

Install Databases on Amazon RDS

Contents:

- *Limitations*
 - *Pre-requisites*
 - *Initialize RDS instance*
 - *Configure the Trifacta platform for RDS*
 - *Install the Databases*
 - *Logging*
-

As needed, the Trifacta® databases can be installed as PostgreSQL DBs on Amazon RDS. Amazon Relational Database Service (Amazon RDS) makes it easy to set up, operate, and scale a relational database in the cloud.

Limitations

- SSL connectivity is not supported.

Pre-requisites

NOTE: You can use the suggested defaults below for sizing your RDS instance. If you have questions or concerns about sizing recommendations, please contact *Trifacta Support*.

- Admin access to an Amazon RDS account

Initialize RDS instance

Steps:

1. In your RDS dashboard, click **Launch a DB instance**.

NOTE: The RDS instance must be launched in the same Amazon region as the Trifacta node.

2. For Select Engine: Select **PostgreSQL**.
3. For Production?: Choose **Yes** if you are deploying the database for a production instance of the Trifacta platform. Otherwise, select **No**.
4. DB Engine: `postgres`
5. For the DB details, see below:

NOTE: Except as noted below, properties should be specified according to your enterprise requirements.

a. Instance Specifications:

- i. License Model: `postresql-license`
 - ii. DB Engine Version: For more information on the supported versions of PostgreSQL, see *System Requirements*.
 - iii. Allocated Storage: at least 10 GB
6. For Advanced Settings, please apply the following settings:

a. Network and Security:

- i. VPC security group must allow for access from the Trifacta platform.
- b. **Database Options:**
 - i. Database Name: `trifacta`
 - ii. Database Port: `5432`
 - 1. The port number can be changed as needed. See *System Ports*.
- 7. Populate other properties according to your enterprise requirements.
- 8. To complete the set up click **Launch DB Instance**.

Configure the Trifacta platform for RDS

Please complete the following steps to integrate the Trifacta platform with the DB instance you just created.

Steps:

1. In the RDS console, you must find the Public DNS endpoint for the RDS instance you created:
 - a. Under Instances, expand the name of the instance you created.
 - b. The DNS endpoint should be listed under the name in the Endpoint section.
2. Set the host for each database to the Public DNS endpoint for the RDS instance:

Database	Property
Main database	<code>webapp.db.host</code>
Jobs database	<code>batch-job-runner.db.host</code>
Scheduling database	<code>scheduling-service.database.host</code>
Time-based Trigger database	<code>time-based-trigger-service.database.host</code>

3. To set custom database names, usernames, and passwords:

a. Edit

```
/opt/trifacta/conf/trifacta-conf.json
```

- b. For each database below, you can review the database name, username, and password.

Database	Property
Main database	<code>webapp.db.name</code>
	<code>webapp.db.username</code>
	<code>webapp.db.password</code>
Jobs database	<code>batch-job-runner.db.name</code>
	<code>batch-job-runner.db.username</code>
	<code>batch-job-runner.db.password</code>
Scheduling database	<code>scheduling-service.database.name</code>
	<code>scheduling-service.database.user</code>
	<code>scheduling-service.database.password</code>
Time-Based Trigger database	<code>time-based-trigger-service.database.name</code>
	<code>time-based-trigger-service.database.user</code>
	<code>time-based-trigger-service.database.password</code>

- c. Make changes in the file as needed and save.

Install the Databases

Steps:

1. Run the following script, which builds the four databases and specifies the appropriate roles for each database, based on the parameters you have specified in

```
trifacta-conf.json
```

```
:
```

NOTE: This script must be run as the root user or via sudo superuser.

```
/opt/trifacta/bin/setup-utils/db/trifacta-create-postgres-roles-dbs.sh
```

2. Login to the application.
3. Create a flow and import a dataset into it. If you are able to wrangle the dataset, the integration is working.

Logging

1. To review database logs in RDS, locate the Instance details page in the RDS console.
2. Click **Recent Events and Logs**.
3. If your account has the appropriate permissions, all Trifacta database logs are available here.