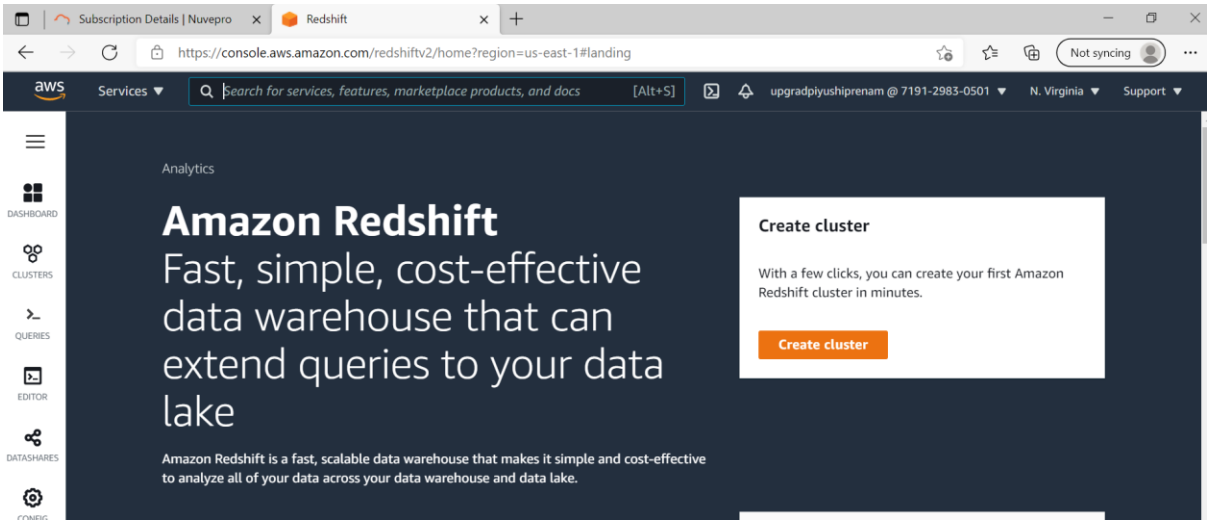
Creation of a RedShift Cluster

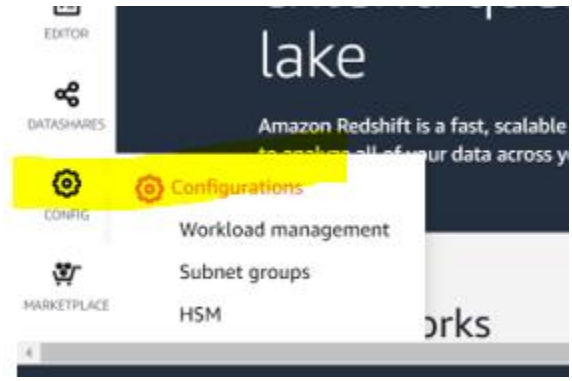
Screenshots of the configuration of the RedShift cluster that you have created:

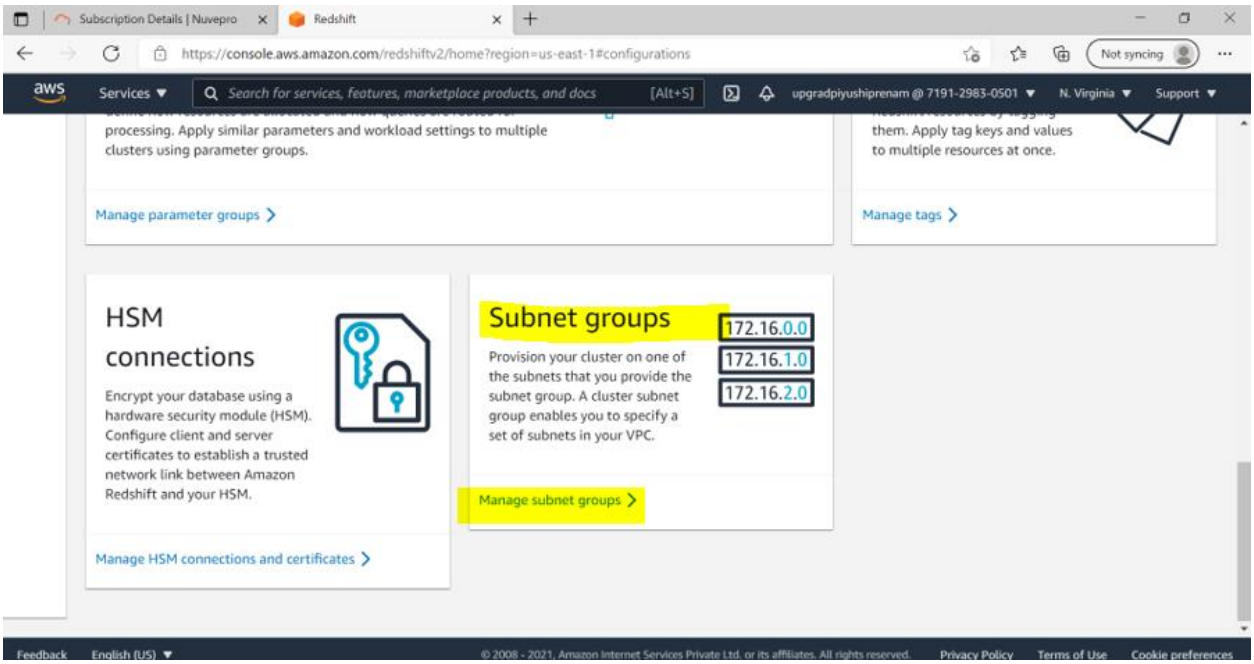
Navigate to aws console and search for Redshift

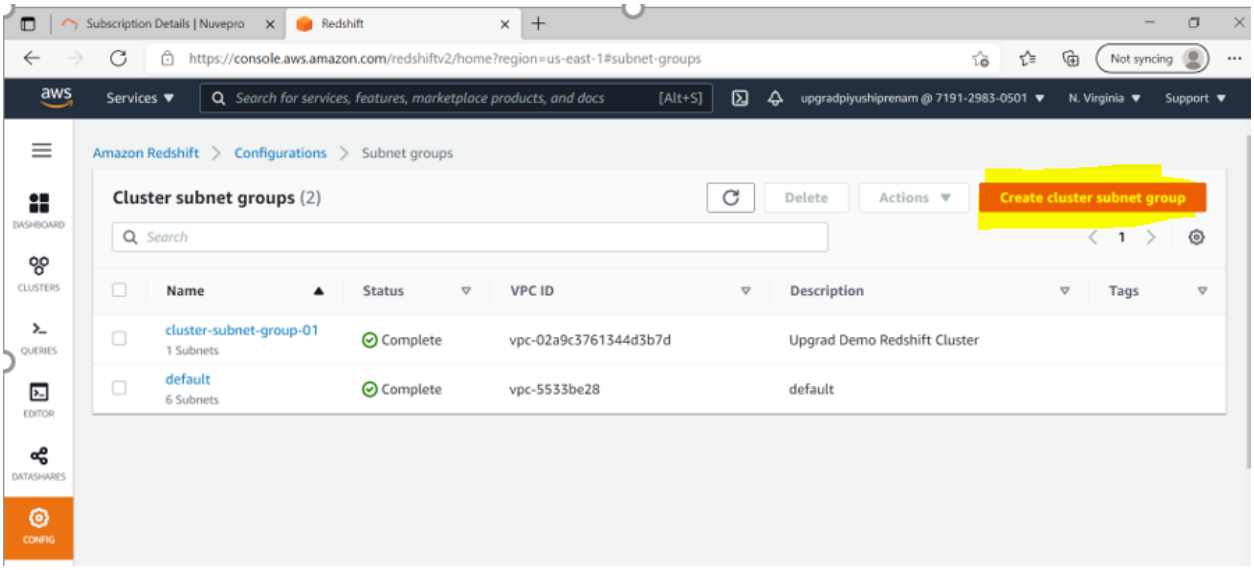


Click on Configurations

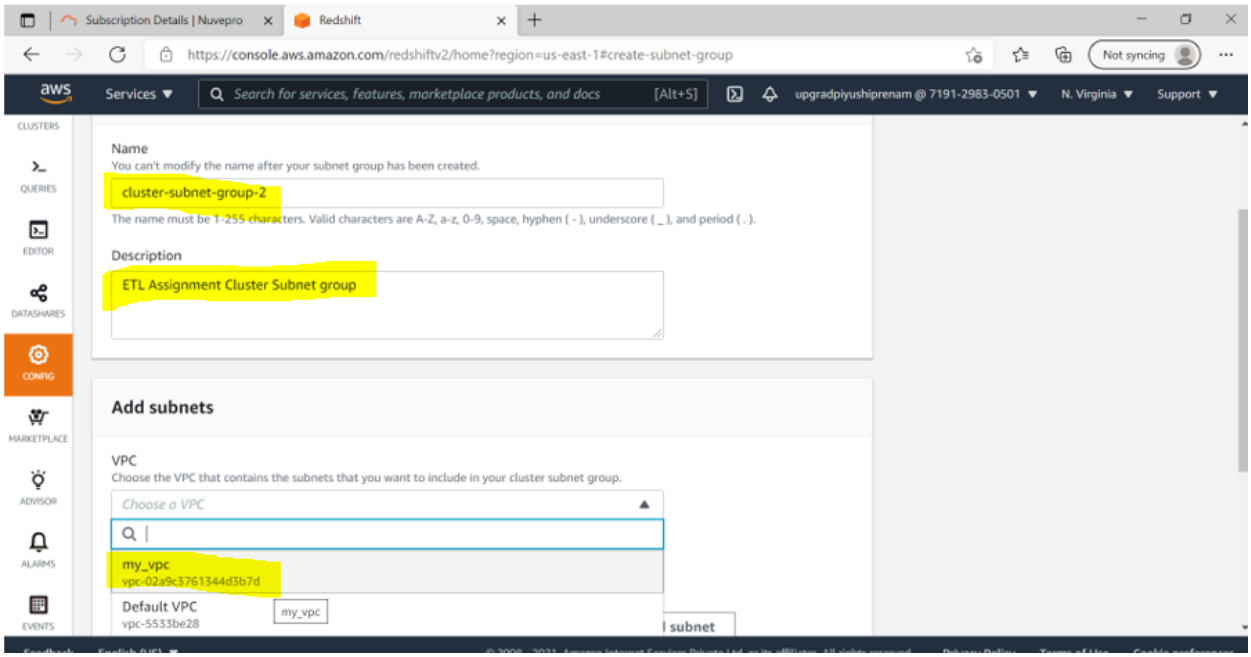
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

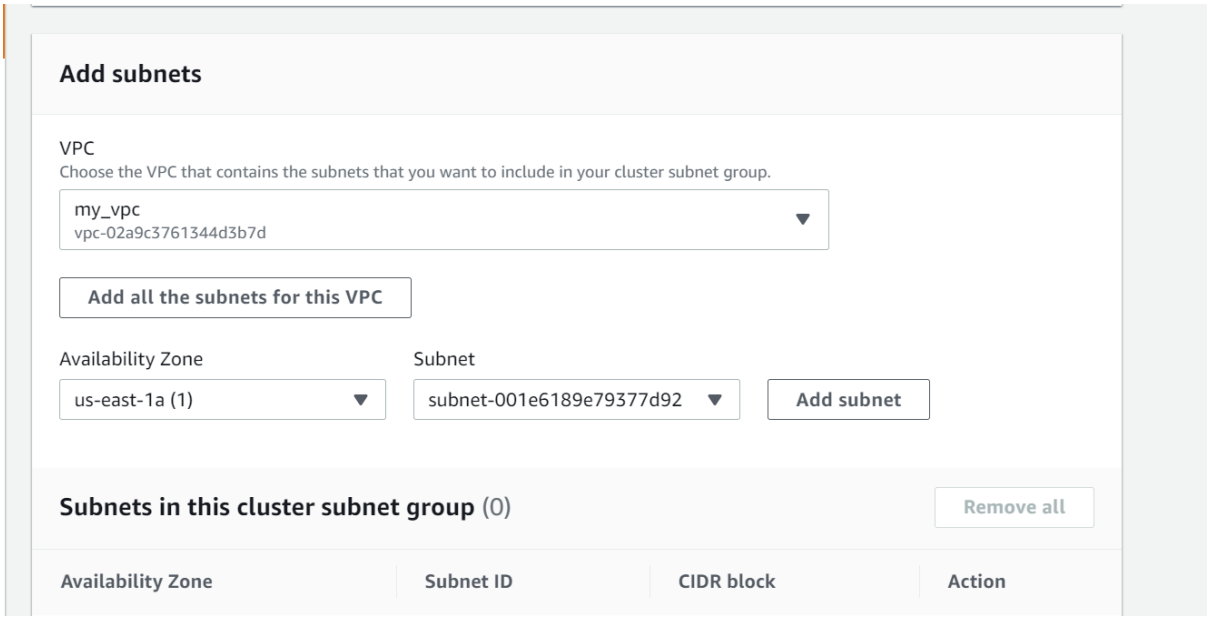




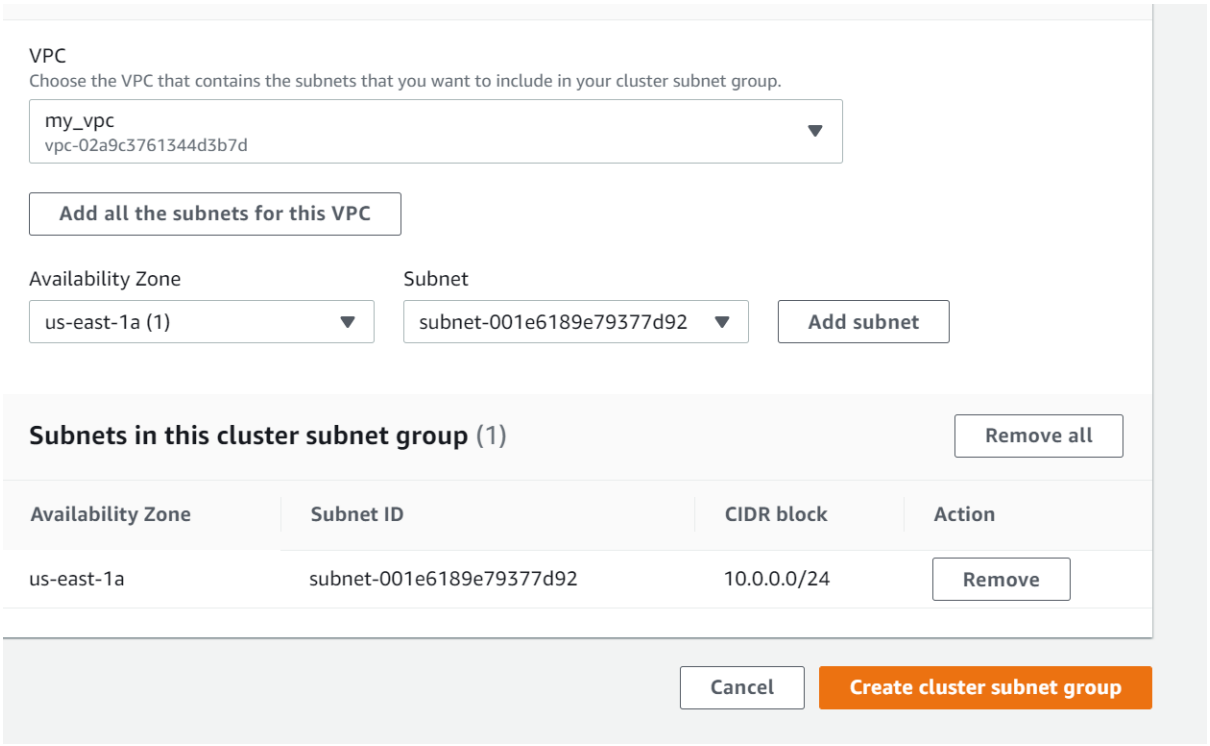


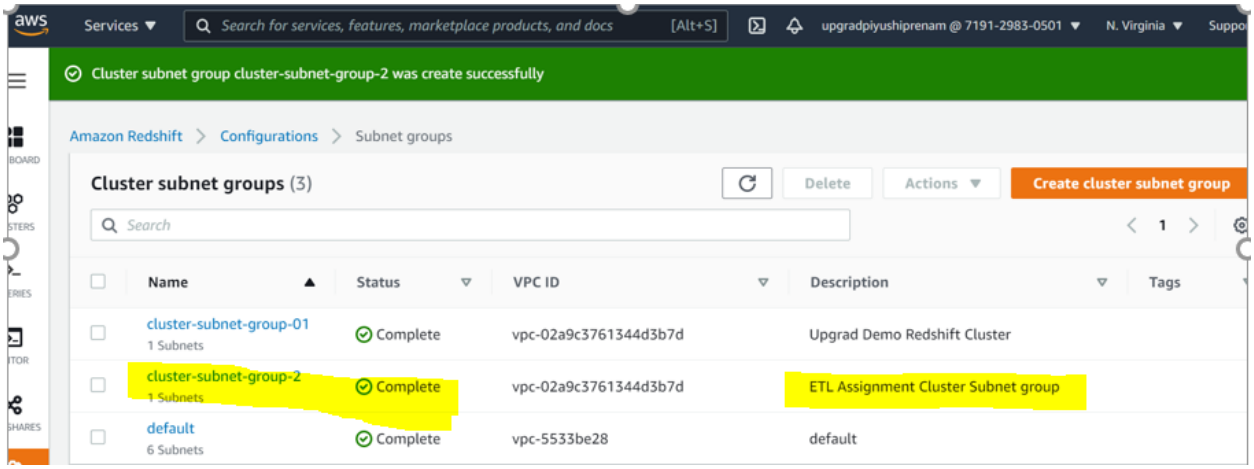
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved



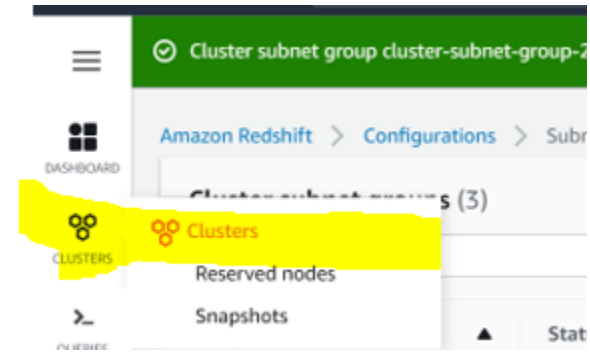


© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

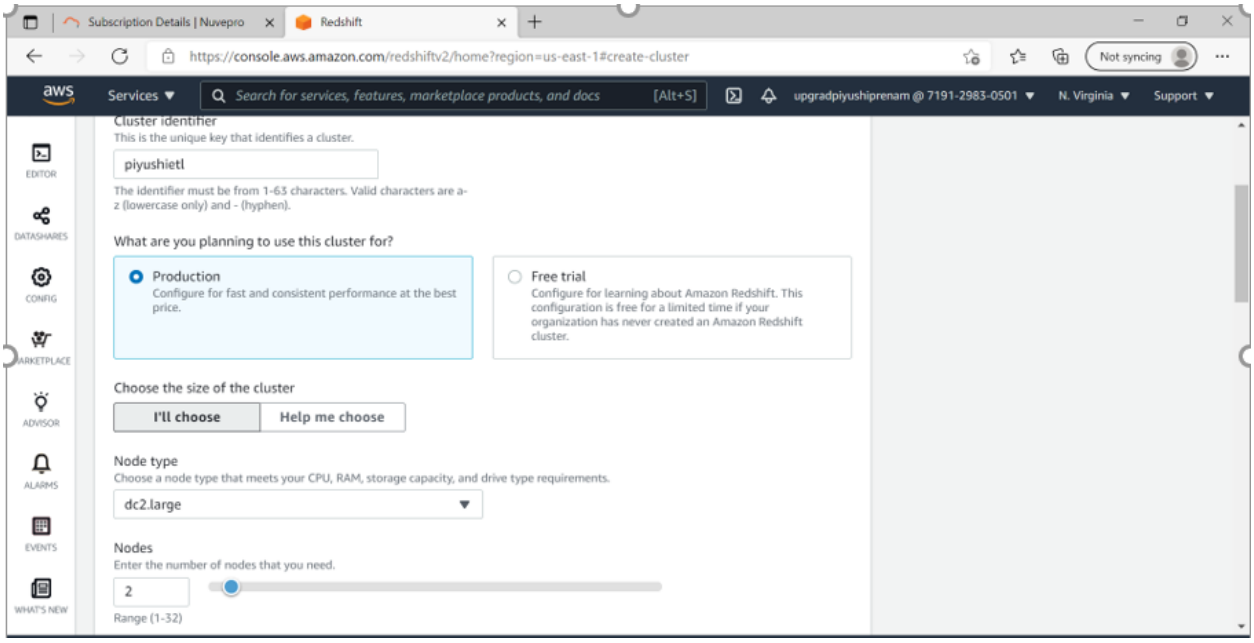


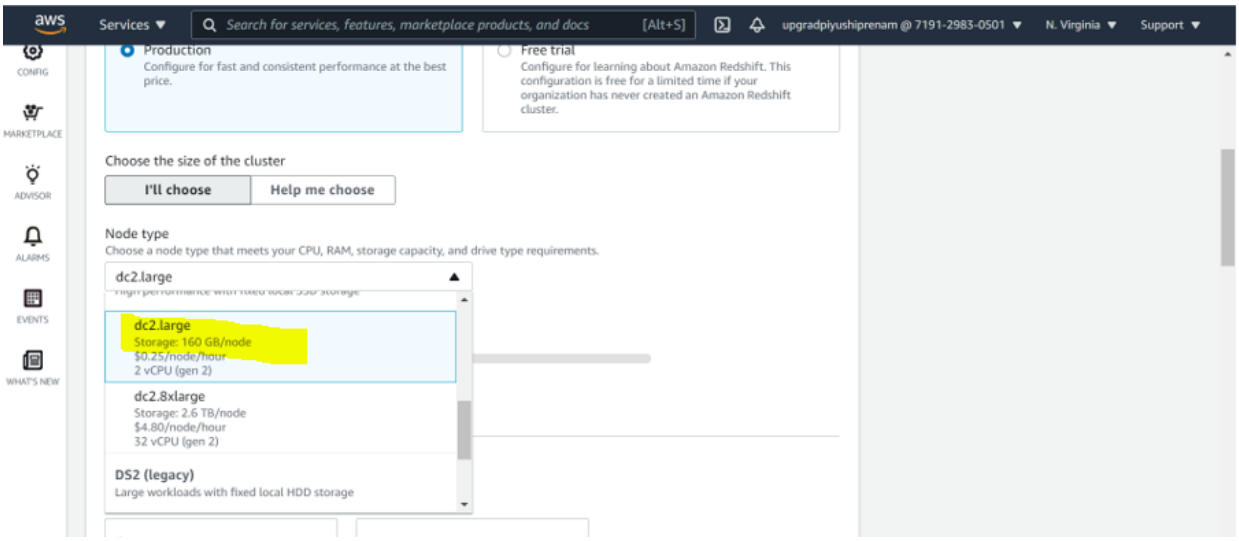


For Cluster creation



© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved





© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

Database configurations

Admin user name

Enter a login ID for the admin user of your DB instance.

awsuser

The name must be 1-128 alphanumeric characters, and it can't be a **reserved word**.

☐ Auto generate password

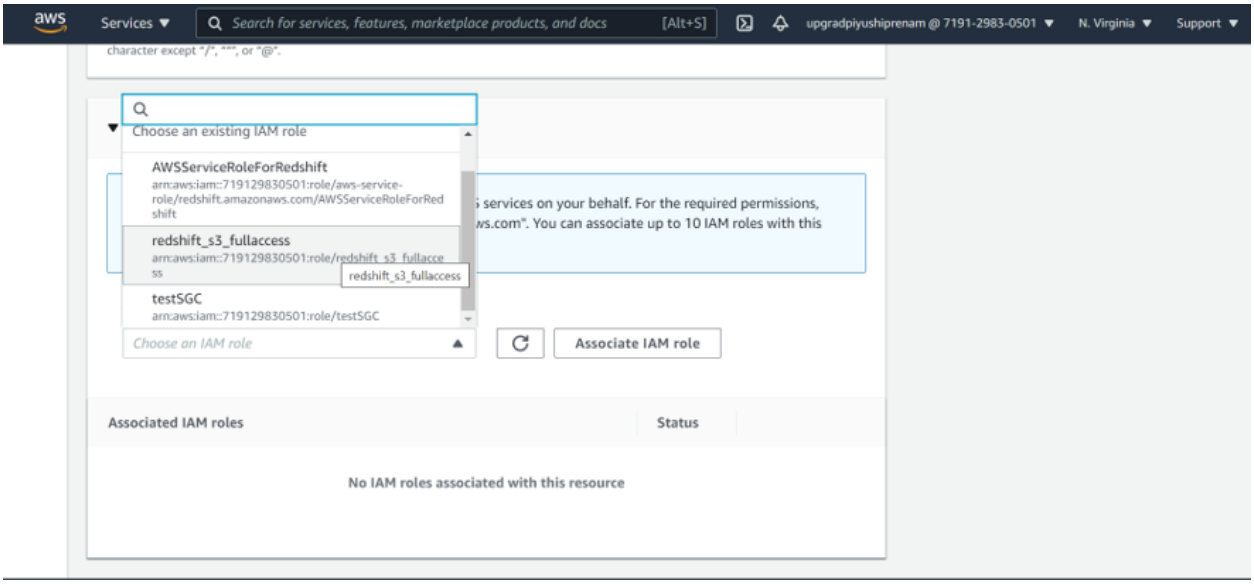
Amazon Redshift can generate a password for you, or you can specify your own password.

Admin user password

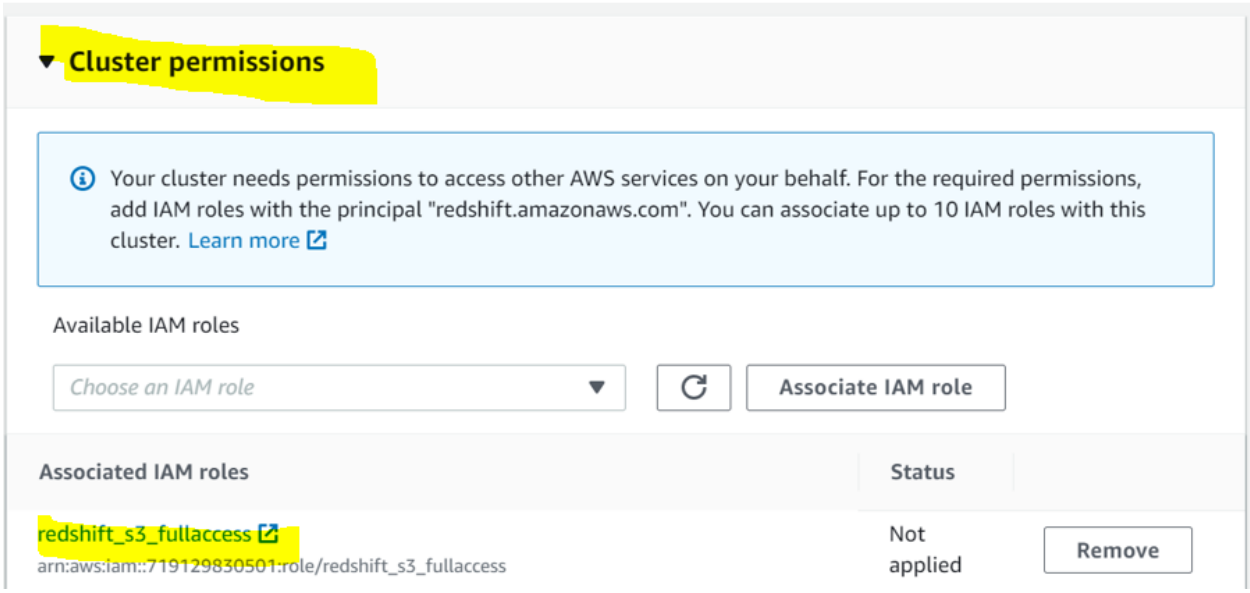
••••••••••••••

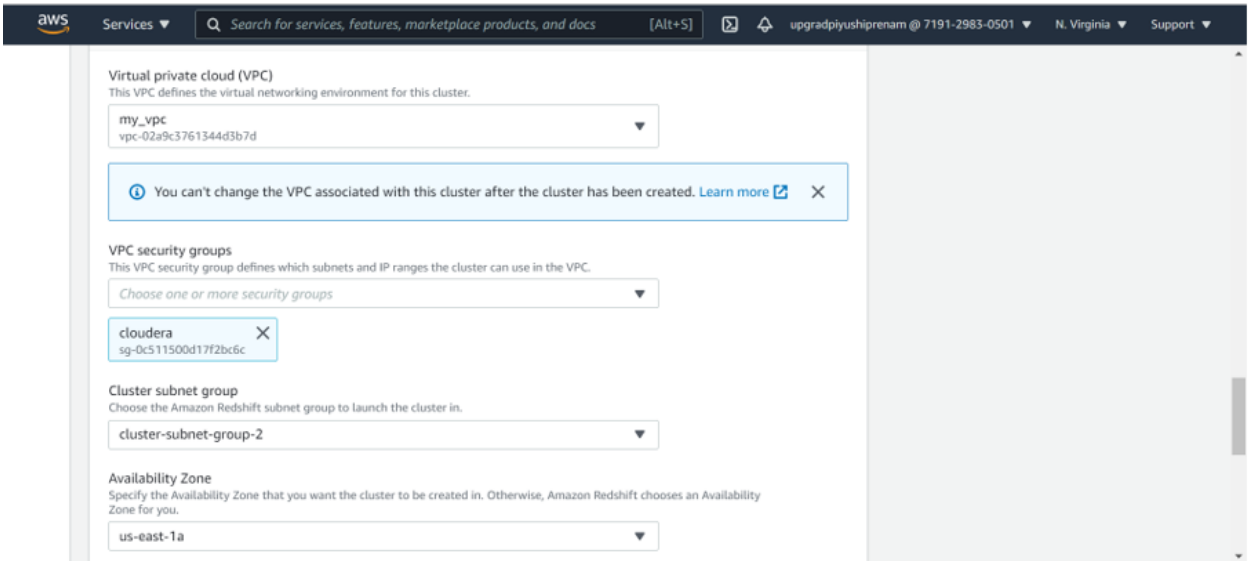
☐ Show password

Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", """, or "@".

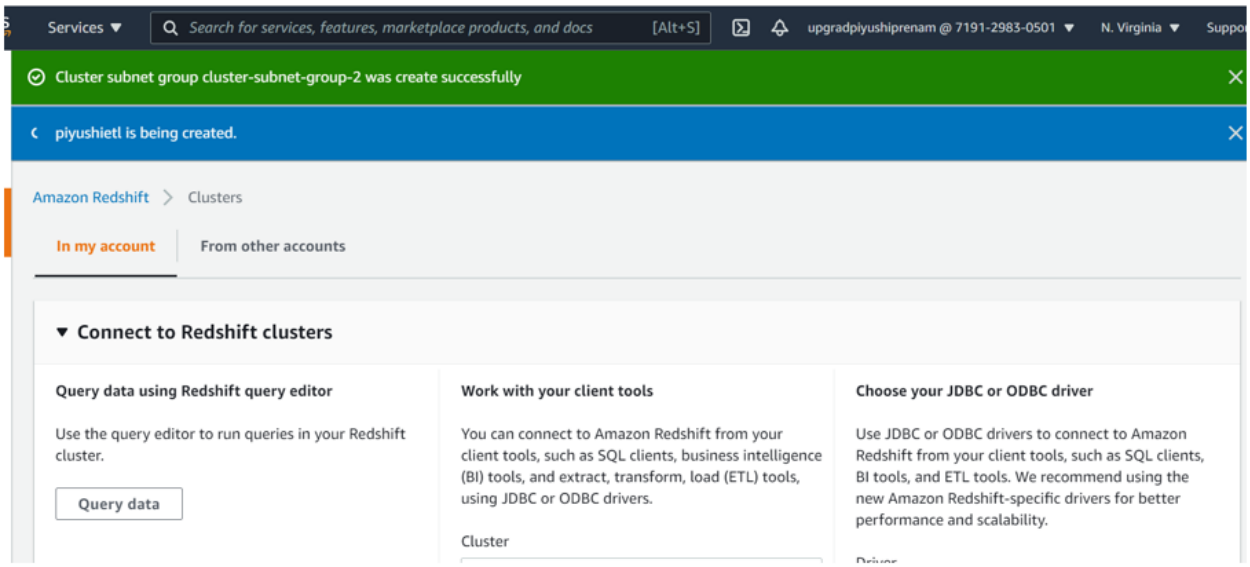


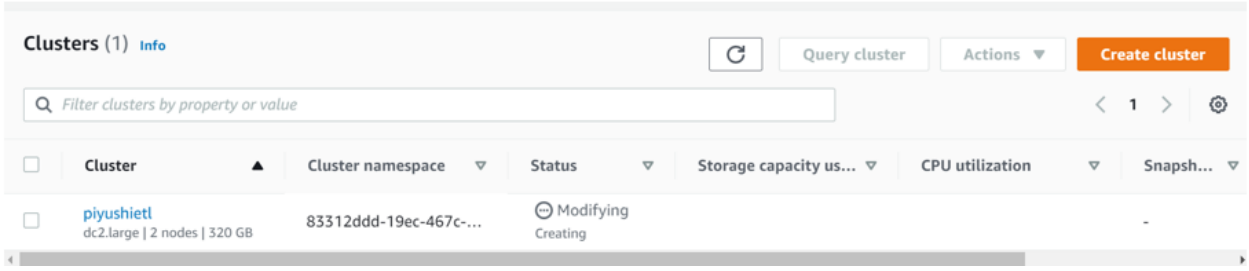
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

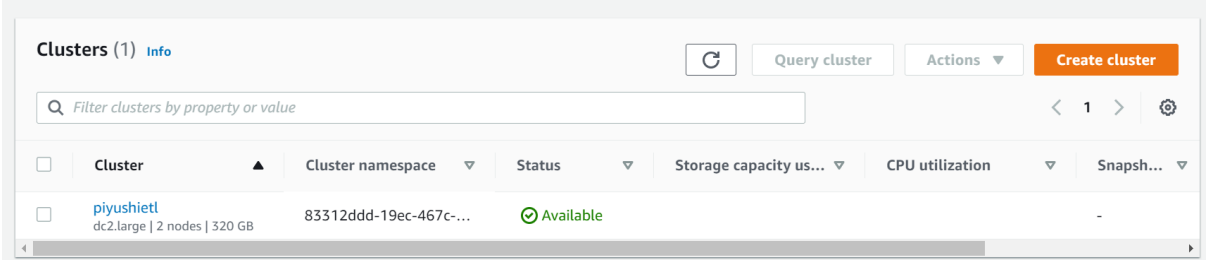




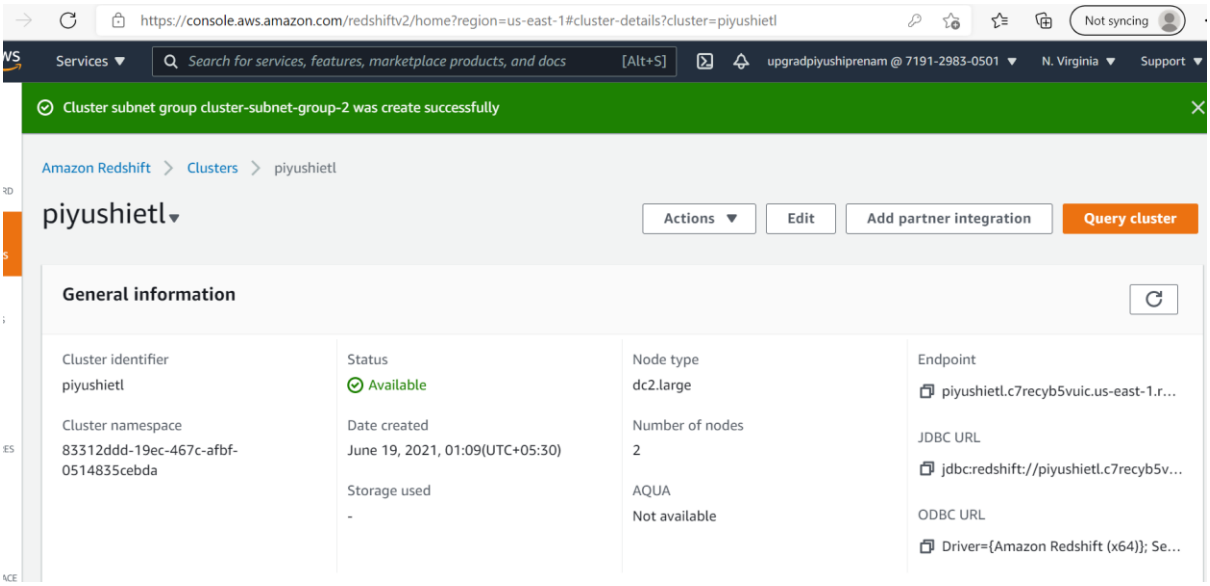
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved



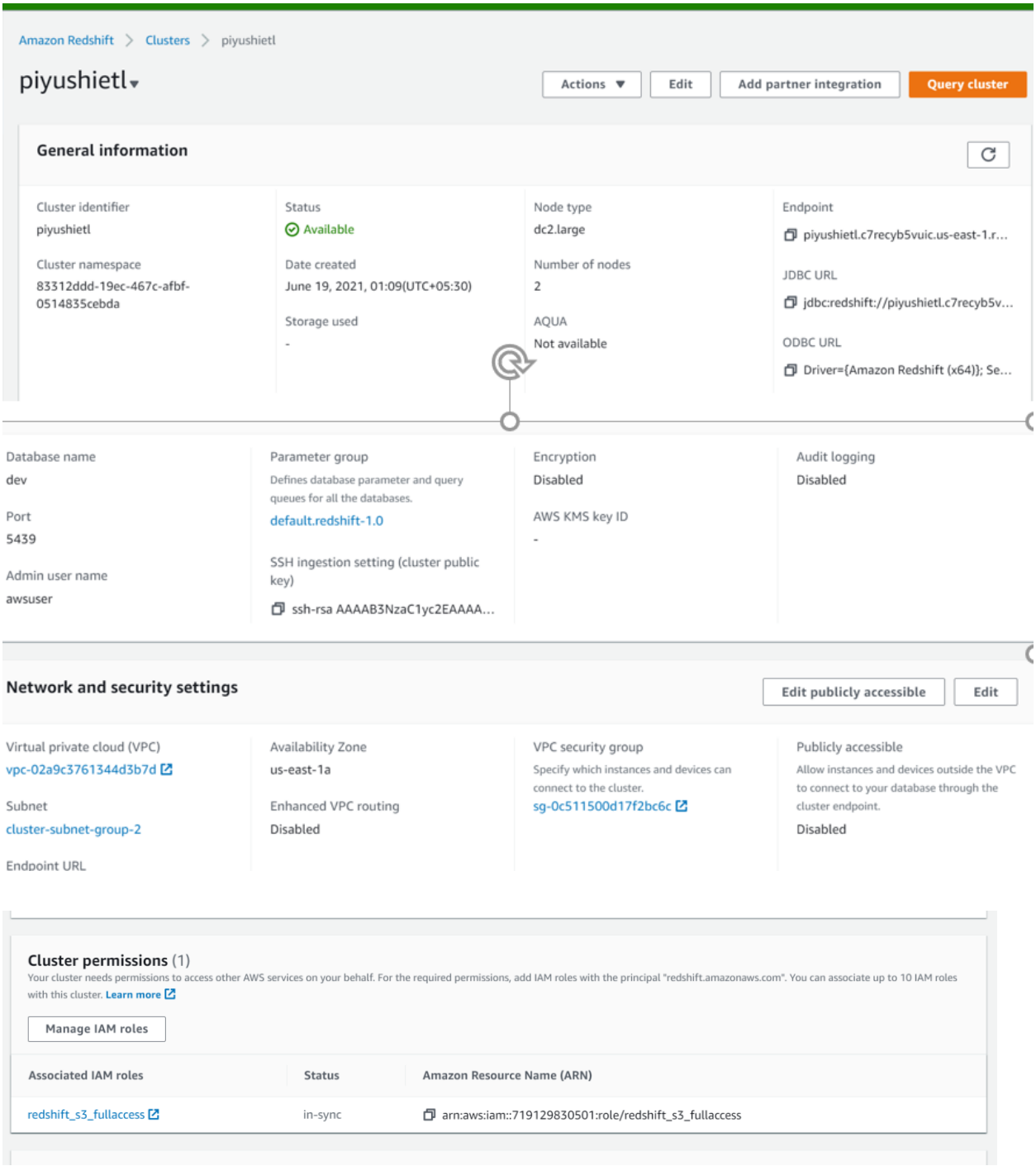




© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

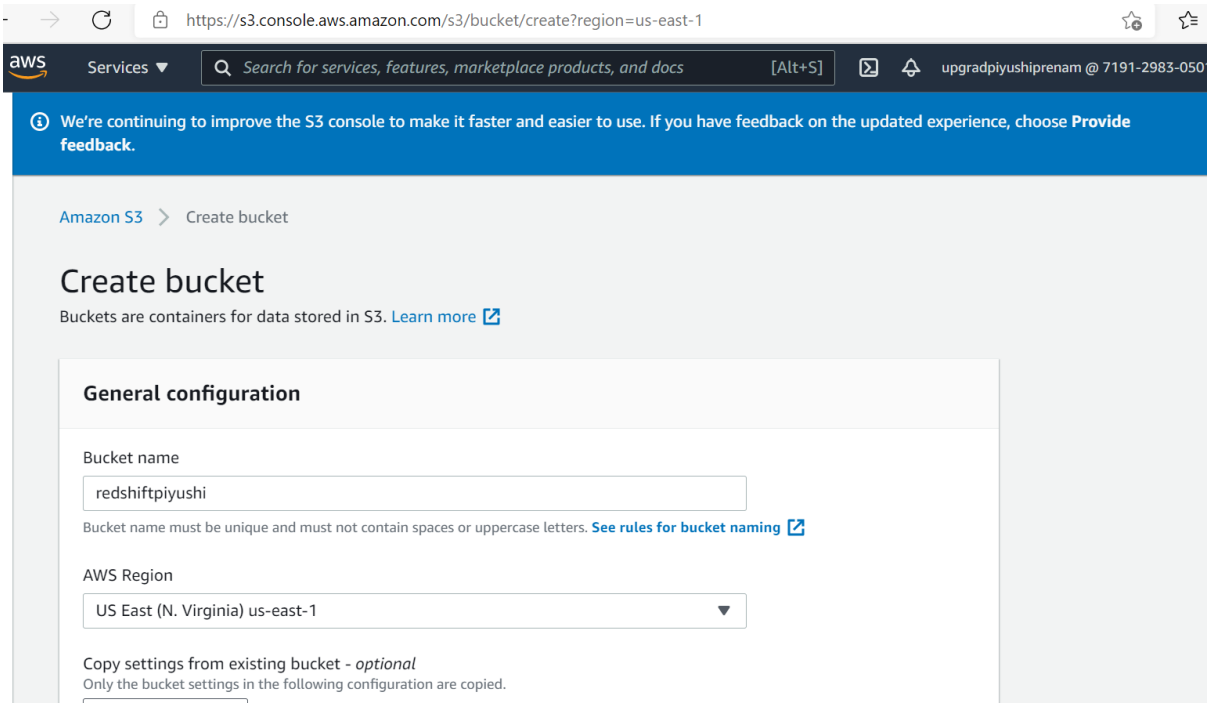


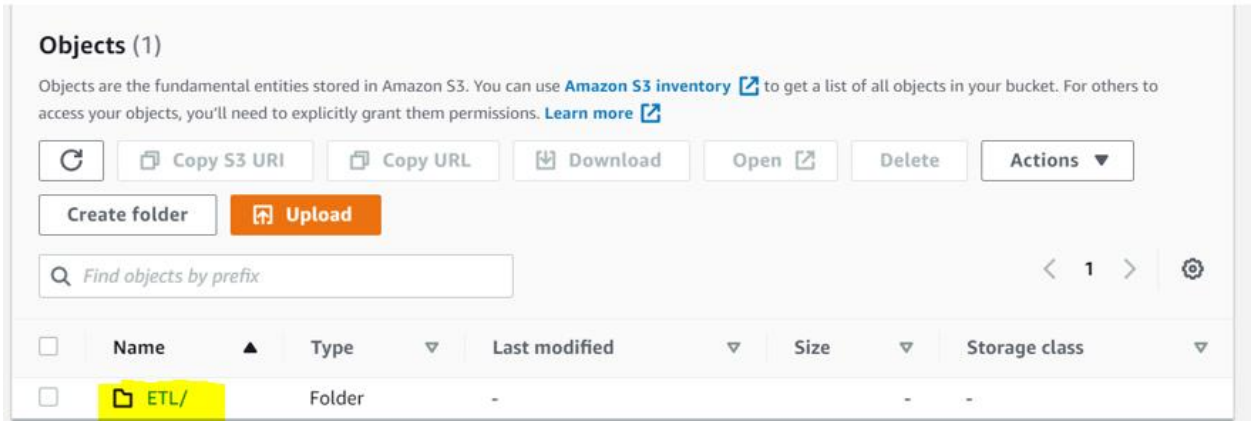
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved



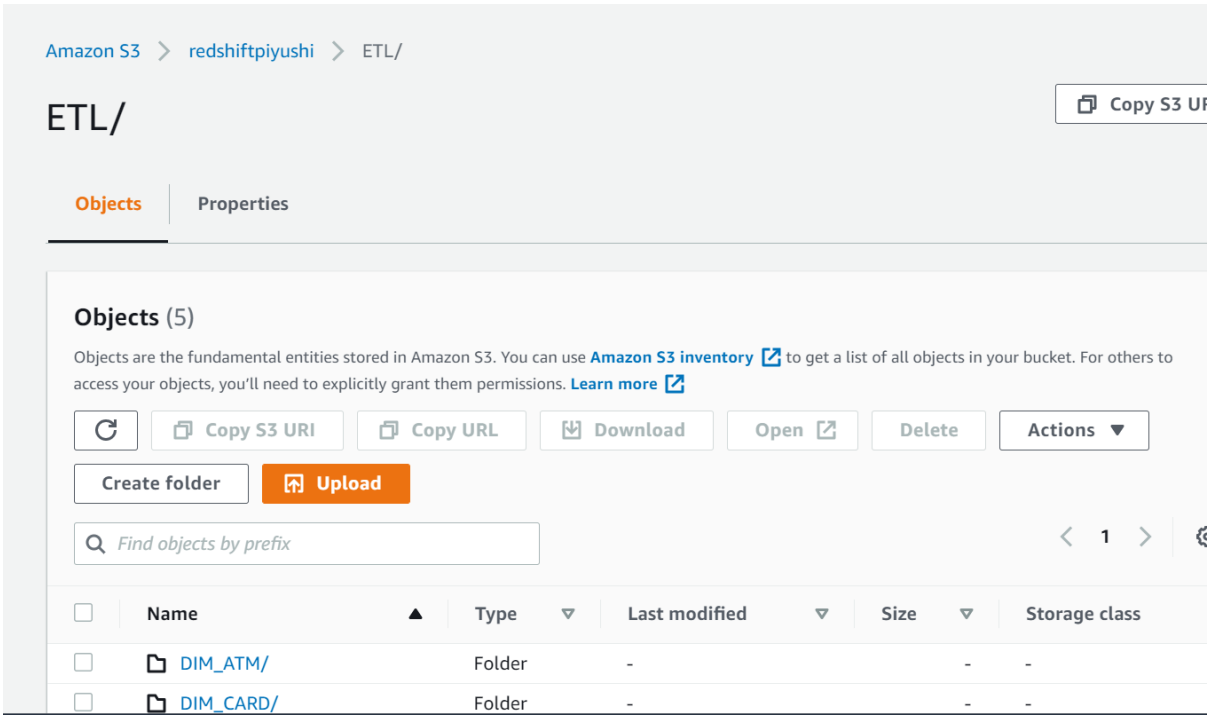
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

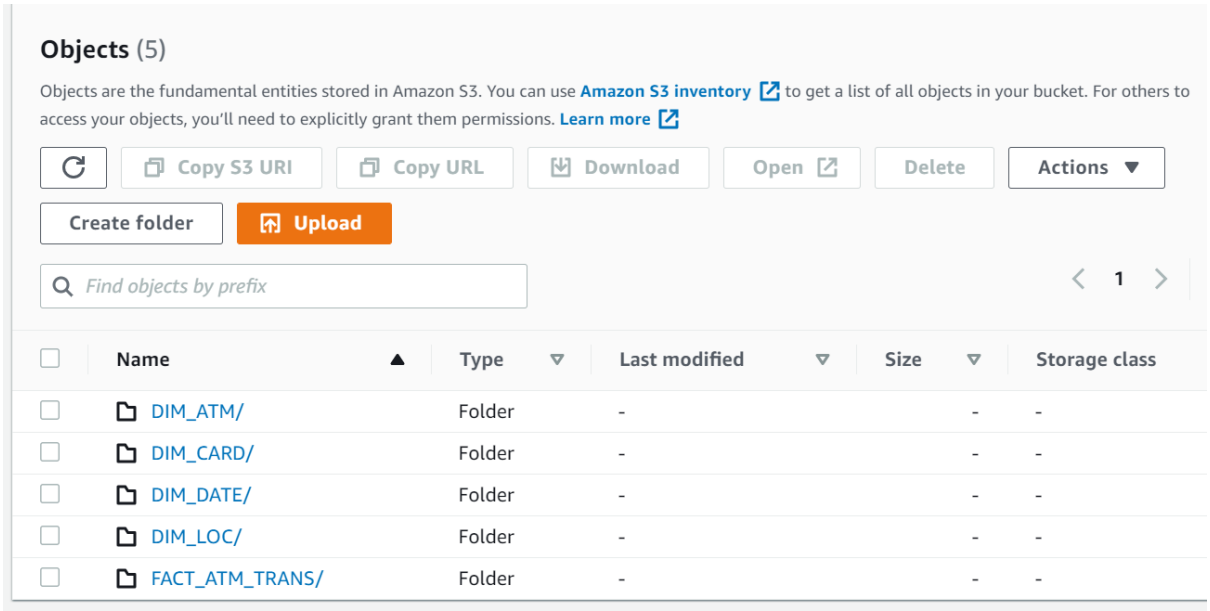
Setting up a database in the RedShift cluster and running queries to create the dimension and fact tables

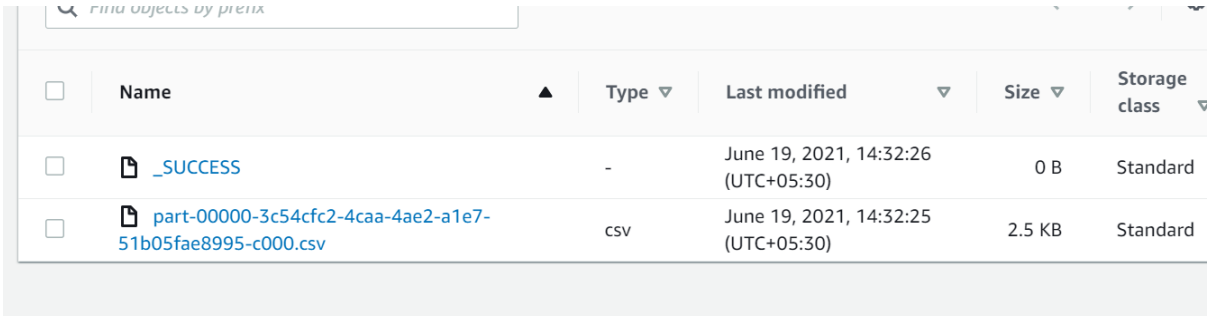




© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

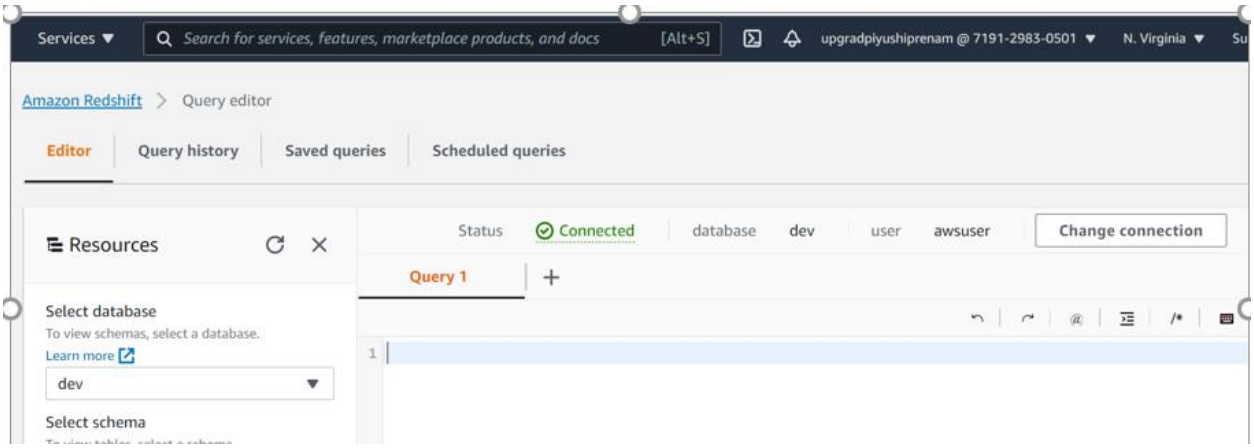


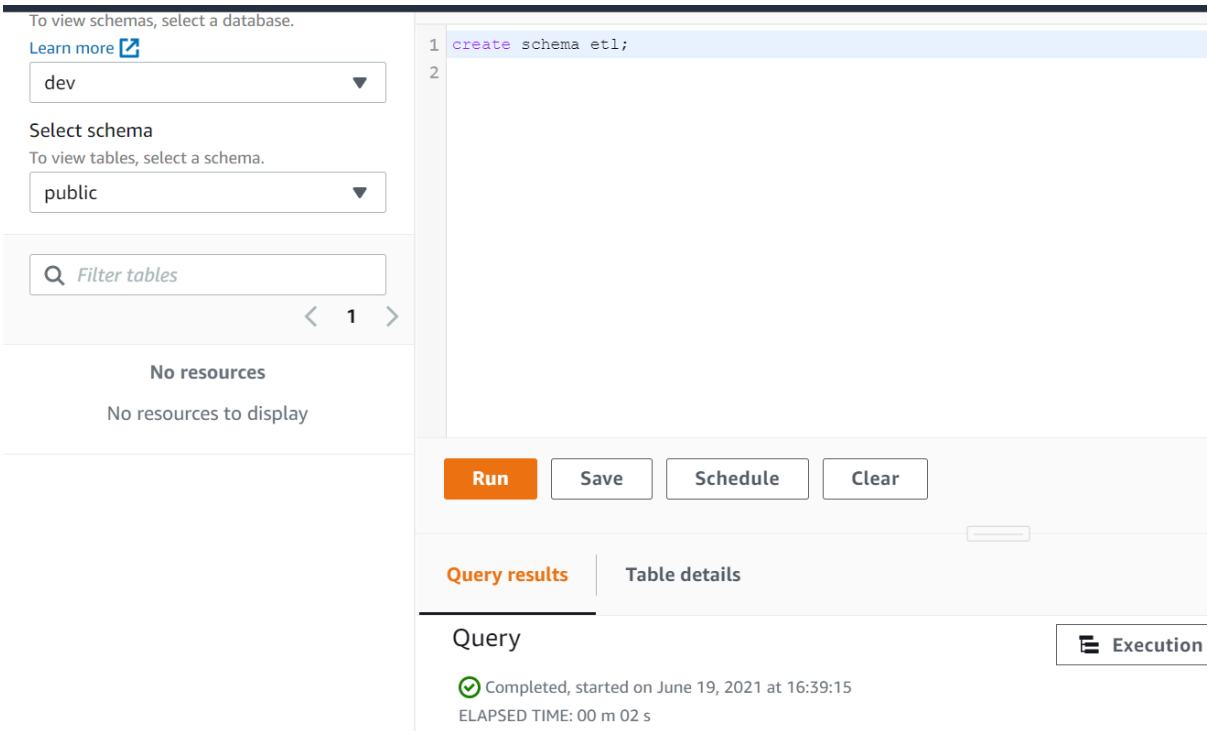




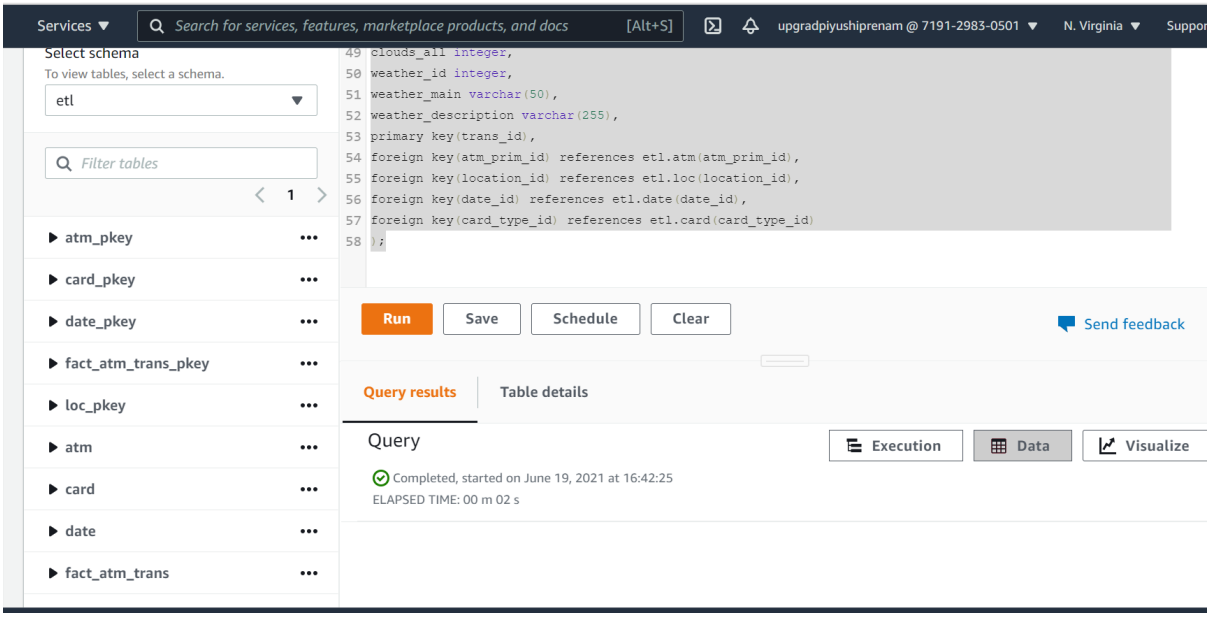
© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

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





© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved




```
create schema etl;

create table etl.loc(
location_id integer not null,
atm_location varchar(50),
atm_streetname varchar(255),
atm_street_number integer,
atm_zipcode integer,
atm_lat NUMERIC(10,3),
atm_lon NUMERIC(10,3),
primary key(location_id));

create table etl.atm(
atm_prim_id integer not null,
atm_id varchar(20),
atm_manufacturer varchar(50),
location_id integer,
primary key(atm_prim_id),
foreign key(location_id) references etl.loc(location_id));

create table etl.date(
year integer,
month varchar(20),
day integer,
hour integer,
weekday varchar(20),
```

© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

```
full_date_time timestamp,
date_id integer,
primary key(date_id));

create table etl.card(
card_type varchar(23),
card_type_id integer,
primary key(card_type_id));

create table etl.FACT_ATM_TRANS(
trans_id BIGINT not null,
atm_prim_id integer,
location_id integer,
date_id integer,
card_type_id integer,
atm_status varchar(20),
currency varchar(10),
service varchar(20),
transaction_amount integer,
message_code varchar(255),
message_text varchar(255),
rain_3h NUMERIC(10,3),
clouds_all integer,
weather_id integer,
weather_main varchar(50),
weather_description varchar(255),
primary key(trans_id),
foreign key(atm_prim_id) references etl.atm(atm_prim_id),
foreign key(location_id) references etl.loc(location_id),
foreign key(date_id) references etl.date(date_id),
foreign key(card_type_id) references etl.card(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

© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved

```
copy etl.atm from
's3://redshiftpiyushi/ETL/DIM_ATM/part-00000-3c54cfc2-4caa-4ae2-a1e7-51b05fae8995-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';

copy etl.loc from
's3://redshiftpiyushi/ETL/DIM_LOC/part-00000-d7f4f92b-ca8f-4101-be12-9e1f16e7d5ae-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';

copy etl.date from
's3://redshiftpiyushi/ETL/DIM_DATE/part-00000-9e8fdccc-8934-4538-b283-939e185b68aa-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
timeformat 'auto'
region 'us-east-1';

copy etl.card from
's3://redshiftpiyushi/ETL/DIM_CARD/part-00000-57508156-8187-4127-8c41-f8f550a90b82-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';

copy etl.FACT_ATM_TRANS from
's3://redshiftpiyushi/ETL/FACT_ATM_TRANS/part-00000-98475c40-758c-455f-ae3f-09300371c62e-c000.csv'
iam_role 'arn:aws:iam::719129830501:role/redshift_s3_fullaccess'
delimiter ',' IGNOREHEADER 1
region 'us-east-1'
TRUNCATECOLUMNS
CSV
```

© Copyright 2020. upGrad Education Pvt. Ltd. All rights reserved