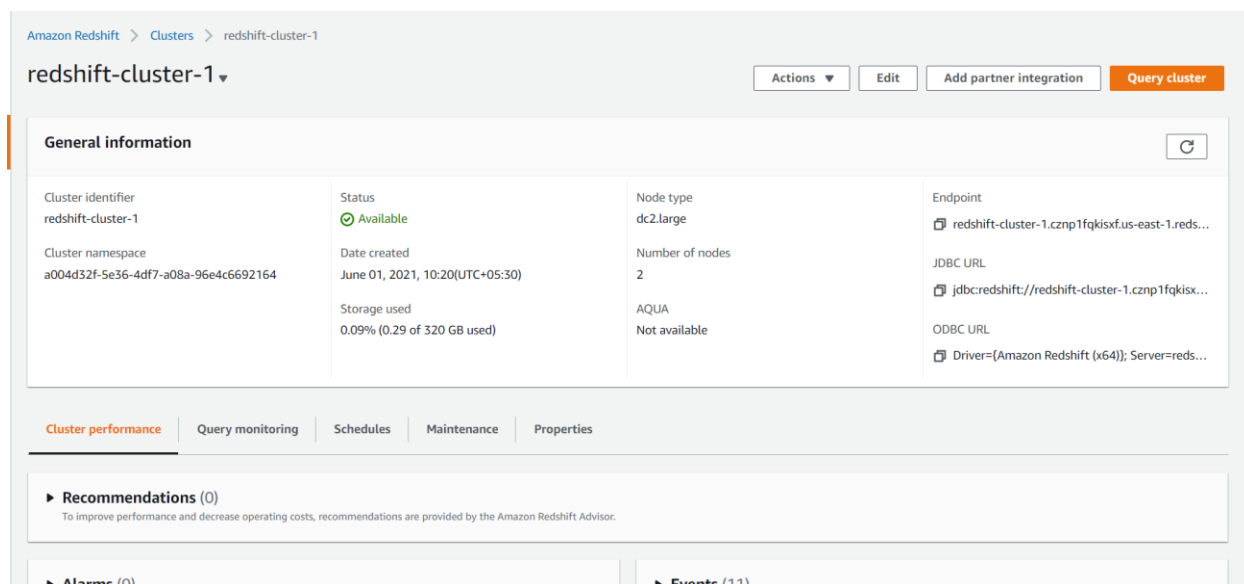


Creation of a RedShift Cluster

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

<Screenshot of the type of machine used along with number of nodes>



The screenshot displays the Amazon RedShift console for a cluster named 'redshift-cluster-1'. The cluster is in an 'Available' status. The configuration details are as follows:

General information			
Cluster identifier redshift-cluster-1	Status Available	Node type dc2.large	Endpoint redshift-cluster-1.czn1f1qkixf.us-east-1.reds...
Cluster namespace a004d32f-5e36-4df7-a08a-96e4c6692164	Date created June 01, 2021, 10:20(UTC+05:30)	Number of nodes 2	JDBC URL jdbc:redshift://redshift-cluster-1.czn1f1qkixf...
	Storage used 0.09% (0.29 of 320 GB used)	AQUA Not available	ODBC URL Driver={Amazon Redshift (x64)}; Server=reds...

Below the general information, there are tabs for 'Cluster performance', 'Query monitoring', 'Schedules', 'Maintenance', and 'Properties'. The 'Cluster performance' tab is currently selected, showing a 'Recommendations (0)' section with a note: 'To improve performance and decrease operating costs, recommendations are provided by the Amazon Redshift Advisor.' At the bottom, there are sections for 'Alarms (0)' and 'Events (11)'.

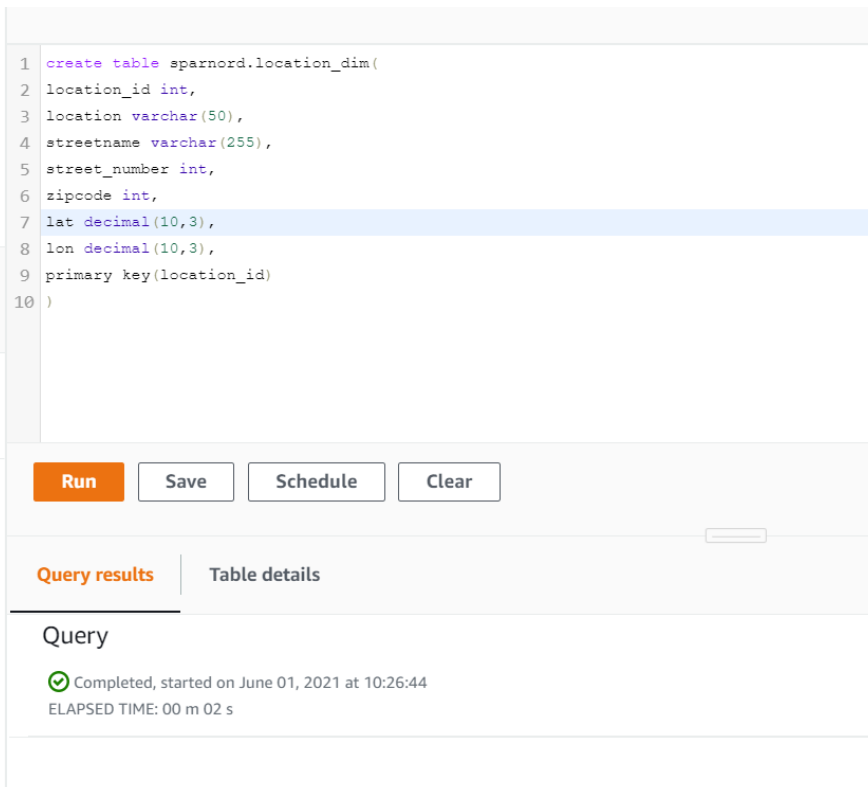
Setting up a database in the RedShift cluster and running queries to create the dimension and fact tables

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

<Queries>

Location dimension table:

```
create table sparnord.location_dim(  
location_id int,  
location varchar(50),  
streetname varchar(255),  
street_number int,  
zipcode int,  
lat decimal(10,3),  
lon decimal(10,3),  
primary key(location_id)  
)
```



The screenshot shows a SQL query editor with a text area containing the following SQL code:

```
1 create table sparnord.location_dim(  
2 location_id int,  
3 location varchar(50),  
4 streetname varchar(255),  
5 street_number int,  
6 zipcode int,  
7 lat decimal(10,3),  
8 lon decimal(10,3),  
9 primary key(location_id)  
10 )
```

Below the text area are four buttons: **Run** (orange), **Save**, **Schedule**, and **Clear**. To the right of the **Run** button is a dropdown menu.

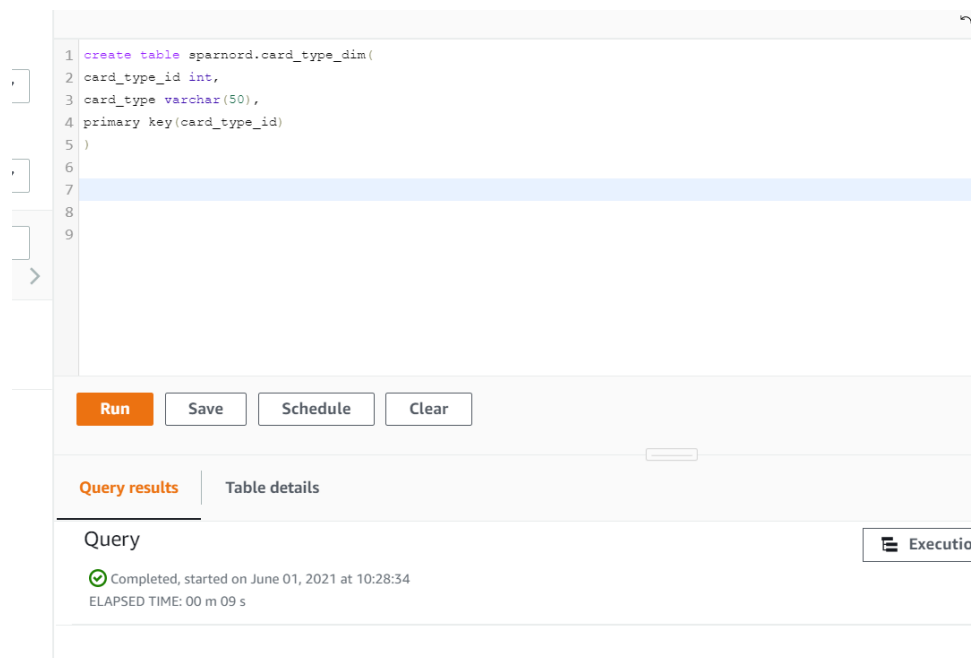
Below the buttons are two tabs: **Query results** (active) and **Table details**. Under the **Query results** tab, the following information is displayed:

Query

✓ Completed, started on June 01, 2021 at 10:26:44
ELAPSED TIME: 00 m 02 s

Card type dimension table:

```
create table sparnord.card_type_dim(  
card_type_id int,  
card_type varchar(50),  
primary key(card_type_id)  
)
```



The screenshot shows a SQL IDE interface. The query editor contains the following SQL code:

```
1 create table sparnord.card_type_dim(  
2 card_type_id int,  
3 card_type varchar(50),  
4 primary key(card_type_id)  
5 )  
6  
7  
8  
9
```

Below the query editor are four buttons: **Run**, **Save**, **Schedule**, and **Clear**. Below these buttons are two tabs: **Query results** and **Table details**. The **Query results** tab is active, showing the following information:

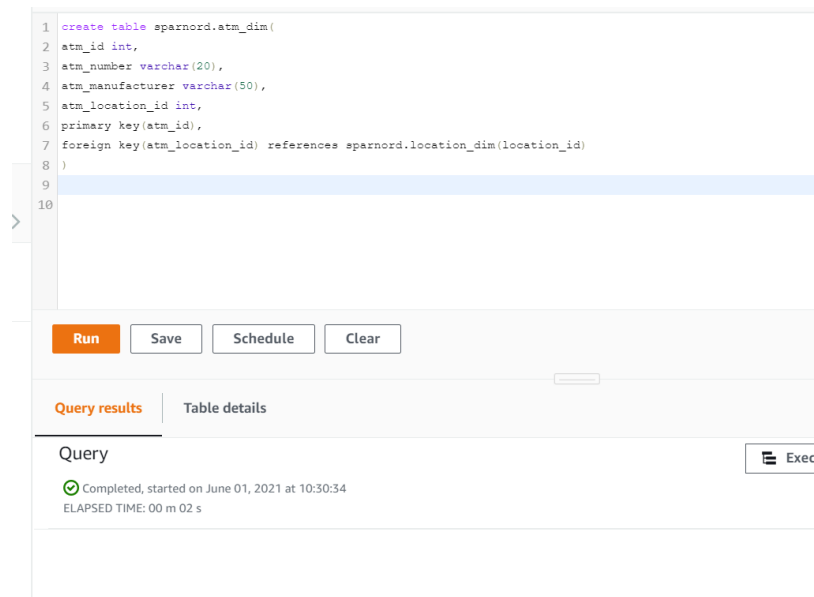
Query

Completed, started on June 01, 2021 at 10:28:34
ELAPSED TIME: 00 m 09 s

On the right side of the **Query results** tab, there is a button labeled **Execution**.

Atm dimension table:

```
create table sparnord.atm_dim(
  atm_id int,
  atm_number varchar(20),
  atm_manufacturer varchar(50),
  atm_location_id int,
  primary key(atm_id),
  foreign key(atm_location_id) references sparnord.location_dim(location_id)
)
```



The screenshot shows a SQL query execution interface. The query is as follows:

```
1 create table sparnord.atm_dim(
2   atm_id int,
3   atm_number varchar(20),
4   atm_manufacturer varchar(50),
5   atm_location_id int,
6   primary key(atm_id),
7   foreign key(atm_location_id) references sparnord.location_dim(location_id)
8 )
9
10
```

Below the query editor, there are buttons for "Run", "Save", "Schedule", and "Clear". The "Run" button is highlighted in orange. Below these buttons, there are tabs for "Query results" and "Table details". The "Query results" tab is active, showing a "Query" section with a green checkmark indicating successful completion. The status message reads: "Completed, started on June 01, 2021 at 10:30:34" and "ELAPSED TIME: 00 m 02 s". There is also an "Exec" button in the top right corner of the results section.

Date dimension table:

```
create table sparnord.date_dim(
  date_id int,
  full_date_time timestamp,
  year int,
  month varchar(20),
  day int,
  hour int,
  weekday varchar(20),
  primary key(date_id)
)
```

```

1 create table sparnord.date_dim(
2   date_id int,
3   full_date_time timestamp,
4   year int,
5   month varchar(20),
6   day int,
7   hour int,
8   weekday varchar(20),
9   primary key(date_id)
10 )
11
12

```

Run Save Schedule Clear

Query results Table details

Query

✓ Completed, started on June 01, 2021 at 10:32:40
ELAPSED TIME: 00 m 02 s

Transaction-fact table:

```

create table sparnord.transaction_fact
(
  trans_id Bigint,
  atm_id int,
  weather_loc_id int,
  date_id int,
  card_type_id int,
  atm_status varchar(20),
  currency varchar(20),
  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 sparnord.location_dim(location_id),
foreign key(atm_id) references sparnord.atm_dim(atm_id),
foreign key(date_id) references sparnord.date_dim(date_id),
foreign key(card_type_id) references sparnord.card_type_dim(card_type_id)
);
```

```
create table sparnord.transaction_fact
(
  trans_id bigint,
  atm_id int,
  weather_loc_id int,
  date_id int,
  card_type_id int,
  atm_status varchar(20),
  currency varchar(20),
  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 sparnord.location_dim(location_id),
  foreign key(atm_id) references sparnord.atm_dim(atm_id),
  foreign key(date_id) references sparnord.date_dim(date_id),
  foreign key(card_type_id) references sparnord.card_type_dim(card_type_id)
);
```

✓ Query 1 ✕
✓ Query 2 ✕
✓ Query 3 ✕
✓ Query 4 ✕
✓ **Query 5** ✕
+

```

13 message_text varchar(255),
14 rain_3h decimal(10,3),
15 clouds_all int,
16 weather_id int,
17 weather_main varchar(50),
18 weather_description varchar(255),
19 primary key(trans_id),
20 foreign key(weather_loc_id) references sparnord.location_dim(location_id),
21 foreign key(atm_id) references sparnord.atm_dim(atm_id),
22 foreign key(date_id) references sparnord.date_dim(date_id),
23 foreign key(card_type_id) references sparnord.card_type_dim(card_type_id)
24 );
25

```

Run Save Schedule Clear

Query results Table details

Query

✓ Completed, started on June 01, 2021 at 10:34:37
 ELAPSED TIME: 00 m 02 s

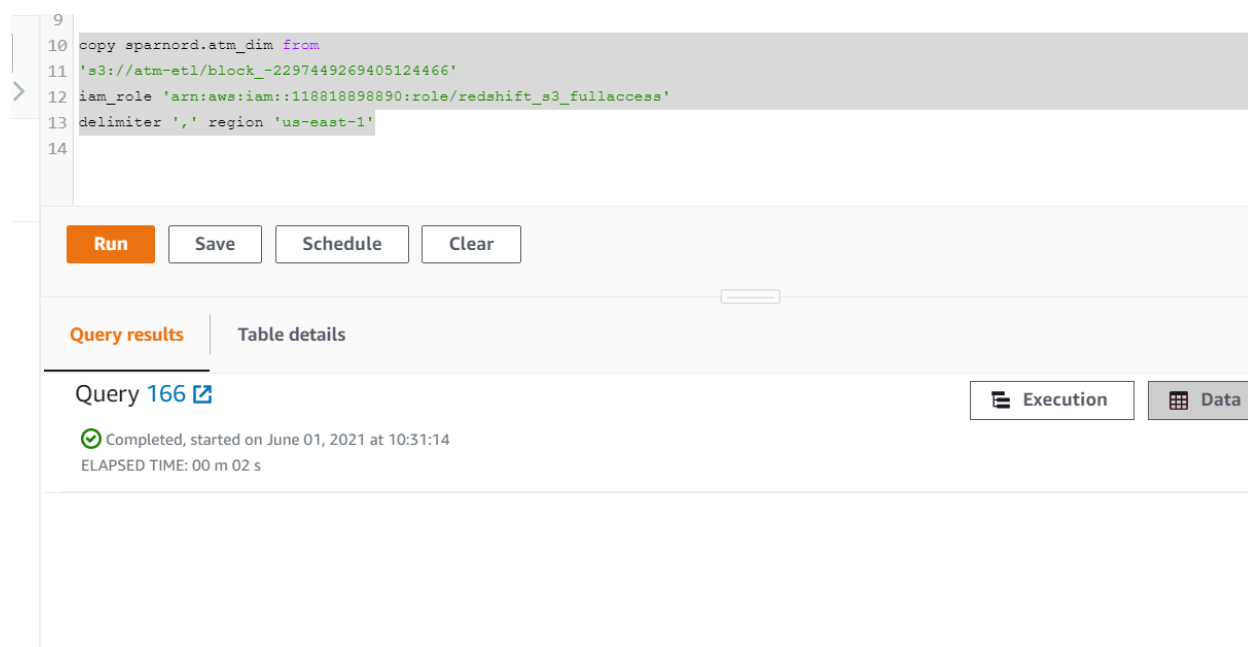
Exec

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

Atm dimension table:

```
copy sparnord.atm_dim from
's3://atm-etl/block_-2297449269405124466'
iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```



The screenshot shows a query editor with the following SQL code:

```
9
10 copy sparnord.atm_dim from
11 's3://atm-etl/block_-2297449269405124466'
12 iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
13 delimiter ',' region 'us-east-1'
14
```

Below the editor are buttons for **Run**, **Save**, **Schedule**, and **Clear**.

The interface shows the **Query results** tab selected. The query is identified as **Query 166**. The status is **Completed**, started on **June 01, 2021 at 10:31:14**, with an **ELAPSED TIME: 00 m 02 s**. There are tabs for **Execution** and **Data**.

Location dimension table:

```
copy sparnord.location_dim from
's3://location-etl/block_154573813621028266'
iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

```

11
12 copy sparnord.location_dim from
13 's3://location-etl/block_154573813621028266'
14 iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
15 delimiter ',' region 'us-east-1'

```

Run Save Schedule Clear [Send feedback](#)

Query results Table details

Query [114](#)

Completed, started on June 01, 2021 at 10:27:34
ELAPSED TIME: 00 m 08 s

Execution Data Visualize

Date dimension table:

```

copy sparnord.date_dim from
's3://date-etl/block_-4833838643168932492'
iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
timeformat as 'auto'

```

```

12
13 copy sparnord.date_dim from
14 's3://date-etl/block_-4833838643168932492'
15 iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
16 delimiter ',' region 'us-east-1'
17 timeformat as 'auto'
18

```

Run Save Schedule Clear [Send feedback](#)

Query results Table details

Query [192](#)

Completed, started on June 01, 2021 at 10:33:40
ELAPSED TIME: 00 m 08 s

Execution Data Visualize

Card type dimension table:

```

copy sparnord.card_type_dim from
's3://cardtype-etl/block_269436742132970922'
iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'

```


Transaction fact table:

```
copy sparnord.transaction_fact from
's3://transaction-fact-etl/transaction-fact'
iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
```

```
50
51 copy sparnord.transaction_fact from
52 's3://transaction-fact-etl/transaction-fact'
53 iam_role 'arn:aws:iam::118818898890:role/redshift_s3_fullaccess'
54 delimiter ',' region 'us-east-1'
```

Run

Save

Schedule

Clear

Send feedback

Query results

Table details

Query 231

Execution

Data

Visualize

Completed, started on June 01, 2021 at 10:36:27

ELAPSED TIME: 00 m 13 s