

Creation of a Redshift Cluster

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

The screenshot displays the AWS Redshift console interface. The top navigation bar includes the AWS logo, 'Services', a search bar, and user information. The breadcrumb trail shows 'Amazon Redshift > Clusters > spar-nord-data-mart'. The cluster name 'spar-nord-data-mart' is prominently displayed at the top of the main content area, along with buttons for 'Actions', 'Edit', 'Add partner integration', and 'Query data'.

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

Cluster identifier	Status	Node type	Endpoint
spar-nord-data-mart	Available	dc2.large	spar-nord-data-mart.cuwsc7ivpb3.us-east-1.redshift.amazonaws.com:5440/atm-data-mart
Custom domain name - new	Date created	Number of nodes	JDBC URL
-	September 05, 2023, 13:44 (UTC+05:30)	2	jdbc:redshift://spar-nord-data-mart.cuwsc7ivpb3.us-east-1.redshift.amazonaws.com:5440/atm-data-mart
Cluster ARN	Storage used		ODBC URL
arn:aws:redshift:us-east-1:041999195106:namespace:f70ec3d5-7bd5-4158-baba-ea772f8ba0b1	-		Driver={Amazon Redshift (x64)}; Server=spar-nord-data-mart.cuwsc7ivpb3.us-east-1.redshift.amazonaws.com; Database=atm-data-mart
Cluster configuration	Multi-AZ		
Production	No		

The bottom section of the console shows the 'Database configurations' tab, which includes buttons for 'Edit admin credentials', 'Rotate encryption keys', and 'Edit'. The configuration details are as follows:

Database name	Parameter group	Encryption	Audit logging
atm-data-mart	Defines database parameter and query queues for all the databases.	Disabled	Disabled
Port	SSH ingestion setting (cluster public key)	AWS KMS key ID	
5440	default.redshift-1.0	-	
Admin user name			
awsuser	ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQCDYHyy4dWxVqFQ6N5Cd1WVvAWfhfRNRrA+mMUD24fLs2kswE17pR5z7mgVs3+T6Qf3ovL3rMU7XTY11544IM2LTAdFYr96E++8pOkn2x/EmY5CWllGwrRik4A4suDZJ3rr8Cb5pqnr/PcH7MByp14bjsZTtAJyi65UJslmJyh2CxJ5DMmuL yCU75qYsiz+OG+pbPCWYaxQtHcEPf4S3q5PLQqkRzFgOLltcsEh0hCEziurZEaeJfLdBhzWaaAWmwvp4mbbMIRI7ydb4G6uGCdhqIsbNWHtOpY3GlR5k4doMcr9DNHEJRDllj8upQXDDIh8PIRiEetGPJaomOI/dqEZ Amazon-Redshift		

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 atm_data;
```

```
create table if not exists 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)
);
```

```
CREATE table if not exists 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)
);
```

```
CREATE table if not exists 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 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 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_ATM(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

<Queries>

```

copy atm_data.dim_location from 's3://bank-atm-etl/dim_location/part-00000-12ec4699-ef22-48cc-8164-004231ee5404-c000.csv'
iam_role 'arn:aws:iam::041999195106:role/myRedshiftRole'
delimiter ';' region 'us-east-1'

```

CSV;

copy atm_data.dim_atm from 's3://bank-atm-etl/dim_atm/part-00000-b0630db1-ace8-47a3-8e2d-4193e94b59c9-c000.csv'

iam_role 'arn:aws:iam::041999195106:role/myRedshiftRole'

delimiter ',' region 'us-east-1'

CSV;

copy atm_data.dim_date from 's3://bank-atm-etl/dim_date/part-00000-5bdd2bf9-a677-4841-a672-de1b0f762d5d-c000.csv'

iam_role 'arn:aws:iam::041999195106:role/myRedshiftRole'

delimiter ',' region 'us-east-1'

CSV

TIMEFORMAT 'auto';

copy atm_data.dim_card_type from 's3://bank-atm-etl/dim_card_type/part-00000-89f8244d-b5f4-4988-9525-cfba36bc683f-c000.csv'

iam_role 'arn:aws:iam::041999195106:role/myRedshiftRole'

delimiter ',' region 'us-east-1'

CSV;

copy atm_data.FACT_ATM_TRANS from 's3://bank-atm-etl/fact_atm_trans/part-00000-06b945fc-9033-4d84-81cb-9f5376c814ac-c000.csv'

iam_role 'arn:aws:iam::041999195106:role/myRedshiftRole'

delimiter ',' region 'us-east-1'

CSV;