

Creation of a Redshift Cluster

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

Create VPC																			
Configure the Security Group																			
Create IAM Rule	<p>etl-project-s3-access Delete</p> <p>Allows Redshift clusters to call AWS services on your behalf etl-project-s3-access</p> <p>Summary Edit</p> <table> <tbody> <tr> <td>Creation date</td><td>September 25, 2022, 12:02 (UTC+05:30)</td><td>ARN</td><td>arn:aws:iam::914450217410:role/etl-project-s3-access</td></tr> <tr> <td>Last activity</td><td>9 hours ago</td><td>Maximum session duration</td><td>1 hour</td></tr> </tbody> </table> <p>Permissions Trust relationships Tags Access Advisor Revoke sessions</p> <p>Permissions policies (1) You can attach up to 10 managed policies.</p> <table> <thead> <tr> <th>Policy name</th> <th>Type</th> <th>Description</th> </tr> </thead> <tbody> <tr> <td>AmazonS3FullAccess</td> <td>AWS managed</td> <td>Provides full access to all bucket</td> </tr> </tbody> </table>	Creation date	September 25, 2022, 12:02 (UTC+05:30)	ARN	arn:aws:iam::914450217410:role/etl-project-s3-access	Last activity	9 hours ago	Maximum session duration	1 hour	Policy name	Type	Description	AmazonS3FullAccess	AWS managed	Provides full access to all bucket				
Creation date	September 25, 2022, 12:02 (UTC+05:30)	ARN	arn:aws:iam::914450217410:role/etl-project-s3-access																
Last activity	9 hours ago	Maximum session duration	1 hour																
Policy name	Type	Description																	
AmazonS3FullAccess	AWS managed	Provides full access to all bucket																	
Create Subnet Group	<p>etlproject Delete Modify</p> <p>Cluster subnet group details</p> <table> <tbody> <tr> <td>VPC ID</td><td>vpc-03484ef7db4bc2875</td><td>Description</td><td>etlproject</td><td>Status</td><td>Complete</td></tr> <tr> <td>Attached clusters</td><td colspan="5">This subnet group has no attached clusters</td></tr> </tbody> </table> <p>Subnets (1)</p> <table> <thead> <tr> <th>Availability Zone</th> <th>Subnet ID</th> <th>CIDR block</th> </tr> </thead> <tbody> <tr> <td>us-east-1a</td><td>subnet-0fab519a02ef4dd98</td><td>172.31.16.0/20</td></tr> </tbody> </table>	VPC ID	vpc-03484ef7db4bc2875	Description	etlproject	Status	Complete	Attached clusters	This subnet group has no attached clusters					Availability Zone	Subnet ID	CIDR block	us-east-1a	subnet-0fab519a02ef4dd98	172.31.16.0/20
VPC ID	vpc-03484ef7db4bc2875	Description	etlproject	Status	Complete														
Attached clusters	This subnet group has no attached clusters																		
Availability Zone	Subnet ID	CIDR block																	
us-east-1a	subnet-0fab519a02ef4dd98	172.31.16.0/20																	

Create Redshift Cluster	<h3>Cluster configuration</h3> <p>Cluster identifier This is the unique key that identifies a cluster.</p> <input type="text" value="etlproject"/> <p>The identifier must be from 1-63 characters. Valid characters are a-z (lowercase only) and - (hyphen).</p> <p>What are you planning to use this cluster for?</p> <div style="display: flex; justify-content: space-between;"> <div style="width: 45%;"> <p><input checked="" type="radio"/> Production Configure for fast and consistent performance at the best price.</p> </div> <div style="width: 45%;"> <p><input type="radio"/> 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.</p> </div> </div> <p>Choose the size of the cluster</p> <div style="display: flex; justify-content: space-around;"> I'll choose Help me choose </div> <p>Node type <small>Info</small> Choose a node type that meets your CPU, RAM, storage capacity, and drive type requirements.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;">dc2.large</div> <p>Number of nodes Enter the number of nodes that you need.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;">2</div> <p>Range (1-32)</p>
Assign the Database username password	<h3>Database configurations</h3> <p>Admin user name Enter a login ID for the admin user of your DB instance.</p> <input type="text" value="awsuser"/> <p>The name must be 1-128 alphanumeric characters, and it can't be a reserved word.</p> <p><input type="checkbox"/> Auto generate password Amazon Redshift can generate a password for you, or you can specify your own password.</p> <p>Admin user password <input type="password" value="*****"/></p> <p><input type="checkbox"/> Show password Must be 8-64 characters long. Must contain at least one uppercase letter, one lowercase letter and one number. Can be any printable ASCII character except "/", "", or "@".</p>

The screenshot shows a modal window titled "Associate IAM roles". At the top, it says "IAM roles (1/2)" and "Choose from existing IAM roles. You can associate up to 50 IAM roles with this cluster." Below is a search bar labeled "Search for IAM role to associate". A table lists IAM roles: "AWSServiceRoleForRedshift" (unchecked) and "etl-project-s3-access" (checked). At the bottom are "Cancel" and "Associate IAM roles" buttons.

Create The cluster:

Open the query editor for the rest of the program.

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

```
CREATE SCHEMA atm_transaction;  
CREATE SCHEMA atm_transaction;
```

Run

Save

Schedule

Clear



Query results

Table details

Query

Execution

Data

 Completed, started on September 26, 2022 at 09:34:02

ELAPSED TIME: 00 m 03 s

```
CREATE TABLE atm_transaction.DIM_LOCATION (
    location_id INT NOT NULL
    ,location VARCHAR(50) NULL
    ,streetname VARCHAR(255) NULL
    ,street_number INT NULL
    ,zipcode INT NULL
    ,lat DECIMAL(10, 3) NULL
    ,lon DECIMAL(10, 3) NULL
    ,PRIMARY KEY (location_id)
);
```

```
1 CREATE TABLE atm_transaction.DIM_LOCATION (
2     location_id INT NOT NULL
3     ,location VARCHAR(50) NULL
4     ,streetname VARCHAR(255) NULL
5     ,street_number INT NULL
6     ,zipcode INT NULL
7     ,lat DECIMAL(10, 3) NULL
8     ,lon DECIMAL(10, 3) NULL
9     ,PRIMARY KEY (location_id)
10 );
11
```

Run

Save

Schedule

Clear



Query results

Table details

Query

Execution

Data

✓ Completed, started on September 26, 2022 at 09:34:50

ELAPSED TIME: 00 m 13 s

```
CREATE TABLE atm_transaction.DIM_ATM (
    atm_id INT NOT NULL
    ,atm_number VARCHAR(20) NULL
    ,atm_manufacturer VARCHAR(50) NULL
    ,atm_location_id INT NULL
    ,PRIMARY KEY (atm_id)
    ,CONSTRAINT atm_location_id FOREIGN KEY (atm_location_id) REFERENCES
atm_transaction.dim_location(location_id)
);
```

```
1 CREATE TABLE atm_transaction.DIM_ATM (
2     atm_id INT NOT NULL
3     ,atm_number VARCHAR(20) NULL
4     ,atm_manufacturer VARCHAR(50) NULL
5     ,atm_location_id INT NULL
6     ,PRIMARY KEY (atm_id)
7     ,CONSTRAINT atm_location_id FOREIGN KEY (atm_location_id)
REFERENCES atm_transaction.dim_location(location_id)
8 );
9
```

Run**Save****Schedule****Clear** **Send feedback****Query results****Table details****Query** **Execution** **Data** **Visual** Completed, started on September 26, 2022 at 09:35:10

ELAPSED TIME: 00 m 03 s

```
CREATE TABLE atm_transaction.DIM_DATE (
    date_id INT NOT NULL
    ,full_date_time VARCHAR(45) NULL
    ,year INT NULL
    ,month VARCHAR(20) NULL
    ,day INT NULL
    ,hour INT NULL
    ,weekday VARCHAR(20) NULL
    ,PRIMARY KEY (date_id)
);

1 CREATE TABLE atm_transaction.DIM_DATE (
2     date_id INT NOT NULL
3     ,full_date_time VARCHAR(45) NULL
4     ,year INT NULL
5     ,month VARCHAR(20) NULL
6     ,day INT NULL
7     ,hour INT NULL
8     ,weekday VARCHAR(20) NULL
9     ,PRIMARY KEY (date_id)
10    );
11
```

Run

Save

Schedule

Clear

Query results

Table details

Query

Execution

Data

✓ Completed, started on September 26, 2022 at 09:35:29

ELAPSED TIME: 00 m 03 s

```
CREATE TABLE atm_transaction.DIM_CARD_TYPE (
    card_type_id INT NOT NULL
    ,card_type VARCHAR(30) NULL
    ,PRIMARY KEY (card_type_id)
);
```

```
1 CREATE TABLE atm_transaction.DIM_CARD_TYPE (
2     card_type_id INT NOT NULL
3     ,card_type VARCHAR(30) NULL
4     ,PRIMARY KEY (card_type_id)
5 );
6
```

Run

Save

Schedule

Clear



Query results

Table details

Query

Execution

Data

✓ Completed, started on September 26, 2022 at 09:35:55

ELAPSED TIME: 00 m 03 s

```
CREATE TABLE atm_transaction.FACT_ATM_TRANS (
    trans_id BIGINT NOT NULL
    ,atm_id INT NULL
    ,weather_loc_id INT NULL
    ,date_id INT NULL
    ,card_type_id INT NULL
    ,atm_status VARCHAR(20) NULL
    ,currency VARCHAR(10) NULL
    ,service VARCHAR(20) NULL
    ,transaction_amount INT NULL
    ,message_code VARCHAR(255) NULL
    ,message_text VARCHAR(255) NULL
    ,rain_3h DECIMAL(10, 3) NULL
    ,clouds_all INT NULL
    ,weather_id INT NULL
    ,weather_main VARCHAR(50) NULL
    ,weather_description VARCHAR(255) NULL
    ,PRIMARY KEY (trans_id)
    ,FOREIGN KEY (weather_loc_id) REFERENCES
atm_transaction.DIM_LOCATION(location_id)
    ,FOREIGN KEY (atm_id) REFERENCES atm_transaction.DIM_ATM(atm_id)
    ,FOREIGN KEY (date_id) REFERENCES atm_transaction.DIM_DATE(date_id)
    ,FOREIGN KEY (card_type_id) REFERENCES
atm_transaction.DIM_CARD_TYPE(card_type_id)
);
```

```

1 CREATE TABLE atm_transaction.FACT_ATM_TRANS (
2     trans_id BIGINT NOT NULL
3     ,atm_id INT NULL
4     ,weather_loc_id INT NULL
5     ,date_id INT NULL
6     ,card_type_id INT NULL
7     ,atm_status VARCHAR(20) NULL
8     ,currency VARCHAR(10) NULL
9     ,service VARCHAR(20) NULL
10    ,transaction_amount INT NULL
11    ,message_code VARCHAR(255) NULL
12    ,message_text VARCHAR(255) NULL
13    ,rain_3h DECIMAL(10, 3) NULL
14    ,clouds_all INT NULL
15    ,weather_id INT NULL
16    ,weather_main VARCHAR(50) NULL
17    ,weather_description VARCHAR(255) NULL
18    ,PRIMARY KEY (trans_id)
19    ,FOREIGN KEY (weather_loc_id) REFERENCES atm_transaction.DIM_LOCATION(location_id)
20    ,FOREIGN KEY (atm_id) REFERENCES atm_transaction.DIM_ATM(atm_id)
21    ,FOREIGN KEY (date_id) REFERENCES atm_transaction.DIM_DATE(date_id)
22    ,FOREIGN KEY (card_type_id) REFERENCES atm_transaction.DIM_CARD_TYPE(card_type_id)
23 );
24

```


Send
 Query results
 Table details
Query

✔ Completed, started on September 26, 2022 at 09:38:48

ELAPSED TIME: 00 m 18 s

Loading data into a Redshift cluster from Amazon S3 bucket

```
copy atm_transaction.DIM_LOCATION
FROM 's3://etlprojectdipanwita/DIM_LOCATION/part-00000-ee69c149-d1d7-4883-
a4a9-952d2b934f4d-c000.csv'
iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
delimiter ','
region 'ap-northeast-1'
IGNOREHEADER 1;
```

```
1 copy atm_transaction.DIM_LOCATION
2 FROM 's3://etlprojectdipanwita/DIM_LOCATION/part-00000-ee69c149-d1d7-4883-a4a9-952d2b934f4d-c000.csv'
3 iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
4 delimiter ','
5 region 'ap-northeast-1'
6 IGNOREHEADER 1;
7
```

Run

Save

Schedule

Clear

 Send feedback

Query results

Table details

Query 309 

 Execution

 Data

 Visualize

 Completed, started on September 26, 2022 at 09:39:30

ELAPSED TIME: 00 m 15 s

```
copy atm_transaction.dim_atm
FROM 's3://etlprojectdipanwita/DIM_ATM/part-00000-238866fd-5d8e-4ea0-87fe-
00d6bbb45e76-c000.csv'
iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
delimiter ','
region 'ap-northeast-1'
IGNOREHEADER 1;
```

```
1 copy atm_transaction.dim_atm
2 FROM 's3://etlprojectdipanwita/DIM_ATM/part-00000-238866fd-5d8e-4ea0-87fe-00d6bbb45e76-c000.csv'
3 iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
4 delimiter ','
5 region 'ap-northeast-1'
6 IGNOREHEADER 1;
7
```

RunSaveScheduleClear[!\[\]\(eaac180de418db4eae4b4cefebda75e8_img.jpg\) Send feedback](#)[Query results](#)[Table details](#)[Query 336 !\[\]\(173968034f6ca6c36e25dcb8a274badd_img.jpg\)](#)[!\[\]\(43fda5baa5446493352974e4b4060607_img.jpg\) Execution](#)[!\[\]\(0df0bdc1e09cbc2587d9dd4511cb0c27_img.jpg\) Data](#)[!\[\]\(d538389f939343cdedbb759655cf0521_img.jpg\) Visualize](#)

 Completed, started on September 26, 2022 at 09:40:03

ELAPSED TIME: 00 m 09 s

```
copy atm_transaction.DIM_DATE
FROM 's3://etlprojectdipanwita/DIM_DATE/part-00000-abdf84dd-8c8a-4dd7-8f08-
67f090e2dc8f-c000.csv'
iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
delimiter ','
region 'ap-northeast-1'
IGNOREHEADER 1;
```

```
1 copy atm_transaction.DIM_DATE
2 FROM 's3://etlprojectdipanwita/DIM_DATE/part-00000-abdf84dd-8c8a-4dd7-8f08-67f090e2dc8f-c000.csv'
3 iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
4 delimiter ','
5 region 'ap-northeast-1'
6 IGNOREHEADER 1;
7
```

Run

Save

Schedule

Clear

Send feedback

Query results

Table details

Query 360 

 Completed, started on September 26, 2022 at 09:40:26

ELAPSED TIME: 00 m 08 s

 Execution

 Data

 Visualize

```
copy atm_transaction.DIM_CARD_TYPE
FROM 's3://etlprojectdipanwita/DIM_CARD_TYPE/part-00000-b894dd46-d885-4e5a-
98f3-892d8b736fa3-c000.csv'
iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
delimiter ','
region 'ap-northeast-1'
IGNOREHEADER 1;
```

```
1 copy atm_transaction.DIM_CARD_TYPE
2 FROM 's3://etlprojectdipanwita/DIM_CARD_TYPE/part-00000-b894dd46-d885-4e5a-98f3-892d8b736fa3-c000.csv'
3 iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
4 delimiter ','
5 region 'ap-northeast-1'
6 IGNOREHEADER 1;
7
```

RunSaveScheduleClearSend feedback[Query results](#)[Table details](#)[Query 380](#)[Execution](#)[Data](#)[Visualize](#)

⌚ Completed, started on September 26, 2022 at 09:40:47

ELAPSED TIME: 00 m 08 s

```
copy atm_transaction.FACT_ATM_TRANS
FROM 's3://etlprojectdipanwita/FACT_ATM_TRANS/part-00000-4ad71947-3fe9-4058-
b7bd-5a1797e1cae1-c000.csv'
iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
delimiter ','
region 'ap-northeast-1'
IGNOREHEADER 1
TRUNCATECOLUMNS CSV;
```

```
1 copy atm_transaction.FACT_ATM_TRANS
2 FROM 's3://etlprojectdipanwita/FACT_ATM_TRANS/part-00000-4ad71947-3fe9-4058-b7bd-5a1797e1cae1-c000.csv'
3 iam_role 'arn:aws:iam::914450217410:role/etl-project-s3-access'
4 delimiter ','
5 region 'ap-northeast-1'
6 IGNOREHEADER 1
7 TRUNCATECOLUMNS CSV;
8
```

RunSaveScheduleClear Send feedback[Query results](#)[Table details](#)[Query 415](#)[Execution](#)[Data](#)[Visualize](#) Completed, started on September 26, 2022 at 09:41:37

ELAPSED TIME: 00 m 20 s