

# Creation of Redshift 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?



#### 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

I'll choose

Help me choose

### Node type [Info](#)

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

dc2.large



### Number of nodes

Enter the number of nodes that you need.

2

Range (1-32)

Cluster permissions

Create an IAM role as the default for this cluster that has the [AmazonRedshiftAllCommandsFullAccess](#) policy attached. This policy includes permissions to run SQL commands to COPY, UNLOAD, and query data with Amazon Redshift. The policy also grants permissions to run SELECT statements for related services, such as Amazon S3, Amazon CloudWatch logs, Amazon SageMaker, and AWS Glue.

Manage IAM roles

Create, associate, or remove an IAM role. You can associate up to 50 IAM roles. You can also choose an IAM role and set it as the default for this cluster.

Associated IAM roles (1) [Info](#)

Set default ▼

Manage IAM roles ▼

Q

Search for associated IAM role by name, status, or role type

< 1 >

<input type="checkbox"/>	<div>IAM roles <a href="#">🔗</a></div>	▼	Status ▼	Role type ▼
<input type="checkbox"/>	<a href="#">redshift_s3_fullaccess</a>		Not applied	--

## Additional configurations ☒ Use defaults




These configurations are optional, and default settings have been defined to help you get started with your cluster. Turn off "Use defaults" to modify these settings now.

### ▼ Network and security

#### Virtual private cloud (VPC)

This VPC defines the virtual networking environment for this cluster.

suru\_vpc  
vpc-02c62d1d9d830b888

 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

default  
sg-02a166a3c095d4037

cloudera  
sg-0a2edd306e88d625c

#### Cluster subnet group

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


cluster-subnet-group-1

#### 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

Enabling this option forces network traffic between your cluster and data repositories through a VPC, instead of the internet. [Learn more](#) 

☒ Disabled

☐ Enabled

## ▼ Database configurations

### Database name

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

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).

### Parameter groups

Defines database parameter and query queues for all the databases.

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)


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

```
create schema atm_new_trans;

drop table if exists atm_new_trans.dim_date;
create table atm_new_trans.dim_date
(
    date_id integer not null,
    year integer,
    month varchar(20),
    day integer,
    weekday varchar(20),
    hour integer,
    full_date timestamp,
    constraint PK_dim_date primary key (date_id)
);
```

```
2
3 drop table if exists atm_new_trans.dim_date;
4 create table atm_new_trans.dim_date
5 (
6     date_id integer not null,
7     year integer,
8     month varchar(20),
9     day integer,
10    weekday varchar(20),
11    hour integer,
12    full_date timestamp,
13    constraint PK_dim_date primary key (date_id)
14 );
15
```

**Run** Save Schedule Clear

 Send feedback

Query results

Table details

Query

 Execution

 Data

 Visualize

✔ Completed, started on January 17, 2022 at 01:27:11  
ELAPSED TIME: 00 m 22 s

```
drop table if exists atm_new_trans.dim_location;
create table atm_new_trans.dim_location
(
    location_id integer not null,
    location varchar(50),
    streetname varchar(255),
    street_number integer,
    zipcode integer,
    lat decimal(10, 3),
    lon decimal(10, 3),
    constraint PK_dim_location primary key (location_id)
);
```


```
16 drop table if exists atm_new_trans.dim_location;
17 create table atm_new_trans.dim_location
18 (
19     location_id integer not null,
20     location varchar(50),
21     streetname varchar(255),
22     street_number integer,
23     zipcode integer,
24     lat decimal(10, 3),
25     lon decimal(10, 3),
26     constraint PK_dim_location primary key (location_id)
27 );
28
29 drop table if exists atm_new_trans.dim_atm;
```

Run

Save

Schedule

Clear


 Send feedback


Query results


Table details

Query

 Execution

 Data

 Visualize

 Completed, started on January 17, 2022 at 01:28:32  
ELAPSED TIME: 00 m 05 s

```
drop table if exists atm_new_trans.dim_atm;
create table atm_new_trans.dim_atm
(
    atm_id integer not null,
    atm_number varchar(20),
    atm_manufacturer varchar(255),
    location_id integer,
    constraint PK_dim_atm primary key (atm_id),
    constraint FK_dim_atm foreign key (location_id) references
atm_new_trans.dim_location(location_id)
);
```


```
28
29 drop table if exists atm_new_trans.dim_atm;
30 create table atm_new_trans.dim_atm
31 (
32     atm_id integer not null,
33     atm_number varchar(20),
34     atm_manufacturer varchar(255),
35     location_id integer,
36     constraint PK_dim_atm primary key (atm_id),
37     constraint FK_dim_atm foreign key (location_id) references atm_new_trans.dim_location(location_id)
38 );
39
40
```

Run

Save

Schedule


Clear

 Send feedback


Query results

Table details

Query

 Execution

 Data

 Visualize

✓ Completed, started on January 17, 2022 at 01:29:16

ELAPSED TIME: 00 m 04 s



```
drop table if exists atm_new_trans.dim_card_type;
create table atm_new_trans.dim_card_type
(
    card_type_id integer not null,
    card_type varchar(50),
    constraint PK_dim_card_type primary key (card_type_id)
);
```

```
40
41 drop table if exists atm_new_trans.dim_card_type;
42 create table atm_new_trans.dim_card_type
43 (
44     card_type_id integer not null,
45     card_type varchar(50),
46     constraint PK_dim_card_type primary key (card_type_id)
47 );
48
49
```

Run

Save

Schedule


Clear


 Send feedback


Query results

Table details

Query

 Execution

 Data

 Visualize

✓ Completed, started on January 17, 2022 at 01:29:58  
ELAPSED TIME: 00 m 04 s

```

drop table if exists atm_new_trans.transaction_fact cascade;
create table atm_new_trans.transaction_fact
(
    trans_id bigint not null,
    date_id integer,
    atm_status varchar(20),
    atm_id integer,
    currency varchar(10),
    card_type_id integer,
    transaction_amount integer,
    service varchar(50),
    message_code varchar(255),
    message_text varchar(255),
    weather_loc_id integer,
    rain_3h decimal(10, 3),
    clouds_all integer,
    weather_id integer,
    weather_main varchar(255),
    weather_description varchar(255),
    constraint PK_transaction_fact primary key (trans_id),
    constraint FK_atm foreign key (atm_id) references
atm_new_trans.dim_atm(atm_id),
    constraint FK_card_type foreign key (card_type_id)
references atm_new_trans.dim_card_type(card_type_id),
    constraint FK_location foreign key (weather_loc_id)
references atm_new_trans.dim_location(location_id),
    constraint FK_date foreign key (date_id) references
atm_new_trans.dim_date(date_id)
);

```

```
49
50 drop table if exists atm_new_trans.transaction_fact cascade;
51 create table atm_new_trans.transaction_fact
52 (
53     trans_id bigint not null,
54     date_id integer,
55     atm_status varchar(20),
56     atm_id integer,
57     currency varchar(10),
58     card_type_id integer,
59     transaction_amount integer,
60     service varchar(50),
61     message code varchar(255),
```

[Run](#)[Save](#)[Schedule](#)[Clear](#)[Send feedback](#)[Query results](#)[Table details](#)

## Query

[Execution](#)[Data](#)[Visualize](#)

✓ Completed, started on January 17, 2022 at 01:30:37  
ELAPSED TIME: 00 m 09 s

## Loading data into a RedShift cluster from Amazon S3 bucket

```
copy atm_new_trans.dim_atm
from 's3://atmtransdata/atmtrans_v2/dim_atm.csv'
iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```

The screenshot displays the Amazon Redshift console interface. At the top, a SQL query is entered in a text area with line numbers 75 through 83. The query is a COPY statement that loads data from an S3 bucket into a Redshift table. Below the query area, there are four buttons: 'Run' (orange), 'Save', 'Schedule', and 'Clear'. To the right of these buttons is a 'Send feedback' link. Below the buttons, there are two tabs: 'Query results' (active) and 'Table details'. Under the 'Query results' tab, the query is identified as 'Query 855'. To the right of the query name are three buttons: 'Execution', 'Data', and 'Visualize'. Below the query name, a green checkmark icon indicates the query is 'Completed', with a timestamp 'started on January 17, 2022 at 01:31:22' and an 'ELAPSED TIME: 00 m 09 s'.

```
75 copy atm_new_trans.dim_atm
76 from 's3://atmtransdata/atmtrans_v2/dim_atm.csv'
77 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
78 delimiter ',' region 'us-east-1'
79 CSV;
80
81 copy atm_new_trans.dim_location
82 from 's3://atmtransdata/atmtrans_v2/dim_location.csv'
83 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
```

**Run** **Save** **Schedule** **Clear** [Send feedback](#)

**Query results** **Table details**

Query **855** [↗](#) **Execution** **Data** **Visualize**

✓ Completed, started on January 17, 2022 at 01:31:22  
ELAPSED TIME: 00 m 09 s

```
copy atm_new_trans.dim_location
from 's3://atmtransdata/atmtrans_v2/dim_location.csv'
iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```


```
80
81 copy atm_new_trans.dim_location
82 from 's3://atmtransdata/atmtrans_v2/dim_location.csv'
83 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
84 delimiter ',' region 'us-east-1'
85 CSV;
86
87 copy atm_new_trans.dim_card_type
88 from 's3://atmtransdata/atmtrans_v2/dim_card_type.csv'
```

Run

Save

Schedule

Clear


 Send feedback


Query results

Table details

Query 864 [🔗](#)

 Execution

 Data

 Visualize

✓ Completed, started on January 17, 2022 at 01:31:54  
ELAPSED TIME: 00 m 06 s

```
copy atm_new_trans.dim_card_type
from 's3://atmtransdata/atmtrans_v2/dim_card_type.csv'
iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```

```
86
87 copy atm_new_trans.dim_card_type
88 from 's3://atmtransdata/atmtrans_v2/dim_card_type.csv'
89 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
90 delimiter ',' region 'us-east-1'
91 CSV;
92
93 copy atm_new_trans.dim_date
94 from 's3://atmtransdata/atmtrans_v2/dim_date.csv'
95 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
```

Run

Save

Schedule

Clear

 Send feedback

Query results

Table details

Query **870** [🔗](#)

 Execution

 Data

 Visualize

✔ Completed, started on January 17, 2022 at 01:32:21  
ELAPSED TIME: 00 m 02 s

```
copy atm_new_trans.dim_date
from 's3://atmtransdata/atmtrans_v2/dim_date.csv'
iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
timeformat 'YYYY-MM-DDTHH:MI:SS'
CSV;
```

```
93 copy atm_new_trans.dim_date
94 from 's3://atmtransdata/atmtrans_v2/dim_date.csv'
95 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
96 delimiter ',' region 'us-east-1'
97 timeformat 'YYYY-MM-DDTHH:MI:SS'
98 CSV;
99
100 copy atm_new_trans.transaction_fact
101 from 's3://atmtransdata/atmtrans_v2/transaction_fact.csv'
102 iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
```

[Run](#)[Save](#)[Schedule](#)[Clear](#)[Send feedback](#)[Query results](#)[Table details](#)Query **879** [🔗](#)[Execution](#)[Data](#)[Visualize](#)

✓ Completed, started on January 17, 2022 at 01:33:11  
ELAPSED TIME: 00 m 22 s

```
copy atm_new_trans.transaction_fact
from 's3://atmtransdata/atmtrans_v2/transaction_fact.csv'
iam_role 'arn:aws:iam::823888918855:role/redshift_s3_fullaccess'
delimiter ',' region 'us-east-1'
CSV;
```

99

100 copy atm\_new\_trans.transaction\_fact  
101 from 's3://atmtransdata/atmtrans\_v2/transaction\_fact.csv'  
102 iam\_role 'arn:aws:iam::823888918855:role/redshift\_s3\_fullaccess'  
103 delimiter ',' region 'us-east-1'  
104 CSV;


▼

Run

Save

Schedule


Clear


 Send feedback


Query results


Table details

Query 895 [🔗](#)

 Execution

 Data

 Visualize

 Completed, started on January 17, 2022 at 01:33:51  
ELAPSED TIME: 00 m 19 s