

Creation of a Redshift Cluster

Screenshots of the configuration of the redshift cluster that have been created

[Amazon Redshift](#) > [Clusters](#) > Create cluster

Create cluster

Cluster configuration

Cluster identifier

This is the unique key that identifies a cluster.

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?


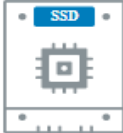
☒ **Production**
Configure for fast and consistent performance at the best price.

☐ **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

Node type

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

Recommended	
RA3 High performance with scalable managed storage	DC2 High performance with fixed local SSD storage
<input type="radio"/> ra3.4xlarge Managed storage: up to 64 TB/node \$3.26/node/hour \$0.024/GB/month	<input checked="" type="radio"/> dc2.large Storage: 160 GB/node \$0.25/node/hour
<input type="radio"/> ra3.16xlarge Managed storage: up to 64 TB/node \$13.04/node/hour \$0.024/GB/month	<input type="radio"/> dc2.8xlarge Storage: 2.6 TB/node \$4.80/node/hour
 <div>ra3.4xlarge 12 vCPU (gen 3)</div>	 <div>dc2.large 2 vCPU (gen 2)</div>

► [Show legacy dense storage node types](#)

Nodes

Enter the number of nodes that you need.

Range (1-32)

Database configurations

Database name (optional)

Specify a database name to create an additional database.

The name must be 1-64 alphanumeric characters (lowercase only), and it can't be a **reserved word**.

Database port (optional)

Port number where the database accepts inbound connections. You can't change the port after the cluster has been created.

The port must be numeric (1150-65535).

Master user name

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


The name must be 1-128 alphanumeric characters, and it can't be a **reserved word**.

Master user password

☐ Show password

- The master password must be 8 - 64 characters.
- The value must contain at least one uppercase letter.
- The value must contain at least one lowercase letter.
- The value must contain at least one number.
- The master password can only contain ASCII characters (ASCII codes 33-126), except ' (single quotation mark), " (double quotation mark), /, \, or @.

▼ Cluster permissions (optional)

Your cluster needs permissions to access other AWS services on your behalf. For the required permissions, add IAM roles with the principal "redshift.amazonaws.com". You can associate up to 10 IAM roles with this cluster. [Learn more](#) 

Available IAM roles



Attached IAM roles

Status

[etlprojectrole](#) 

[arn:aws:iam::757859277641:role/etlprojectrole](#)




Not applied

▼ Network and security

Virtual private cloud (VPC)

This VPC defines the virtual networking environment for this cluster. Choose a VPC that has a subnet group. Only valid VPCs are enabled in the list.

my_vpc
vpc-0a02c0ed680be85d6

 You can't change the VPC associated with this cluster after the cluster has been created. [Learn more](#)  

VPC security groups

This VPC security group defines which subnets and IP ranges the cluster can use in the VPC.

Choose one or more security groups

cloudera
sg-0f4fd44d360cf8286

Cluster subnet group

Choose the Amazon Redshift subnet group to launch the cluster in.

redshift-testcluster

Availability Zone

Specify the Availability Zone that you want the cluster to be created in. Otherwise, Amazon Redshift chooses an Availability Zone for you.

No preference

Enhanced VPC routing

Forces cluster traffic through a VPC.

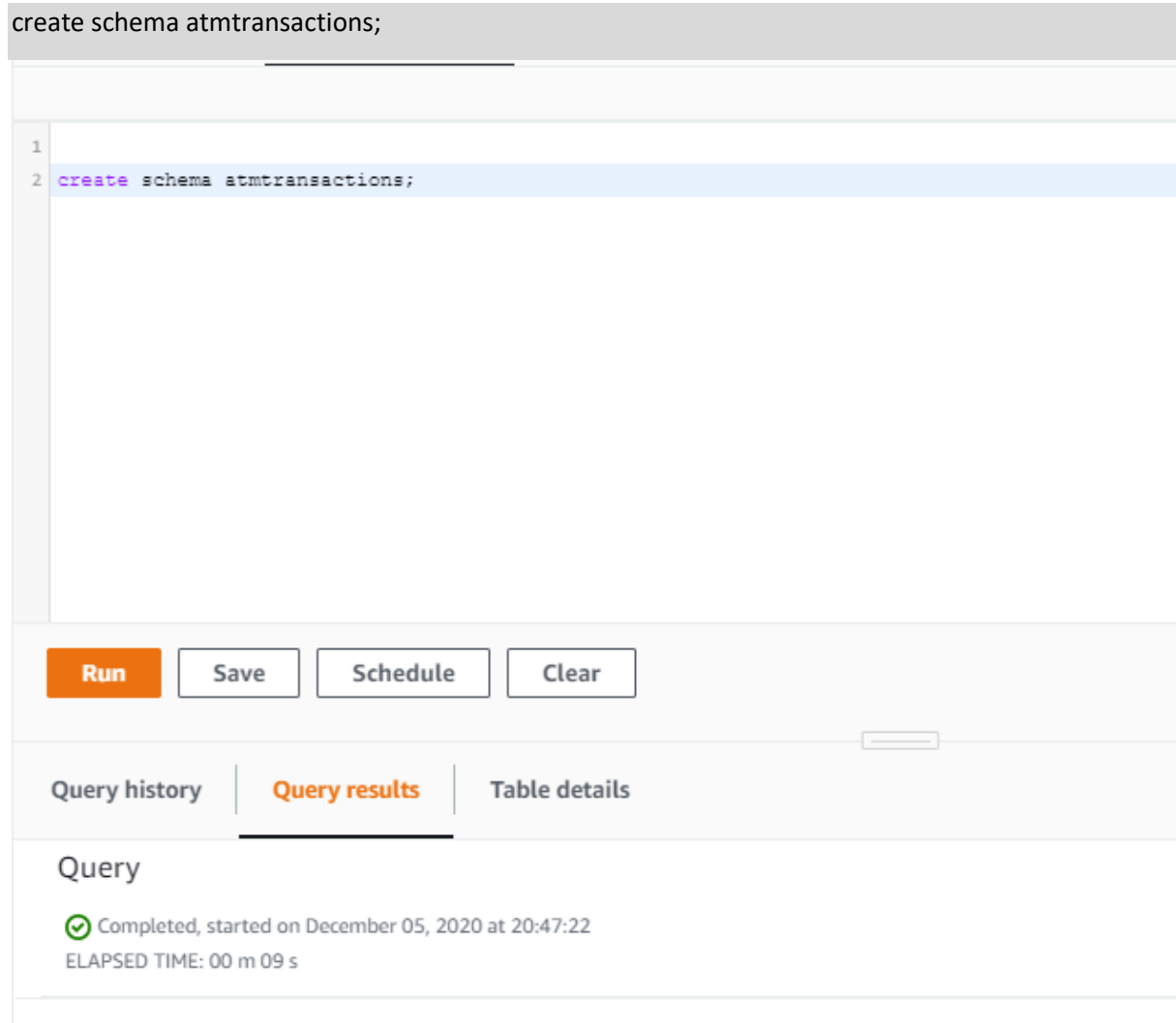
☒ Disabled

☐ Enabled

Queries used for creating the Dimension and Fact tables on the Redshift cluster along with screenshots of the successful status of the query

Schema creation:

```
create schema atmtransactions;
```



1

2 `create schema atmtransactions;`

Run Save Schedule Clear

Query history **Query results** Table details

Query

✓ Completed, started on December 05, 2020 at 20:47:22
ELAPSED TIME: 00 m 09 s

DIM_LOCATION

```
create table atmtransactions.dim_location(  
location_id INT not null,  
location VARCHAR(50),  
streetname VARCHAR(255),  
street_number INT,  
zipcode INT,  
lat DECIMAL(10,3),  
lon DECIMAL(10,3),  
PRIMARY KEY (location_id));
```

Select a schema to view data tables.

atmtransactions ▼

Q Filter tables

< 1 >

- dim_location ...
- dim_location_pkey ...

```

1 create table atmtransactions.dim_location(
2 location_id INT not null,
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 );
11

```

Run Save Schedule Clear

Query history Query results Table details

Query

✓ Completed, started on December 05, 2020 at 20:48:20
ELAPSED TIME: 00 m 10 s

DIM_ATM

```

create table atmtransactions.dim_atm(
atm_id INT not null,
atm_number VARCHAR(20),
atm_manufacturer VARCHAR(50),
atm_location_id INT,
PRIMARY KEY (atm_id),
FOREIGN KEY(atm_location_id) REFERENCES atmtransactions.dim_location(location_id)
);

```

Select a schema to view data tables.

atmtransactions ▼

Q Filter tables

< 1 >

- dim_atm ...
- dim_atm_pkey ...
- dim_location ...
- dim_location_pkey ...

```

1 create table atmtransactions.dim_atm(
2 atm_id INT not null,
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 atmtransactions.dim_location(location_id)
8 );
9
10

```

Run Save Schedule Clear

Query history Query results Table details

Query

✓ Completed, started on December 05, 2020 at 20:49:02
ELAPSED TIME: 00 m 09 s

DIM_DATE

```
create table atmtransactions.dim_date(  
  
date_id INT NOT NULL,  
  
full_date_time TIMESTAMP,  
  
year INT,  
  
month VARCHAR(20),  
  
day INT,  
  
hour INT,  
  
weekday VARCHAR(20),  
  
PRIMARY KEY (date_id)  
  
);
```

The screenshot shows a database management interface. On the left, a sidebar displays a list of tables under the 'atmtransactions' schema: dim_atm, dim_atm_pkey, dim_date, dim_date_pkey, dim_location, and dim_location_pkey. The 'dim_date' table is selected. The main area shows the SQL query for creating the 'dim_date' table, which is the same as the one in the first block. Below the query editor, there are buttons for 'Run', 'Save', 'Schedule', and 'Clear'. The 'Run' button is highlighted. Below these buttons, there are tabs for 'Query history', 'Query results', and 'Table details'. The 'Query results' tab is active, showing a 'Query' section with a green checkmark indicating the query was completed successfully on December 05, 2020 at 20:49:36, with an elapsed time of 00 m 09 s.

DIM_CARD_TYPE

```
create table atmtransactions.dim_card_type(  
  
card_type_id INT NOT NULL,  
  
card_type VARCHAR(20),  
  
PRIMARY KEY (card_type_id)  
  
);
```

Select a schema to view data tables.

atmtransactions ▼

Q Filter tables

< 1 >

- ▶ dim_atm ...
- ▶ dim_atm_pkey ...
- ▶ dim_card_type ...
- ▶ dim_card_type_pkey ...
- ▶ dim_date ...
- ▶ dim_date_pkey ...
- ▶ dim_location ...
- ▶ dim_location_pkey ...

```

1 create table atmtransactions.dim_card_type(
2   card_type_id INT NOT NULL,
3   card_type VARCHAR(20),
4   PRIMARY KEY (card_type_id)
5 );
6
7
8

```

Run Save Schedule Clear

Query history Query results Table details

Query

✓ Completed, started on December 05, 2020 at 20:50:06
ELAPSED TIME: 00 m 09 s

FACT TABLE

```

create table atmtransactions.fact_atm_trans(
trans_id BIGINT NOT NULL,
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(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 atmtransactions.dim_location (location_id),
FOREIGN KEY (atm_id) REFERENCES atmtransactions.dim_atm (atm_id),
FOREIGN KEY (date_id) REFERENCES atmtransactions.dim_date (date_id),
FOREIGN KEY (card_type_id) REFERENCES atmtransactions.dim_card_type (card_type_id)
);

```

Select a schema to view data tables.

atmtransactions

Q Filter tables

< 1 >

- dim_atm ...
- dim_atm_pkey ...
- dim_card_type ...
- dim_card_type_pkey ...
- dim_date ...
- dim_date_pkey ...
- dim_location ...
- dim_location_pkey ...
- fact_atm_trans ...
- fact_atm_trans_pkey ...

```

1 create table atmtransactions.fact_atm_trans (
2 trans_id BIGINT NOT NULL,
3 atm_id INT,
4 weather_loc_id INT,
5 date_id INT,
6 card_type_id INT,
7 atm_status VARCHAR(20),
8 currency VARCHAR(10),
9 service VARCHAR(20),
10 transaction_amount INT,
11 message_code VARCHAR(255),
12 message_text VARCHAR(255),
13 rain_3h DECIMAL(10,3),
14 clouds_all INT,
15 weather_id INT,

```

Run Save Schedule Clear

Query history Query results Table details

Query

✓ Completed, started on December 05, 2020 at 20:50:44
ELAPSED TIME: 00 m 13 s

Queries used for loading the data into the Dimension and Fact tables in the Redshift cluster from the S3 bucket along with screenshots of the successful status of the query

```

copy atmtransactions.dim_location from
's3://etlproject-testbucket/DIM_LOCATION/part-00000-6c532654-3696-4b9a-9aa1-a95d2174c6fe-c000.csv'
iam_role 'arn:aws:iam::757859277641:role/etlprojectrole'
delimiter ',' region 'us-east-1' CSV
NULL AS 'NULL'
timeformat 'auto';

```


Query history

Query results

Table details

Query 230

Completed, started on December 05, 2020 at 20:51:43

ELAPSED TIME: 00 m 14 s

Select count(1) from atmtransactions.dim_location;

Query 250

Completed, started on December 05, 2020 at 20:52:30

ELAPSED TIME: 00 m 10 s

Rows returned (1)

Search rows

count
109

```
copy atmtransactions.dim_atm from
's3://etlproject-testbucket/DIM_ATM/part-00000-00954a5c-df97-476c-a3f3-5a71cd2e1ad6-
c000.csv'
iam_role 'arn:aws:iam::757859277641:role/etlprojectrole'
delimiter ',' region 'us-east-1' CSV
NULL AS 'NULL'
timeformat 'auto';
```

Query history

Query results

Table details

Query 267

Completed, started on December 05, 2020 at 20:53:19

ELAPSED TIME: 00 m 14 s

```
Select count(1) from atmtransactions.dim_atm;
```

Query [287](#)

✓ Completed, started on December 05, 2020 at 20:53:49
ELAPSED TIME: 00 m 09 s

Rows returned (1)

Search rows

count
156

```
copy atmtransactions.dim_card_type from
```

```
's3://etlproject-testbucket/DIM_CARD_TYPE/part-00000-a4684899-eaf0-474b-b714-e0e35e10e8c8-c000.csv'
```

```
iam_role 'arn:aws:iam::757859277641:role/etlprojectrole'
```

```
delimiter ',' region 'us-east-1' CSV;
```

Query history

Query results

Table details

Query [322](#)

✓ Completed, started on December 05, 2020 at 20:56:15
ELAPSED TIME: 00 m 09 s

```
Select count(1) from atmtransactions.dim_card_type;
```

Query 338 [🔗](#)

✅ Completed, started on December 05, 2020 at 20:56:57
ELAPSED TIME: 00 m 09 s

Rows returned (1)

🔍 *Search rows*

count

12

```
copy atmtransactions.dim_date from
's3://etlproject-testbucket/DIM_DATE/part-00000-03ede4ed-670a-4992-a3d9-4b64da2bb187-
c000.csv'
iam_role 'arn:aws:iam::757859277641:role/etlprojectrole'
delimiter ',' region 'us-east-1' CSV
NULL AS 'NULL'
timeformat 'auto';
```

Query history

Query results

Table details

Query 357 [🔗](#)

✅ Completed, started on December 05, 2020 at 20:57:59
ELAPSED TIME: 00 m 14 s

```
Select count(1) from atmtransactions.dim_date;
```

Query history	Query results	Table details
Query 371 🔗 <div> ✓ Completed, started on December 05, 2020 at 20:58:25 ELAPSED TIME: 00 m 08 s </div>		
Rows returned (1) <div> <input type="text" value="Search rows"/> </div>		
count		
8685		

```
copy atmtransactions.fact_atm_trans from
's3://etlproject-testbucket/FACT_ATM_TRANS/part-00000-11f70dc6-3acb-4a5e-8bb1-
06ac6fb509a9-c000.csv'
iam_role 'arn:aws:iam::757859277641:role/etlprojectrole'
delimiter ',' region 'us-east-1' CSV
NULL AS 'NULL'
timeformat 'auto';
```

Query history	Query results	Table details
Query 388 🔗 <div> ✓ Completed, started on December 05, 2020 at 20:59:33 ELAPSED TIME: 00 m 28 s </div>		

```
Select count(1) from atmtransactions.fact_atm_trans;
```

Query history	Query results	Table details
Query 438 🔗 <div> ✓ Completed, started on December 05, 2020 at 21:00:56 ELAPSED TIME: 00 m 08 s </div>		
Rows returned (1) <div> <input type="text" value="Search rows"/> </div>		
count		
2468572		