

## 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>

Amazon Redshift > Clusters > etl-redshift-ujjawal

etl-redshift-ujjawal

Actions Edit Add partner integration Query data

**General information**

Cluster identifier etl-redshift-ujjawal	Status Available	Node type dc2.large	Endpoint etl-redshift-ujjawal.c4emlc0wyaz.us-east-1.re...
Cluster namespace 6be13c3c-5a32-49b1-b7cb-3562f5d564cc	Date created May 23, 2022, 02:16 (UTC+05:30)	Number of nodes 4	JDBC URL jdbc:redshift://etl-redshift-ujjawal.c4emlc0wy...
	Storage used 0.01% (0.10 of 640 GB used)	AQUA Not available	ODBC URL Driver={Amazon Redshift (x64)}; Server=etl-red...

Amazon Redshift > Clusters > Create cluster

Create cluster

**Cluster configuration**

**Cluster identifier**  
This is the unique key that identifies a cluster.  
etl-redshift-ujjawal  
The identifier must be from 1-63 characters. Valid characters are a-z (lowercase only) and - (hyphen).

**What are you planning to use this cluster for?**

☒ **Production**  
Configure for fast and consistent performance at the best price.

☐ **Free trial**  
Configure for learning about Amazon Redshift. This configuration is free for a limited time if your organization has never created an Amazon Redshift cluster.

**Choose the size of the cluster**

I'll choose Help me choose

**Node type**  
Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.  
dc2.large

**Number of nodes**  
Enter the number of nodes that you need.  
2

## Database configurations

### Admin user name

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

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 `@`.

## 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.

### Encryption

Encrypt all data on your cluster.

- ☒ Disabled
- ☐ Use AWS Key Management Service (AWS KMS)
- ☐ Use a hardware security module (HSM)

## ▼ Maintenance

### Maintenance window

Choose the start of the weekly time range when you want pending notifications or maintenance applied to the cluster.

- ☒ Use default maintenance window

### Maintenance track

The maintenance track controls which cluster version is applied during a maintenance window.

- ☒ **Current**  
Use the most current approved cluster version.
- ☐ **Trailing**  
Use the cluster version before the current version.
- ☐ **Preview**  
Use the cluster version with beta releases of new features.

## ▼ Monitoring

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 atm_data;
```

- **Creating location dimension table**

```
create table atm_data.DIM_LOCATION (  
  
location_id int not null DISTKEY SORTKEY, location varchar(50),  
streetname varchar(255),  
street_number int,  
  
zipcode int,  
lat decimal(10,3),  
lon decimal(10,3), PRIMARY KEY(location_id)  
  
);
```

- **Creating atm dimension table**

```
create table 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 atm_data.DIM_LOCATION(location_id) );
```

- **Creating date dimension table**

```
create table 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)  
  
);
```

- **Creating card type dimension table**

```
create table atm_data.DIM_CARD_TYPE (  
  
card_type_id int not null DISTKEY SORTKEY, card_type varchar(30)  
PRIMARY KEY(card_type_id)  
  
);
```

- **Creating atm transactions fact table**

```
create table 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(225),  
message_text varchar(225),  
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 atm_data.DIM_LOCATION(location_id), FOREIGN  
KEY(atm_id) references atm_data.DIM_DATA(atm_id),  
FOREIGN KEY(date_id) references atm_data.DIM_DATE(date_id),  
FOREIGN KEY(card_type_id) references 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

#### • Copying the data to dim\_location table

```
copy atm_data.dim_location from 's3://etlassignmentujjawal/dim_location/part-00000-4f4b02d0-919a-442e-9134-f459cbdb7909-c000.csv'
iam_role 'arn:aws:iam::464886120274:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

CSV;

#### • Copying the data to dim\_atm table

```
copy atm_data.dim_atm from 's3://etlassignmentujjawal/dim_atm/part-00000-c4425605-e626-4cd2-adb2-cef68f7cb1b9-c000.csv'
iam_role 'arn:aws:iam::464886120274:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

CSV;

#### • Copying the data to dim\_date table

```
copy atm_data.dim_date from 's3://etlassignmentujjawal/dim_date/part-00000-7a7ef505-bc12-476c-a0a6-e9e8b544fe44-c000.csv'
iam_role 'arn:aws:iam::464886120274:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

CSV;

#### • Copying the data to dim\_card\_type table

```
copy atm_data.dim_card_type from 's3://etlassignmentujjawal/dim_card_type/part-00000-
```

```
b9c7eb07-29c6-4445-ba0f-98de14834601-c000.csv'
```

```
iam_role 'arn:aws:iam::464886120274:role/redshift_s3_fullaccess' delimiter ',' region 'us-east-1'
CSV;
```

#### • Copying the data to fact\_atm\_trans table

```
copy atm_data.fact_atm_trans from 's3://etlassignmentujjawal/fact_atm_trans/part-00000-978dd709-2ef2-4145-8ab5-9981558a8c60-c000.csv'
iam_role 'arn:aws:iam::464886120274:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

CSV;