Configure for Cloudera

Contents:

This section provides additional configuration requirements for integrating the Trifacta® platform with the Cloudera platform.

- This section applies only to the versions of CDH that Trifacta supports. For more information, see *Supported Deployment Scenarios for Hortonworks*.

> **NOTE:** Except as noted, the following configuration items apply to the latest supported version of Cloudera platform.

Pre-requisites

Before you begin, it is assumed that you have completed the following tasks:

1. Successfully installed a supported version of Cloudera platform into your enterprise infrastructure.
2. Installed the Trifacta software in your environment. For more information, see *Install Software*.
3. Reviewed the mechanics of platform configuration. See *Required Platform Configuration*.
4. Configured access to the Trifacta database. See *Configure the Databases*.
5. Performed the basic cluster integration configuration. See *Configure for Hadoop*.
6. You have access to platform configuration through the Trifacta node or through the Admin Settings page. You can apply this change through the *Admin Settings Page* (recommended) or

    ```
    trifacta-conf.json
    ```

    . For more information, see *Platform Configuration Methods*.

Configure Trifacta platform

Configure Hive Locations

If you are enabling an integration with Hive on the cluster, there are some distribution-specific parameters that must be set. For more information, see *Configure for Hive*.

Configure SSL for Hive

CDH supports two methods of enabling SSL communications with Hive:

1. **SASL-QOP method:** Enable encryption between Hive JDBC and HiveServer 2 using SASL-QOP. This method is available by default with the Trifacta platform.
2. **TLS/SSL method:** Use TLS/SSL encryption for JDBC connections to HiveServer 2.

To determine the method in use:

1. In Cloudera Manager configuration, search for: `tls`.
2. If the options for TLS/SSL are enabled, please complete the following configuration steps.
3. If these options are not enabled, the cluster can still use the SASL-QOP method. For more information on this method, see *Configure for Hive* .

Enable TLS/SSL Method

Steps:

1. The default Hive JDBC driver provided with your Trifacta installation must be replaced with the drive provided by Cloudera. Please complete the following commands, noting the wildcards (*) in the JAR path:

   > **NOTE:** The current driver must be removed or replaced in the working directory. Do not leave it in the directory.

   ```
   cd /opt/trifacta/services/data-service/build/dependencies
   rm *hive*jdbc*
   cp /opt/cloudera/parcels/CDH-5.8*/jars/hive-jdbc-1.1.0-cdh5.8.0*jar .
   ```

2. Enable the Hive connection. For the Hive connection string options, you must specify something like the following:

   ```
   "connectStrOpts": ";ssl=true;sslTrustStore=</path/to/truststore>;
   trustStorePassword=<storePassword>"
   ```

   > **NOTE:** The truststore specified above must exist on the Trifacta node and be accessible to the Trifacta user through the listed password. This truststore must contain the certificate for the Hive server.

3. Save the parameters file. For more information on creating the connection, see *Configure for Hive*.
4. Restart the platform. See *Start and Stop the Platform*.
5. Verify that you can read from a Hive source through the application. See *Hive Browser*.

Configure for Cloudera Navigator

If you are enabling metadata publication through Cloudera Navigator, additional configuration is required. See *Configure Publishing to Cloudera Navigator*.

Restart

To apply your changes, restart the platform. See *Start and Stop the Platform.*