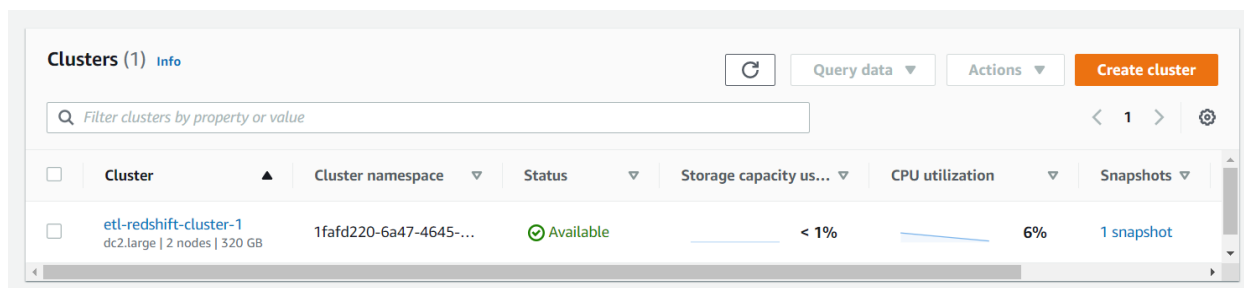


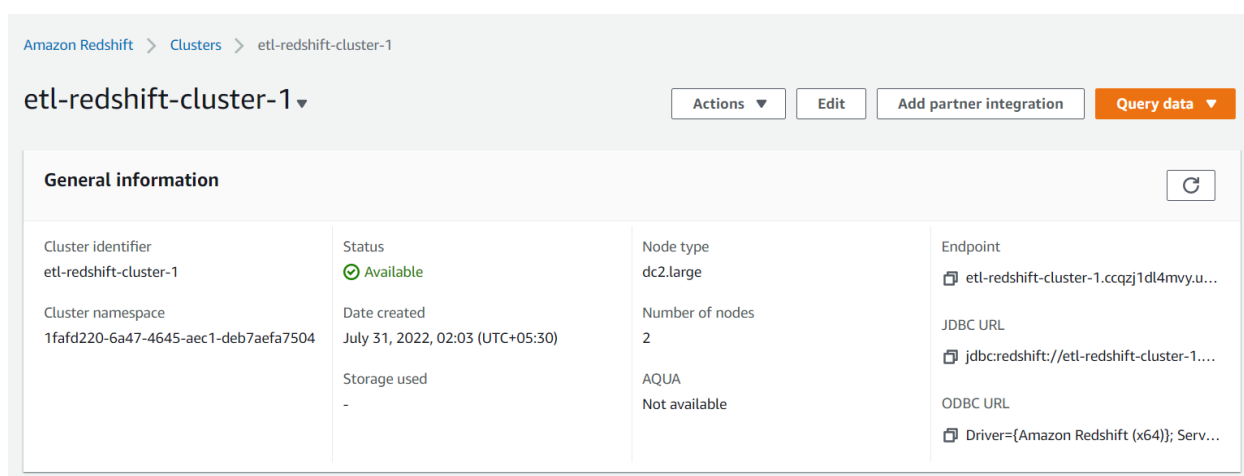
Creation of a Redshift Cluster

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



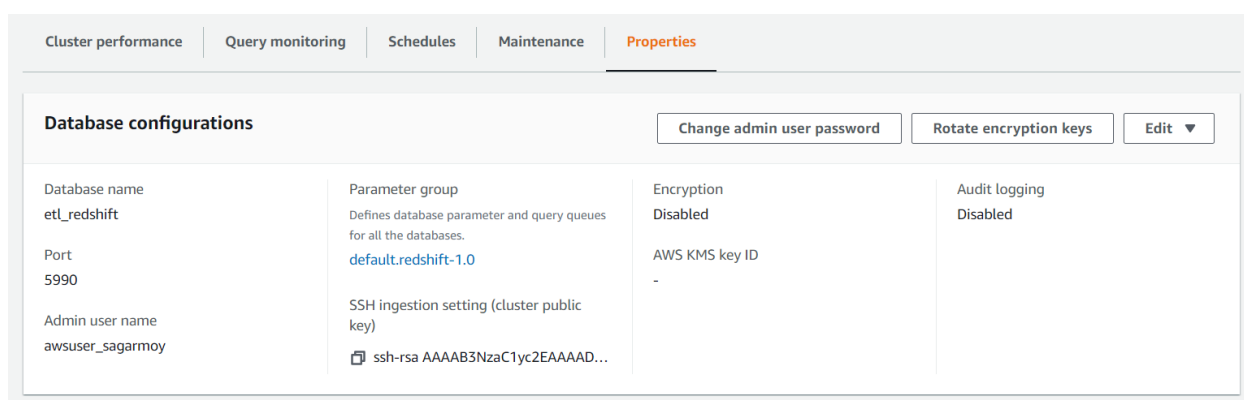
The screenshot shows the 'Clusters (1)' page in the Amazon Redshift console. It includes a search bar, a refresh button, and buttons for 'Query data', 'Actions', and 'Create cluster'. A table lists the cluster details:

Cluster	Cluster namespace	Status	Storage capacity us...	CPU utilization	Snapshots
etl-redshift-cluster-1 dc2.large 2 nodes 320 GB	1fafd220-6a47-4645-...	Available	< 1%	6%	1 snapshot



The screenshot shows the details page for the 'etl-redshift-cluster-1' cluster. It includes buttons for 'Actions', 'Edit', 'Add partner integration', and 'Query data'. The 'General information' section provides the following details:

Cluster identifier	Status	Node type	Endpoint
etl-redshift-cluster-1	Available	dc2.large	etl-redshift-cluster-1.ccqz1dl4mvy.u...
Cluster namespace	Date created	Number of nodes	JDBC URL
1fafd220-6a47-4645-aec1-deb7aefa7504	July 31, 2022, 02:03 (UTC+05:30)	2	jdbc:redshift://etl-redshift-cluster-1....
Storage used		AQUA	ODBC URL
-		Not available	Driver={Amazon Redshift (x64)}; Serv...



The screenshot shows the 'Properties' tab for the 'etl-redshift-cluster-1' cluster. It includes buttons for 'Change admin user password', 'Rotate encryption keys', and 'Edit'. The 'Database configurations' section provides the following details:

Database name	Parameter group	Encryption	Audit logging
etl_redshift	Defines database parameter and query queues for all the databases. default.redshift-1.0	Disabled	Disabled
Port	SSH ingestion setting (cluster public key)	AWS KMS key ID	
5990	ssh-rsa AAAAB3NzaC1yc2EAAAAD...	-	
Admin user name			
awsuser_sagarmoy			

Network and security settings Edit

Virtual private cloud (VPC) vpc-0d43812e47fa275d5	Availability Zone us-east-1d	VPC security group Specify which instances and devices can connect to the cluster. sg-03997c516bd10ba24	Publicly accessible Allow instances and devices outside the VPC to connect to your database through the cluster endpoint. Disabled
Subnet upgrad-redshift-cluster-subnet-group-1	Enhanced VPC routing Disabled		
Endpoint URL -			

Cluster permissions

1 Create an IAM role as the default for this cluster that has the [AmazonRedshiftAllCommandsFullAccess](#) policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift. The policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue.

Associated IAM roles (1) Info
Set default Manage IAM roles

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default for this cluster.

Search for associated IAM role by name, status, or role type

<input type="checkbox"/>	IAM roles	Status	Role type
<input type="checkbox"/>	upGrad_Redshift_S3_role	in-sync	--

Cluster subnet groups (1)

Search

Refresh Delete Actions Create cluster subnet group

<input type="checkbox"/>	Name	Status	VPC ID	Description	Tags
<input type="checkbox"/>	upgrad-redshift-cluster-subr 1 Subnets	Complete	vpc-0d43812e47fa275d5	Redshift cluster Upgrad	

Amazon Redshift > Configurations > Subnet groups > Subnet group

upgrad-redshift-cluster-subnet-group-1 Delete Modify

Cluster subnet group details

VPC ID vpc-0d43812e47fa275d5	Description Redshift cluster Upgrad	Status Complete
Attached clusters etl-redshift-cluster-1		

Subnets (1)

Search

Availability Zone	Subnet ID	CIDR block
us-east-1d	subnet-0561fa21c9fc5c186	10.0.0.0/24

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:

```
create schema etl_atm;
```

```
create table etl_atm.location(  
    location_id integer not null,  
    location varchar(50),  
    streetname varchar(255),  
    street_number integer,  
    zipcode integer,  
    lat numeric(10,3),  
    lon numeric(10,3),  
    primary key(location_id));
```

```
create table etl_atm.atm(  
    atm_id integer not null,  
    atm_number varchar(20),  
    atm_manufacturer varchar(50),  
    atm_location_id integer,  
    primary key(atm_id),  
    foreign key(atm_location_id) references etl_atm.location(location_id));
```

```
create table etl_atm.date(  
    date_id integer not null,  
    full_date_time timestamp,  
    year integer,  
    month varchar(20),  
    day integer,  
    hour integer,  
    weekDay varchar(20),  
    primary key(date_id));
```

```
create table etl_atm.card_type(  
    card_type_id integer not null,  
    card_type varchar(30),  
    primary key(card_type_id));
```

```
create table etl_atm.atm_trans(  
    trans_id BIGINT not null,  
    atm_id integer,  
    weather_loc_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(weather_loc_id) references etl_atm.location(location_id),  
    foreign key(atm_id) references etl_atm.atm(atm_id),  
    foreign key(date_id) references etl_atm.date(date_id),  
    foreign key(card_type_id) references etl_atm.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

```
copy etl_atm.location from  
's3://etl-bucket-sagarmoy/dim_location/part-00000-6353ab28-56be-4286-b8d8-f68e19b40a2a-c  
000.csv'  
iam_role 'arn:aws:iam::113744691613:role/upGrad_Redshift_S3_role'  
delimiter ',' IGNOREHEADER 1  
region 'us-east-1';
```

```
copy etl_atm.atm from  
's3://etl-bucket-sagarmoy/dim_atm/part-00000-43aa1aeb-a56d-42fe-b793-98ee0906d81c-c000.  
csv'  
iam_role 'arn:aws:iam::113744691613:role/upGrad_Redshift_S3_role'  
delimiter ',' IGNOREHEADER 1  
region 'us-east-1';
```

```
copy etl_atm.date from
's3://etl-bucket-sagarmoy/dim_date/part-00000-8742b834-89b9-452d-a5de-e25d43b4582f-c000
.csv'
iam_role 'arn:aws:iam::113744691613:role/upGrad_Redshift_S3_role'
delimiter ',' IGNOREHEADER 1
timeformat 'auto'
region 'us-east-1';
```

```
copy etl_atm.card_type from
's3://etl-bucket-sagarmoy/dim_card_type/part-00000-3061e921-f5b9-443f-9b0a-572028d168c7-
c000.csv'
iam_role 'arn:aws:iam::113744691613:role/upGrad_Redshift_S3_role'
delimiter ',' IGNOREHEADER 1
region 'us-east-1';
```

```
copy etl_atm.atm_trans from
's3://etl-bucket-sagarmoy/fact_atm_trans/part-00000-f3f4a12c-4e3b-42cf-9448-92fb67dfe736-c0
00.csv'
iam_role 'arn:aws:iam::113744691613:role/upGrad_Redshift_S3_role'
delimiter ',' IGNOREHEADER 1
region 'us-east-1'
TRUNCATECOLUMNS
CSV;
```