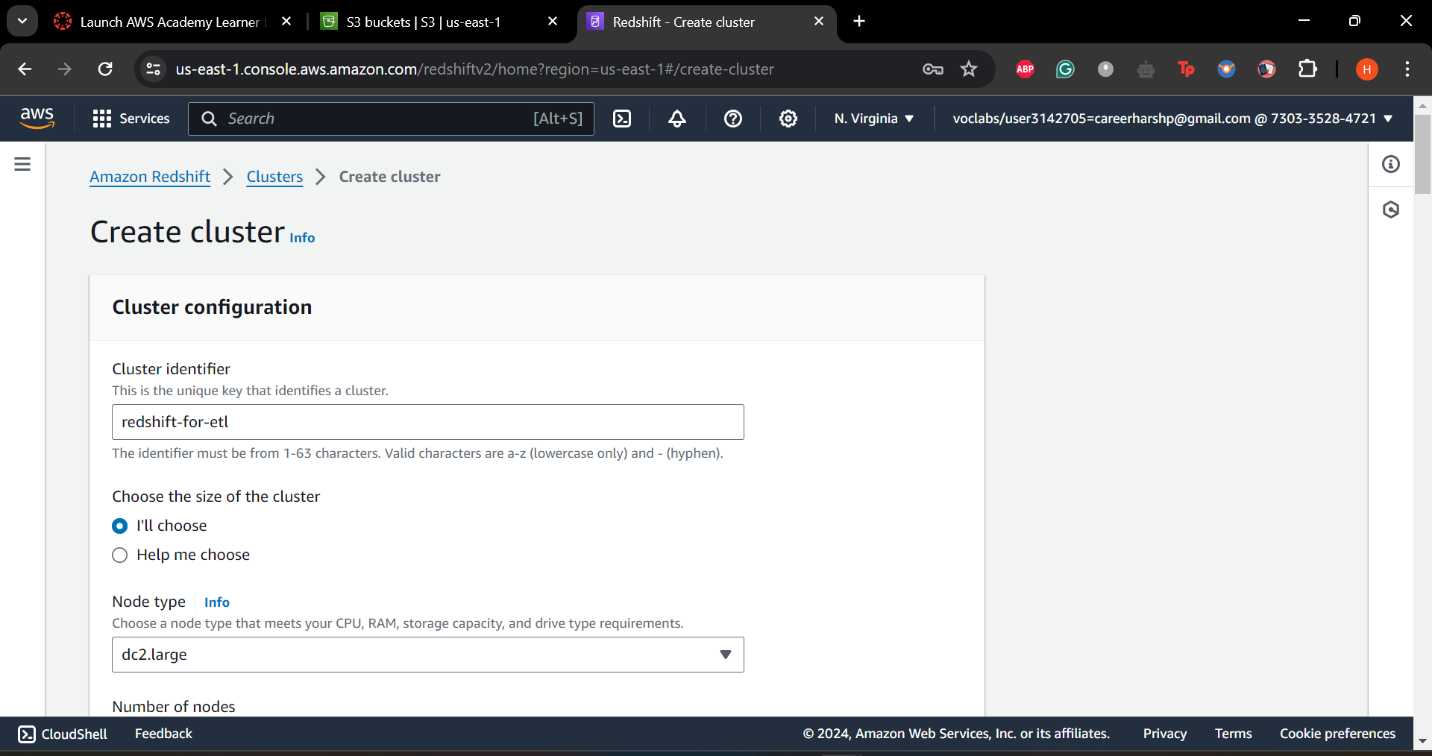
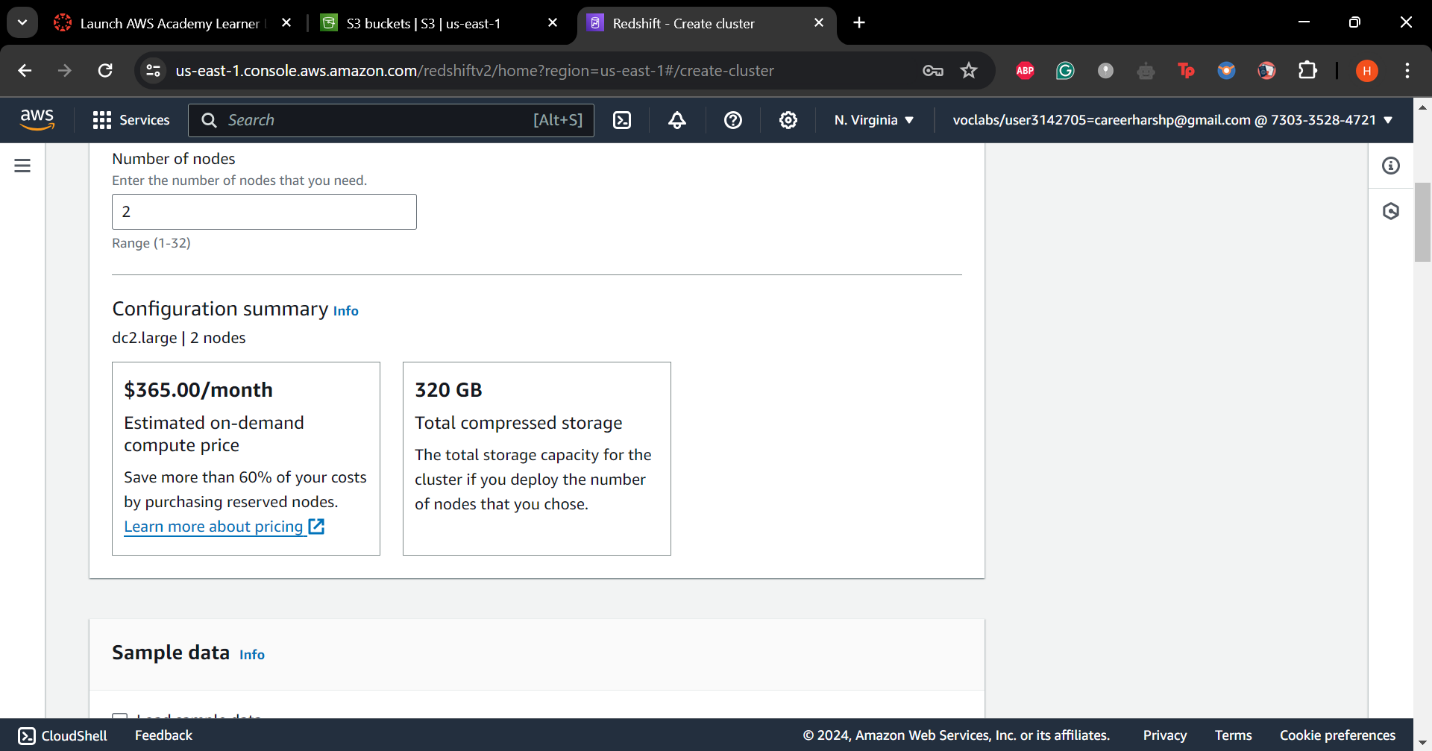
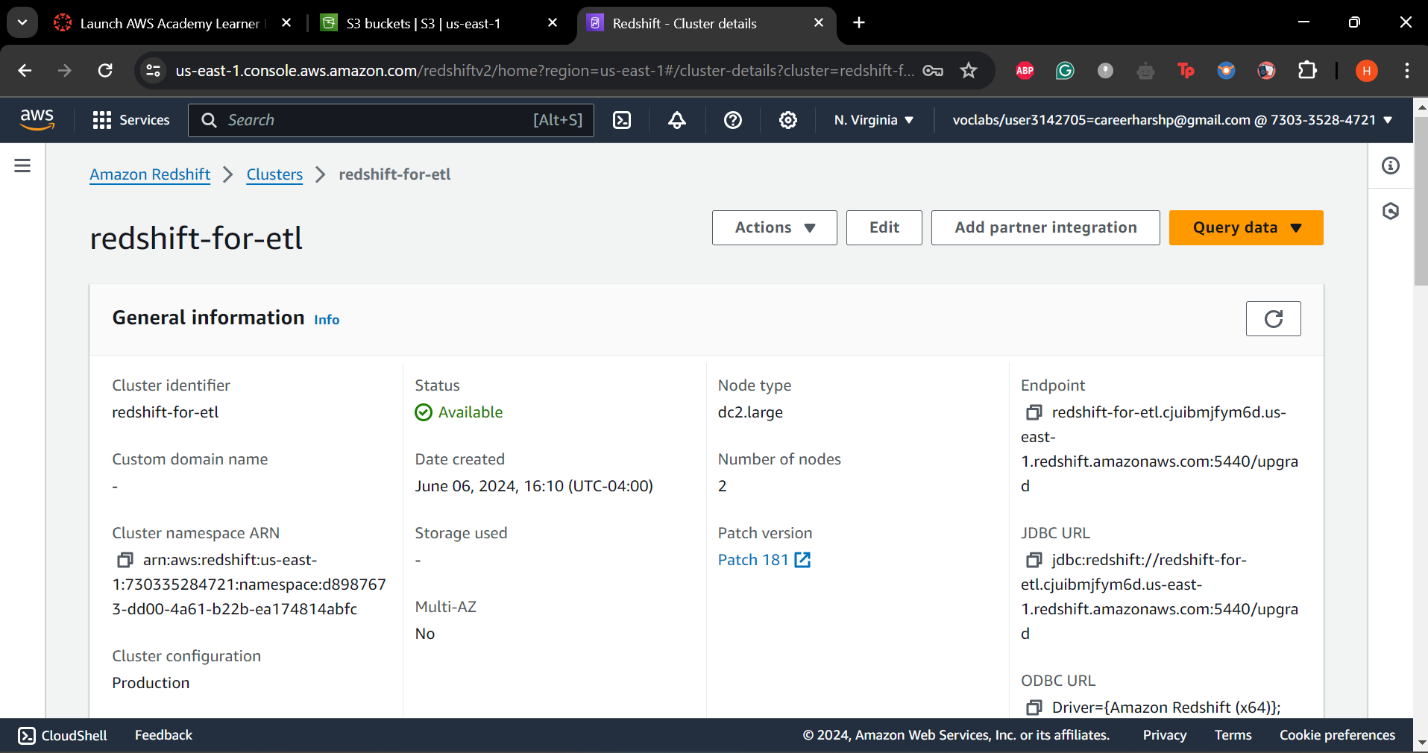
### Creation of a Redshift Cluster

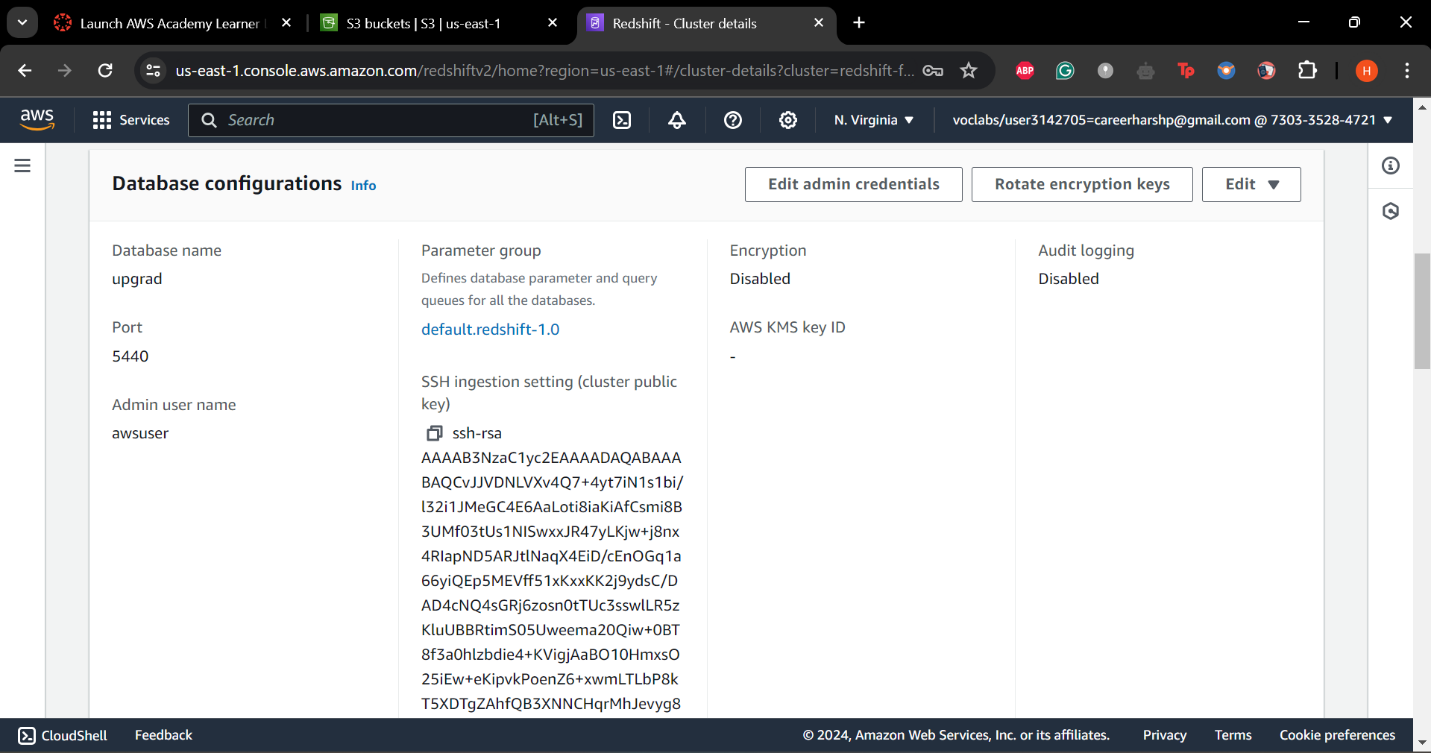
**Screenshots of the configuration of the Redshift cluster that you have created:**

<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 foreign keys:**

<Queries>

create schema etl\_atm\_data;

create table if not exists etl\_atm\_data.DIM\_LOCATION(

location\_id int not null sortkey distkey,

location varchar(50),

streetname varchar(255),

street\_number int,

zipcode int,

lat decimal(10,3),

lon decimal(10,3),

PRIMARY KEY(location\_id)

);

create table if not exists etl\_atm\_data.DIM\_ATM(

atm\_id int not null distkey sortkey,

atm\_number varchar(20),

atm\_manufacturer varchar(50),

atm\_location\_id int,

PRIMARY KEY(atm\_id),

FOREIGN KEY(atm\_location\_id) REFERENCES etl\_atm\_data.DIM\_LOCATION(location\_id)

);

create table if not exists etl\_atm\_data.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),

PRIMARY KEY(date\_id)

);

create table if not exists etl\_atm\_data.DIM\_CARD\_TYPE(

card\_type\_id int not null distkey sortkey,

card\_type varchar(30),

PRIMARY KEY(card\_type\_id)

);

create table if not exists etl\_atm\_data.FACT\_ATM\_TRANS(

trans\_id bigint not null distkey sortkey,

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 etl\_atm\_data.DIM\_LOCATION(location\_id),

FOREIGN KEY(atm\_id) REFERENCES etl\_atm\_data.DIM\_ATM(atm\_id),

FOREIGN KEY(date\_id) REFERENCES etl\_atm\_data.DIM\_DATE(date\_id),

FOREIGN KEY(card\_type\_id) REFERENCES etl\_atm\_data.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>

copy etl\_atm\_data.DIM\_LOCATION from 's3://etl-project-upgrad/dim\_location/part-00000-4497a930-8930-45db-93a5-705c7d58d299-c000.csv' iam\_role 'arn:aws:iam::730335284721:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV;

copy etl\_atm\_data.DIM\_ATM from 's3://etl-project-upgrad/dim\_atm/part-00000-dc0c784f-40a6-4330-a08d-553d23e030a5-c000.csv' iam\_role 'arn:aws:iam::730335284721:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV;

copy etl\_atm\_data.DIM\_DATE from 's3://etl-project-upgrad/dim\_date/part-00000-339f7738-7b0f-47ba-8fed-c81f7c1579e2-c000.csv' iam\_role 'arn:aws:iam::730335284721:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV timeformat 'auto';

copy etl\_atm\_data.DIM\_CARD\_TYPE from 's3://etl-project-upgrad/dim\_card\_type/part-00000-666ce7dc-4bc4-4312-99db-c75d5c570b8f-c000.csv' iam\_role 'arn:aws:iam::730335284721:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV;

copy etl\_atm\_data.FACT\_ATM\_TRANS from 's3://etl-project-upgrad/fact\_atm\_trans/part-00000-f625ca03-24bb-4275-9a39-fb54a1eeb281-c000.csv' iam\_role 'arn:aws:iam::730335284721:role/myRedshiftRole' delimiter ',' region 'us-east-1' CSV;