



Talend Open Studio for Data Integration Installation and Upgrade Guide for Linux

7.3.1

Contents

| | |
|--|-----------|
| Copyleft..... | 3 |
| Talend Open Studio for Data Integration: Prerequisites..... | 4 |
| About this installation guide..... | 4 |
| Preparing your installation..... | 4 |
| Hardware requirements..... | 5 |
| Software requirements..... | 6 |
| Installing the XULRunner package..... | 11 |
| Setting up JAVA_HOME..... | 12 |
| Installing your Talend Open Studio for Data Integration manually..... | 13 |
| Installing and configuring your Talend Studio..... | 13 |
| Upgrading your Talend products..... | 21 |
| Backing up the environment..... | 21 |
| Upgrading the Talend projects in Talend Studio..... | 21 |
| Appendices..... | 22 |
| Supported Third-Party System/Database/Business Application Versions..... | 22 |

Copyleft

Adapted for 7.3.1. Supersedes previous releases.

The content of this document is correct at the time of publication.

However, more recent updates may be available in the online version that can be found on [Talend Help Center](#).

This documentation is provided under the terms of the Creative Commons Public License (CCPL).

For more information about what you can and cannot do with this documentation in accordance with the CCPL, please read: <http://creativecommons.org/licenses/by-nc-sa/2.0/>.

Notices

Talend is a trademark of Talend, Inc.

All brands, product names, company names, trademarks and service marks are the properties of their respective owners.

License Agreement

The software described in this documentation is licensed under the Apache License, Version 2.0 (the "License"); you may not use this software except in compliance with the License. You may obtain a copy of the License at <http://www.apache.org/licenses/LICENSE-2.0.html>. Unless required by applicable law or agreed to in writing, software distributed under the License is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the License for the specific language governing permissions and limitations under the License.

This product includes software developed at AOP Alliance (Java/J2EE AOP standards), ASM, Amazon, AntLR, Apache ActiveMQ, Apache Ant, Apache Axiom, Apache Axis, Apache Axis 2, Apache Batik, Apache CXF, Apache Chemistry, Apache Common Http Client, Apache Common Http Core, Apache Commons, Apache Commons Bcel, Apache Commons JXPath, Apache Commons Lang, Apache Derby Database Engine and Embedded JDBC Driver, Apache Geronimo, Apache Hadoop, Apache Hive, Apache HttpClient, Apache HttpComponents Client, Apache JAMES, Apache Log4j, Apache Lucene Core, Apache Neethi, Apache POI, Apache ServiceMix, Apache Tomcat, Apache Velocity, Apache WSS4J, Apache WebServices Common Utilities, Apache Xml-RPC, Apache Zookeeper, Box Java SDK (V2), CSV Tools, DataStax Java Driver for Apache Cassandra, Ehcache, Ezmorph, Ganymed SSH-2 for Java, Google APIs Client Library for Java, Google Gson, Groovy, Guava: Google Core Libraries for Java, H2 Embedded Database and JDBC Driver, Hector: A high level Java client for Apache Cassandra, Hibernate Validator, HighScale Lib, HsqlDB, Ini4j, JClouds, JLine, JSON, JSR 305: Annotations for Software Defect Detection in Java, JUnit, Jackson Java JSON-processor, Java API for RESTful Services, Java Agent for Memory Measurements, Jaxb, Jaxen, Jettison, Jetty, Joda-Time, Json Simple, LightCouch, MetaStuff, Mondrian, OpenSAML, Paracel JDBC Driver, PostgreSQL JDBC Driver, Resty: A simple HTTP REST client for Java, Rocoto, SL4J: Simple Logging Facade for Java, SQLite JDBC Driver, Simple API for CSS, SshJ, StAX API, StAXON - JSON via StAX, The Castor Project, The Legion of the Bouncy Castle, W3C, Woden, Woodstox: High-performance XML processor, Xalan-J, Xerces2, XmlBeans, XmlSchema Core, Xmlsec - Apache Santuario, Zip4J, atinject, dropbox-sdk-java: Java library for the Dropbox Core API, google-guice. Licensed under their respective license.

Talend Open Studio for Data Integration: Prerequisites

About this installation guide

This guide explains how to install and configure your Talend product. You can install your product by using the Talend Installer, by manually installing the Talend modules, or with the Red Hat Package Manager (RPM). Before you begin, we recommend that you read the Preparing your installation section, and verify that you meet the hardware and software requirements for your installation.

Note: Talend Support will investigate issues related to third-party components and databases if they are required for the Talend product to function, but Talend cannot provide patches on behalf of third-party components or databases.

Preparing your installation

Software packages

This page lists the software packages you need to download to install your Talend product.

For the software package file names in the tables below:

- YYYYMMDD_HHmm corresponds to the package timestamp.
- A.B.C. corresponds to package version number (major.minor.patch.).

Note: The software modules must be the same version on both the client and server side. When downloading software packages, make sure the timestamps and version numbers are the same.

Manual installation software packages

| File name | Description |
|--|---|
| Talend-Studio-YYYYMMDD_HHmm-VA.B.C.zip | Studio IDE (GUI) To download it, go to this page |

Community and Support

There are several ways to get help and support for your Talend installation:

- [Official Talend Documentation](#). Here you can find everything to help you install and use your Talend product.
- [Talend Community](#). This is the place where you can ask questions to the community, and get answers.

Hardware requirements

Before installing your Talend product, make sure the machines you are using meet the following hardware requirements recommended by Talend.

Memory and disk usage heavily depends on the size and nature of your Talend projects. However, in summary, if your Jobs include many transformation components, you should consider upgrading the total amount of memory allocated to your servers, based on the following recommendations.

Memory usage

| Product | Client/Server | Memory requirements (minimum-recommended) |
|---------------|---------------|--|
| Talend Studio | Client | 3GB – 4GB |

Note: Depending on the number of executed processes running on a module, you may need to increase the available memory. If you have several products installed on the same host, Talend recommends to use an i7 CPU with 8 logical processors.

Disk space requirements

| Product | Client or Server | Required disk space for installation | Required disk space for use |
|---------------|------------------|--------------------------------------|-----------------------------|
| Talend Studio | Client | 3GB | 3GB+ recommended |

1 For example, 5 million records = 10 GB required space on the disk. Talend recommends you double the required size to avoid problems during high transactions.

2 These requirements do not take the MongoDB metadata size into account.

3 Recommended for a campaign that counts 50,000 tasks, each task having 50 attributes.

ulimit settings on Unix systems

To improve Talend server modules and Unix system performance, you can configure the system resources (ulimit) according to the needs of the user or group. These settings are defined in the `/etc/security/limits` file.

ulimit syntax

```
ulimit <limit type> <item> <value>
```

There are two ulimit types: hard and soft.

- The soft limit is the effective resource limit. The user can increase the soft limit up to the value of the hard limit.
- The hard limit is the maximum resource limit. This value is set by the superuser and cannot be exceeded.

Note: If you do not specify a limit type, the hard limit type is used by default.

The following ulimit settings are important for your Talend deployment.

| Item | Description | Flag | Value |
|--------|-------------------------------------|------|---------|
| fsize | Maximum file size | -f | KB |
| nofile | Maximum number of open files | -n | |
| stack | Maximum stack size | -s | KB |
| cpu | Maximum CPU time | -t | minutes |
| nproc | Maximum number of processes/threads | -u | |

Note: You can list all available ulimit settings with the following command: `ulimit -a`

Example

```
ulimit -H -n 2000
```

This command sets a hard limit of 2000 open files per process.

For complete details on the ulimit settings, see the [SS64 reference guide for ulimit](#).

Software requirements

Compatible Operating Systems

This page details the recommended and supported Operating Systems for Talend products.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Talend Studio

Table 1: Compatible operating systems for Talend Studio

| Operating system family (64 bit) | Operating system | Version | Support type |
|----------------------------------|---------------------------------|-----------|--------------|
| Linux | Ubuntu | 18.04 LTS | Recommended |
| | Red Hat Enterprise Linux Server | 8 | Supported |
| | | 7 | Supported |

| Operating system family (64 bit) | Operating system | Version | Support type |
|----------------------------------|------------------|-------------------|--------------|
| | CentOS | 8 | Supported |
| | | 7 | Supported |
| Microsoft | Windows | 10 | Recommended |
| | Windows Server | 2019 | Supported |
| | | 2016 (RTM) | Supported |
| | | 2012 | Supported |
| Mac | Apple MacOS | Catalina 10.15 | Supported |
| | | Mojave 10.14 | Supported |
| | | High Sierra 10.13 | Supported |

Compatible Java Environments

The following tables provide information on the recommended Java Environment you should download and install to use your Talend product.

The Compiler Compliance Level corresponds to the Java version used for the Job code generation. This option can be changed in the Studio preferences. For more information, see the Talend Studio User Guide.

Note: All Talend products and associated third-party applications, such as the Hadoop cluster, should use the same Java version for compliance. Before you install or upgrade any associated third-party application, Talend recommends that you check which Java version they support.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Studio Java environments

Table 2: Compatible Java environments for Talend Studio

| Java platform | Java version ^{1,2,3} | Support type |
|---------------|-------------------------------|--------------|
| Open JDK | 11 | Recommended |
| | 8 | Supported |
| Oracle | 11 | Recommended |
| | 8 | Supported |

- 1 When running Oracle 8, Studio should conform to JDK Compiler Compliance Level 1.8 (default). If the installed JDK version is lower than 1.8.0_161, you need to also install the additional resource [Java Cryptography Extension \(JCE\) Unlimited Strength Jurisdiction Policy Files](#). This is also applicable to the execution of standalone Jobs.
- 2 When running Oracle 11, Studio should conform to JDK Compiler Compliance Level 11 (default).
- 3 The recommended distribution for Open JDK is [Zulu](#).

Compatible Apache software and JMS Brokers for Talend ESB

The following tables provide information on the compatible Apache software and JMS Brokers for Talend ESB.

Supported Apache software

| Software | More information |
|--------------------------------------|-------------------------------|
| Apache Karaf 4.2.7 ¹ | Release notes |
| Apache CXF 3.3.4 ¹ | Release notes |
| Apache Camel 2.24.2 ² | Release notes |
| Apache ActiveMQ 5.15.10 ¹ | Release notes |

1 Service release upgrade.

2 Minor release upgrade.

Supported Messaging Brokers for SOAP/JMS

| Software | More information |
|-------------------------|-------------------------------|
| Apache ActiveMQ 5.15.10 | Release notes |
| IBM WebSphere MQ 9.1 | |
| IBM WebSphere MQ 9.0 | |
| IBM WebSphere MQ 8.5 | |

Compatible web application servers

The following tables provide information on the recommended and supported Web application servers for the Talend server modules.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Compatible containers

The following tables provide information on the recommended and supported containers for the Talend server modules.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Talend ESB

| Runtime Containers | Version | Support type |
|-------------------------------|---------------------|--------------|
| Talend Runtime (Apache Karaf) | 7.3 ² | Recommended |
| Apache Tomcat | 9.0.30 ¹ | Recommended |
| | 9.0.30 ³ | Supported |

1 Recommended version for Talend Identity Management.

2 Not recommended for Talend Identity Management.

3 Only for CXF Services, Camel Routes, Service Activity Monitoring, Talend Identity Management and Security Token Service.

Compatible Web browsers

The following table provides information on the recommended and supported Web browsers you should use to take the most of your Talend products.

The minimum supported screen resolution is 1366 x 768 (px). Browser and system settings, such as scaling, zooming, and window size, will affect browser compatibility.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

| Web browser | Support type |
|--|--------------|
| Mozilla Firefox ESR 68 | Recommended |
| Mozilla Firefox up to latest available browser version | Supported |
| Microsoft Internet Explorer 11 or later | Supported |
| Microsoft Edge up to latest available browser version | Supported |

| Web browser | Support type |
|--|--------------|
| Apple Safari 12 or later | Supported |
| Google Chrome up to latest available browser version | Supported |

Note: Talend recommends that you ensure you are running the latest version of your browser. Talend supports the latest versions of the browsers listed above, unless a specific version is listed.

Compatible databases

The following tables provide information on the recommended and supported databases you can use with Talend server modules.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Compatible messaging systems

The following tables provide information on the recommended messaging systems you can use with Talend server modules.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

Compatible artifact repository

The following table provides information on the supported artifact repository you can use with Talend server modules.

In the following documentation:

- Recommended: designates an environment recommended by Talend based on our experiences and customer usage.
- Supported: designates a supported environment for use with the listed component or service.
- Supported with limitations: designates an environment that is supported by Talend but with certain conditions explained in notes.

| Artifact repository | Version | Support type |
|---------------------|-------------------|--------------|
| JFrog Artifactory | SaaS | Recommended |
| | 6.12 ¹ | Recommended |

| Artifact repository | Version | Support type |
|---------------------|--------------|--------------|
| Sonatype Nexus | 3.15 to 3.18 | Supported |

1 Latest at the date of release – September 23, 2019.

Compatible execution servers

Use the following table to ensure that your execution server version is compatible with Talend Administration Center, Talend CommandLine and Talend Studio versions.

Note: The information contained in this section is valid at the date of publication but may be subject to change at a later date.

Job Servers (Talend JobServer and Job server in Talend Runtime)

| Talend Administration Center, Talend CommandLine and Talend Studio version | Compatible Talend JobServer versions |
|--|---|
| 7.3.x | 6.4.x, 6.5.x, 7.0.x, 7.1.x, 7.2.x and 7.3.x |

Warning:

When activating SSL and token authentication in Talend Administration Center, only Job Servers that support SSL / token authentication can be used. Talend Administration Center is not able to monitor older Job Servers that do not support encrypted communication.

For more information, see *Configuring SSL transport and authentication*.

Installing the XULRunner package

On Linux, the XULRunner package is required to run the Studio. The XULRunner package version that is recommended is XULRunner v1.9.2.28.

The supported versions are v1.8.x - 1.9.x and v3.6.x.

Procedure

1. Download XULRunner v1.9.2.28 from [this location](#).
2. Unpack the archive file in the same directory where you unpacked the studio archive, but do not unpack it within the Studio folder.
3. Add the following line at the end of the Studio `.ini` file that corresponds to your Linux architecture:

```
-Dorg.eclipse.swt.browser.XULRunnerPath=</usr/lib/xulrunner>
```

where `</usr/lib/xulrunner>` is the XULRunner installation path.

Example

For example, if you have unpacked the Studio in a directory under your user home directory /home/<user>/Talend/, you need to add the following to the .ini file: -Dorg.eclipse.swt.browser.XULRunnerPath=/home/<user>/Talend/xulrunner/

Setting up JAVA_HOME

In order for your Talend product to use the Java environment installed on your machine, you must set the JAVA_HOME environment variable.

Procedure

1. Find the folder where Java is installed.

For example:

- /usr/lib/jvm/java-x-oracle
- /usr/lib/jvm/zulu-8/bin

2. Open a terminal.

3. Use the export command to set the JAVA_HOME and Path variables.

For example:

- ```
export JAVA_HOME=/usr/lib/jvm/jre1.8.0_65
export PATH=$JAVA_HOME/bin:$PATH
```

- ```
export JAVA_HOME=/usr/lib/jvm/<zulu_jdk>
export PATH=/<JAVA_HOME>/bin:$PATH
```

4. Add these lines at the end of the global profiles in the /etc/profile file or in the user profiles in the ~/.profile file.

After changing one of these files you have to log on again.

Installing your Talend Open Studio for Data Integration manually

Installing and configuring your Talend Studio

Unzip the archive

Procedure

1. Download your product from [this page](#).
2. Unzip it.

Editing the memory and JVM settings

To gain in performance at runtime and when launching Talend Studio, proceed as follows: you can edit the memory settings in the `.ini`.

Procedure

1. Edit the `TOS_DI-linux-gtk-x86_64.ini` file.
2. Edit the memory attributes. For example:

```
-vmargs -Xms512m -Xmx1536m -XX:MaxMetaspaceSize=512m
```

Tip: For big projects, you may need to increase `Xmx` to 4096m.

For more details, see <http://www.oracle.com/technetwork/java/hotspotfaq-138619.html>.

Launching your Talend Studio

Procedure

Double-click the `TOS_DI-linux-gtk-x86_64` executable to launch your Talend Studio.

You can also launch your Talend Studio from the terminal using the `TOS_DI-linux-gtk-x86.sh` file.

If needed, add the execution rights to it using the following command:

```
chmod +x TOS_DI-linux-gtk-x86.sh
```

Installing external modules

Talend Studio requires specific third-party Java libraries or database drivers to be installed to connect to sources and targets.

Those libraries or database drivers, known as external modules, may be required by some of Talend components or by some connection wizards or by both. Due to license restrictions, Talend may not be able to ship some of these external modules within Talend Studio. You need to install them for your Studio to functional properly.

Warning: Make sure that the `-Dtalend.disable.internet` parameter is not present in the Studio `.ini` file or is set to `false`.

When to install external modules

Your Talend Studio will let you know when you need to install external modules and what external modules you need to install.

Your Talend Studio notify you about required external modules in several ways.

- The **Additional Talend packages** wizard opens when you launch your Talend Studio if any additional packages, including external modules, need to be installed for any features to function in the Studio.

Tip: The **Additional Talend packages** wizard also opens when you select **Help > Install Additional Packages** from the Studio menu.

- On your design workspace, if a component requires the installation of external modules before it can work properly, a red error indicator appears on the component. With your mouse pointer over the error indicator, you can see a tooltip message showing which external modules are required for that component to work.
- When you open the **Basic settings** or **Advanced settings** view of a component for which one or more external modules are required, you will see a piece of highlighted information about external modules, followed by an **Install** button. Clicking the **Install** button opens a wizard that will show you the external modules to be installed.
- The **Modules** view lists all the modules required for the Studio to work properly, including those Java libraries and drivers that you must install.

If the **Modules** view is not shown under your design workspace, go to **Window > Show View... > Talend** and then select **Modules** from the list.

| Status | Context | Module | Maven URI |
|---------------|--------------------------|----------------|---|
| Not installed | tSugarCRMInput | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tVtigerCRMOutput | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | plugin:org.talend.lib... | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tWebService | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tPOP | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tSageX3Output | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tMicrosoftCrmOutput | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tFileInputMail | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tMicrosoftCrmInput | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |
| Not installed | tWebServiceOutput | activation.jar | mvn:org.talend.libraries/activation/6.0.0/jar |

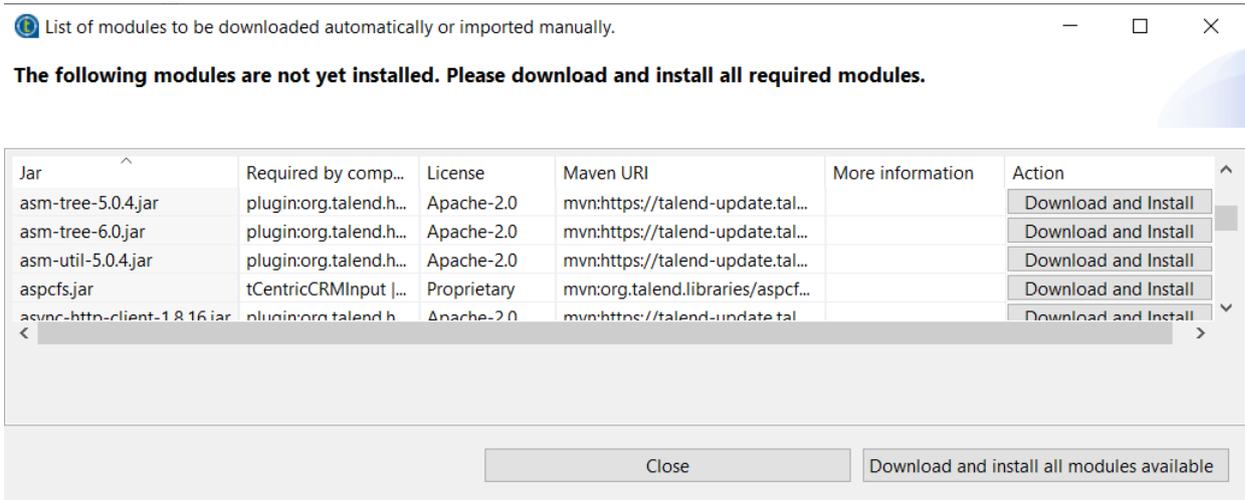
In this view:

| Item | Description |
|-------------------|--|
| Filter text field | Allows you to search external modules based on the status, the context, the module file name, and the Maven URI. |
| Status | Points out if a module is installed or not installed on your system. |

| Item | Description |
|---|--|
| | <p>The  icon indicates that the module is not necessarily required for the corresponding component or Metadata connection listed in this column.</p> <p>The  icon indicates that the module is absolutely required for the corresponding component or Metadata connection.</p> |
| Context | Gives the name of the component or Metadata connection using the module. If this column is empty, the module is then required for the general use of your Talend Studio. |
| Module | Gives the exact name of the module. |
| Maven URI | <p>Uniquely identifies the module deployment in Maven.</p> <p>You can customize the Maven URI of a module by clicking the Maven URI field and then clicking the [...] button that appears. For more information, see Customizing the Maven URI for external module deployment on page 16.</p> |
|  | <p>Refreshes this view to reflect the latest module installation status.</p> <p>In case of collaborative work, once a required module is installed in one user's studio, the other users can simply refresh their Modules view to add this module to their own Talend Studio.</p> |
|  | Imports custom Maven settings from a local file. |
|  | Exports custom Maven settings into a local file. |
|  | Allows you to install an already downloaded external module into your Studio. For details, see Installing external modules manually using the Modules view on page 19. |
|  | Opens the Jar download and installation wizard, which will list all the selected external modules that are not integrated in Talend Studio. |

- A Jar installation wizard appears when you:
 - drop a component from the **Palette** if one or more external modules required for that component to work are missing in the Studio.
 - click the **Check** button in a Metadata connection setup wizard in the Studio if one or more external modules required for the connection are missing in the Studio.
 - click the **Guess schema** button in the **Component** view of a component if one or more external modules required for that component to work are missing in the Studio.
 - click **Install** on the top of the **Basic settings** or **Advanced settings** view of a component for which one or more required external modules are missing.
 - run a Job that involves components or Metadata connections for which one or more required external modules are missing.

- select one or more modules that are not integrated in the Studio and click the  button in the **Modules** view.



This wizard:

- lists the external modules to be installed and the licenses under which they are provided
- provides the default Maven URIs that identify the deployment of the modules
- provides the official websites where you can learn more about the modules
- lets you download and install automatically all the modules available in the Talend repository
- allows you to install those not available in the Talend repository manually.

When you drop a component, set up a connection, or guess the schema of a database, that requires an external module for which neither the Jar file nor its download URL information is available on the Talend website, the Jar installation wizard does not appear, but the **Error Log** view will present an error message informing you that the download URL for that module is not available. You can try to find and download it by yourself, and then install it manually into the Studio.

Tip: To show the **Error Log** view on the tab system, go to **Window > Show views**, then expand the **General** node and select **Error Log**.

Customizing the Maven URI for external module deployment

In Talend Studio, each external module is given a default URI to identify its deployment in Maven. When needed, you can change the Maven URI.

For example, when replacing an installed database driver with a new version, you need to specify another Maven URI for it.

Note:

Changing the Maven URI for an external module will affect all the components and metadata connections that use that module within the project.

When working on a remote project, your custom Maven URI settings will be automatically synchronized to the Talend Artifact Repository and will be used when other users working on the same project install the external module.

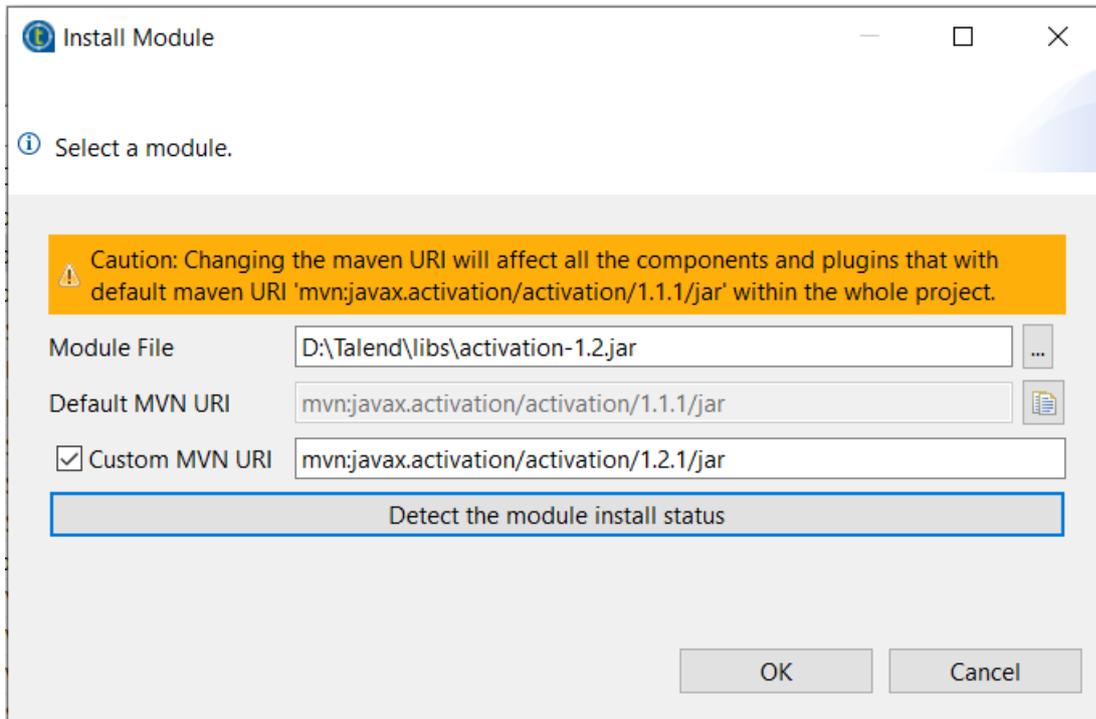
Procedure

1. In the **Modules** view, click the Maven URI you want to customize and then click the [...] button that appears.

The **Install Module** dialog box opens.

2. If you want to install another version of the external module, specify the full path to the module file in the **Module File** field, or click the [...] button to browse in your local file system.
3. Select the **Custom MVN URI** check box and enter a new URI in the field.

You can click the  button next to the Default MVN URI field to copy the default Maven URI, then paste it in the **Custom MVN URI** field and modify it.



4. Click the **Detect the module install status** button and then **OK** to validate the custom URI and close the dialog box.

Results

The new Maven URI takes effect and is displayed in the **Modules** view, from which can export all your Maven URI changes into a local JSON file.

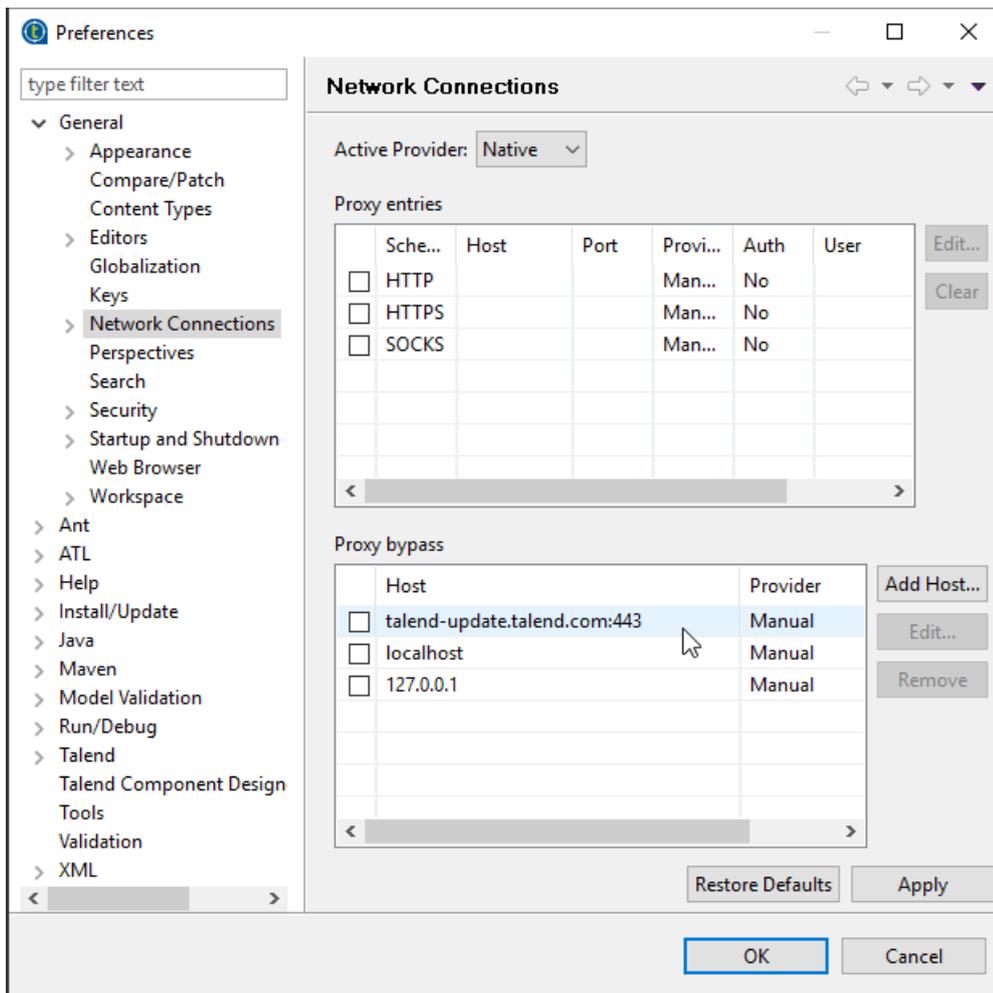
Installing external modules from within the Studio

You can download and automatically install most external modules using the wizard provided by your Talend Studio.

Before you begin

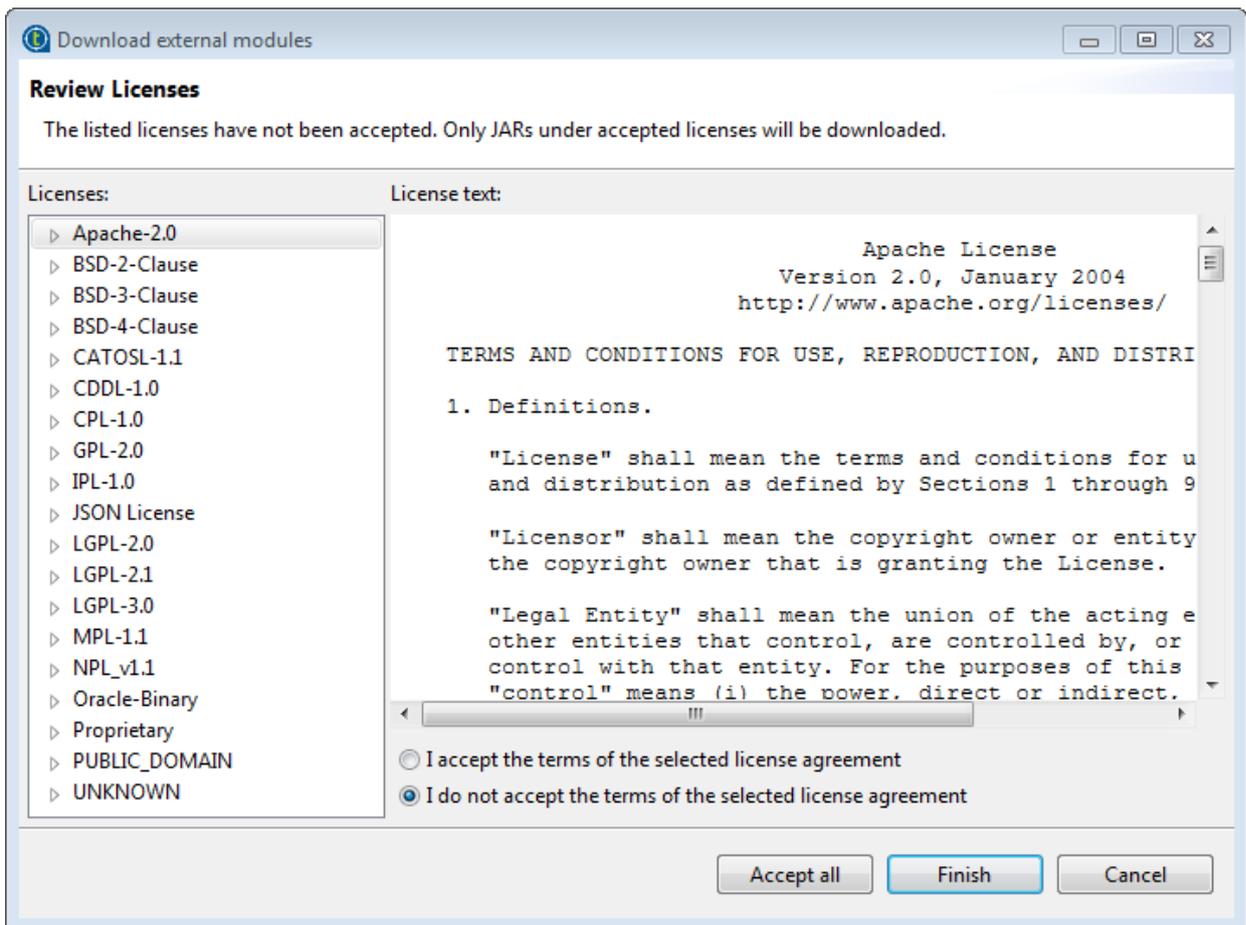
Make sure your Talend Studio has a secure Internet connection.

If you are working behind a network proxy, make sure you have correctly set up your proxy and add the web site <http://talend-update.talend.com> and the port 443 to your whitelist. To access the proxy settings, select **Window > Preferences** from the menu to open the **Preferences** window, then expand the **General** node and click **Network Connections**.



Procedure

1. Do the following to open the **Download external modules** dialog box:
 - In the **Additional Talend Packages** wizard, select the **Required third-party libraries** and/or **Optional third-party libraries** check boxes and click **Finish**.
 - In the Jar installation wizard, click the **Download and Install** button to install a particular module, or click the **Download and install all modules available** button to install all the available modules.



2. Accept the license terms and start the download and installation process:

- To download and install the external module(s) provided under a particular license, select that license from the **Licenses** pane, review the license terms, select the **I accept the terms of the license agreement** option, and click **Finish**.
- To download and install all external modules provided under all the listed licenses, click the **Accept all** button.

Results

When the installation process is completed, the chosen external module or modules are installed into your Talend Studio, and you can use Talend Studio features that depend on these modules.

Installing external modules manually using the Modules view

If you have already downloaded external modules, you can install them manually into your Talend Studio.

Before you begin

If you are going to install the JDBC driver for Oracle 9i into your Talend Studio, change the file name from `ojdbc14.jar` to `ojdbc14-9i.jar` first.

Procedure

1. Click the  button in the upper right corner of the **Modules** view or in the Jar installation wizard to browse your local file system.

2. In the **Open** dialog box of your file system, browse to the module you want to install, double-click the `.jar` file, or select it and then click **Open** to install it into your Talend Studio.

Results

The dialog box closes and the selected module is installed in the library folder of the current Talend Studio.

Installing external modules manually for Talend Web applications

Some modules required for a Talend Web application to work are not available on the Talend website but can be downloaded directly from external websites. Once downloaded, these modules must be placed in specific folders.

Procedure

- For the Talend MDM Server, place the downloaded JDBC drivers for the Oracle and MySQL databases in the following folder:

```
<TomcatPath>/webapps/talendmdm/WEB-INF/lib
```

- For Talend Administration Center, place the downloaded modules in the following folder:

```
<TomcatPath>/webapps/org.talend.administrator/WEB-INF/lib
```

Disabling Internet access for the Studio

About this task

You can disable Internet access for your Talend Studio by editing the Studio `.ini` file.

Warning: Do this only if you have no needs of accessing the Internet to download and install custom components, third-party libraries, and so on.

Procedure

1. Open the Studio `.ini` file corresponding to your operating system, and add the following line to it:

```
-Dtalend.disable.internet=true
```

2. Restart your Talend Studio.
When launched again, the Studio will not show:
 - The **Exchange** link on the toolbar
 - The **Talend > Exchange** node in the **Preferences** dialog
 - The options for installing third party libraries in the **Additional Packages** dialog box
 - The **Talend News** link on the welcome screen

Upgrading your Talend products

Backing up the environment

Before backing up and migrating the data of each Talend solution, you need to make sure your environment is correctly backed up.

The environment backup process includes the following mandatory steps:

Note: You need to complete these steps in the following order.

1. Saving the local projects

For details, refer to [Backing up the environment](#) in the Talend Migration Guide.

Saving the local projects

Procedure

1. Launch the Studio.
2. Click the  icon and export your local projects to an archive file.

Upgrading the Talend projects in Talend Studio

Importing your local projects

Procedure

1. Launch the new Talend Studio you have just installed.
2. In the login windows, select **Import** then import the archive file containing your local projects.

Results

The local projects are displayed in the **Project** list and appear on the Talend Studio **Repository** view.

For more information on how to export local projects to an archive file, see [Saving the local projects](#) on page 21.

Appendices

Supported Third-Party System/Database/Business Application Versions

This document provides the information about the versions of the systems or databases or business applications supported by Talend Studio.

Supported systems, databases and business applications by Talend components

The access to these systems, databases and business applications varies depending on the Studio you are using.

| Systems/Databases | Versions | OS |
|-------------------------------------|--------------------------------|-----------------|
| Access ¹ | 2003 | Windows |
| | 2007 | Windows |
| Amazon Aurora | MySQL edition v5 (5.6 and 5.7) | |
| Amazon RDS for Microsoft SQL Server | | |
| Amazon Redshift | 1.x | |
| AS/400 | V7R1 to V7R3 | |
| Bonita | 6.5.2 | |
| | 7.2.4 | |
| | 7.9.0 | |
| Cassandra | 3.0 to 3.4 | Windows + Linux |
| CouchBase | 5.x | Windows |
| | 6.0 | Windows |
| CouchDB | 1.0.2 | Windows |
| DB2 | 10.5 | |
| DB Generic | ODBC | Windows |
| DynamoDB | | |
| Elasticsearch | 5.6.x | |
| | 6.4.x | |
| EXASolution | 6.0 and earlier | Windows |

| Systems/Databases | Versions | OS |
|------------------------------------|-----------------------|-------------------------------|
| Excel | | |
| eXist-db | 1.4.0 | |
| FireBird | 2.1 to 3.0 | Windows + Linux |
| FTP | | |
| Greenplum | 4.3.x | Windows (client only) + Linux |
| | 5.x | Windows (client only) + Linux |
| HSQLDb | 1.8.0 to 2.4 | |
| IBM DB2 and IBM DB2 Z/OS | 10.5 | Windows + Linux |
| | 11.1 | Windows + Linux |
| Informix | 11.50 | Windows + Linux |
| Ingres | 10.2 | Windows + Linux |
| | 11 | Windows + Linux |
| Interbase | | |
| JavaDB | 6 | Windows + Linux |
| JDBC | | |
| JSON | | |
| Kafka ² | 0.8.2.0 | Windows + Linux |
| | 0.9.0.1 | Windows + Linux |
| | 0.10.0.1 | Windows + Linux |
| | 1.1.0 | Windows + Linux |
| | 2.2.1 | Windows + Linux |
| LDAP | No version limitation | Windows + Linux |
| MapRDB | | |
| MarkLogic | V9 | |
| MaxDB | 7.6 | |
| Microsoft Azure Blob Storage | | |
| Microsoft Azure SQL Data Warehouse | | |

| Systems/Databases | Versions | OS |
|-----------------------------------|------------------------|-----------------|
| Microsoft AX | Dynamics AX 4.0 | |
| | Dynamics AX 2012 | |
| Microsoft CRM | 2011 | |
| | 2015 | |
| | 2016 | |
| Microsoft CRM Online | 2011 | |
| | 2016 | |
| | 2018 | |
| Microsoft SQL Server ³ | 2014 to latest version | Windows + Linux |
| MongoDB | 3.6.x | Windows + Linux |
| | 4.0.x | Windows + Linux |
| | 4.2.x | |
| MySQL | MySQL 5.x | Windows + Linux |
| | MySQL 8.x | Windows + Linux |
| | MariaDB | Windows + Linux |
| | Amazon RDS | Windows + Linux |
| | Google Cloud SQL | Windows + Linux |
| MOM | | |
| Neo4j | 1.x.x | Linux |
| | 2.x.x / 2.2.x / 2.3 | Linux |
| | 3.2.x | Linux |
| | 3.5.x | Linux |
| Netezza | 7.0.x | Windows + Linux |
| | 7.1.x | Windows + Linux |
| | 7.2.x | Windows + Linux |
| NetSuite | 2018 | Windows + Linux |
| OleDb | 2000 | |
| | 2003 | |

| Systems/Databases | Versions | OS |
|-----------------------------|----------------------------|-----------------|
| | 2005 | |
| | 2007 | |
| | 2010 | |
| Oracle | Oracle 12c Release 1 | Windows + Linux |
| | Oracle 12c Release 2 | Windows + Linux |
| | Oracle 18c | Windows + Linux |
| | Oracle 19c | Windows + Linux |
| Palo | Open source version 5 | |
| ParAccel | 3.1 | |
| | 3.5 | |
| PostgreSQL | v7.2 to v8.x | Windows + Linux |
| | v9.x / v10.x / v11.x | Windows + Linux |
| | Amazon RDS | Windows + Linux |
| | Google Cloud SQL | Windows + Linux |
| PostgresPlus | v7.2 to v8.x | Windows + Linux |
| | v9.x | Windows + Linux |
| Red Hat BRMS | 6.1 | Windows + Linux |
| REST Service | | Windows + Linux |
| Salesforce | V46 and earlier | Windows + Linux |
| SAP | 4.6 | |
| SAP Business Suite (ERP) | Netweaver: From 7.3 to 7.5 | Windows |
| | ERP6.0, From EhP6 to EhP8 | Windows |
| SAP Business Warehouse (BW) | Netweaver: From 7.3 to 7.5 | Windows |
| SAP HANA ⁴ | No version limitations | Windows |
| SAS | 9.1 | Windows + Linux |
| | 9.2 | Windows + Linux |
| SOAP Service | | |
| SQLite | 3.6.7 | Windows + Linux |

| Systems/Databases | Versions | OS |
|-------------------|----------------|-----------------|
| SugarCRM | 5.2 | Windows + Linux |
| Sybase | 12.5 | Windows + Linux |
| | 12.7 | Windows + Linux |
| | 15.2 | Windows + Linux |
| | 15.5 | Windows + Linux |
| | 15.7 | Windows + Linux |
| | 16.0 | Windows + Linux |
| SybaseIQ | 12.5 | Windows + Linux |
| | 12.7 | Windows + Linux |
| | 15.2 | Windows + Linux |
| | 16.0 | Windows + Linux |
| Teradata | 12 to 16 | Windows + Linux |
| VectorWise | 2 | Windows + Linux |
| Vertica | 9.0.x to 9.3.1 | Windows + Linux |
| VtigerCRM | Vtiger 5.0 | |
| | Vtiger 5.1 | |

1 When working with Java 8, only the General collation mode is supported.

2 The Kerberos kinit option and the Kerberos keytab option are both supported. For information about the security options supported by the Kafka components, see [Talend Help Center](#).

3 Microsoft SQL Server support is provided through the Microsoft SQL JDBC driver. For more information, see the [Download Microsoft JDBC Driver for SQL Server](#) page.

4 Supported through SAP JDBC driver.

Messaging brokers supported by Talend messaging components

| Supported messaging brokers / standards | Component |
|---|--------------------|
| JMS standard 1.1 | tJMSInput |
| | tJMSOutput |
| MicrosoftMQ 3.0 | tMicrosoftMQInput |
| | tMicrosoftMQOutput |

| Supported messaging brokers / standards | Component |
|---|------------|
| JBoss Messaging 1.4.4 | tMomInput |
| | tMomOutput |
| WebSphere MQ 8.0 | tMomInput |
| | tMomOutput |
| ActiveMQ 5.13.2 | tMomInput |
| | tMomOutput |

Supported Big Data platforms

In general, Talend certifies a specific release version for a given Big Data (Hadoop) Distribution vendor. These are typically what is recommended for use for that vendor. For incremental upgrades and service packs by a given vendor, Talend relies on the vendors' compatibility statements to ensure the proper running and execution of the Talend software. Where compatibility is stated, Talend also supports that version under our Support SLA. If an incompatibility should be verified by the Hadoop vendor, then Talend considers that a re-test and upgrade may be necessary.

If the Hadoop distribution you want to use is not yet supported and available in your Talend Studio, it may be available through an update. You can search for support information on the [Talend Help Center](#).

For details, search for adding support for the latest Hadoop distribution on Talend Help Center

For more information about the versions of all the supported third-party systems/databases, see [Supported systems, databases and business applications by Talend components](#) on page 22.

Supported Big Data platform distribution versions for Talend Jobs

Regular Hadoop distributions

To find the compatibility between regular Hadoop distributions and Talend-supported Big Data platforms, click on a Big Data platform below.

- [HBase](#)
- [HCatalog](#)
- [HDFS](#)
- [Hive](#)
- [Sqoop](#)
- [Spark](#)
- [Azure Data Lake Storage Gen2](#)
- [Kafka in Spark Streaming Jobs](#)

Old versions of the supported Big Data platforms are being retired by their vendors. Talend ceases to support a version once this version reaches its date of end of support set by its vendor.

Talend and its community provide you with the convenience to keep using a version its vendor ceases to support in Talend products. For this reason, this version could be still listed in the following tables and available in the products but Talend stops providing support for this version.

Talend supports the minor versions of the distribution versions listed in the following tables.

Table 3: Supported Hadoop distributions with HBase

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|--------------------|------------------------------------|
| HDP | v2.5.0 | Yes |
| | v2.6.0 | Yes |
| | v2.6.0.3-8 | Yes |
| | v3.14.12-1 | Yes |
| Cloudera | 5.5 (YARN mode) | Yes |
| | 5.6 (YARN mode) | Yes |
| | 5.7 (YARN mode) | Yes |
| | 5.8 (YARN mode) | Yes |
| | 5.10 (YARN mode) | Yes |
| | 5.12.1 (YARN mode) | Yes |
| | 5.13.0 | Yes |
| | 6.1.1 | Yes |
| MapR ^{1,2} | 5.0.0 (YARN mode) | Yes |
| | 5.1.0 (YARN mode) | Yes |
| | 5.2.0 (YARN mode) | Yes |

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.

2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 4: Supported Hadoop distributions with HCatalog

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|-----------------|------------------------------------|
| HDP | v2.5.0 | Yes |
| | v2.6.0 | Yes |
| | v2.6.0.3-8 | Yes |
| | v3.14.12-1 | Yes |
| Cloudera | 5.5 (YARN mode) | Yes |
| | 5.6 (YARN mode) | Yes |
| | 5.7 (YARN mode) | Yes |

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|----------------------|---------------------------|------------------------------------|
| | 5.8 (YARN mode) | Yes |
| | 5.10 (YARN mode) | Yes |
| | 5.12.1 (YARN mode) | Yes |
| | 5.13.0 | Yes |
| | 6.1.1 | Yes |
| MapR ^{1, 2} | 5.0.0 (YARN mode) | Yes |
| | 5.1.0 (YARN mode) | Yes |
| | 5.2.0 (YARN mode) | Yes |
| | 6.0.0 (YARN mode) | Yes |
| | 6.0.1 MEP 5.0 (YARN mode) | Yes |
| | 6.1 / MEP 6.1 | Yes |

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.

2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 5: Supported Hadoop distributions with HDFS

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|--------------------|------------------------------------|
| HDP | v2.5.0 | Yes |
| | v2.6.0 | Yes |
| | v2.6.0.3-8 | Yes |
| | v3.14.12-1 | Yes |
| Cloudera | 5.5 (YARN mode) | Yes |
| | 5.6 (YARN mode) | Yes |
| | 5.7 (YARN mode) | Yes |
| | 5.8 (YARN mode) | Yes |
| | 5.10 (YARN mode) | Yes |
| | 5.12.1 (YARN mode) | Yes |
| | 5.13.0 | Yes |
| | 6.1.1 | Yes |

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|---------------------------|------------------------------------|
| MapR ^{1,2} | 5.0.0 (YARN mode) | Yes |
| | 5.1.0 (YARN mode) | Yes |
| | 5.2.0 (YARN mode) | Yes |
| | 6.0.0 (YARN mode) | Yes |
| | 6.0.1 MEP 5.0 (YARN mode) | Yes |
| | 6.1 / MEP 6.1 | Yes |

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.

2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 6: Supported Hadoop distributions with Hive

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|---------------------------|------------------------------------|
| HDP | v2.5.0 | Yes |
| | v2.6.0 | Yes |
| | v2.6.0.3-8 | Yes |
| | v3.14.12-1 | Yes |
| Cloudera | 5.5 (YARN mode) | Yes |
| | 5.6 (YARN mode) | Yes |
| | 5.7 (YARN mode) | Yes |
| | 5.8 (YARN mode) | Yes |
| | 5.10 (YARN mode) | Yes |
| | 5.12.1 (YARN mode) | Yes |
| | 5.13.0 | Yes |
| | 6.1.1 | Yes |
| MapR ^{1,2} | 5.0.0 (YARN mode) | Yes |
| | 5.1.0 (YARN mode) | Yes |
| | 5.2.0 (YARN mode) | Yes |
| | 6.0.0 (YARN mode) | Yes |
| | 6.0.1 MEP 5.0 (YARN mode) | Yes |

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|---------------|------------------------------------|
| | 6.1 / MEP 6.1 | Yes |

Note: The Profiling perspective does not support the Embedded connection mode on Hive distributions. This mode is available mainly for test purposes done by Hadoop developers. The studio may not be able to run correctly with the embedded mode.

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.
2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 7: Supported Hadoop distributions with Sqoop

| Hadoop distribution | Version | Supports Kerberos Kinit and Keytab |
|---------------------|---------------------------|------------------------------------|
| HDP | v2.5.0 | Yes |
| | v2.6.0 | Yes |
| | v2.6.0.3-8 | Yes |
| | v3.1.4.12-1 | Yes |
| Cloudera | 5.5 (YARN mode) | Yes |
| | 5.6 (YARN mode) | Yes |
| | 5.7 (YARN mode) | Yes |
| | 5.8 (YARN mode) | Yes |
| | 5.10 (YARN mode) | Yes |
| | 5.12.1 (YARN mode) | Yes |
| | 5.13.0 | Yes |
| | 6.1.1 | Yes |
| MapR ^{1,2} | 5.0.0 (YARN mode) | Yes |
| | 5.1.0 (YARN mode) | Yes |
| | 5.2.0 (YARN mode) | Yes |
| | 6.0.0 (YARN mode) | Yes |
| | 6.0.1 MEP 5.0 (YARN mode) | Yes |
| | 6.1 / MEP 6.1 | Yes |

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.

2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 8: Supported Hadoop distributions with Spark

| Hadoop distribution | Version | Works with Spark Stand-alone | Works with Spark YARN | Supports Kerberos Kinit and Keytab |
|---------------------|---------------------------|------------------------------|--------------------------|------------------------------------|
| HDP | v2.5.0 | | v1.6 (Deprecated) | Yes (YARN only) |
| | v2.6.0 | | v1.6 (Deprecated) / v2.1 | Yes (YARN only) |
| | v2.6.0.3-8 | | v1.6 (Deprecated) / v2.1 | Yes (YARN only) |
| | v3.1.4.12-1 | | v2.3 | Yes (YARN only) |
| Cloudera | 5.5 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) |
| | 5.6 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) |
| | 5.7 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) | Yes (YARN only) |
| | 5.8 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) / v2.1 | Yes (YARN only) |
| | 5.10 (YARN mode) | v1.6 (Deprecated) / 2.1 | v1.6 (Deprecated) / v2.1 | Yes (YARN only) |
| | 5.12.1 (YARN mode) | v2.2 | v2.2 | Yes (YARN only) |
| | 5.13.0 | v2.2 | v2.2 | Yes (YARN only) |
| | 6.1.1 | v2.4 | v2.4 | Yes (YARN only) |
| MapR ^{1,2} | 5.0.0 (YARN mode) | v1.3 (Deprecated) | v1.3 (Deprecated) | Yes (YARN only) |
| | 5.1.0 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) |
| | 5.2.0 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) | Yes (YARN only) |
| | 6.0.0 (YARN mode) | v2.1 | v2.1 | Yes (YARN only) |
| | 6.0.1 MEP 5.0 (YARN mode) | v2.2 | v2.2 | Yes (YARN only) |
| | 6.1 / MEP 6.1 | v2.3 | v2.3 | Yes (YARN only) |

1. For all MapR versions prior to v6.0, the MapR security ticket mechanism is supported by the Studio.

2. In MapR v6.0, HBase support has been removed in favor of MapR-DB. For more information, see https://maprdocs.mapr.com/home/ReleaseNotes/install_upgrade_notes.html.

Table 9: Supported Hadoop distributions with Azure Data Lake Storage Gen2 (ADLS Gen2)

| Hadoop distribution | Version | Works with Spark Stand-alone | Works with Spark YARN | Supports Kerberos Kinit and Keytab |
|---------------------|-------------|------------------------------|-----------------------|------------------------------------|
| HDP | v3.1.4.12-1 | | v2.3 | Yes (YARN only) |
| Cloudera | 6.1.1 | v2.4 | v2.4 | Yes (YARN only) |

If you need information about the Big Data Cloud platforms supported by Talend with ADLS Gen2, see the following section called *Supported Cloud Big Data platform distribution versions for Talend Jobs* of your installation guide.

Table 10: Supported Hadoop distributions with Kafka in Spark Streaming Jobs

| Hadoop distribution | Version | Works with Spark Stand-alone | Works with Spark YARN | Supports Kerberos Kinit and Keytab | Kafka versions |
|---------------------|--------------------|------------------------------|--------------------------|------------------------------------|----------------|
| HDP | v2.5.0 | | v1.6 (Deprecated) | Yes (YARN only) | v0.10 |
| | v2.6.0 | | v1.6 (Deprecated) / v2.1 | Yes (YARN only) | v0.10 |
| | v2.6.0.3-8 | | v1.6 (Deprecated) / v2.1 | Yes (YARN only) | v1.x |
| | v3.1.4.12-1 | | v2.3 | Yes (YARN only) | v2.x |
| Cloudera | 5.5 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) | v0.10 |
| | 5.6 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) | v0.10 |
| | 5.7 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) | Yes (YARN only) | v0.10 |
| | 5.8 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) / v2.1 | Yes (YARN only) | v0.10 |
| | 5.10 (YARN mode) | v1.6 (Deprecated) / 2.1 | v1.6 (Deprecated) / v2.1 | Yes (YARN only) | v0.10 |
| | 5.12.1 (YARN mode) | v2.2 | v2.2 | Yes (YARN only) | v0.10 |
| | 5.13.0 | v2.2 | v2.2 | Yes (YARN only) | v0.10 |
| | 6.1.1 | v2.4 | v2.4 | Yes (YARN only) | v2.x |
| MapR ^{1,2} | 5.0.0 (YARN mode) | v1.3 (Deprecated) | v1.3 (Deprecated) | Yes (YARN only) | v0.10 |
| | 5.1.0 (YARN mode) | v1.5 (Deprecated) | v1.5 (Deprecated) | Yes (YARN only) | v0.10 |

| Hadoop distribution | Version | Works with Spark Stand-alone | Works with Spark YARN | Supports Kerberos Kinit and Keytab | Kafka versions |
|---------------------|---------------------------|------------------------------|-----------------------|------------------------------------|----------------|
| | 5.2.0 (YARN mode) | v1.6 (Deprecated) | v1.6 (Deprecated) | Yes (YARN only) | v0.10 |
| | 6.0.0 (YARN mode) | v2.1 | v2.1 | Yes (YARN only) | v0.10 |
| | 6.0.1 MEP 5.0 (YARN mode) | v2.2 | v2.2 | Yes (YARN only) | v0.10 |
| | 6.1 / MEP 6.1 | v2.3 | v2.3 | Yes (YARN only) | v1.x |

Supported Cloud Big Data platform distribution versions for Talend Jobs

Cloud Hadoop distributions

Talend supports the following cloud platforms for Big Data. Click your cloud platform to see the Big data support information.

- [Amazon EMR](#)
- [Microsoft HDInsight](#)
- [Google Dataproc](#)
- [Databricks on AWS](#)
- [Databricks on Azure](#)
- [Cloudera on AWS](#)
- [Cloudera on Azure](#)
- [Quoble on AWS](#)

Old versions of the supported Big Data platforms are being retired by their vendors. Talend ceases to support a version once this version reaches its date of end of support set by its vendor.

Talend and its community provide you with the convenience to keep using a version its vendor ceases to support in Talend products. For this reason, this version could be still listed in the following tables and available in the products but Talend stops providing support for this version.

Talend supports the minor versions of the platform versions listed in the following tables.

Table 11: Amazon EMR

| Amazon EMR version | Supported frameworks | Supported Hadoop elements in Spark batch | Supported Hadoop elements in Spark streaming | Supported Hadoop elements in Standard |
|-----------------------|-------------------------------------|--|--|--|
| v4.5.0 (Apache 2.7.2) | Standard Spark v1.6 (Deprecated) | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |

| Amazon EMR version | Supported frameworks | Supported Hadoop elements in Spark batch | Supported Hadoop elements in Spark streaming | Supported Hadoop elements in Standard |
|---------------------------------------|--|--|--|--|
| v4.6.0 (Apache 2.7.2) | Standard Spark v1.6 (Deprecated) | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |
| v5.0.0 (Apache 2.7.2) | Standard Spark v2.0 | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |
| v5.5.0 (Apache 2.7.2) | Standard Spark v2.1 | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |
| v5.8.0 (Apache 2.7.2) | Standard Spark v2.2 | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |
| v5.15.0 (Hadoop 2.8.3) (Apache 2.7.2) | Standard Spark v2.3 | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive | HBase HDFS HCatalog Hive Sqoop |

The supported Amazon EMR versions for the tAmazonEMRManage component are 4.5.0, 4.6.0, 4.9.2, 5.11.0 and 5.15.0.

Table 12: Google Dataproc for Big Data

| Google Dataproc version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|-------------------------|------------------------|---|--|--------------------------------|
| v1.1 | Standard Spark v2.0 | Hive BigQuery Google Storage Avro Delimited Parquet Positional XML JSON | Hive BigQuery Google Storage Avro Delimited Parquet Positional XML JSON Google PubSub | Hive |
| v1.4 | Standard Spark v2.4 | Hive BigQuery Google Storage Avro Delimited Parquet Positional XML JSON | Hive BigQuery Google Storage Avro Delimited Parquet Positional XML JSON Google PubSub | Hive |

Table 13: Databricks on Azure for Big Data

| Databricks on Azure version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|-----------------------------|------------------------|--|--|--------------------------------|
| 3.5 LTS | Standard Spark v2.2 | Hive Azure Blob Storage ADLS Gen1 | Hive Azure Blob Storage ADLS Gen1 | DBFS |
| 5.5 LTS | Standard Spark v2.4 | Hive Azure Blob Storage ADLS Gen1 ADLS Gen2 Snowflake DeltaLake MongoDB TDM components as technical preview tDataprepRun | Hive Azure Blob Storage ADLS Gen1 ADLS Gen2 Snowflake DeltaLake MongoDB TDM components as technical preview tDataprepRun | DBFS |

Table 14: Databricks on AWS for Big Data

| Databricks on AWS version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|---------------------------|------------------------|---|--|--------------------------------|
| 3.5 LTS | Standard Spark v2.2 | Hive S3 DynamoDB | Hive S3 DynamoDB Kinesis | DBFS |
| 5.5 LTS | Standard Spark v2.4 | Hive S3 DynamoDB Snowflake MongoDB TDM components as technical preview tDataprepRun | Hive S3 DynamoDB Kinesis Snowflake MongoDB TDM components as technical preview tDataprepRun | DBFS |

Table 15: Qubole on AWS for Big Data

| Qubole version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|----------------|------------------------|-----------------------------------|---------------------------------------|--------------------------------|
| Qubole Spark 2 | Standard Spark v2.2 | Redshift S3 DynamoDB | Redshift S3 DynamoDB Kinesis | S3 Hive |

Table 16: Cloudera Altus on Azure for Big Data

| Cloudera version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|------------------|----------------------|---|---|--------------------------------|
| CDH5.11 | Spark v2.1 | ADLS Gen1 Azure Blob Storage HDFS | ADLS Gen1 Azure Blob Storage HDFS | |

Table 17: Cloudera Altus on AWS for Big Data

| Cloudera version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|------------------|----------------------|-----------------------------------|---------------------------------------|--------------------------------|
| CDH5.11 | Spark v2.1 | S3 Redshift DynamoDB | S3 Kinesis Redshift DynamoDB | |

Table 18: Microsoft HD Insight for Big Data

| Microsoft HD Insight version | Supported frameworks | Supported elements in Spark batch | Supported elements in Spark streaming | Supported elements in Standard |
|------------------------------|-------------------------|-----------------------------------|---------------------------------------|--------------------------------|
| 3.4 | Spark v1.6 (deprecated) | Hive | Hive | |
| 3.6 | Spark v2.1 | Hive | Hive | |

Supported Cloudera Navigator versions for Talend Jobs

The support for Cloudera Navigator is available to the Spark Jobs you are creating in the Studio, which means you must be using a subscription-based Talend Big Data solution.

Cloudera Navigator uses a Cloudera [SDK library](#) to provide functionalities and must be compatible with the version of this SDK library. The version of your Cloudera Navigator is determined by the Cloudera Manager installed with your Cloudera distribution and the compatible SDK is automatically used based on the version of your Navigator.

However, not all the Cloudera Navigator versions have their compatible SDK versions. For more details about the Cloudera SDK versions and their compatible Navigator versions, see the Cloudera documentation about [Cloudera Navigator SDK Version Compatibility](#).

In the following documentation:

- supported: Talend went through a complete QA validation process.
- compatible: Talend did not go through a complete QA validation process but the feature should work as part of Cloudera backward compatibility on Cloudera V5.X branches.

| Studio version | Cloudera Navigator version | Related Cloudera version | Support type |
|----------------|----------------------------|--------------------------|--------------|
| 7.3 | 6.1.1 | 6.1.1 | Supported |
| | 2.4 | 5.5 to 5.8 | Supported |
| | 2.12.0 | 5.11 to 5.14 | Supported |
| | 2.5 to 2.7 | 5.5 to 5.8 | Compatible |
| | 2.9.3 to 2.9.x | 5.11 to 5.14 | Compatible |
| | 2.10.3 to 2.10.x | 5.11 to 5.14 | Compatible |
| | 2.11.2 to 2.11.x | 5.11 to 5.14 | Compatible |
| | 2.12.1 to 2.12.x | 5.11 to 5.14 | Compatible |

Supported databases for profiling data

The table below lists the databases supported from the Profiling perspective of Talend Studio. For a complete list about supported third-party systems, see [Supported systems, databases and business applications by Talend components](#) on page 22.

| Database name | Database version |
|---------------------------------------|--|
| Amazon Aurora | Amazon RDS for Aurora |
| Amazon Redshift | Initial release of Amazon Redshift |
| AS/400 | V7R1 to V7R3 |
| | V6R1 to V7R2 |
| Hive | See, Supported Hive distributions for profiling data on page 41. |
| IBM DB2 and IBM DB2 Z/OS ¹ | 11.1 |
| | 10.5 |
| Impala (a sub-module of Cloudera) | CDH5.1 and above |
| | MapR 6.1.0 (YARN mode) |
| | MapR 6.0.1 (YARN mode) |
| | MapR 6.0.0 (YARN mode) |
| | MapR 5.2.0 (YARN mode) |
| | MapR 5.1.0 (YARN mode) |
| Informix | 11.50 |
| Ingres | 10.2 |
| Microsoft SQL Server | Amazon RDS for SQL Server |
| | Azure SQL Database |
| | 2017 |
| | 2016 |
| | 2014 |
| MySQL | Amazon RDS for MySQL |
| | Amazon RDS for MariaDB |
| | Azure Database for MySQL |
| | MySQL 8.0 |
| | MySQL 5.1/5.5/5.6 |
| | MariaDB |
| Netezza | 7.2 |
| | 6 |

| Database name | Database version |
|--------------------------|-------------------------------|
| Oracle with SID | Amazon RDS for Oracle |
| | Oracle 19c |
| | Oracle 18c |
| | Oracle 12c Release 1 |
| Oracle with service name | Amazon RDS for Oracle |
| | Oracle 19c |
| | Oracle 18c |
| | Oracle 12c Release 1 |
| PostgreSQL | Amazon RDS for PostgreSQL |
| | Azure Database for PostgreSQL |
| | 12.1 |
| | 10 |
| | 9.1+ |
| | 8.3 |
| SQLite | 3.6.7 |
| Sybase (ASE and IQ)QLite | 16.0 |
| | 15.7 |
| | 15.2 |
| | 12.7 |
| | 12.5 |
| Teradata | 16 |
| | 15 |
| | 14 |
| | 13 |
| | 12 |
| Vertica | 9.x |

1 Binary large objects (BLOBs) are not supported.

Supported Hive distributions for profiling data

The table below shows the compatibility between big data distributions and the HiveServer.

Note: The Hive embedded mode is available for test purposes for Hadoop developers. When in embedded mode, the studio may not run correctly.

| Big data distribution | | HiveServer 1 | HiveServer2 |
|-----------------------|---------------------------|--------------------------------------|--------------------------------------|
| HortonWorks | HDP 1.0.0 (deprecated) | Embedded and Standalone | |
| | HDP 1.2 | Embedded and Standalone | Embedded and Standalone |
| | HDP 1.3 | Embedded and Standalone | Embedded and Standalone |
| | HDP 2.0 | Embedded (Linux only) and Standalone | Embedded (Linux only) and Standalone |
| | HDP 2.1 | Embedded (Linux only) and Standalone | Embedded (Linux only) and Standalone |
| | HDP 2.2 | Embedded (Linux only) and Standalone | Embedded (Linux only) and Standalone |
| | HDP 2.3.2 | | Standalone |
| | HDP 2.4.0 | | Standalone |
| | HDP 2.5.0 | | Standalone |
| | HDP 2.6.0 | | Standalone |
| Cloudera ¹ | CDH4 | Embedded and Standalone | Embedded and Standalone |
| | CDH5 | Embedded and Standalone | Embedded and Standalone |
| | CDH5.1 MR1 | | Standalone |
| | CDH5.4 YARN | | Standalone |
| | CDH5.5 (and above) YARN | | Standalone |
| MapR | MapR 1.2 (deprecated) | Standalone | |
| | MapR 2.0 | Embedded and Standalone | |
| | MapR 2.1.2 | Embedded and Standalone | |
| | MapR 3.0.1 | Embedded and Standalone | Embedded and Standalone |
| | MapR 3.1.0 | Embedded and Standalone | Embedded and Standalone |
| | MapR 4.0.1 YARN | Embedded and Standalone | Embedded and Standalone |
| | MapR 5.0 (and above) YARN | | Standalone |
| Apache | Apache 1.0.0 (Hive 0.9.0) | Embedded and Standalone | |

| Big data distribution | | HiveServer 1 | HiveServer2 |
|------------------------------|-----------------------------|--------------------------------------|---|
| | Apache 0.20.23 (Hive 0.7.1) | Standalone | |
| Pivotal HD | Pivotal HD 1.0.1 | Standalone | |
| | Pivotal HD 2.0 (deprecated) | Embedded (Linux only) and Standalone | Embedded (Linux only) and Standalone (Linux only) |

1 Kerberos authentication is supported.