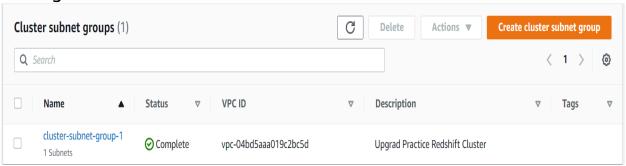
Redshift Setup

■ Creating a SUBNET GROUP



Creating Cluster

Cluster configuration

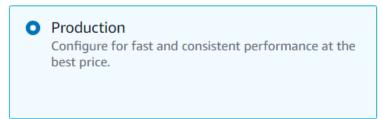
Cluster identifier

This is the unique key that identifies a cluster.

redshift-cluster-1

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?



Free trial

Configure for learning about Amazon Redshift configuration is free for a limited time if your organization has never created an Amazon Red cluster.

Choose the size of the cluster

I'll choose	Help me choose
-------------	----------------

Choosing Node type – dc2.large

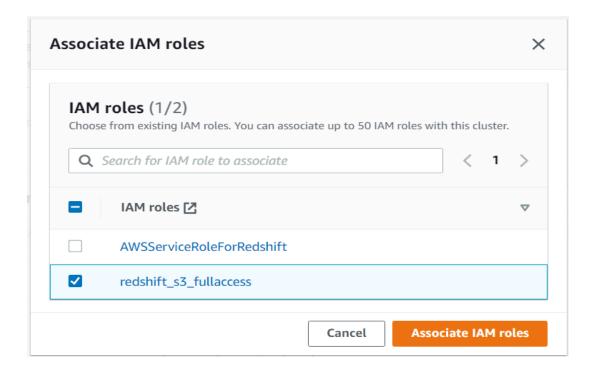
I'll choose	Help me choose	
Node type Info		
Choose a node type that me	eets your CPU, RAM, storag	ge capacity, and drive type requireme
dc2.large		▼
Number of nodes		
Enter the number of nodes	that you need.	
2		
Range (1-32)		

Database configuration – username and password

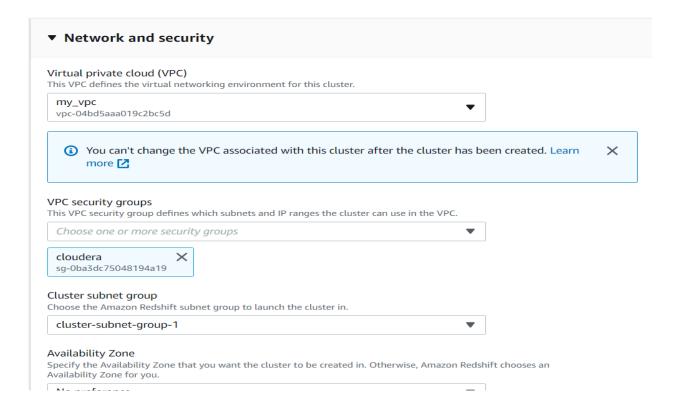
Database configurations	
Admin user name Enter a login ID for the admin user of your DB instance.	
awsuser	
The name must be 1-128 alphanumeric characters, and it can't be a reserved word Auto generate password Amazon Redshift can generate a password for you, or you can specify your own password.	
Admin user password	
•••••	
Show password	
Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and printable ASCII character except "/", """, or "@".	d one number. Can be a

Attaching 'IAM' Role(redshift_s3_fullaccess)

For s3 access to cluster.



Network and Security Configuration



Database configurations – under additional configurations and here, set a database name. The default port is 5439, which is known globally. So, the best practice to change the port number and set it to any one between 1150 and 65535. After this you can click on the Create Cluster button

▼ Database configurations	
Database name Specify a database name to create an additional database.	
upgard	
The name must be 1-64 alphanumeric characters (lowercase only), and it can't be a reserved word .	
Database port Port number where the database accepts inbound connections. You can't change the port after the cluster has been created	d.
5440	
The port must be numeric (1150-65535).	
Parameter groups Defines database parameter and query queues for all the databases.	
default.redshift-1.0 Default parameter group for redshift-1.0	
Encryption Encrypt all data on your cluster.	
Disabled	
○ Use AWS Key Management Service (AWS KMS)	
 Use a hardware security module (HSM) 	

Schema and Tables Creation on RedShift

create schema ATM;

- Creating table Location Dimension table

```
create table ATM.DIM_LOCATION(
location_id VARCHAR(20) distkey sortkey primary key,
location VARCHAR(50),
streetname VARCHAR(255),
street_number INT,
zipcode INT,
lat DECIMAL(10,3),
lon DECIMAL(10,3)
);
```

Loading data from S3 bucket

copy ATM.DIM_LOCATION from

's3://etl-upgrad/DIM_LOCATION/'

iam_role 'arn:aws:iam::730892717471:role/redshift_s3_fullaccess'

delimiter ',' region 'us-east-1' IGNOREHEADER 1 removequotes;

Q Search rows					<	1 > @
location_id ▽	location \triangledown	streetname ∇	street_number ▽	zipcode ▽	lat ▽	lon ▽
location10	Aars	Himmerlandsgade	70	9600	56.803	9.518
location100	Taars	Bredgade	91	9830	57.385	10.116
location102	Vadum	Ellehammersvej	43	9430	57.118	9.861
location106	Vinderup	$S ilde{A}f\hat{A}$, ndergade	5	7830	56.481	8.779
location11	Abildgaard	Hj $\tilde{A}f\hat{A}$, rringvej	144	9900	57.447	10.506
location16	Brugsen i Breum	Aakj $ ilde{A}^f\hat{A}^!$ rsvej	1	7870	56.688	9.069
location22	Durup	Torvet	4	7870	56.745	8.949
location26	Frederiksberg	Gammel Kongevej	157	1850	55.677	12.537
location3	Aalborg Storcenter Afd	Hobrovej	452	9200	57.005	9.876
location30	Glyng $ ilde{A}f\hat{A}$, re	FÃ <i>f</i> ¦rgevej	1	7870	56.762	8.867

- Creating Atm dimension table

```
create table ATM.DIM_ATM(
atm_id INT distkey sortkey primary key,
atm_number varchar(20),
atm_manufacturer VARCHAR(50),
atm_location_id VARCHAR(50),
foreign key(atm_location_id) references ATM.DIM_LOCATION(location_id)
);
```

Loading data from S3 bucket

copy ATM.DIM_ATM from

's3://etl-upgrad/DIM_ATM/'

iam_role 'arn:aws:iam::730892717471:role/redshift_s3_fullaccess'

delimiter ',' region 'us-east-1' IGNOREHEADER 1 removequotes;

dim_atm atm.dim_atm			3 Show schema
Q Search ro	ws		
atm_id	▼ atm_number	atm_manufacturer	
5	5	NCR	location69
15	15	NCR	location104
22	22	NCR	location79
23	23	Diebold Nixdorf	location107
24	24	NCR	location40
27	27	NCR	location34
28	28	NCR	location66
33	33	NCR	location102
35	35	NCR	location1
39	39	NCR	location95 Ac

- Creating Card type Dimension table

```
create table ATM.DIM_CARD_TYPE(
card_type_id VARCHAR(20) distkey sortkey primary key,
card_type VARCHAR(30)
);
```

Loading data from S3 bucket

copy ATM.DIM_CARD_TYPE from

's3://etl-upgrad/DIM_CARD_TYPE/'

iam_role 'arn:aws:iam::730892717471:role/redshift_s3_fullaccess'

delimiter', region 'us-east-1' IGNOREHEADER 1 removequotes;

dim_card_type

atm.dim_card_type

Q Search rows	
card_type_id	▽ card_type
card1	CIRRUS
card10	Visa Dankort
card6	Maestro
card7	MasterCard
card11	Visa Dankort - on-us
card4	$H ilde{A} f ilde{A}'_{I} vekort$
card5	$ extstyle{H} ilde{A}f\hat{A}^{I}_{I}vekort$ - on-us
card8	Mastercard - on-us
card9	VISA
card12	VisaPlus

- Creating Date dimension table

create table ATM.DIM_DATE(
date_id VARCHAR(20) distkey sortkey primary key,
full_date_time TIMESTAMP,
year INT,
month VARCHAR(20),
day INT,
hour INT,
weekday VARCHAR(20)
);

Loading data from S3 Bucket

copy ATM.DIM_DATE from

's3://etl-upgrad/DIM_DATE/'

iam_role 'arn:aws:iam::730892717471:role/redshift_s3_fullaccess'

timeformat 'YYYY-MM-DDTHH:MI:SS'

delimiter', region 'us-east-1' IGNOREHEADER 1 removequotes;

date_id	\triangledown	full_date_time	▽	year	▽	month	▽	day	▽	hour	▽	weekday
date1009		2017-08-13 05:00:00		2017		August		13		5		Sunday
date1010		2017-08-13 06:00:00		2017		August		13		6		Sunday
date1013		2017-08-13 09:00:00		2017		August		13		9		Sunday
date1018		2017-08-13 14:00:00		2017		August		13		14		Sunday
date1021		2017-08-13 17:00:00		2017		August		13		17		Sunday
date1028		2017-08-14 00:00:00		2017		August		14		0		Monday
date1029		2017-08-14 01:00:00		2017		August		14		1		Monday
date1031		2017-08-14 03:00:00		2017		August		14		3		Monday
date104		2017-04-05 08:00:00		2017		April		5		8		Wednesday
date1045		2017-08-14 17:00:00		2017		August		14		17		Monday

- Creating Transaction Fact table

```
create table ATM.FACT_ATM_TRANS(
trans_id VARCHAR(20) distkey sortkey primary key,
atm_id INT,
weather_loc_id VARCHAR(20),
date_id VARCHAR(20),
card_type_id VARCHAR(20),
atm_status VARCHAR(20),
currency VARCHAR(10),
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),
foreign key(weather_loc_id) references ATM.DIM_LOCATION(location_id),
foreign key(atm_id) references ATM.DIM_ATM(atm_id),
foreign key(date_id) references ATM.DIM_DATE(date_id),
foreign key(card_type_id) references ATM.DIM_CARD_TYPE(card_type_id)
);

    Loading data from S3 Bucket

copy ATM.FACT_ATM_TRANS from
's3://etl-upgrad/FACT_ATM_TRANS/'
iam_role 'arn:aws:iam::730892717471:role/redshift_s3_fullaccess'
delimiter', region 'us-east-1' IGNOREHEADER 1 removeguotes;
```

trans_id ▽	atm_id ▽	weather_loc_i d ▽	date_id ▽	card_type_i d ▽	atm_status ▽	currency ▽	service
transaction100	35	location1		card10	Active	DKK	Withdrav
transaction10000	35	location1	date2447	card11	Active	DKK	Withdrav
transaction100000 4	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100000 5	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100000 6	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100001	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100001	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100001 5	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100001	60	location37	date4361	card11	Active	DKK	Withdrav
transaction100001	60	location37	date4386	card11	Active	Activate Win	