talend | Connecting the Data-Driven Enterprise

Talend Open Studio for Big Data

Installation and Upgrade Guide for Linux

6.4.1

Adapted for v6.4.1. Supersedes previous releases.

Publication date: June 29, 2017

Copyleft

Notices

License Agreement

Table of Contents

Preface

1. General information

1.1. Purpose

This Installation Guide explains how to install configure and upgrade the *Talend* modules and related applications. For detailed explanation on how to use and fine-tune the *Talend* applications, please refer to the appropriate Administrator or User Guides of the *Talend* solutions.

Information presented in this document applies to *Talend* products **6.4.1**.

1.2. Audience

This guide is for administrators and users of the *Talend* products.

The layout of GUI screens provided in this document may vary slightly from your actual GUI.

1.3. Typographical conventions

This guide uses the following typographical conventions:

- text in **bold:** window and dialog box buttons and fields, keyboard keys, menus, and menu options,
- text in **[bold]:** window, wizard, and dialog box titles,
- text in `courier`: system parameters typed in by the user,
- text in *italics*: file, schema, column, row, and variable names,
- The  icon indicates an item that provides additional information about an important point. It is also used to add comments related to a table or a figure,
- The  icon indicates a message that gives information about the execution requirements or recommendation type. It is also used to refer to situations or information the end-user needs to be aware of or pay special attention to.
- `Any command is highlighted with a grey background or code typeface.`

2. Feedback and Support

Your feedback is valuable. Do not hesitate to give your input, make suggestions or requests regarding this documentation or product and find support from the **Talend** team, on **Talend Community** at:

https://community.talend.com/

Chapter 1. Before installing your Talend product

These pages present and list everything you need to know before installing your *Talend* product:

- *Preparing your installation*
- *Hardware requirements*
- *Software requirements*

1.1. Preparing your installation

These pages provide information about:

- *Files to download*
- *Community and Support*

1.1.1. Files to download

Here are the files you need to download to install your *Talend* product:

- the software packages. For more information, see *Software package*.

1.1.1.1. Software package

This page details the software package you need to download to install your *Talend* product.

In this page:

- YYYYMMDD_HHmm corresponds to the package timestamp
- A.B.C. corresponds to package version number (Major. Minor. Patch.)

Table 1.1. Manual installation software package

Zip/jar file name	Description
Talend-Studio-YYYYMMDD_HHmm-**V**A.B.C**.zip**	Studio IDE (GUI)

To download it, go to this page.

1.1.2. 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.

1.2. Hardware requirements

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

Memory 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.

Table 1.2. Memory usage

Product	Client/Server	Recommended alloc. memory
Studio	Client	3 GB minimum, 4 GB recommended

The same requirements also apply for disk usage. It also depends on your projects but can be summarized as:

Table 1.3. Disk usage

Product	Client/Server	Required disk space for installation	Required disk space for use
Studio	Client	3 GB	3+ GB

Ulimit settings on Unix systems

- To make the most out of the *Talend* server modules and improve performance on Unix systems, you should set the limit of system resources (ulimit) to unlimited.

1.3. Software requirements

These pages contain the exhaustive list of the databases and third party software that are compatible and supported with the 6.4.1 version of your *Talend* product.

- *Compatible Operating Systems*
- *Java*
- *Installing the XULRunner package*

1.3.1. Compatible Operating Systems

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.

The information contained in the following table is applicable for the 6.4.1 version of your Talend product at the time of its release. For updated information on the latest supported versions of the third-party systems, see the online version of this page on Talend Help Center.

These tables provide a summary of the supported Operating Systems.

Table 1.4. Talend Studio

Support type	Operating System (64-bit)	
Recommended	Linux	Ubuntu 16.04 LTS
	Windows	Microsoft Windows Professional 7
Supported	Linux	Ubuntu 17.04
		Ubuntu 14.04 LTS
		Red Hat Enterprise Linux Server/CentOS 7.3

Support type	Operating System (64-bit)	
		Red Hat Enterprise Linux Server/CentOS 7.2
		Red Hat Enterprise Linux Server/CentOS 7.1
		Red Hat Enterprise Linux Server/CentOS 6.8
		Red Hat Enterprise Linux Server/CentOS 6.7
	Windows	Microsoft Windows 10
		Microsoft Windows 8.1
		Microsoft Windows Server 2016 RTM
		Microsoft Windows Server 2012 RTM
	Mac	OS X 10.12 Sierra
		OS X 10.11 El Capitan
		OS X 10.10 Yosemite

1.3.1.1. Statement regarding Virtualization and Docker deployments

In general, *Talend* supports deployment on virtual machines. For Virtualization Systems, *Talend* relies on the vendors' operating-system compatibility statements.

Talend does not deliver prepackaged Docker Images for the *Talend* Servers, and cannot maintain a standard setup for customer-based Docker environments, so standard Service Level Agreements do not apply.

For any customer issue which also can be reproduced in a non-Docker environment on a supported platform, Support Service Level can be applied as usual. For any issue which only occurs in a customer-composed Docker environment, *Talend* will only provide best effort to address any issues that arise.

1.3.2. Java

In order to use your *Talend* product, Java must be installed on your machine. If you install your *Talend Studio* using the *Talend Studio Installer*, you do not need to set up a Java Environment as it is embedded in the Installer.

These pages list:

- *Compatible Java environments*
- *Setting up JAVA_HOME*

1.3.2.1. Compatible Java environments

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.

The information contained in the following table is applicable for the 6.4.1 version of your Talend product at the time of its release. For updated information on the latest supported versions of the third-party systems, see the online version of this page on Talend Help Center.

These tables provide a summary of the supported Java Runtime environments.

In this table:

- ✅ **(R)** means that this combination is recommended;
- ✅ means that this combination is supported;
- ❌ means that this combination is not supported.

Note that only the 64-bit versions of the compatible Java Runtime environments are supported.

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

Table 1.5. Studio Java environments

Support Type	JRE Version	Studio JDK Compiler Compliance Level	Notes
Recommended	Oracle 8	1.8 (default)	
Supported	Oracle 8	1.7	Needs to be switched to manually. Only supported for Big Data Distributions requiring it. Routes are not supported with JDK Compiler Compliance level 1.7.

Depending on the license you have, the available Execution Servers may differ.

For example, the recommended combination is:

- Oracle 8 installed on the machine running the Studio;
- The **Compiler Compliance Level** set to 1.7 in the project settings of the Studio;
- Oracle 8 installed on the machine(s) running the Execution Server(s) and the Talend Server Application(s);
- Big Data Distributions compatible with Java 1.8 used.

For more information on Java specificities (version, Operating Systems compatibility), see Talend Help Center and Talend Community.

1.3.2.2. 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.

To do so, proceed as follows:

1. Find the folder where Java is installed, usually */usr/lib/jvm/java-x-oracle*.
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
```

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.

1.3.3. 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.

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.

   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/
   ```

Chapter 2. Installing Talend Studio for the first time

We strongly encourage you to read the *Before installing your Talend product* before starting this chapter.

This chapter details the procedures required to install Talend Studio.

2.1. Downloading and installing Talend Studio

Download

1. Download the product on this page.

   Note that the *.zip* file contains binaries for ALL platforms (Linux/Unix, Windows and MacOS).

2. Once the download is complete, extract the archive file on your hard drive.

   It is recommended to avoid spaces and long names in the target installation directory path.

Configure the memory settings

- If you want to tune the memory allocation for your JVM, you only need to edit the *TOS_BD-linux-gtk-x86_64.ini* file.

  The default values are:

  ```
  -vmargs -Xms40m -Xmx500m -XX:MaxMetaspaceSize=128m
  ```

  If you only have 512MB of memory on your computer, you can specify the memory allocation as following, for example:

  ```
  -vmargs -Xms40m -Xmx256m -XX:MaxMetaspaceSize=64m
  ```

  Learn more on http://www.oracle.com/technetwork/java/hotspotfaq-138619.html

2.2. Launching your Talend Studio

The following procedures describe how to launch the Studio, how to log in the Studio and how to create your first project.

2.2.1. Launching your Studio

Launch the Studio

1. If you want to use the *.sh* script file to launch your **Talend Studio**, add the execution rights to that file using the following command:

   ```
   $ chmod +x TOS_BD-linux-gtk-x86.sh
   $ ./TOS_BD-linux-gtk-x86.sh
   ```

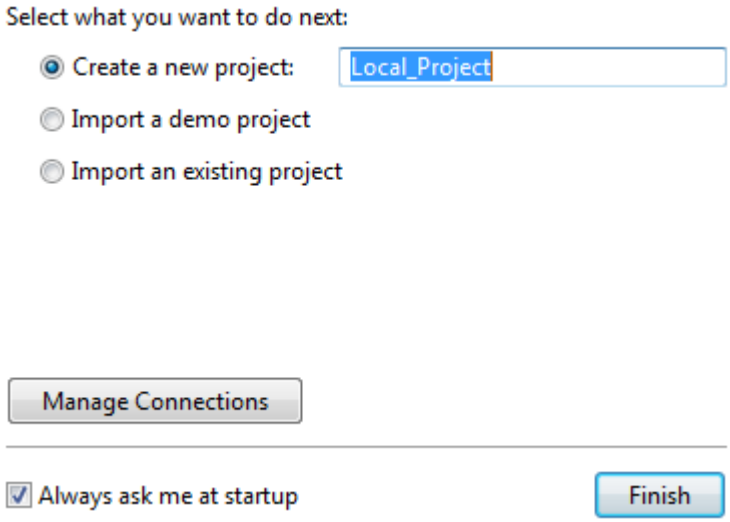
2. Double-click the *TOS_BD-linux-gtk-x86_64* executable file to launch your *Talend Studio* or use the TOS_BD-*linux-gtk-x86_64.sh* file.

Public license

- First screen is a license screen. In the **[License]** window that appears, read and accept the terms of the license agreement to proceed to the next step.

Login and first project

1. As first time user, you need to set up a new project or you can also import a Demo project which gathers numerous job samples.



2. To import a demo project, select **Import a demo project** and click **Finish**. In the dialog box that opens, select the project you want to import.

   To create a new project, select **Create a new project** and enter the name of your project in the corresponding field.

3. Click **Finish** when complete, to open a welcome window and launch the Studio.

2.3. Installing external modules

Talend Studio requires specific third-party Java libraries or database drivers (*.jar* files) to be installed to connect to sources and targets. Those libraries or drivers, known as external modules, can be required by some of *Talend* components. Due to license restrictions, *Talend* may not be able to ship certain external modules within *Talend Studio*.

2.3.1. Identify required external modules

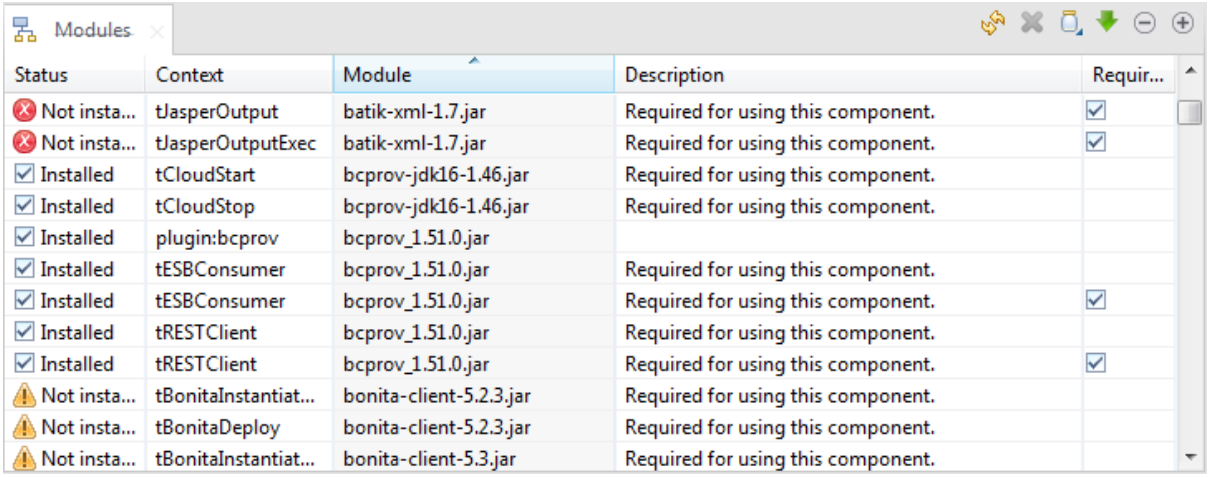
When you launch *Talend Studio* or select **Help** > **Install Additional Packages** in the *Talend Studio* menu, if any external modules are found missing for any features, the **[Additional Talend packages]** wizard opens, showing the **Optional** and **Required third-party libraries** check boxes. Make sure these check boxes are selected and click **Finish** to open the **[Download external modules]** dialog box, which lists all the available external modules, displays the license terms under which the external modules are provided, and lets you install all the modules at a single click. For more information, see *Install external modules*.

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 to use the components embedded in the Studio, including those Java libraries and drivers that you must install to get the relevant components working.

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



The table below describes the information presented in the **Modules** view.

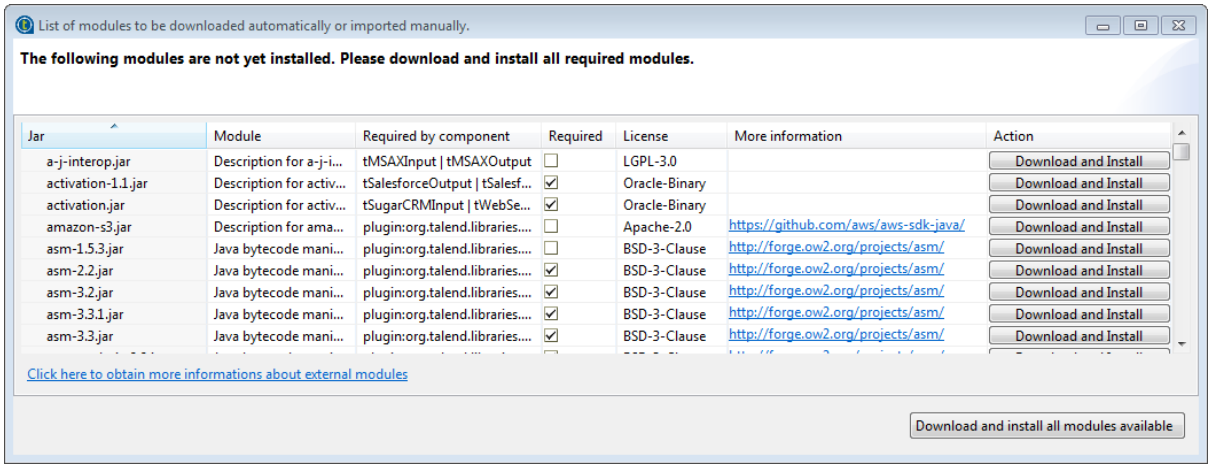
Column	Description
Status	points out if a module is installed or not installed on your system. The icon indicates that the module is not necessarily required for the corresponding component listed in the **Context** column. The icon indicates that the module is absolutely required for the corresponding component.
Context	lists the name of **Talend** component using the module. If this column is empty, the module is then required for the general use of *Talend Studio*. This column lists any external libraries added to the routines you create and save in the Studio library folder. For more information, see the *Talend Studio User Guide*.
Module	lists the module exact name.
Description	explains why the module/library is required.
Required	the selected check box indicates that the module is required.

In addition to the **Modules** view, the Studio provides a mechanism that enables you to easily identify, download and install most of the required third-party modules from the **Talend** website and directs you to valid websites for the rest.

A Jar installation wizard appears whenever any required external module is found missing for any feature in the Studio, including 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, or
- 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 for which one or more required external modules are missing, or
- click the button in the **Modules** view.

When you click this button, the wizard that appears will list all the required external modules that are not integrated in the Studio.



The table below describes the information presented in the wizard.

Item	Description
Jar	The file name of the external module.
Module	A short description about the nature of the module.
Required by component	Lists the components that require the external module.
Required	The selected check box indicates that the module is required.
License	The license under which the module is provided.
More information	Provides the URL of the valid website where you can find more information about this module and download the module manually.
Action	Download and Install: Click to open the **[Download external modules]** dialog box to download and install the module, which is available on the **Talend** website; Open in browser: Click the link to open the valid website to download the module, which is not available on the **Talend** website, and then click the jar button to import the downloaded module into your studio; : You need to find and download the module yourself and click the jar button to import it into the your studio.
Download and install all modules available	Click to open the **[Download external modules]** dialog box to download and install all the required modules that are available on the **Talend** website.
Do not show again	Select to prevent the wizard from appearing again unless you click the button in the **Modules** tab view. This check box shows only when you drop a component, or guess the schema of a database, that requires an external module, or click the **Install** button on the **Component** tab of a component that requires an external module.
Click here to obtain more information about external modules	Click to go to **Talend** online documentation on installing third-party modules.

This wizard lists the external modules to be installed, the licenses under which they are provided, and the URLs of the valid websites where they are downloadable, and allows you to download and install automatically all the modules available on the Talend website and download those not available on the **Talend** website by following the links provided in the **Action** column and then install them into your Studio manually.

When you drop a component, 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.

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

2.3.2. Install external modules

To download and install modules in the Studio

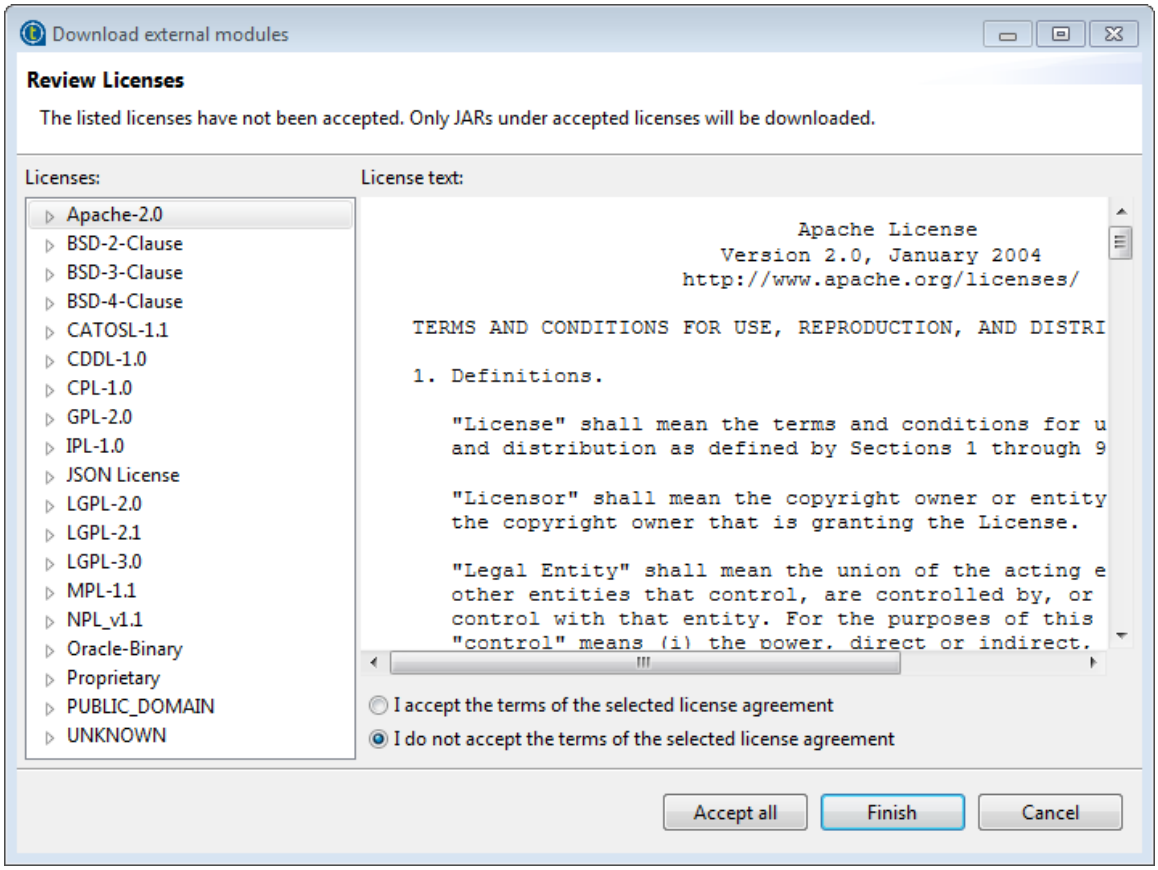
To download and install external modules automatically, do the following:

*If you are working behind a network proxy, make sure you have correctly set up your proxy before you can download and install external modules in your Studio. 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**.*

*Add the web site **http://talend-update.talend.com** and the port 443 to the whitelist.*

1. 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, or select **Help** > **Install Additional Packages** from the menu to open the **[Additional Talend Packages]** wizard. From this wizard, make sure the Optional and Required third-party libraries check boxes are selected and click **Finish**. The **[Download external modules]** dialog box opens.

   This **[Additional Talend Packages]** wizard appears automatically when you launch *Talend Studio* if any additional packages, including external modules, need to be installed for any features to function in the Studio.



2. 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 start the download and installation process.

To download and install all external modules provided under all the listed licenses, click the **Accept all** button to start the download and installation process.

Upon installation of the chosen external module or modules, a dialog box appears to notify you about the number of modules successfully installed and/or about the modules failed to install, if any.

To install manually an external module you already have in your local file system, do the following:

Talend Open Studio for Big Data does not come with the JDBC drivers for Oracle databases due to Apache license restrictions. For Oracle9*i*, the required JDBC driver downloadable from Oracle website is named *ojdbc14.jar*, the same as that for Oracle 10g. To enable the JDBC driver for Oracle9*i* you have downloaded to work in **Talend Open Studio for Big Data**, you have to change the file name to *ojdbc14-9i.jar* before installing it into the Studio.

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

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

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.

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

   You can now use the component dependent on this module in any of your Job designs.

To install modules downloaded from external websites

Some modules 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.

- For the studio, the downloaded modules must be placed in the following folder:

  <StudioPath>/configuration/.m2

Chapter 3. Upgrading your Talend products

This chapter describes the various operations required to migrate version of the *Talend* solutions.

We assume that you have installed and configured these solutions as described in the *Installing Talend Studio for the first time*.

The migration and upgrade process includes the following mandatory steps:

These steps usually need to be completed in the following order.

1. Backing up the environment, see the *Backing up the environment*.
2. Upgrading the Talend projects in the Studio, see the *Upgrading the Talend projects in the Studio*.

3.1. Backing up the environment

Before you start migrating your *Talend* solutions, make sure your environment is correctly backed up.

The environment backup process includes the following mandatory steps:

These steps usually need to be completed in the following order.

1. Saving the local projects, see Saving the local projects.

Saving the local projects

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

3.2. Upgrading the Talend projects in the Studio

Importing your local projects

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

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

   For more information on how to export local projects to an archive file, see Saving the local projects.

Appendix A. Appendices

The following appendices contain complementary information to go further with your *Talend* product:

- *Supported Third-Party System/Database/Business Application Versions*

A.1. 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.

A.1.1. Supported systems, databases and business applications by Talend components

The information contained in the following table is applicable for the 6.4.1 version of your Talend product at the time of its release. For updated information on the latest supported versions of the third-party systems, see the online version of this page on Talend Help Center.

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

Systems/Databases	Versions	OS	Available with...
Alfresco	2.1	N/A[1]	All **Talend** products
Amazon Redshift	Initial release of Amazon Redshift	N/A[1]	All **Talend** products
AS/400	V5R2 to V5R4 V5R3 to V6R1 V6R1 to V7R2	N/A[1]	All **Talend** products
Access[2]	2003 2007	Windows	**Talend** products with Data Integration (DI), Master Data Management (MDM), Enterprise Service Bus (ESB) or Big Data
Bonita	5.2.3 5.3.1 5.6.1 5.10.1 6.5.2 7.2.4	N/A[1]	All **Talend** products
Cassandra	2.0.0 3.0/3.1/3.2/3.3/3.4 (Deprecated versions: 1.1.2/1.2.2)	Windows + Linux	**Talend** products with Big Data
CouchBase	2.0	Windows	**Talend** products with Big Data
CouchDB	1.0.2	Windows	**Talend** products with Big Data
DB Generic	ODBC	Windows	All **Talend** products
DB2	10.5 10.1	Windows + Linux	**Talend** components with all products. **Talend** products with MDM or ESB.
DynamoDB	No specified version	N/A[1]	**Talend** products with Big Data
EXASolution	6.0 and earlier	Windows	**Talend** products with DI, MDM, ESB or Big Data
Elasticsearch	Until 2.3.X	N/A[1]	**Talend** products with Big Data
FireBird	2.1	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data

Systems/Databases	Versions	OS	Available with...
Greenplum	4.2.1.0	Windows (client only) + Linux	**Talend** products with DI, MDM, ESB or Big Data
HSQLDb	1.8.0	N/A[1]	**Talend** products with DI, MDM, ESB or Big Data
Informix	11.50	Windows + Linux	All **Talend** products
Ingres	9.2	Windows + Linux	All **Talend** products
Interbase	7 and above	N/A[1]	**Talend** products with DI, MDM, ESB or Big Data
JavaDB	6	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data
Kafka	0.8.2.0 0.9.0.1[3] 0.10.0.1[3]	Windows + Linux	**Talend** products with Big Data
LDAP	No version limitation	Windows + Linux	All **Talend** products
Microsoft AX	Dynamics AX 4.0 Dynamics AX 2012	N/A[1]	All **Talend** products
Microsoft CRM	4.0 2011 2013 2015 2016	N/A[1]	All **Talend** products
MS SQL Server	2000 2003 2005 2008 2012 2014 [4] 2016 [4]	Windows + Linux	All **Talend** products
MaxDB	7.6	N/A[1]	**Talend** products with DI, MDM, ESB or Big Data
MongoDB	2.5.X 2.6.X 3.0.X 3.2.X	Windows + Linux	**Talend** products with Big Data
MySQL	Mysql4 Mysql5 MariaDB	Windows + Linux	All **Talend** products
Netezza	7.2	Windows + Linux	All **Talend** products
NetSuite	2014 2016	Windows + Linux	All **Talend** products
Neo4j	1.X.X 2.X.X/2.2.X/2.3	Linux	**Talend** products with Big Data

Systems/Databases	Versions	OS	Available with...
OleDb	2000 2003 2005 2007 2010	N/A[1]	All **Talend** products
Oracle	Oracle 8i/9i/10g/11g/11g (11.6)/12c	Windows + Linux	All **Talend** products
ParAccel	3.1 3.5	N/A[1]	**Talend** products with DI, MDM, ESB or Big Data
PostgreSQL	9.X	Windows + Linux	All **Talend** products
PostgresPlus	9.X	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data
Red Hat BRMS	6.1	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data
Salesforce	V39 and earlier	Windows + Linux	All **Talend** products
SAP	ECC 6.0 EhP6	Windows	All **Talend** products
SAP BW	7.3 7.4 7.5	Windows	All **Talend** products
SAP Hana	1.0	Windows	All **Talend** products
SAS	9.1 9.2	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data
SQLite	3.6.7	Windows + Linux	All **Talend** products
Sybase	12.5 12.7 15.2 15.5 15.7 16.0	Windows + Linux	All **Talend** products
SybaseIQ	12.5 12.7 15.2 16.0	Windows + Linux	All **Talend** products
Teradata	12 13 14 15	Windows + Linux	All **Talend** products
VectorWise	2	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data
Vertica	3 3.5	Windows + Linux	**Talend** products with DI, MDM, ESB or Big Data

Systems/Databases	Versions	OS	Available with...
	4 4.1 5.0 5.1 6.0 6.1.X 7.0.X 7.1.X		
VtigerCRM	Vtiger 5.0 Vtiger 5.1	N/A[1]	All **Talend** products

1. The test information is not available yet.

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

3. For information about the security options supported by the Kafka components, see Talend Help Center.

4. No new feature introduced by MS SQL Server 2014/2016 is supported.

A.1.2. Supported Hadoop distribution versions for Talend Jobs

- ✅: officially supported.
- ❌: not officially supported.
- 🔒: the Kerberos kinit option is supported by the Studio.
- 🔒🔒: the Kerberos kinit option and the Kerberos keytab option are both supported by the Studio.

The information contained in the following table is applicable for the 6.4.1 version of your Talend product at the time of its release. For updated information on the latest supported versions of the third-party systems, see the online version of this page on Talend Help Center.

If support for the Hadoop distribution you want to use is not yet available in your *Talend Studio*, it may be available via an update. For related information, see Talend Help Center.

		HBase	HCatalog	HDFS	Hive		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 [2]	Hive2 [2]			Standalone	YARN		
Google Dataproc	V1.1	❌	❌	❌	❌	✅	❌	❌	❌	✅	❌	✅
HDP	V1.2.0 (Deprecated)	✅ 🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	❌	❌	✅ 🔒🔒	✅ 🔒🔒
	V1.3.0 (Deprecated)	✅ 🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	❌	❌	✅ 🔒🔒	✅ 🔒🔒
	V2.0.0 (Deprecated)	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	✅ 🔒🔒	❌	❌	✅ 🔒🔒	✅ 🔒🔒

		HBase	HCatalog	HDFS	Hive		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 [2]	Hive2 [2]			Standalone	YARN		
	V2.1.0 (Deprecated)[3]	✅	✅	✅	✅	✅	✅	✅	❌	✅	✅	✅
	V2.2.0 (Deprecated)[3]	✅	✅	✅	✅	✅	✅	✅	❌	❌	✅	✅
	V2.3.2 (Deprecated)	✅	✅	✅	❌	✅	✅	✅	❌	✅	✅	✅
	V2.4.0	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	V2.5.0	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	V2.6.0	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
Apache	1.0.0 (deprecated)	✅ 0.92.0	❌	✅	✅ 0.9.0	❌	❌	✅	❌	❌	❌	✅
Cloudera	CDH4 (deprecated)	✅	❌	✅	✅	✅	✅	✅	❌	❌	✅	✅
	CDH4.3 + (deprecated)	✅	❌	✅	✅	✅	✅	✅	❌	❌	✅	✅
	CDH 5.0 (deprecated)	✅	✅	✅	✅	✅	✅	✅	❌	❌	✅	✅
	CDH 5.1 (deprecated)	✅	✅	✅	✅	✅	✅	✅	❌	❌	✅	✅
	CDH 5.1 (deprecated)	✅	✅	✅	✅	✅	✅	✅	❌	❌	✅	✅
	CDH 5.4 (deprecated)	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	CDH 5.5 (YARN mode)	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	CDH 5.6 (YARN mode)	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	CDH 5.7 (YARN mode)	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅
	CDH 5.8 (YARN mode)	✅	✅	✅	❌	✅	✅	✅	✅	✅	✅	✅

		HBase	HCatalog	HDFS	Hive		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 [2]	Hive2 [2]			Standalone	YARN		
	CDH 5.10 (YARN mode)	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
MapR	2.0.0 (deprecated)	✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	✓
	2.1.2 (deprecated)	✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	✓
	2.1.3 (deprecated)	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓
	3.0.1 (deprecated)	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓
	3.1.0 (deprecated)	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓
	4.0.1 (deprecated)[4]	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓
	4.1.0 (deprecated)[4]	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	5.0.0 (YARN mode)[4]	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
	5.1.0 (YARN mode)[4]	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
	5.2.0 (YARN mode)[4]	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
Amazon EMR	Apache 1.0.3 (deprecated)	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗	✓
	Apache 2.4.0 (deprecated)	✓	✗	✓	✓	✗	✗	✓	✗	✗	✓	✓
	EMR 4.0.0 (deprecated)	✗	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓
	EMR 4.5.0 (Apache 2.7.2)	✗	✓	✓	✗	✓	✓	✓	✗	✓	✓	✓
	EMR 4.6.0 (Apache 2.7.2)	✓	✓	✓	✗	✓	✓	✓	✗	✓	✓	✓
	EMR 5.0.0 (Apache 2.7.2)	✓	✓	✓	✗	✓	✓	✓	✗	✓	✓	✓
	EMR 5.4/5.5	✓	✓	✓	✗	✓	✓	✓	✗	✓	✓	✓
Pivotal HD	1.0.1 (deprecated)	✓	✗	✓	✓	✗	✗	✓	✗	✗	✓	✓
	2.0 (deprecated)	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✓

		HBase	HCatalog	HDFS	Hive		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 [2]	Hive2 [2]			Standalone	YARN		
Microsoft HD Insight	3.1 (deprecated)	❌	❌	❌	✅	✅	❌	✅	❌	❌	❌	✅
	3.2 (deprecated)	❌	❌	❌	✅	✅	❌	✅	✅	❌	❌	✅
	3.4	❌	❌	❌	✅	✅	❌	✅	✅	✅	❌	✅
Custom[1]												

1. This enables the connection between the Studio and a custom Hadoop distribution not yet officially supported in the Studio. For further information, see the sections describing how to connect to a custom Hadoop distribution of the Talend Big Data Getting Started Guide or the documentation of any related component that creates the connection to a Hadoop distribution, such as tHDFSConnection.

2. In the Standalone mode, Hive 1 does not support the Kerberos security.

3. This option also allows you to connect to a Teradata Hadoop platform. For further information about the version of the Hortonworks Data Platform used in the Teradata platform you are using, see Teradata's related documentation.

4. For this MapR version, the MapR security ticket mechanism is supported by the Studio.

For further information about the versions of all the supported third-party systems/databases, see *Supported systems, databases and business applications by Talend components*.