

Backup and Recovery

Contents:

- *Stop All Services*
- *Backup Platform Files*
 - *Configuration*
 - *License*
- *Backup Databases*
 - *Location of backup and recovery tools*
 - *Backup commands*
 - *Scheduling*
- *Restart*
- *Recovery*

This section provides overview information on the key data and metadata that should be managed by your enterprise backup and recovery policies.

NOTE: This section covers how to perform a basic cold backup of the product. Hot backups are not supported.

All backups should be performed in accordance with your enterprise's backup and recovery policies.

Stop All Services

Before you begin, the Trifacta platform and databases should be stopped. See *Start and Stop the Platform*.

Backup Platform Files

The following directories on the Trifacta node should be backed up on a regular basis:

Configuration

You can back up all key configuration files into the `/tmp` directory using the following commands:

```
cp /opt/trifacta/conf/trifacta-conf.json /tmp/trifacta-conf.json
cp /opt/trifacta/conf/env.sh /tmp/env.sh
cp /etc/init.d/trifacta /tmp/trifacta.service
```

License

You should backup your license key:

```
/opt/trifacta/license
```

See *License Key*.

Backup Databases

The Trifacta platform utilizes the following databases as part of normal operations. These databases should be backed up on a regular basis:

Database Name	Databaseld	Description
Main DB	trifacta	Stores users and metadata for flows, including datasets, and recipes.
Jobs DB	trifacta-activiti	Stores and maintains job execution status and details.
Scheduling DB	trifactaschedulingservice	Stores metadata for scheduled jobs.
Time-based Trigger DB	trifactatimebasedtriggerse rvice	Additional database required for scheduled jobs.
Configuration Service DB	trifactaconfigurationservi ce	Stores configuration settings for the workspace.
Artifact Storage Service DB	trifactaartifactstorageser vice	Stores feature usage data such value mappings for the standardization feature.
Job Metadata Service DB	trifactajobmetadataservice	Stores metadata on job execution.

For more information on setting up these databases, see *Install Databases*.

Location of backup and recovery tools

PostgreSQL

Depending on your operating system, you can find the backup tools in the following location.

NOTE: These locations apply to PostgreSQL 9.6.

CentOS/RHEL:

```
/usr/pgsql-9.6/bin/pg_dump  
/usr/pgsql-9.6/bin/psql
```

Ubuntu:

```
/usr/lib/postgresql/9.6/bin/pg_dump  
/usr/lib/postgresql/9.6/bin/psql
```

MySQL

Please locate the following programs in your MySQL distribution:

```
mysqldump  
mysql
```

Backup commands

The following commands can be used to back up the databases.

PostgreSQL

For more information on command options, see <https://www.postgresql.org/docs/9.6/static/backup.html>.

NOTE: These commands must be executed as the `trifacta` user.

NOTE: The following commands are for PostgreSQL 9.6 for all supported operating systems. For specific commands for other versions, please see the database documentation.

Tip: You may see performance improvements by backing up and restoring using `.TAR` files. However, there is a risk that `.TAR` support could change in the future. For more information, please see the PostgreSQL documentation.

Trifacta DB:

```
pg_dump trifacta > trif_triDB_bkp_<date>.sql
```

Jobs DB:

```
pg_dump trifacta-activiti > trif_actDB_bkp_<date>.sql
```

Scheduling DB:

```
pg_dump trifactaschedulingservice > trif_schDB_bkup_<date>.sql
```

Time-Based Trigger DB:

```
pg_dump trifactatimebasedtriggerservice > trif_tbtsDB_bkup_<date>.sql
```

Configuration Service DB:

```
pg_dump trifactaconfigurationservice > trif_confservDB_bkup_<date>.sql
```

Artifact Storage DB:

```
pg_dump trifactaartifactstorageservice > trif_artifactstorageservDB_bkup_<date>.sql
```

Job Metadata Service DB:

```
pg_dump trifactajobmetadataservice > trif_jobmetadataservDB_bkup_<date>.sql
```

MySQL

For more information on command options, see <https://dev.mysql.com/doc/refman/5.7/en/mysqldump-sql-format.html>.

```
su - mysql
```

NOTE: The following commands are for MySQL 5.7 for all supported operating systems. For specific commands for other versions, please see the database documentation.

Trifacta DB:

```
mysqldump trifacta > trif_triDB_bkp_<date>.sql
```

Jobs DB:

```
mysqldump trifacta-activiti > trif_actDB_bkp_<date>.sql
```

Scheduling DB:

```
mysqldump trifactaschedulingservice > trif_schDB_bkup_<date>.sql
```

Time-Based Trigger DB:

```
mysqldump trifactatimebasedtriggerservice > trif_tbtsDB_bkup_<date>.sql
```

Configuration Service DB:

```
mysqldump trifactaconfigurationservice > trif_confservDB_bkup_<date>.sql
```

Artifact Storage DB:

```
mysqldump trifactaartifactstorageservice > trif_artifactstorageservDB_bkup_<date>.sql
```

Job Metadata Service DB:

```
mysqldump trifactajobmetadataservice > trif_jobmetadataservDB_bkup_<date>.sql
```

Scheduling

You can schedule nightly execution of these backups using a third-party scheduler such as cron.

Restart

You can restart the Trifacta platform now. See *Start and Stop the Platform*.

Recovery

See *Platform Rollback*.