

Install Databases for PostgreSQL

Contents:

- *Limitations*
 - *Pre-requisites*
 - *Select Configuration File*
 - *Database Install*
 - *Acquire Port Information*
-

The Trifacta® platform requires access to the following databases.

- **Main database:** storage of users and metadata about your datasets, including completed jobs.
- **Jobs database:** storage of job tracking information. Jobs are purged upon completion or job timeout. Failed jobs are purged periodically.
- **Configuration Service database:** storage of parameter settings at the workspace level.
- **Artifact Storage Service database:** storage for feature usage data such value mappings for the standardization feature.
-

The scheduling feature is enabled by default. If it's enabled, the following databases are also required:

- **Scheduling database:** Storage of schedules, including datasets to execute
- **Time-based Trigger database:** Storage of triggering information.
- For more information, see *Configure Automator*.

This section describes how to install the database server, after which you can create and initialize the databases and their users.

Limitations

- You must install a supported version of the database. For more information on supported versions of this database type, see *System Requirements*.
- SSL connections are not supported.

Pre-requisites

- The installing user must have write permissions to the directory from which the commands are executed.
- The installing user must have sudo privileges.

Select Configuration File

By default, the Trifacta platform assumes that you are installing the databases in a PostgreSQL instance. This configuration file is stored here:

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

Database Install

NOTE: The following distributions and commands are for PostgreSQL 9.6.

O/S Distribution	URL	Package Name
CentOS 6	https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-6-x86_64/pgdg-centos96-9.6-3.noarch.rpm	postgresql96-server
CentOS 7	https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-7-x86_64/pgdg-centos96-9.6-3.noarch.rpm	postgresql96-server
RHEL 6	https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-6-x86_64/pgdg-centos96-9.6-3.noarch.rpm	postgresql96-server
RHEL 7	https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-7-x86_64/pgdg-centos96-9.6-3.noarch.rpm	postgresql96-server

For CentOS 6.x:

```
wget https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-6-x86_64/pgdg-centos96-9.6-3.noarch.rpm
sudo yum -y install pgdg-centos96-9.6-3.noarch.rpm
sudo yum -y install postgresql96-server
```

For CentOS 7.x:

```
wget https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-7-x86_64/pgdg-centos96-9.6-3.noarch.rpm
sudo yum -y install pgdg-centos96-9.6-3.noarch.rpm
sudo yum -y install postgresql96-server
```

For Red Hat Enterprise Linux 6.x:

```
wget https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-6-x86_64/pgdg-redhat96-9.6-3.noarch.rpm
sudo yum -y install pgdg-redhat96-9.6-3.noarch.rpm
sudo yum -y install postgresql96-server
```

For Red Hat Enterprise Linux 7.x:

```
wget https://download.postgresql.org/pub/repos/yum/9.6/redhat/rhel-7-x86_64/pgdg-redhat96-9.6-3.noarch.rpm
sudo yum -y install pgdg-redhat96-9.6-3.noarch.rpm
sudo yum -y install postgresql96-server
```

For Ubuntu 14.04:

Add the repository's archive key to your apt-key keyring:

```
wget --quiet -O - https://www.postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add -
```

Create a file named `/etc/apt/sources.list.d/pgdg.list`, containing the following:

```
deb http://apt.postgresql.org/pub/repos/apt/ trusty-pgdg main
deb-src http://apt.postgresql.org/pub/repos/apt/ trusty-pgdg main
```

Run the following command:

```
sudo apt-get update
sudo apt-get install -y postgresql-9.6
```

For Ubuntu 16.04:

Add the repository's archive key to your apt-key keyring:

```
wget --quiet -O - https://www.postgresql.org/media/keys/ACCC4CF8.asc | sudo apt-key add -
```

Create a file named `/etc/apt/sources.list.d/pgdg.list`, containing the following:

```
deb http://apt.postgresql.org/pub/repos/apt/ xenial-pgdg main
```

Run the following command:

```
sudo apt-get update  
sudo apt-get install -y postgresql-9.6
```

Acquire Port Information

After you have completed the installation, you must acquire the port information for each database from the following locations on the Trifacta node. These port numbers need to be applied inside the Trifacta platform.

i NOTE: By default, PostgreSQL and the platform use port 5432 for communication. If that port is not available at install/upgrade time, the next available port is used, which is typically 5433. This change may occur if a previous version of PostgreSQL is on the same server. When a non-default port number is used, the platform must be configured to use it. For more information, see *Change Database Port*.

CentOS/RHEL (PostgreSQL 9.6):

```
/var/lib/pgsql/9.6/data/postgresql.conf
```

Ubuntu (PostgreSQL 9.6) :

```
/etc/postgresql/9.6/main/postgresql.conf
```