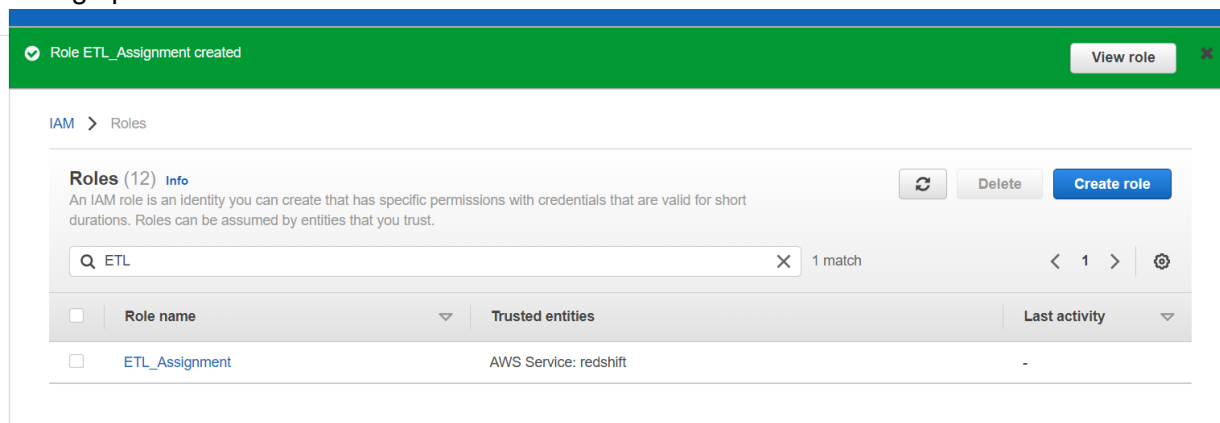


Creation of a Redshift Cluster

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

Setting up a new Full Access IAM role



Role ETL_Assignment created

IAM > Roles

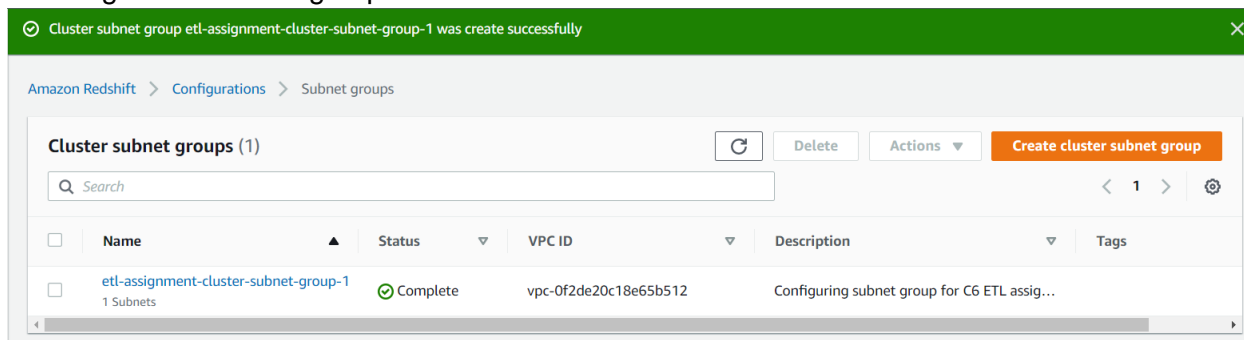
Roles (12) Info

An IAM role is an identity you can create that has specific permissions with credentials that are valid for short durations. Roles can be assumed by entities that you trust.

Search: ETL 1 match

<input type="checkbox"/>	Role name	Trusted entities	Last activity
<input type="checkbox"/>	ETL_Assignment	AWS Service: redshift	-

Creating a new subnet group



Cluster subnet group etl-assignment-cluster-subnet-group-1 was create successfully

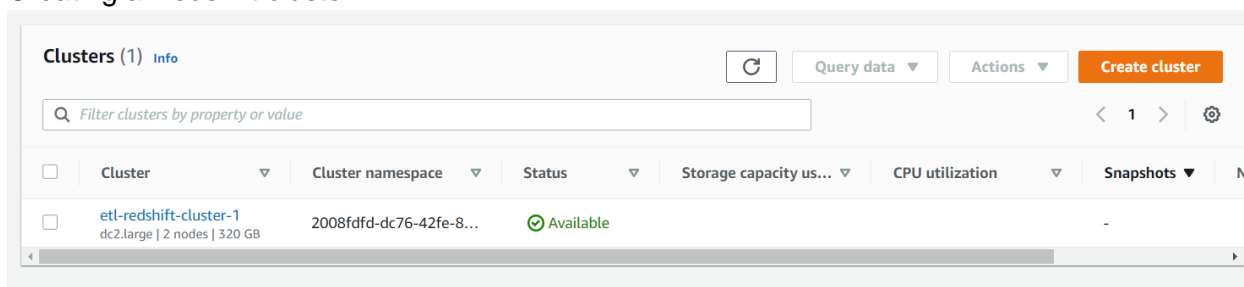
Amazon Redshift > Configurations > Subnet groups

Cluster subnet groups (1)

Search: Search

<input type="checkbox"/>	Name	Status	VPC ID	Description	Tags
<input type="checkbox"/>	etl-assignment-cluster-subnet-group-1 1 Subnets	Complete	vpc-0f2de20c18e65b512	Configuring subnet group for C6 ETL assig...	

Creating a Redshift cluster



Clusters (1) Info

Filter clusters by property or value

<input type="checkbox"/>	Cluster	Cluster namespace	Status	Storage capacity us...	CPU utilization	Snapshots
<input type="checkbox"/>	etl-redshift-cluster-1 dc2.large 2 nodes 320 GB	2008fdfd-dc76-42fe-8...	Available			-

Added the IAM role

Associated IAM roles (1) Info				Set default ▼		Manage IAM roles ▼	
<div><div>Q</div><div>Search for associated IAM role by name, status, or role type</div></div>				< 1 >			
<input type="checkbox"/>	<div>IAM roles ↗</div>	▼	Status	▼	Role type		▼
<input type="checkbox"/>	ETL_Assignment		<div><div>✔</div><div>in-sync</div></div>		--		

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;
```

```
create table etl.dim_location
```

```
(  
    location_id int primary key,  
    location varchar(50),  
    streetname varchar(255),  
    street_number int,  
    zipcode int,  
    lat numeric(10,3),  
    lon numeric(10,3)  
);
```

```
create table etl.dim_atm
```

```
(  
    atm_id int primary key,  
    atm_number varchar(20),  
    atm_manufacturer varchar(50),  
    atm_location_id int references etl.dim_location(location_id)  
);
```

```
create table etl.dim_date
```

```
(  
    date_id int primary key,  
    full_date_time timestamp,  
    year int,  
    month varchar(20),  
    day int,  
    hour int,  
    weekday varchar(20)  
);
```

```
create table etl.dim_card
```

```
(  
    card_type_id int primary key,  
    card_type varchar(30)  
);
```

```
create table etl.fact_atm_trans
(
    trans_id bigint primary key,
    atm_id int references etl.dim_atm(atm_id),
    weather_loc_id int references etl.dim_location(location_id),
    date_id int references etl.dim_date(date_id),
    card_type_id int references etl.dim_card(card_type_id),
    atm_status varchar(20),
    currency varchar(10),
    service varchar(20),
    transaction_amount int,
    message_code varchar(255),
    message_text varchar(255),
    rain_3h numeric(10,3),
    clouds_all int,
    weather_id int,
    weather_main varchar(50),
    weather_description varchar(255)
);
```

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 Location Dimension data and verify number of records
*/
copy etl.dim_location
from 's3://etlanalytics/dimensions/DIM_LOCATION/part-00000-4c755f65-41e7-4d10-b1ea-
260da3880c6e-c000.csv'
iam_role 'arn:aws:iam::752744469700:role/ETL_Assignment'
delimiter ','
ignoreheader 1
region 'us-east-1'
csv;

select count(*) from etl.dim_location;

/*
* Copy ATM Dimension data and verify number of records
*/
```

```
copy etl.dim_atm
from 's3://etlanalytics/dimensions/DIM_ATM/part-00000-c5e4c21e-e82a-404e-be1e-
f53981a309f6-c000.csv'
iam_role 'arn:aws:iam::752744469700:role/ETL_Assignment'
delimiter ','
ignoreheader 1
region 'us-east-1'
csv;
```

```
select count(*) from etl.dim_atm;
```

```
/*
* Copy Date Dimension data and verify number of records
*/
```

```
copy etl.dim_date
from 's3://etlanalytics/dimensions/DIM_DATE/part-00000-0f099929-54a2-43fe-b785-
0b099e821178-c000.csv'
iam_role 'arn:aws:iam::752744469700:role/ETL_Assignment'
delimiter ','
timeformat 'YYYY-MM-DDTHH:MI:SS'
ignoreheader 1
region 'us-east-1'
csv;
```

```
select count(*) from etl.dim_date;
```

```
/*
* Copy Card Type Dimension data and verify number of records
*/
```

```
copy etl.dim_card
from 's3://etlanalytics/dimensions/DIM_CARD_TYPE/part-00000-bccd4fa3-840a-4d8a-b2ce-
07e7a340c80a-c000.csv'
iam_role 'arn:aws:iam::752744469700:role/ETL_Assignment'
delimiter ','
ignoreheader 1
region 'us-east-1'
csv;
```

```
select count(*) from etl.dim_card;
```

```
/*
* Copy ATM Transaction Fact data and verify number of records
*/
```

```
copy etl.fact_atm_trans
```

```
from 's3://etlanalytics/facts/FACT_ATM_TRANS/part-00000-77277577-44f7-4218-98d1-ccc47458d6cf-c000.csv'
iam_role 'arn:aws:iam::752744469700:role/ETL_Assignment'
delimiter ','
ignoreheader 1
truncatecolumns
region 'us-east-1'
csv;
```

```
select count(*) from etl.fact_atm_trans;
```

```
/*
```

```
* Query to view any query execution errors
```

```
*/
```

```
select * from stl_load_errors;
```