Configure for Azure

Contents:

Please complete the following steps in the listed order to configure your installed instance of the Trifacta® platform to integrate with the running environment cluster.

Pre-requisites

1. Deploy running environment cluster and Trifacta node.

   > **NOTE:** The running environment cluster can be deployed as part of the installation process. You can also integrate the platform with a pre-existing cluster. Details are below.

2. Install Trifacta platform on the node.

For more information, see *Install for Azure*.

Configure Azure

Create registered application

You must create an Azure Active Directory (AAD) application and grant it the desired access permissions, such as read/write access to resources and read/write access to the Azure Key Vault secrets.

> **NOTE:** If you are integrating with Azure Databricks and are Managed Identities for authentication, please skip this section. That configuration is covered in a later step.

This service principal is used by the Trifacta platform for access to all Azure resources. For more information, see *https://docs.microsoft.com/en-us/azure/azure-resource-manager/resource-group-create-service-principal-portal*.

After you have registered, acquire the following information:

Azure Property	Location	Use

Application ID	Acquire this value from the Registered app blade of the Azure Portal.	Applied to Trifacta platform configuration: `azure.applicationid`.
Service User Key	Create a key for the Registered app in the Azure Portal.	Applied to Trifacta platform configuration: `azure.secret`. **NOTE:** If you are using Azure AD to integrate with an Azure Databricks cluster, the Azure AD secret value stored in `azure.secret` must begin with an alphanumeric character. This is a known issue.
Directory ID	Copy the Directory ID from the Properties blade of Azure Active Directory.	Applied to Trifacta platform configuration: `azure.directoryId`.

To create an Azure Active Directory (AAD) application, please complete the following steps in the Azure console.

Steps:

1. Create registered application:
    a. In the Azure console, navigate to **Azure Active Directory > App Registrations**.
    b. Create a New App. Name it `trifacta`.

    NOTE: Retain the Application ID and Directory ID for configuration in the Trifacta platform.

2. Create a client secret:
    a. Navigate to **Certificates & secrets**.
    b. Create a new Client secret.

    NOTE: Retain the value of the Client secret for configuration in the Trifacta platform.

3. Add API permissions:
    a. Navigate to **API Permissions**.
    b. Add Azure Key Vault with the `user_impersonation` permission.

These properties are applied later in the configuration process.

Configure the Platform

Configure for Azure Databricks

You can integrate the Trifacta platform with Azure Databricks. For more information, see *Configure for Azure Databricks*.

Configure base storage layer

For Azure installations, you can set your base storage layer to be HDFS or WASB.

NOTE: The base storage layer must be set after installation. After it has been configured, it cannot be modified.

Azure storage	webapp.storageProtocol setting
WASB	`wasbs`

ADLS Gen2	`abfss`
ADLS Gen1	`adl`

See *Set Base Storage Layer*.

Configure for Key Vault

For authentication purposes, the Trifacta platform must be integrated with an Azure Key Vault keystore. See *Configure Azure Key Vault*.

Configure for SSO

If needed, you can integrate the Trifacta platform with Azure AD for Single-Sign On to the platform. See *Configure SSO for Azure AD*.

Configure for ADLS Gen2

Enable read-only or read-write access to ADLS Gen2. For more information, see *ADLS Gen2 Access*.

Configure for ADLS Gen1

Enable read-only or read-write access to ADLS Gen1. For more information, see *ADLS Gen1 Access*.

Configure for WASB

Enable read-only or read-write access to WASB. For more information on integrating with WASB, see *WASB Access*.

Configure for Azure Gov Cloud

To enable use of the Azure Gov Cloud, please perform the following configuration steps.

> **NOTE:** Managed Identities is not supported for Azure Gov Cloud.

Steps:

1. You can apply this change through the *Admin Settings Page* (recommended) or `trifacta-conf.json`. For more information, see *Platform Configuration Methods*.
2. Locate the following parameter and set it to `US_GOV`:

   ```
   "azure.environment": "US_GOV",
   ```

3. Save your changes and restart the platform.

Configure relational connections

If you are integrating Trifacta Self-Managed Enterprise Edition with relational datastores, please complete the following configuration sections.

Create encryption key file

An encryption key file must be created on the Trifacta node. This key file is shared across all relational connections. See *Create Encryption Key File*.

Create Azure SQL Database connection

For more information, see *Azure SQL Database Connections*.

Create Azure SQL DW connection

For more information, see *Microsoft SQL Data Warehouse Connections*.

Testing

1. Load a dataset from the cluster.
2. Perform a few simple steps on the dataset.
3. Click **Run** in the Transformer page.
4. When specifying the job:
    a. Click the Profile Results checkbox.
    b. Select **Spark**.
5. When the job completes, verify that the results have been written to the appropriate location.