

# Create SFTP Connections

## Contents:

- *Limitations*
  - *Pre-requisites*
    - *SSH Keys*
  - *Enable*
    - *Configure file storage protocols and locations*
    - *Java VFS service*
  - *Create Connection*
    - *Create through application*
    - *Create through APIs*
- 

You can create connections to SFTP servers to upload your datasets to the Trifacta® platform.

- Linux- and Windows-based SFTP servers are supported.

## Limitations

- Ingest of over 500 files through SFTP at one time is not supported.
- You cannot run jobs using Avro or Parquet sources uploaded via SFTP.
- When you specify a parameterized output as part of your job execution, the specified output location may include additional unnecessary information about the SFTP connection identifier. None of this information is sensitive. This is a known issue.
- Jobs can be executed from SFTP sources on the following running environments:
  - Trifacta Photon
  - HDFS-based Spark, which includes Cloudera and Hortonworks
  
- You cannot publish TDE format to SFTP destinations.
- You cannot publish compressed Snappy files to SFTP destinations.

## Pre-requisites

- Acquire user credentials to access the SFTP server. You can use username/password credentials or SSH keys. See below.
- Verify that the credentials can access the proper locations on the server where your data is stored. Initial directory of the user account must be accessible.

## SSH Keys

If preferred, you can use SSH keys to for authentication to the SFTP server.

**NOTE:** SSH keys must be private RSA keys. If you have OpenSSH keys, you can use the `ssh-keygen` utility to convert them to private RSA keys.

## Enable

By default, this connection type is automatically enabled.

**NOTE:** You must provide the protocol identifier and storage locations for the SMTP server. See below.

## Configure file storage protocols and locations

The Trifacta platform must be provided the list of protocols and locations for accessing SFTP.

### 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 parameters and set their values according to the table below:

```
"fileStorage.whitelist": ["sftp"],  
"fileStorage.defaultBaseUris": ["sftp:///"],
```

Parameter	Description
filestorage.whitelist	A comma-separated list of protocols that are permitted to access SFTP.  <b>NOTE:</b> The protocol identifier "sftp" must be included in this list.
filestorage.defaultBaseUris	For each supported protocol, this param must contain a top-level path to the location where Trifacta platform files can be stored. These files include uploads, samples, and temporary storage used during job execution.  <b>NOTE:</b> A separate base URI is required for each supported protocol. You may only have one base URI for each protocol.  <b>NOTE:</b> For SFTP, three slashes at the end are required, as the third one is the path value.

3. Save your changes and restart the platform.

## Java VFS service

Use of SFTP connections requires the Java VFS service in the Trifacta platform.

**NOTE:** This service is enabled by default.

For more information on configuring this service, see *Configure Java VFS Service*.

## Create Connection

### Create through application

Any user can create a SFTP connection through the application.

**NOTE:** Only an administrator can make a connection available for all users.

## Steps:

1. Login to the application.
2. In the menu, select **Settings menu > Settings > Connections**. See *Connections Page*.
3. In the Connections page, click **Create Connection**. See *Create Connection Window*.
4. In the Create Connection window, click the SFTP connection card.
5. Specify the properties for your SFTP server.

Property	Description
Host	The hostname of the FTP server to which you are connecting. Do not include any protocol identifier ( <code>sftp://</code> ).
Port	The port number to use to connect to the server. Default port number is 22.
Credential Type	Select one of the following:  <code>basic</code> - authenticate via username and password  <code>SSH Key</code> - authenticate via username and SSH key
User Name	The username to use to connect.
Password	(Basic credential type) The password associated with the username.
SSH Key	(SSH Key credential type) The SSH key that applies to the username.
Test Connection	Click this button to test the connection that you have specified.
Default Directory	Absolute path on the SFTP server where users of the connection can begin browsing.
Block Size (Bytes)	Fetch size in bytes for each read from the SFTP server.  <div style="border: 1px solid #ccc; padding: 5px;"><b>NOTE:</b> Raising this value may increase speed of read operations. However, if it is raised too high, resources can become overwhelmed, and the read can fail.</div>
Connection Name	The name of the connection as you want it to appear in the application.
Description	This description is displayed in the application.

For more information, see *Create Connection Window*.

6. Click **Save**.

## Create through APIs

1. Acquire the vendor and type information. See *Connection Types*.
2. Create the connection through the APIs. See *API Connections Create v4*.