



Talend Real-time Big Data Platform

Installation Guide for Mac

6.4.1

Adapted for v6.4.1. Supersedes previous releases.

Publication date: June 29, 2017

Copyright © 2017 Talend Inc. All rights reserved.

Notices

Talend, Talend Integration Factory, Talend Service Factory, and Talend ESB are trademarks of Talend, Inc.

Apache CXF, CXF, Apache Karaf, Karaf, Apache Camel, Camel, Apache Maven, Maven, Apache Syncope, Syncope, Apache ActiveMQ, ActiveMQ are trademarks of The Apache Foundation. Eclipse Equinox is a trademark of the Eclipse Foundation, Inc. SoapUI is a trademark of SmartBear Software. Hyperic is a trademark of VMware, Inc. Nagios is a trademark of Nagios Enterprises, LLC.

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

End User License Agreement

The software described in this documentation is provided under **Talend's** End User License Agreement (EULA) for commercial products. By using the software, you are considered to have fully understood and unconditionally accepted all the terms and conditions of the EULA.

To read the EULA now, visit <http://www.talend.com/legal-terms/us-eula>.

Table of Contents

Preface	v
1. General information	v
1.1. Purpose	v
1.2. Audience	v
1.3. Typographical conventions	v
2. Feedback and Support	v
Chapter 1. Before installing your Talend product	1
1.1. Preparing your installation	2
1.1.1. Installation modes	2
1.1.2. Files to download	2
1.1.3. Community and Support	3
1.2. Hardware requirements	3
1.3. Software requirements	4
1.3.1. Compatible Operating Systems	4
1.3.2. Java	6
1.3.3. Compatible Apache software and JMS Brokers for Talend ESB	8
1.3.4. Compatible web application servers and containers	8
1.3.5. Compatible Web browsers	10
1.3.6. Compatible version control systems	10
1.3.7. Compatible Databases	11
1.3.8. Compatible Messaging Systems	15
1.3.9. Compatible Nexus Artifact Repository	15
1.3.10. Compatible execution servers	15
1.3.11. Talend Data Preparation and Talend Administration Center compatibility matrix	16
1.3.12. Port information	17
Chapter 2. Installing your Talend product using Talend Installer (recommended)	19
2.1. Installing Talend Studio with the Talend Studio Installer	20
Chapter 3. Installing your Talend product manually	21
3.1. Setting up your version control system	22
3.1.1. Installing and configuring an Apache Subversion (SVN) server	22
3.1.2. Installing and configuring Git	23
3.2. Installing your Talend Studio	23
3.2.1. Installing and launching your Talend Studio	23
3.2.2. Setting up a Talend Administration Center connection in Talend Studio	24
3.2.3. Configuring Nexus in Talend Studio	25
3.2.4. Installing external modules	27
Appendix A. Appendices	33
A.1. Introduction to the Talend products	34
A.1.1. Apache Tomcat Server	34
A.1.2. Database	35
A.1.3. Version control system	35
A.1.4. Artifact Repository	35
A.1.5. Talend Runtime	39
A.1.6. JobServer	40
A.1.7. Talend Studio	40
A.1.8. Talend Activity Monitoring Console log database	40
A.2. Architecture of the Talend products	41
A.3. Talend High Availability	43
A.4. H2 Database Administration & Maintenance	44
A.4.1. About H2 embedded database	45
A.4.2. Administrating the H2 database through the Web console	45
A.4.3. Setting up the H2 database for access from other machines	47
A.5. Supported Third-Party System/Database/Business Application Versions	48
A.5.1. Supported systems, databases and business applications by Talend components	48
A.5.2. Supported Hadoop distribution versions for Talend Jobs	51
A.5.3. Supported Hadoop distribution versions for Talend Data Preparation with Big Data	54
A.5.4. Supported databases for profiling data	54
A.5.5. Supported Hive distributions for profiling data	55

Preface

1. General information

1.1. Purpose

This Installation Guide explains how to install and configure 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:

- [Installation modes](#)
- [Files to download](#)
- [Community and Support](#)

1.1.1. Installation modes

There are two different methods to install your *Talend* product:

- the automatic mode, using *Talend Installer*. It is the recommended way of installing your *Talend* product. For more information, see [Installing your Talend product using Talend Installer \(recommended\)](#).
- the manual installation. This method allows you to customize every step of your installation. For more information, see [Installing your Talend product manually](#).

1.1.2. Files to download

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

- the licence file. For more information, see [License key](#).
- the software packages. For more information, see [Software package](#).

1.1.2.1. License key

You should have received an email from *Talend* including your personal license key in a file with no extension.

The license key is mandatory to be able to access each module of *Talend*. **Keep this file at hand in a safe place.**

1.1.2.2. 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.)

The software modules must be all in the same versions/revisions. This means that both YYYYMMDD_HHmm and A.B.C must match on both client side and server side.

Table 1.1. Manual installation software package

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

The links to download these packages are listed in your licence email.

1.1.3. 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.
- [Talend Professional Support](#). If you are a *Talend* subscription customer, you can open a ticket to the *Talend* Support.
- [Talend Consulting Portal](#). If you are a *Talend* subscription customer, you can ask for a consultant to help through the installation of your *Talend* product.

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
<i>Studio</i>	Client	3 GB minimum, 4 GB recommended
<i>Talend Data Preparation</i>	Server	2 GB minimum, 4 GB recommended
<i>Talend Data Stewardship</i>	Server	1 GB minimum, 2 GB recommended
<i>Talend Dictionary Service</i>	Server	1 GB minimum, 2 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
<i>Studio</i>	Client	3 GB	3+ GB
<i>Talend Data Preparation</i>	Server	300 MB	1 GB + datasets size These requirements do not take the MongoDB metadata size into account.
<i>Talend Data Stewardship</i>	Server	3 GB	100 MB for a campaign that counts 50,000 tasks, each task having 50 attributes.

Product	Client/Server	Required disk space for installation	Required disk space for use
<i>Talend Dictionary Service</i>	Server	1 GB	1+ GB

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](#)
- [Compatible version control systems](#)
- [Compatible execution servers](#)
- [Talend Data Preparation and Talend Administration Center compatibility matrix](#)
- [Port information](#)

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
		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

Support type	Operating System (64-bit)	
		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

Table 1.5. Talend server modules

Support type	Operating System		Processor
Recommended	Linux	Red Hat Enterprise Linux Server 7.2	64-bit
	Windows	Microsoft Windows Server 2012 R2	64-bit
Supported	Linux	Ubuntu 17.04	64-bit
		Ubuntu 16.04 LTS	64-bit
		Ubuntu 14.04 LTS	64-bit
		Red Hat Enterprise Linux Server/CentOS 7.3	64-bit
		Red Hat Enterprise Linux Server/CentOS 7.1	64-bit
		Red Hat Enterprise Linux Server/CentOS 6.9	64-bit
		Red Hat Enterprise Linux Server/CentOS 6.8	64-bit
		SUSE SLES 12	64-bit
		SUSE SLES 11	64-bit
	Unix	Solaris (SunOs) 11	x86/64-bit ¹
		Solaris (SunOs) 11	Sparc/64-bit ²
		Solaris (SunOs) 10 ³	x86/64-bit ¹
		Solaris (SunOs) 10 ³	Sparc/64-bit ²
	Windows	Microsoft Windows Server 2016	64-bit
		Microsoft Windows Server 2012	64-bit
Deprecated	Unix	AIX 7.1	64-bit (IBM Java 8 only) ⁴

1. Only supported for Talend Administration Center, CommandLine, JobServer, Talend ESB and Talend Runtime.

2. Only supported for Talend ESB and Talend Runtime.

3. At least patch level 9 should be installed.

4. Only supported for CommandLine, JobServer, Talend ESB and Talend Runtime.

The server modules include Artifact Repository, CommandLine, JobServer, Talend Activity Monitoring Console, Talend Administration Center, Talend Data Preparation, Talend Data Quality Portal, Talend Data Stewardship, Talend ESB, Talend Log Server, Talend Repository Manager, Talend Runtime.

Note that *Talend Metadata Bridge* is only supported on Windows.

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 1.6. 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.

Table 1.7. Server Java environments

JRE Version	JobServer	MDM Server	ESB/Talend Runtime	ESB/ Microservices	Big Data Distributions	Talend Server Applications ¹	Comment/ Limitation
Oracle 7 (running alongside Oracle 8)	✗	✗	✗	✗	Compatible with Java 1.7 only	✗	Compatible with Studio JDK Compiler Compliance Level 1.7 only
Oracle 8 (recommended)	✓ (R)	✓ (R)	✓ (R)	✓ (R)	Compatible with Java 1.8	✓ (R)	Compatible with Studio JDK Compiler Compliance Level 1.7 (default/ recommended) or 1.8
IBM 8 (deprecated)	✓	✗	✗	✗	✗	✗	Only for AIX and SLES

1: Talend server applications include Artifact Repository, CommandLine, Talend Activity Monitoring Console, Talend Administration Center, Talend Data Preparation, Talend Data Quality Portal, Talend Data Stewardship, Talend Log Server, Talend Repository Manager.

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 `/Library/Java/Home`.
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. Compatible Apache software and JMS Brokers for Talend ESB

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](#).

Some of the ESB tools use Apache software components. Talend ESB contains the following Apache Project versions:

Table 1.8. Supported Apache software

Software	Notes	More information
Apache Karaf 4.1.1	Service release upgrade.	Release notes: https://issues.apache.org/jira/secure/ReleaseNote.jspa?projectId=12311140&version=12339244
Apache CXF 3.1.11	Service release upgrade.	Release notes and Migration Guide: http://cxf.apache.org/cxf-3111-release-notes.html
Apache Camel 2.17.6	Minor release upgrade.	Release notes: http://camel.apache.org/camel-2176-release.html
Apache ActiveMQ 5.14.5	Minor release upgrade.	Release notes: http://activemq.apache.org/activemq-5145-release.html

Talend ESB supports the following JMS Brokers.

Note that ESB Java Consumer & Provider using SOAP/JMS-based messaging are supported for the following (JMS) Message-Brokers.

Table 1.9. Supported Messaging Brokers for SOAP/JMS

Software	More information
Apache ActiveMQ 5.14.5	Release notes: http://activemq.apache.org/activemq-5145-release.html
IBM WebSphere MQ 7.5	Release notes: http://www-01.ibm.com/support/docview.wss?uid=swg27043190

Table 1.10. Supported Containers

Software	Product
Jetty 9.3.14	ESB Runtime - OSGi Container
Apache Tomcat 8.0.44	ESB Microservices

1.3.4. Compatible web application servers and containers



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](#).

Please refer to the following grids for a summary of supported web application servers and runtime containers.

Table 1.11. Talend Administration Center, Talend Activity Monitoring Console and Talend Repository Manager

Support type	Web application servers
Recommended	Apache Tomcat 8.0 ¹
Supported	Apache Tomcat 7 ¹
	Pivotal tc Server 3.2
Deprecated	JBoss EAP 6.4
	IBM Websphere 8.5
	Weblogic 12c

1. TLS 1.2 is supported. For more information, see <https://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>.

Table 1.12. Talend Data Quality Portal

Support type	Web application servers
Recommended	Apache Tomcat 8.0 ¹
Supported	Apache Tomcat 7 ¹
	Pivotal tc Server 3.2

1. TLS 1.2 is supported. For more information, see <https://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>.

Table 1.13. Talend Data Stewardship

Support type	Web application servers
Recommended	Apache Tomcat 8.0 ¹

1. TLS 1.2 is supported. For more information, see <https://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>.

Table 1.14. Talend Dictionary Service

Support type	Web application servers
Recommended	Apache Tomcat 8.0 ¹

1. TLS 1.2 is supported. For more information, see <https://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>.

Table 1.15. Talend ESB

Support type	Runtime Containers
Recommended	Talend Runtime (Apache Karaf) 6.4 ¹
	Apache Tomcat 8.0 ²⁵
Supported with limitations	Apache Tomcat 7 and 8.0 ³⁵
Deprecated	JBoss EAP 6.4 ⁴
	Weblogic 12c ⁴
	IBM Websphere 8.5 ⁴

1. Except for Talend Identity Management Service, where Tomcat Apache 8 is recommended.
2. Only for Talend Identity Management
3. Only for CXF Services, Camel Routes, Service Activity Monitoring, Talend Identity Management and Security Token Service.
4. Only for CXF Services and Camel Routes.
5. TLS 1.2 is supported. For more information, see <https://tomcat.apache.org/tomcat-8.0-doc/ssl-howto.html>.

1.3.5. Compatible Web browsers



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](#).

Please refer to the table below for a summary of supported Web browsers.

Talend Administration Center, Drools, Talend Data Preparation, Talend Data Stewardship	Web browser
recommended	Mozilla Firefox 51 and 53 ¹
supported	Firefox ESR 45 and 52
	Microsoft Internet Explorer 11
	Microsoft Edge
	Apple Safari 10
	Google Chrome 53 to 58 ¹

1. Talend maintenance releases will support the most recent browser version at the time of the release.

1.3.6. Compatible version control 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](#).

The following version control clients are embedded in the *Talend* products:

- Apache Subversion (version 1.8)

- Git (JGit/EGit 3.4.2)

The tables below provide a summary of the supported version control system servers that you can use to store your projects.

Table 1.16. Apache Subversion version control servers

Support type	Version control servers
recommended	VisualSVN Server 3.3 (compatible with Apache Subversion 1.8)
supported	VisualSVN Server 3.5 (compatible with Apache Subversion 1.9)
	Bitnami Subversion Stack (compatible with Apache Subversion 1.8)
	Svnserve (compatible with Apache Subversion 1.8)
	SVNEdge 5.1

Table 1.17. Git version control servers

Support type	Version control servers
recommended	GitHub (SaaS)
supported	BitBucket (SaaS)
	AWS CodeCommit (SaaS)
	GitLab 9.1
	GitLab 8.13
	Gitblit 1.7.1

1.3.7. Compatible Databases



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](#).

Please refer to the following grids for a summary of supported databases.

Table 1.18. Talend Administration Center

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	Azure SQL
	H2 1.4
	MariaDB 10.1
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)

Support type	Databases
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.19. Talend Identity and Access Management

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	Azure SQL
	H2 1.4
	Derby DB > 10.8
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.20. Talend Activity Monitoring Console

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	IBM DB2 10.1
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.21. Service Activity Monitoring (SAM)

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	Derby DB > 10.8
	IBM DB2 10.1
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.22. ESB Service Registry/Authorization/Talend Identity Management/Event Logging

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	Derby DB > 10.8
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.23. Talend Data Quality Portal

Support type	Databases used to store the analyses
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g

Support type	Databases used to store the analyses
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Support type	Database used to store the portal configuration
Supported (embedded in the product)	HSQL 2.3.2



If you are using MySQL, the my.ini file of your MySQL server installation directory must be edited in order to add the following line:

```
max_allowed_packet = 64M
```

The 64M value is only given as example. It represents the maximum size of a query packet the server can handle as well as the maximum query size the server can process. You can replace it with any value superior to 16M.

Table 1.24. Talend Repository Manager

Support type	Databases
Recommended	MySQL 5.7 ¹
	Oracle 12c ¹
Supported	Azure SQL
	H2 1.4
	MariaDB 10.1
	MS SQL Server 2016 ¹
	MS SQL Server 2014
	MS SQL Server 2012 (SP2)
	MySQL 5.6
	Oracle 11g
	PostgreSQL 9.6 ¹²
	PostgreSQL 9.4 and 9.5 ³

1. The corresponding Amazon Relational Database Service (Amazon RDS) version is supported.

2. The corresponding Google Cloud SQL version is supported.

3. PostgreSQL JDBC driver 9.4 is used.

Table 1.25. Talend Data Preparation

Support type	Database
Recommended (embedded in the product)	MongoDB 3.4

Table 1.26. Talend Data Stewardship

Support type	Database
Recommended (embedded in the product)	MongoDB 3.4

Table 1.27. Talend Dictionary Service

Support type	Database
Recommended (embedded in the product)	MongoDB 3.4

1.3.8. Compatible Messaging 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](#).

Please refer to the following grid for a summary of supported messaging systems.

Talend Data Preparation, Talend Data Stewardship, Talend Dictionary Service	Messaging System
recommended	Apache Kafka 0.10.1

1.3.9. Compatible Nexus Artifact Repository



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](#).

Please refer to the table below for the recommended Nexus Artifact Repository version that you can use to publish and store your artifacts.

Support type	Artifact repository
recommended ¹	Sonatype Nexus 2.14

1. Embedded in the Talend Administration Center archive file.

1.3.10. Compatible execution servers

Make sure the execution server version is compatible with the *Talend Administration Center*, *CommandLine* and *Studio* versions.

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

Table 1.28. Job servers (JobServer and job server in Talend Runtime)

	Job Server	Job Server	Job Server	Job Server
Version	6.1.x	6.2.x	6.3.x	6.4.x

Talend Administration Center, Studio, CommandLine	6.1.x	✓			
Talend Administration Center, Studio, CommandLine	6.2.x	✓	✓		
Talend Administration Center, Studio, CommandLine	6.3.x	✓	✓	✓	
Talend Administration Center, Studio, CommandLine	6.4.x	✓	✓	✓	✓

Table 1.29. ESB OSGI servers (Talend Runtime)

		Talend Runtime (ESB OSGI Server)	Talend Runtime (ESB OSGI Server)	Talend Runtime (ESB OSGI Server)	Talend Runtime (ESB OSGI Server)
	Version	6.1.x	6.2.x	6.3.x	6.4.x
Talend Administration Center, Talend ESB Infrastructure Services	6.1.x	✓			
Talend Administration Center, Talend ESB Infrastructure Services	6.2.x	✓	✓		
Talend Administration Center, Talend ESB Infrastructure Services	6.3.x	✓	✓	✓	
Talend Administration Center, Talend ESB Infrastructure Services	6.4.x	✓	✓	✓	✓

1.3.11. Talend Data Preparation and Talend Administration Center compatibility matrix

Make sure that your *Talend Administration Center* version is compatible with the *Talend Data Preparation* version that you are using.

Table 1.30. Talend Data Preparation and Talend Administration Center compatibility

	Talend Data Preparation 1.2	Talend Data Preparation 1.3	Talend Data Preparation 2.0	Talend Data Preparation 2.1
Talend Administration Center 6.2	✓	✓	✗	✗
Talend Administration Center 6.3	✗	✗	✓	✗

Talend Administration Center 6.4				
----------------------------------	---	---	---	---

1.3.12. Port information

This section describes the most important TCP/IP ports the *Talend* products use. Make sure that your firewall configuration is compatible with these ports or change the default ports where needed.

Additionally, add the web site <http://talendforge.org/> and the port 80 to the whitelist. To be able to download patches or external libraries, also add the web site <http://talend-update.talend.com> and the port 443 to the whitelist.

In this table:

Port: a TCP/IP port or a range of ports.

Direction: In (Inbound); Out (Outbound) - related to the communication direction (for example a HTTP Port for a CXF Route or Service we listen on request) will be an 'Inbound' port. For example, a browser which sends a request to port 7080 will have this port as 'Outbound' port in this list.

Usage: which part of the Product component uses this port (for example 1099 is used by the JMX Monitoring component of Talend Runtime).

Config: the file or location where the value can be changed.

Table 1.31. Talend Studio Ports

Port	Direction	Usage	Config
8090	IN	tESBProviderRequest (SOAP Data Server) and tRESTRRequest (REST Data Service default port) components	REST: Preferences / Talend / ESB SOAP: tESBProviderRequest component details



Chapter 2. Installing your Talend product using Talend Installer (recommended)

The following pages contain procedures allowing you to automatically install your *Talend* product using *Talend Installer*.

This method is the recommended way of installing your *Talend* product. However, if you want to fully customize every step of your installation, you can perform a manual installation. For more information, see [Installing your Talend product manually](#).


2.1. Installing Talend Studio with the Talend Studio Installer

Talend Studio Installer is a convenient way of installing your *Talend Studio*.

As it comes with an embedded Java Environment, you can install it without any prerequisite.

To install you *Talend Studio* with the *Talend Studio Installer*, proceed as follows:

1. Download the *TalendStudio-A-B-C-osx-installer.app.tar* file.
2. Extract the *TalendStudio-A-B-C-osx-installer.app.tar* archive file.
3. Double-click the *TalendStudio-A-B-C-osx-installer.app* file to launch *Talend Studio Installer*.
4. Accept the License Agreement.
5. Choose the directory where you want your *Talend* product to be installed.
6. Add your license file.
7. Choose where you want the workspace directory to be located.
8. Launch the installation.



Chapter 3. Installing your Talend product manually

The following pages contain procedures allowing you to manually install your *Talend* product.

The recommended installation method is the automatic installation with *Talend Installer*. The manual installation should only be performed if you want to customize every step of your installation. For more information on the automatic installation, see [Installing your Talend product using Talend Installer \(recommended\)](#).

The manual installation procedures must be executed in the following order:

1. [Setting up your version control system](#)
2. [Installing your Talend Studio](#)

For more information on the functional architecture of your *Talend* product, see [Architecture of the Talend products](#).

3.1. Setting up your version control system

The following pages contain procedures on how to install your version control system:

- [Installing and configuring an Apache Subversion \(SVN\) server](#)
- [Installing and configuring Git](#)

For more information on version control systems, see [Version control system](#).

3.1.1. Installing and configuring an Apache Subversion (SVN) server

This procedure describes how to install and configure an Apache Subversion (SVN) server in order to store all your project data (Jobs, Database connections, Routines, Joblets, etc.) in the shared Repository of the Talend Studio.

Download and install Bitnami Subversion Stack

1. Download the SVN server installer, for example Bitnami Subversion Stack at <http://bitnami.org/stack/subversion>.
2. Install it in the *Applications/Utilities* directory.



Choose a port according to your environment. The default port 8080 clashes with Talend Administration Center.

3. Launch Bitnami in a terminal.

It will provide a step-by-step wizard.

```
MEPCToum:~ ctoum$ cd /Applications/Utilities/subversion-1.6.15-0/
MEPCToum:subversion-1.6.15-0 ctoum$ ll
total 88
drwxr-xr-x  14 ctoum  admin   476 21 oct 11:33 ./
drwxrwxr-x+ 45 root   admin  1530 21 oct 11:45 ../
-rw-r--r--   1 ctoum  admin   8132 21 oct 11:25 README.txt
drwxr-xr-x  16 ctoum  admin   544 21 oct 11:25 apache2/
-rw-r--r--   1 ctoum  admin   773 17 août 02:55 changelog.txt
drwxr-xr-x   8 ctoum  admin   272 21 oct 11:24 common/
-rwxr--r--   1 ctoum  admin  25445 21 oct 11:25 ctlscrip.sh*
drwxr-xr-x   4 ctoum  admin   136 21 oct 11:25 img/
drwxr-xr-x   0 ctoum  admin   272 21 oct 11:24 licenses/
drwxr-xr-x   3 ctoum  admin   102 21 oct 11:25 scripts/
drwxr-xr-x   5 ctoum  admin   170 21 oct 11:25 sqlite/
drwxr-xr-x   8 ctoum  admin   272 21 oct 11:33 subversion/
drwxr-xr-x   3 ctoum  admin   102 21 oct 11:33 uninstall.app/
-rwxr-xr-x   1 ctoum  admin   304 21 oct 11:33 use_subversion*
MEPCToum:subversion-1.6.15-0 ctoum$ ./ctlscrip.sh start
Syntax OK
/Applications/Utilities/subversion-1.6.15-0/apache2/scripts/ctl.sh : httpd started at port 8080
/Applications/Utilities/subversion-1.6.15-0/subversion/scripts/ctl.sh : subversion started at port 3690
MEPCToum:subversion-1.6.15-0 ctoum$
```

Configure Bitnami

1. Create a new repository and name it `<my_repo>`.
2. Create a new user and define his user name and password.
3. Grant read-write authorization to this user.

3.1.2. Installing and configuring Git

This procedure describes how to install and configure Git in order to store all your project data (Jobs, Database connections, Routines, Joblets, etc.) in the shared Repository of the Talend Studio.

For more information on the supported Git servers, see [Compatible version control systems](#).

1. Download the Git version corresponding to your system at <https://git-scm.com/downloads> and follow the installation instructions.
2. Open a terminal instance.
3. Create an SSH key using the following command:

```
ssh-keygen
```

4. Put the generated key files in the `/home/User_Name/.ssh` folder.
5. Add the generated public key to settings of your Git server.
6. Use the following command to create a *known-hosts* file:

```
ssh-keyscan -H git_server_hostname >> known_hosts
```

7. Create a *config* file in your *.ssh* folder.
8. Add the following content and adapt it to your configuration:

```
Hostname git_server_hostname
IdentityFile /home/User_Name/.ssh/id_rsa
```

9. Add the connection information to the *Talend Administration Center* configuration. For more information, see the *Talend Administration Center User Guide*.

3.2. Installing your Talend Studio

To install and configure your *Talend Studio*, follow these procedures:

- [Installing and launching your Talend Studio](#)
- [Setting up a Talend Administration Center connection in Talend Studio](#)
- [Configuring Nexus in Talend Studio](#)
- [Installing external modules](#)

For more information on *Talend Studio*, see [Talend Studio](#) and the *Talend Studio User Guide*.

3.2.1. Installing and launching your Talend Studio

To install and launch your *Talend Studio*, follow these procedures:

- [Installing your Talend Studio](#)
- [Editing the memory and JVM settings](#)
- [Launching your Talend Studio](#)

For more information on how to use the Talend Studio, see the *Talend Studio User Guide*.

3.2.1.1. Installing your Talend Studio

Unzip the archive

1. Copy the archive *Talend-Studio-YYYYYYYYY_YYYY-VA.B.C.zip* to a directory of your choice.
2. Unzip it.
3. Create a file (without extension) named *license* containing your license key (found in your email), and paste the file at the root of the extracted directory.

3.2.1.2. Editing the memory and JVM settings

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

1. Edit the *Talend-Studio-macosx-cocoa.ini* file.
2. Edit the memory attributes. For example:

```
-vmargs -Xms40m -Xmx500m -XX:MaxMetaspaceSize=256m
```

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

3.2.1.3. Launching your Talend Studio

1. Double-click the *Talend-Studio-macosx-cocoa.app* executable to launch your *Talend Studio*.
2. In the **[License setup]** dialog box that appears:
 - paste your license key in the blank area, or
 - click **Browse License...** to browse and select your license file and then click **OK**, or
 - click **Import license...** if you have already set your license and project in Talend Administration Center and want to retrieve this license.

For more information on how to launch the Studio and set up a connection to Talend Administration Center, refer to the *Getting started* chapter in the relevant *Studio User Guide*.

3.2.2. Setting up a Talend Administration Center connection in Talend Studio

To set up a connection to *Talend Administration Center* from *Talend Studio*, do the following:

1. Launch *Talend Studio*.
2. In the *Talend Studio* login window, click the [...] button to define a new connection.
3. In the **[Connections]** window that opens, click the [+] button to create a new connection.
4. Set the **Repository** type as *Remote* and enter a **Name** and **Description** for the connection, the **E-mail** and **Password** for the user you created in *Talend Administration Center*, and the URL for *Talend Administration*

Center (for example, <http://localhost:8080/org.talend.administrator> but, depending on your configuration, you may have to replace <localhost> with the server IP address) in the **Web-app Url** field.



Be careful not to use an existing local workspace. If needed, you can create another folder in the Talend Studio alongside the default workspace folder.

5. Click **OK**.

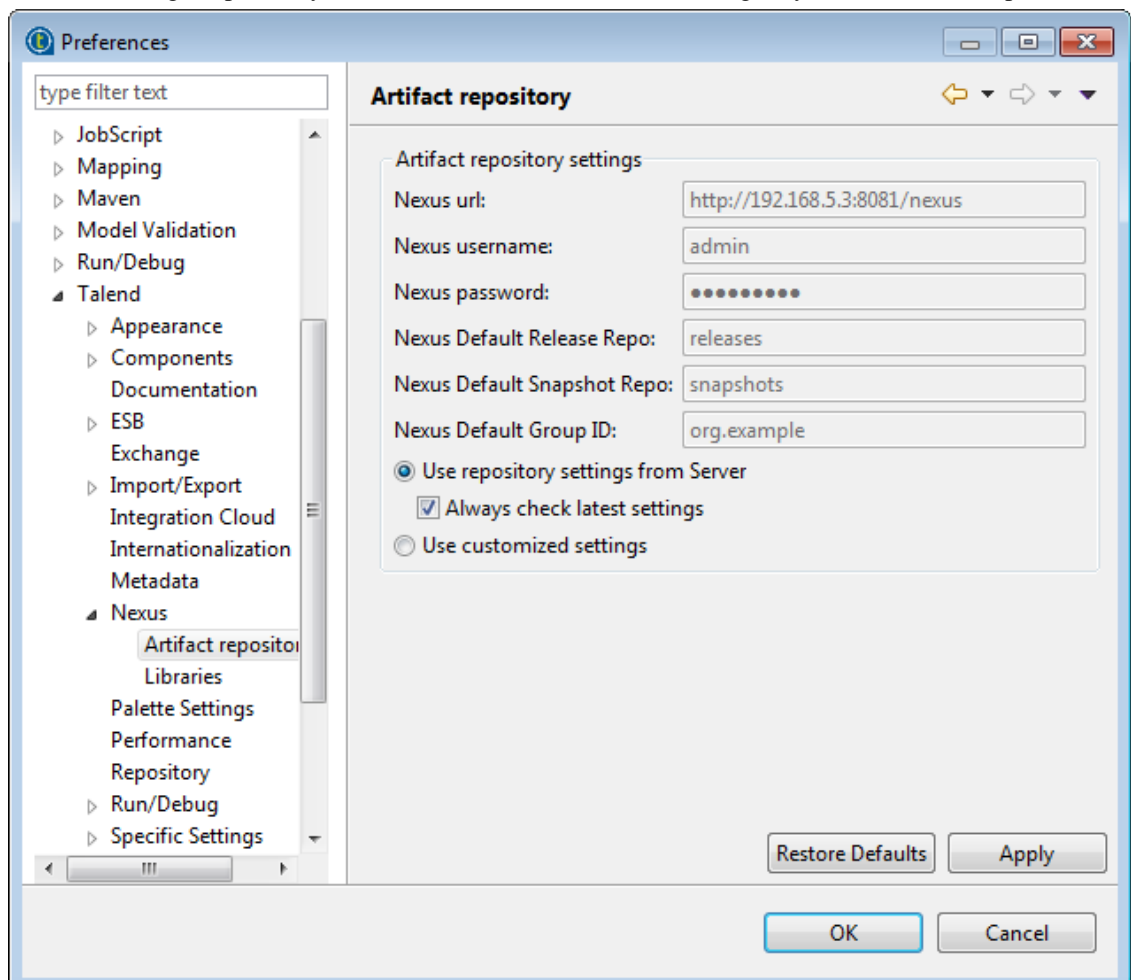
You can now select the newly created connection in the *Talend Studio* login window to connect to a collaborative project.

3.2.3. Configuring Nexus in Talend Studio

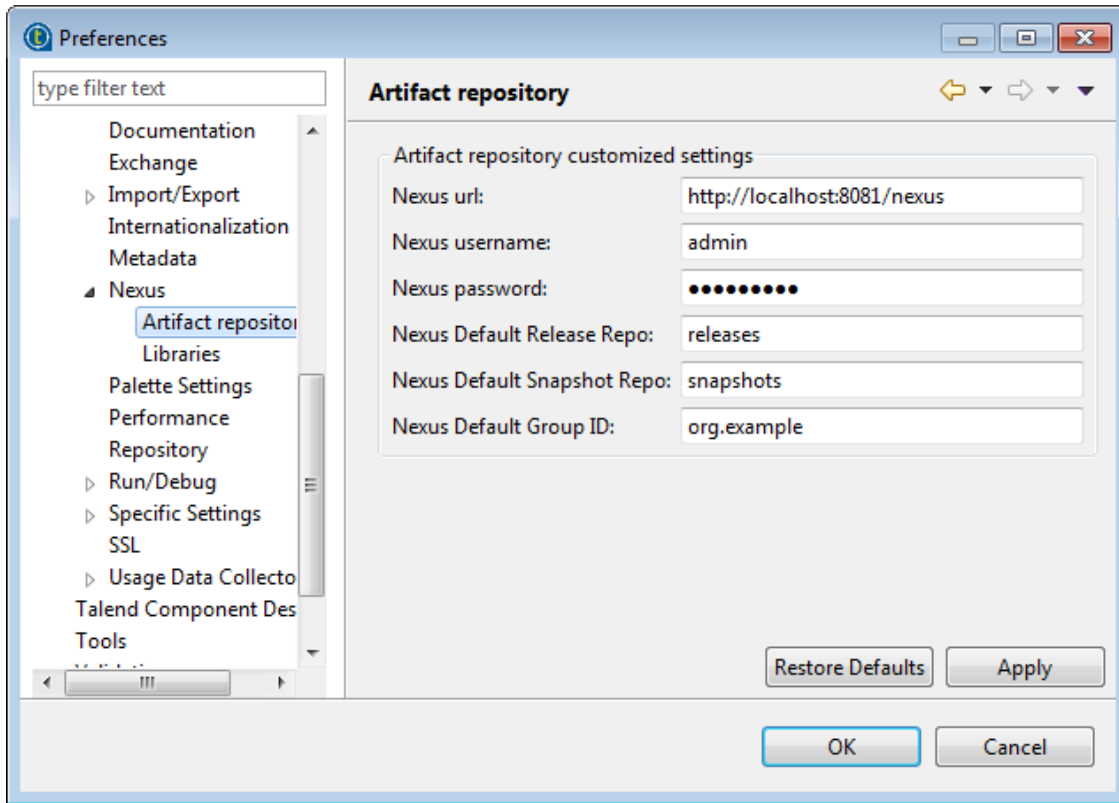
In the Studio, you can now configure the access to your Nexus Artifact Repository in its **Preferences** and publish your Services, Routes and Jobs into the two repositories according to your needs. For more information, see the *Talend Studio User Guide*.

Once you have installed and launched the Nexus Artifact Repository, open your Studio and do the following:

1. Click **Window > Preferences** from the menu bar to open the **[Preferences]** dialog box.
2. In the tree view, expand the **Talend > Nexus** nodes and select **Artifact repository**.
 - If your *Talend Studio* is connected with the *Talend Administration Center*, all the Nexus artifact repository settings are automatically retrieved from the *Talend Administration Center*. You can choose to use the retrieved settings to publish your Services, Routes and Jobs, or configure your own artifact repositories.



- If your Studio is working on a local connection, all the fields are pre-filled with the locally-stored default settings.



3. When connected with the *Talend Administration Center*, by default, the Studio checks the latest artifact repository settings each time it interacts with the artifact server. To disable this, if the artifact repository settings are not subject to frequent changes or if you have a poor internet connection, for example, clear the **Always check latest settings** check box.
4. When connected with the *Talend Administration Center*, if you want to configure your own artifact repositories, select the **Use customized settings** option.
5. Modify the artifact repository settings according to your needs.
 - **Nexus url:** Type in the location URL of your repository.
 - **Nexus username:** Type in the username to connect to your repository.
 - **Nexus password:** Type in the password to connect to your repository.
 - **Nexus Default Release Repo:** Type in the name of the repository into which to publish the Release version of your artifact items by default.
 - **Nexus Default Snapshot Repo:** Type in the name of the repository into which to publish the Snapshot version of your artifact items by default.
 - **Nexus Default Group ID:** Type in the name of the group in which to publish your artifact items by default.
6. Click **Apply** to apply your changes and **OK** to close the wizard.

Now, you will be able to publish your Services, Routes and Jobs onto your Artifact repository. For more information on how to publish Services, Routes and Jobs, see the *Talend Studio User Guide*.

3.2.4. 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 or by some connection wizards or by both. Due to license restrictions, *Talend* may not be able to ship certain external modules within *Talend Studio*.

3.2.4.1. Identify required external modules

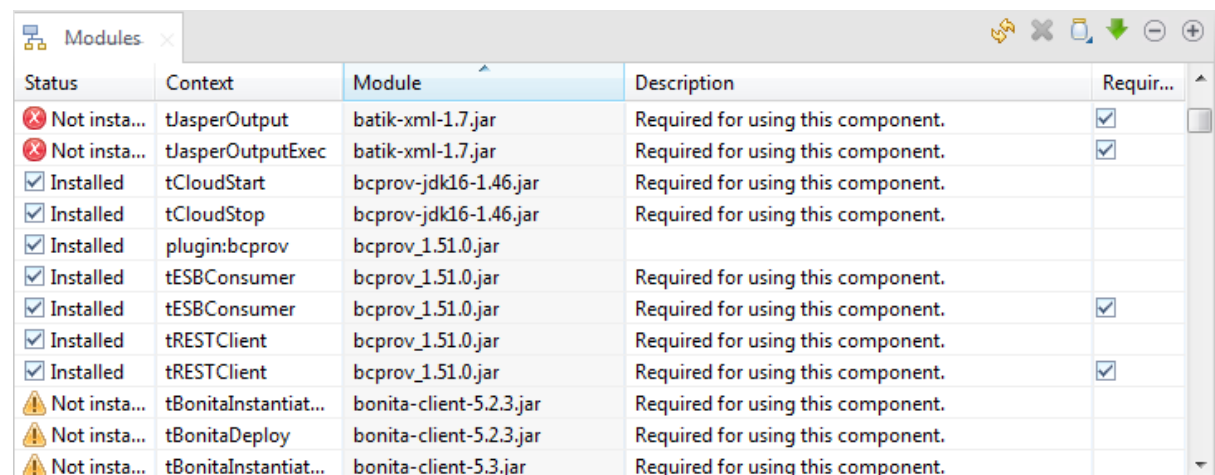
When you launch *Talend Studio* or select **Help > Install Additional Packages** in the *Talend Studio* menu, if any external modules are found missing for any features, the **[Additional Talend packages]** wizard opens, showing the **Optional** and **Required third-party libraries** check boxes. Make 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 or Metadata connection 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.



Status	Context	Module	Description	Required
Not installed	tJasperOutput	batik-xml-1.7.jar	Required for using this component.	<input checked="" type="checkbox"/>
Not installed	tJasperOutputExec	batik-xml-1.7.jar	Required for using this component.	<input checked="" type="checkbox"/>
Installed	tCloudStart	bcprov-jdk16-1.46.jar	Required for using this component.	<input type="checkbox"/>
Installed	tCloudStop	bcprov-jdk16-1.46.jar	Required for using this component.	<input type="checkbox"/>
Installed	plugin:bcprov	bcprov_1.51.0.jar		<input type="checkbox"/>
Installed	tESBConsumer	bcprov_1.51.0.jar	Required for using this component.	<input type="checkbox"/>
Installed	tESBConsumer	bcprov_1.51.0.jar	Required for using this component.	<input checked="" type="checkbox"/>
Installed	tREStClient	bcprov_1.51.0.jar	Required for using this component.	<input type="checkbox"/>
Installed	tREStClient	bcprov_1.51.0.jar	Required for using this component.	<input checked="" type="checkbox"/>
Not installed	tBonitaInstantiat...	bonita-client-5.2.3.jar	Required for using this component.	<input type="checkbox"/>
Not installed	tBonitaDeploy	bonita-client-5.2.3.jar	Required for using this component.	<input type="checkbox"/>
Not installed	tBonitaInstantiat...	bonita-client-5.3.jar	Required for using this component.	<input type="checkbox"/>


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 or Metadata connection listed in the Context column. The  icon indicates that the module is absolutely required for the corresponding component or Metadata connection.
Context	lists the name of Talend component or Metadata connection using the module. If this column is empty, the module is then required for the general use of <i>Talend Studio</i> .

Column	Description
	 This column lists any external libraries added to the routines you create and save in the Studio library folder. For more information, see the <i>Talend Studio User Guide</i> .
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.

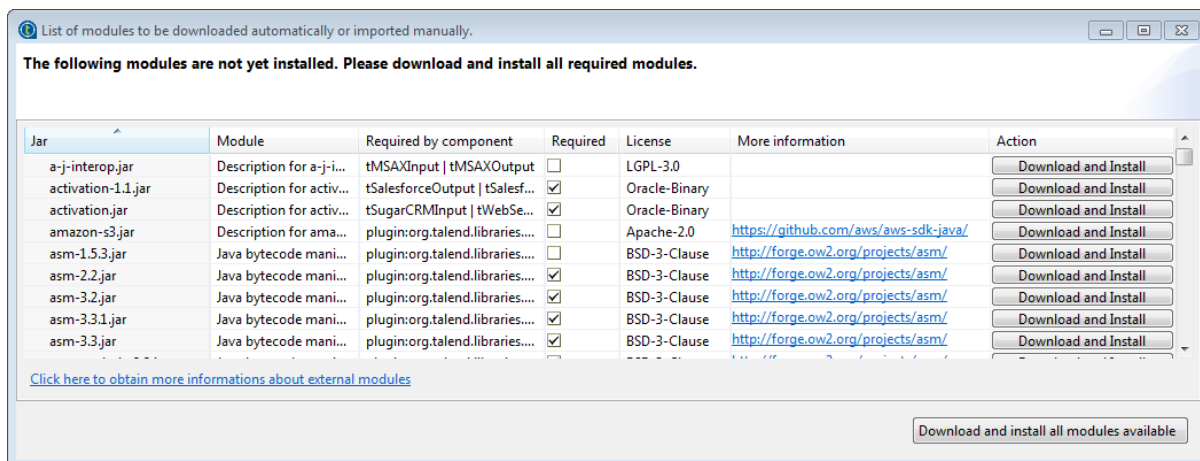
In addition to the **Modules** view, the Studio provides a mechanism that enables you to easily identify, download and install most of the required third-party modules from the **Talend** website and directs you to valid websites for the rest.

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 **Check** button in a Metadata connection setup wizard in *Talend Studio* if one or more external modules required for the connection 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 or Metadata connections for which one or more required external modules are missing, or
- click the  button in the **Modules** view.

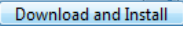


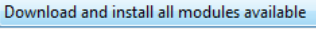



When you click this button, the wizard that appears will list all the required external modules that are not integrated in the Studio.



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.

Item	Description
Action	<p>: Click to open the [Download external modules] dialog box to download and install the module, which is available on the Talend website;</p> <p>: 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;</p> <p>: You need to find and download the module yourself and click the jar button to import it into the your studio.</p>
	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	<p>Select to prevent the wizard from appearing again unless you click the  button in the Modules tab view.</p> <p>This check box shows only when you drop a component, set up a connection, 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.</p>
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, 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.



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

3.2.4.2. Install external modules

To download and install modules in the Studio

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



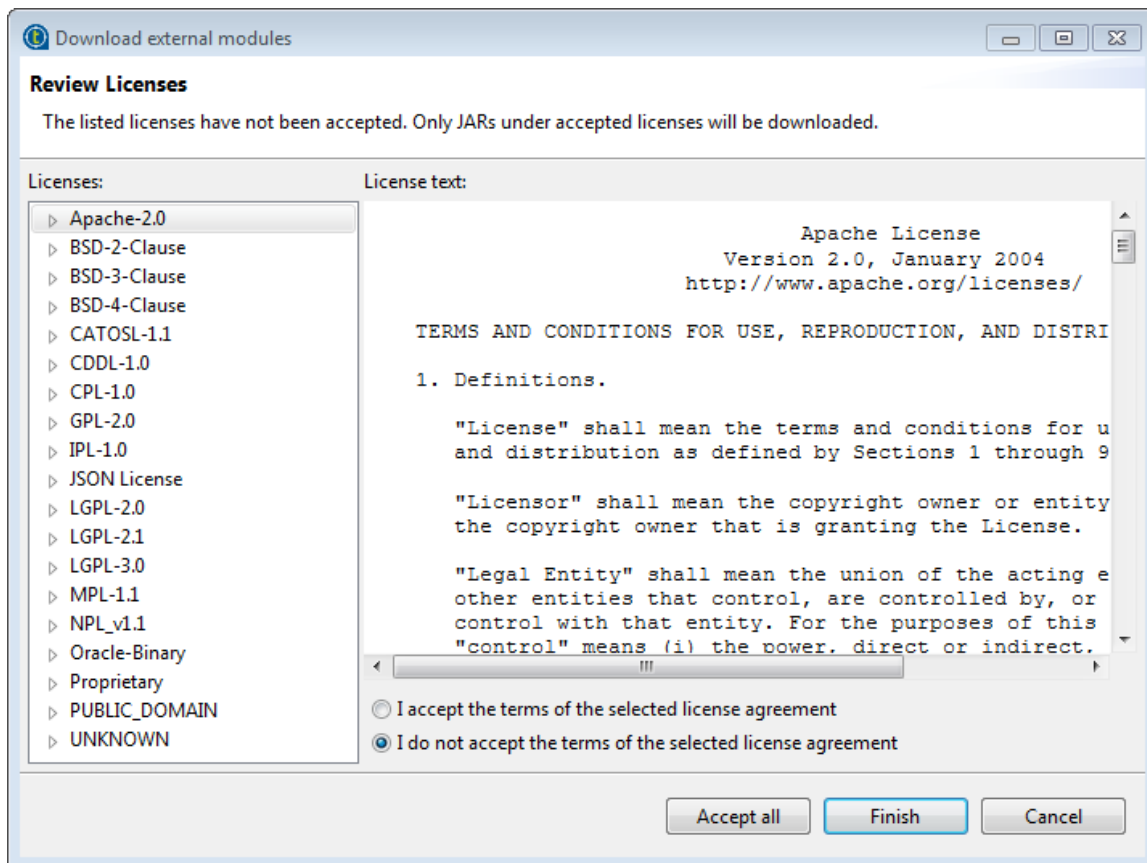
*If you are working behind a network proxy, make sure you have correctly set up your proxy before you can download and install external modules in your Studio. To access the proxy settings, select **Window > Preferences** from the menu to open the **[Preferences]** window, then expand the **General** 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.



- 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 Oracle9i, 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 Oracle9i 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.

- 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.

- 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 or Metadata connection dependent on this module in any of your Job designs.



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 studio(s).

To install modules in CommandLine

If you use the Studio and CommandLine on different machines, you need to retrieve the downloaded `.jar` files and add them in CommandLine.

1. Make sure CommandLine is not started, then download the external modules from the **Modules** view as explained in the previous procedure.
2. Copy the downloaded `.jar` files from `<StudioPath>/configuration/.m2` and paste them into `<CommandLinePath>/configuration/.m2`, where `<StudioPath>` and `<CommandLinePath>` are the installation directories of the Studio and CommandLine respectively.

Since these folders are hidden, make sure your system is configured to show hidden files and folders.

The `<CommandLinePath>/configuration/.m2` folder is not created by default. It is created the first time you start the CommandLine application.

3. Restart CommandLine.

You can now use the component or Metadata connection dependent on these modules.

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`

For Talend Administration Center, the downloaded modules must be placed in the following folder:

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



Appendix A. Appendices

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

- *[Introduction to the Talend products](#)*
- *[Architecture of the Talend products](#)*
- *[Talend High Availability](#)*
- *[H2 Database Administration & Maintenance](#)*
- *[Supported Third-Party System/Database/Business Application Versions](#)*

A.1. Introduction to the Talend products

The present section lists all the elements required for using the Talend products. To ease their management, we recommend that you centralize all the server modules on one single system.



All Talend applications to be installed must be the same version.

- An application server (Apache Tomcat server) that hosts *Talend Administration Center*.
- A database server storing the administration metadata of *Talend Administration Center* (by default, an embedded H2 database is used).
- A database server storing data quality reports and analyses managed by *Talend Data Quality Portal*.
- A database server storing golden or master data records managed by *Talend Data Stewardship Console* (if your license includes the Data Stewardship option).

The *Talend Data Stewardship Console* is deprecated since Talend 6.4. Consider migrating to *Talend Data Stewardship*.

- A version control system for Project metadata.
- A Web browser to access Web application:
 - *Talend Administration Center* where projects, users and processes can be managed and administrated. For more information, see the *Talend Administration Center User Guide*.
 - *Data Quality Portal* that extends the dynamic reporting and monitoring capabilities of Talend Studio. For more information, see the *Talend Data Quality Portal User and Administrator Guide*.
- A Nexus artifact repository in which are stored software updates, external libraries and artifacts.
- Execution servers (JobServers) or Talend Runtime execution containers (based on Apache Karaf) to deploy and execute processes.
- A Studio API to carry out technical processes. For more information, see the *Talend Studio User Guide*.
- A *Talend Activity Monitoring Console*, providing detailed monitoring capabilities that can be used to consolidate the collected log information. For more information, see the *Talend Activity Monitoring Console User Guide*.

Each of these elements is detailed in the following sub-sections.

A.1.1. Apache Tomcat Server

The Apache Tomcat server is an application server that hosts *Talend Administration Center*. This Web application gives access to all management and administration functionalities for an integration project, allowing users to (depending on their role):

- Create and manage projects.
- Create and manage user accounts and roles/rights.
- Access the Publisher to publish and schedule artifact publishing on the Nexus artefact repository.
- Access the Job Conductor to schedule, deploy and execute Jobs.
- Access the ESB Conductor to set the deployment and execution of Services, Routes and Generic OSGi features.
- Access the Monitoring node to monitor the execution of Jobs and visualize the logs.

- Access the ESB Infrastructure to monitor the endpoints activity through Service Locator and monitor all service events through Service Activity Monitoring.



Talend Administration Center can also be hosted by JBoss or Pivotal tc application servers.

A.1.2. Database

The administration database server is used to store administration information and manage the persistence in Talend Administration Center. By default an embedded H2 database is used, but you can also use MySQL, MS SQL Server, or Oracle to store all cross-project data (users, projects, authorization, license, tasks, triggers, monitoring).

The administration database will be named *<talend_administrator>* in the rest of this document.

The *<talend_administrator>* administration database will contain all the data related to project information and administration including: administration data, project declaration, user declaration and authorization, task list, etc.

The tables in this database are automatically created when connecting for the first time to *Talend Administration Center*. The created tables include (among others):

- a *Users* table,
- a *Projects* table,
- a *Rights* table.



These tables are created, populated and managed automatically by Talend, users do not need to take any action.

Additionally, to perform Data Quality operations, a HSQL, MySQL or Oracle database is used to store report and analysis results and a stewardship console database is used to store golden or master data records. For more information about those databases, see the relevant sections in the present guide.

A.1.3. Version control system

A version control system (either Git or SVN) is used to manage the persistence of all the data relating to the objects (Jobs, Services, Routes, Business Models, Metadata, Routines, Documentation, etc.) stored in "git" or "svn" mode in the shared Repository of the Talend Studio.

We recommend you to store several projects per repository, simply in order not to have too many repositories to deal with. However you can choose to store only one project per Git or SVN repository, if you prefer so.

For more information on how to configure your version control systems, see [Setting up your version control system](#).

You can also have several version control repositories each containing several projects. For more information on how to create projects and store them in Git or SVN, see the *Talend Administration Center User Guide*.

A.1.4. Artifact Repository

The artifact repository delivered by *Talend* and based on Sonatype Nexus is a preconfigured application centralizing the management and usage of the **Software Update**, **User libraries** and **snapshots** and **releases** repositories:

- **Software Update** is used to manage application updates (patches) distributed by *Talend*. By default the *talend-updates* repository is embedded within **Software Update** and retrieves the updates published by *Talend*. This repository allows the user to visualize the updates available.

For detailed information, see [Software update repository](#).

- The **User libraries** repository is used to store all external libraries. These libraries are retrieved by Talend Studio at start-up and shared with Talend Administration Center via the *talend-custom-libs* repository.

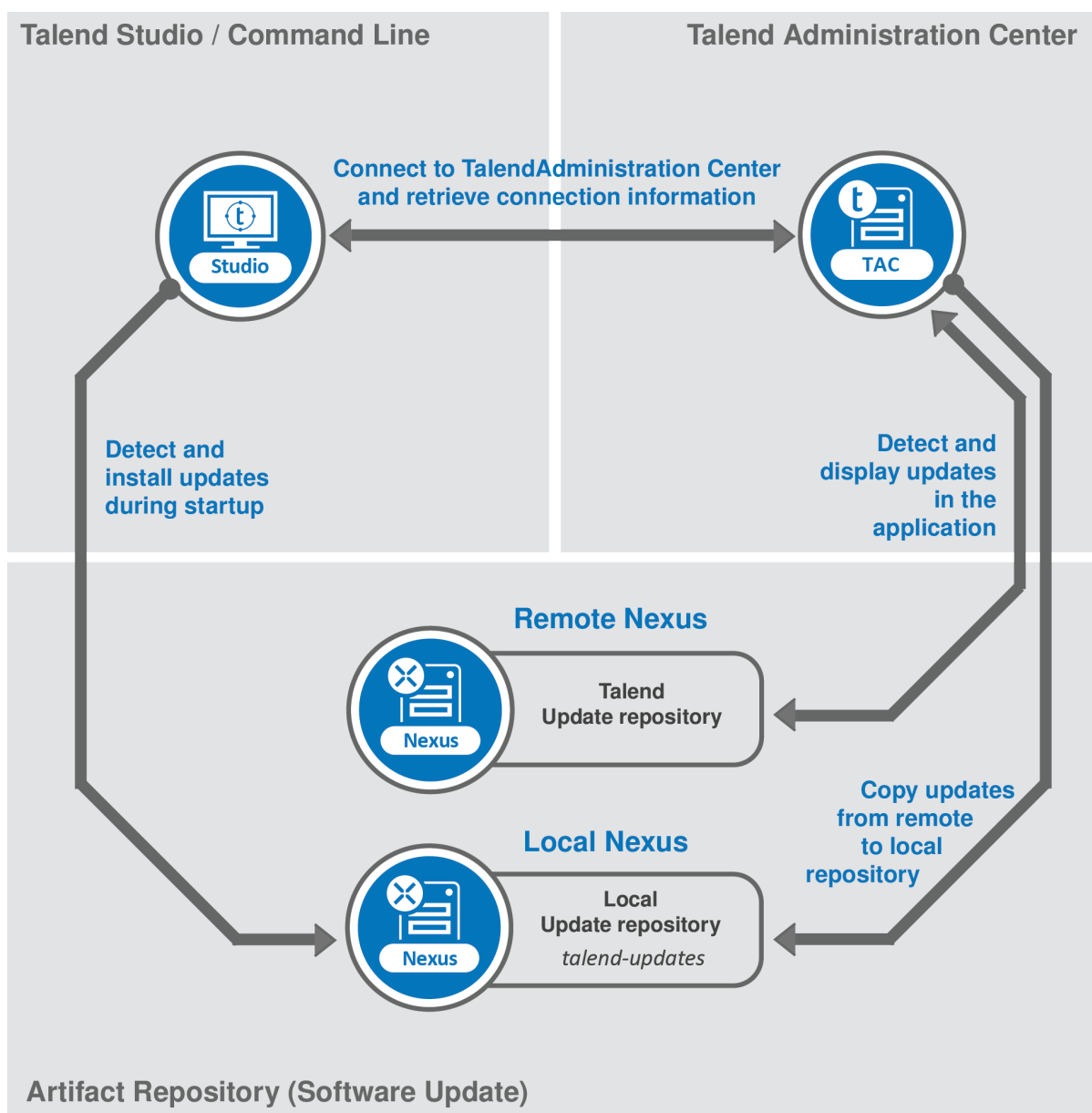
For detailed information, see [User Libraries repository](#).

- The **snapshots** and **releases** repositories are used as a catalog in which all artifacts to be deployed and executed are stored. These artifacts are designed by the user from the Studio or any other Java IDE. By default, the **snapshots** repository is used for development purposes and the **releases** repository is used for production. These repositories make artifacts available for deployment and or execution in an execution server.

For detailed information, see [Snapshots and Releases artifact repositories](#).

A.1.4.1. Software update repository

The following image shows the architecture of **Software Update** linked to *Talend Administration Center* and to the Talend Studio.



To download and install some software updates, you need to connect to **Software Update** (integrated within the Nexus artifact repository) and its embedded repository named *talend-updates*.



The *Artifact-Repository-Nexus-VA.B.C* archive file containing Nexus is embedded in the *Talend Administration Center* archive file that is provided by Talend.

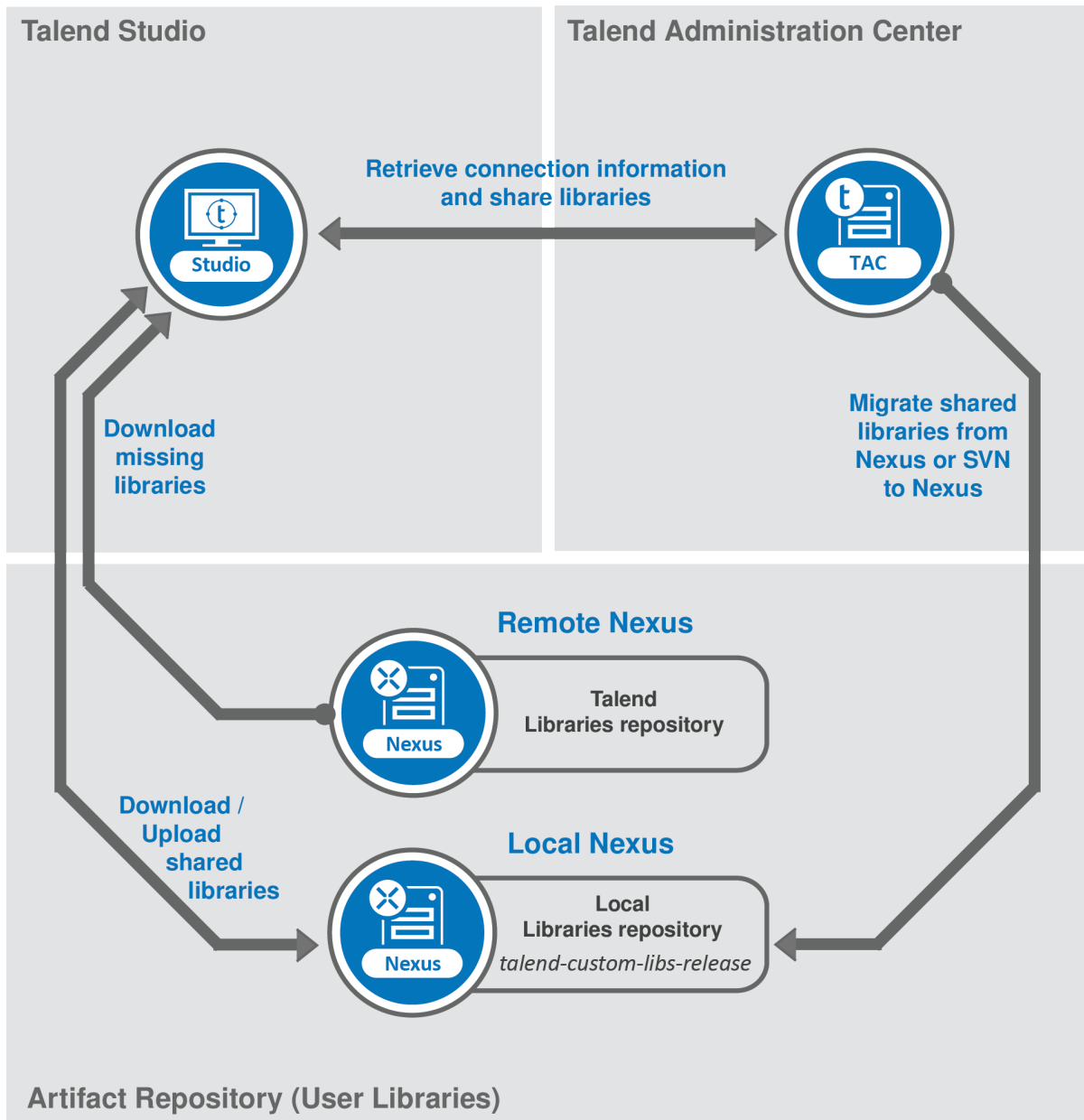
In *Talend Administration Center*, the patches available for the current version that have been copied from the Talend remote repository to the local *talend-updates* repository are detected and the administrator can accept them.

Talend Studio is connected to *Talend Administration Center* to retrieve the repository connection information and the updates are detected and installed automatically.

For more information on how to check updates via these repositories, see the *Talend Administration Center* and *Talend Studio User Guides*.

A.1.4.2. User Libraries repository

The following image shows the architecture of the **User Libraries** repository.



To download and install some specific third-party Java libraries or database drivers that are needed by *Talend Studio*, you need to connect to the **User Libraries** repository (integrated within the Nexus artifact repository) and its embedded repository named *talend-custom-libs-release*.



The *Artifact-Repository-Nexus-VA.B.C* archive file containing Nexus is embedded in the *Talend Administration Center* archive file that is provided by Talend.

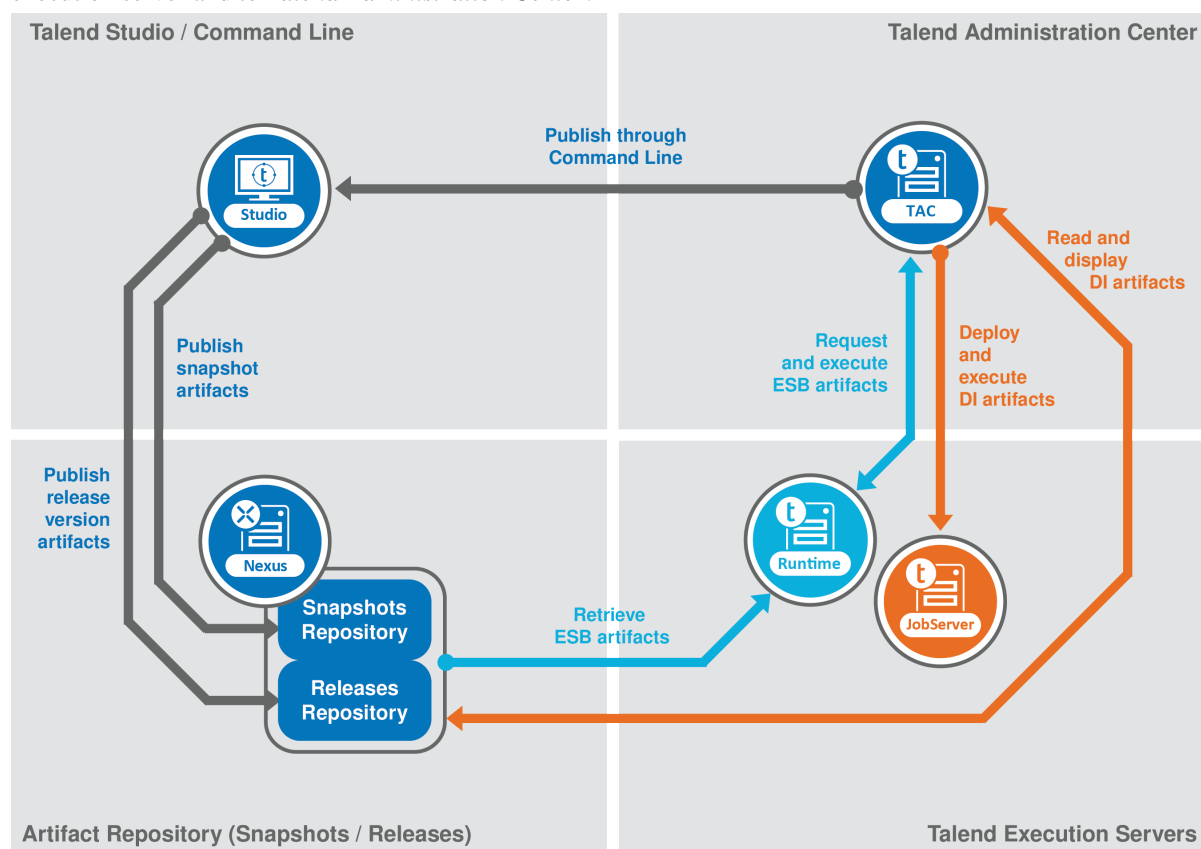
When Talend Studio opens, the external libraries missing from the local *talend-custom-libs-release* repository are detected. You are prompted to download them from the remote Nexus repository, hosted by Talend, and install them.

Talend Administration Center is connected to *Talend Studio* and to the local repository and the installed libraries are shared automatically.

For more information on how to install external libraries, see [Installing external modules](#).

A.1.4.3. Snapshots and Releases artifact repositories

The following image shows the architecture of the Snapshots and Releases repositories linked to the Studio, to an execution server and to *Talend Administration Center*.



The artifact repository is also used to store as artifacts all the Services, Routes and Jobs created in Studio or any Generic OSGi Feature created in any other Java IDE.

From the Studio, you can publish those artifacts in the *snapshots* and *releases* repositories (integrated within Nexus). The artifacts are provided to an execution server and then can be selected through *Talend Administration Center* in order to set their deployment.

Talend Administration Center also fetches the artifact status and URLs and display them in its page **Job Conductor** and **ESB Conductor**.

When the deployment of an artifact is initiated in *Talend Administration Center*, the execution server requests the corresponding artifact in the artifact repository. Then, the artifact can be deployed and executed.

Two embedded repositories are provided to store your artifacts:

- a *snapshots* repository to publish snapshot artifacts for development purposes,
- a *releases* repository to publish stable artifacts for production purposes.



The *Artifact-Repository-Nexus-VA.B.C.D.E* archive file containing Nexus is embedded in the *Talend Administration Center* archive file that is provided by Talend.

A.1.5. Talend Runtime

Talend Runtime (based on Apache Karaf) is an execution container in which you can deploy and execute all your Jobs stored on your Git or SVN repository.

It is also used to deploy and execute as OSGi bundles all your Services, Routes and Generic OSGi features that are retrieved from the **Provisioning** repositories.

A.1.6. JobServer

JobServer is an application that allows a system installed on the same network as *Talend Administration Center* to declare itself as an execution server. These systems must obviously have a working JVM.

A.1.7. Talend Studio

Talend Studio is a rich client that allows the user (such as a project manager, a developer or a DBA) to work on any **Talend** project for which he has authorization.

Talend Studio offers a comprehensive set of tools and functions for all its key capabilities including:

- Integration
- Activity monitoring Console
- Data profiling
- Data quality
- Mediation
- SoapUI

These tools are ALL accessible in different perspectives from one Studio.



The availability of perspectives in your Studio depends either on the license you have when you are working in a local project, or on the type of the remote project itself when you are working in remote projects.

For further information on user authorization on remote project, see the *Talend Administration Center User Guide*.

For further information about the different perspectives available in the studio, see the *Talend Studio User Guide*.

For more information on how to install *Talend Studio*, see [Installing your Talend Studio](#).

A.1.8. Talend Activity Monitoring Console log database

If you want to use the *Talend Activity Monitoring Console*, an <AMC> log database must be created, which can be installed on any server. This <AMC> database will initially be empty. Its name may be modified, but you must take into account this modification in the rest of this document.

The <AMC> database will contain three tables that collect data allowing users to monitor Jobs. The three tables will collect data from the following components:

- **tFlowMeterCatcher**,
- **tLogCatcher**,
- **tStatCatcher**.

Instructions on how to create these tables and their structure is provided in the *Talend Activity Monitoring Console User Guide*.

A corresponding SQL user must be created and thus mapped to have access to this database. This user should be granted the "create" and "update" rights.



The installation of the <AMC> database is optional for the solutions that only hold the ESB option, and is only required when using the **tFlowMeterCatcher**, **tLogCatcher** and **tStatCatcher** components. These components can be used in Jobs (for example, ESB Consumer jobs) for REST and Soap ESB Service Providers. Only the **tLogCatcher** component is supported for now and Routes currently can not use these components.

A.2. Architecture of the Talend products

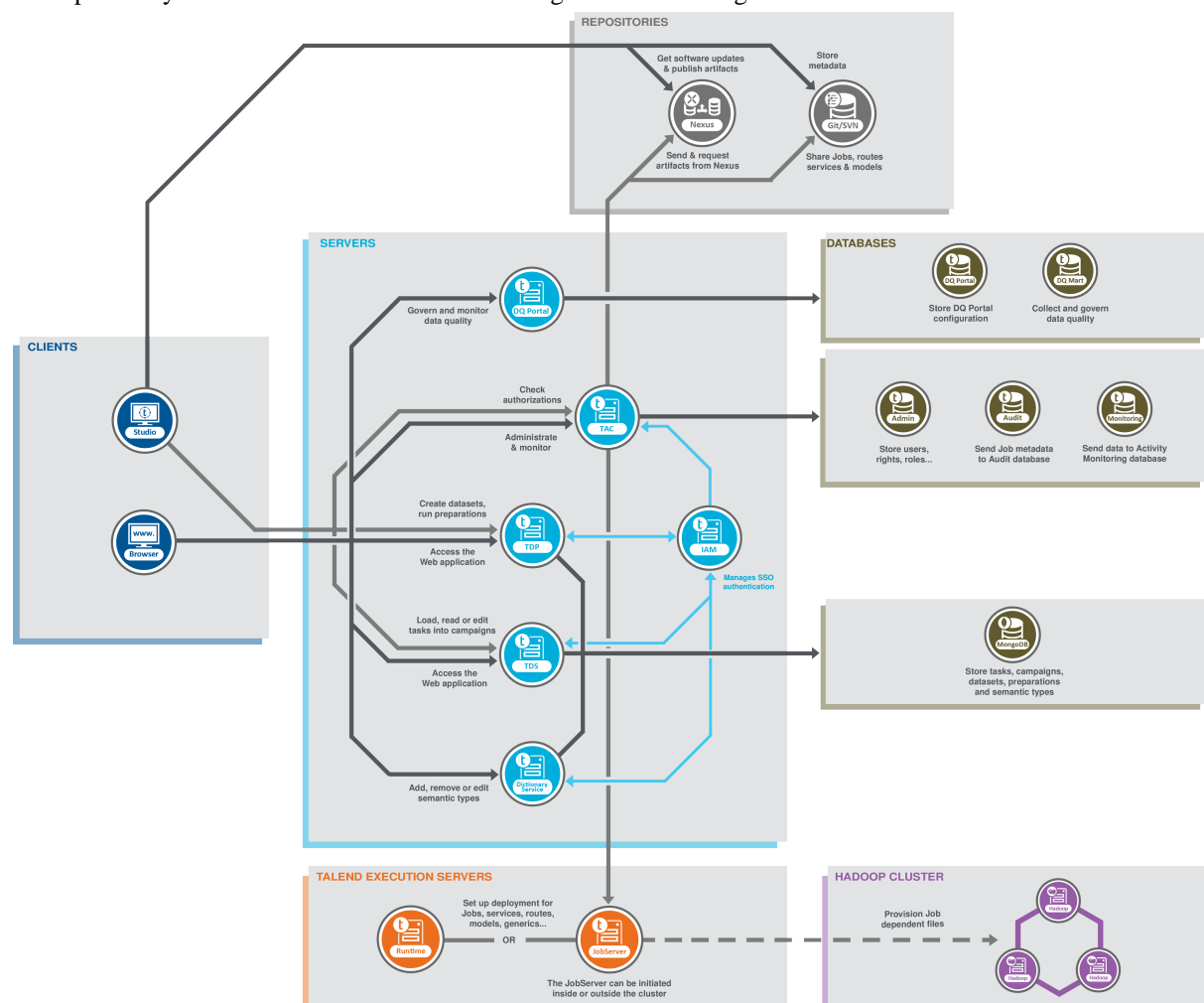
The operating principles of the Talend products could be summarized as briefly as the following topics:

- building technical or business-related processes and data profiling analyses,
- administrating users, projects, access rights and processes and their dependencies,
- deploying and executing technical processes,
- monitoring the execution of technical processes.



Depending on your license, some of the functional blocks may not be available to you.

Each of the above topics can be isolated in different functional blocks and the different types of blocks and their interoperability can be described as in the following architecture diagram :



Building and administrating

- The **CLIENTS** block includes one or more Talend Studio APIs and Web browsers that could be on the same or on different machines.

From the Studio API, end-users can carry out technical processes: data integration or data service processes, mediation routes and services, and publish them on the Artifact Repository and data profiling analyses and reports regardless of data volume and process complexity. The *Talend Studio* allows the user to work on any project for which he has authorization. For more information, see the *Talend Studio User Guide*.

From a Web browser, end-users connect to the remotely based Administration Center through a secured HTTP protocol. The end-user category in this description may include developers, project managers, administrators and any other person involved in building data flows, Web, REST and data services, and mediation routes. Each of these end-users will use either the Studio or the Administration Center or both of them depending on the company policy.

End-users also use a Web browser to connect to *Talend Data Quality Portal* that plugs to the data quality data mart and extends the dynamic reporting and monitoring capabilities of the *Talend Studio*, and to *Talend Data Stewardship Console* to resolve the data issues included in the listed tasks to reach the golden data record (if they subscribed to this option).

Additionally, from the Web Browser you access the *Talend Data Preparation* Web application. This is where you import your data, from local files or other sources, and cleanse or enrich it by creating new preparations on this data. You can also access the *Talend Data Stewardship* Web application. This is where campaign owners and data stewards manage campaigns and tasks. You can optionally access the *Talend Dictionary Service* server to add, remove or edit the semantic types used on data in the Web applications.

- The **TALEND SERVERS** and **DATABASES** blocks and the **Git/SVN** grey circle include a web-based Administration Center (application server) connected to two shared repositories: one based on a Git or SVN server and one based on a database server (**Admin**).

The Administration Center enables the management and administration of all projects. Administration metadata (user accounts, access rights and project authorization for example) is stored in the database server and project metadata (Jobs, Business Models, Routines, Routes, Services for example) is stored in the Git or SVN server (to easily share them between the different end-users).

The Administration Center also enables to configure the tasks that handle job executions and triggers. It also looks after the job generation and deployment to the execution servers. For more information, see the *Talend Administration Center User Guide*.

The Administration Center also includes the servers used by the Talend Web applications, namely *Talend Data Preparation* and *Talend Data Stewardship*, and also *Talend Dictionary Service*. The *Talend Identity and Access Management* server is used to enable Single Sign-On between those applications.

Finally, the Administration Center enables you to access and manage the Routes or Services created from *Talend Studio* and published into the **Artifact Repository**, and set up and monitor their deployment and execution into the **Talend Runtime**. For more information, see the *Talend Administration Center User Guide*.

Deploying and executing

- The **Nexus** grey circle represents the artifact repository that stores all the:
 - **Software Updates** available for download.
 - Routes and Services that are published from the Studio and are ready to be deployed and executed in **Talend Runtime**.
- The **TALEND EXECUTION SERVERS** block represents the execution servers that run technical processes according to the execution scheduling set up in the *Talend Administration Center* Web application. Those execution servers can be of:

- One or more **Talend Runtimes** (execution container) deployed inside your information system. The Talend Runtime deploys and executes the technical processes according to the set up defined in the *Talend Administration Center* Web application. Those processes are Jobs built from the Studio and centralized on the Git or SVN server. Routes and Services retrieved from the artifact repository.

If you have several Talend Runtimes on which to deploy the Service and Route artifacts, you will be able to load balance their execution according to your needs. All instances of Talend Runtime will communicate between each other via the Service Locator to identify the one more likely to deploy and execute the artifact(s) set to deployment in *Talend Administration Center*. The Talend Runtime elected for the deployment will request for the artifact(s) to deploy and execute from the artifact repository and the artifact repository will thus send the artifact(s) requested along with all the dependencies needed for its/their execution to the Talend Runtime, that will deploy and execute it/them.

- One or more JobServers deployed inside your information system that run technical processes (Jobs) according to scheduled time, date or event set in the *Talend Administration Center* Web application.

The end-user can transfer technical processes to a remote execution server directly from the Studio (distant run).



You must install the JobServer files ("Agent"), delivered by **Talend**, on each of the execution servers to become operational.

Monitoring

- The **Monitoring** circle represents the monitoring: the Activity Monitoring Console and the Service Activity Monitoring.

The Activity Monitoring Console allows end-users to monitor the execution of technical processes. It provides detailed monitoring capabilities that can be used to consolidate log information collected, understand the interaction between underlying data flows, prevent faults that could be unexpectedly generated and support system management decisions. For more information on the Activity Monitoring Console, see the *Talend Activity Monitoring Console User Guide*.

The Service Activity Monitoring allows end-users to monitor service calls. It provides monitoring and consolidated event information that can be used to understand the underlying requests and replies that compose the event, monitor faults that may be unexpectedly generated and support the system management decisions. For more information on the Service Activity Monitoring, see the *Talend Administration Center User Guide*.

A.3. Talend High Availability

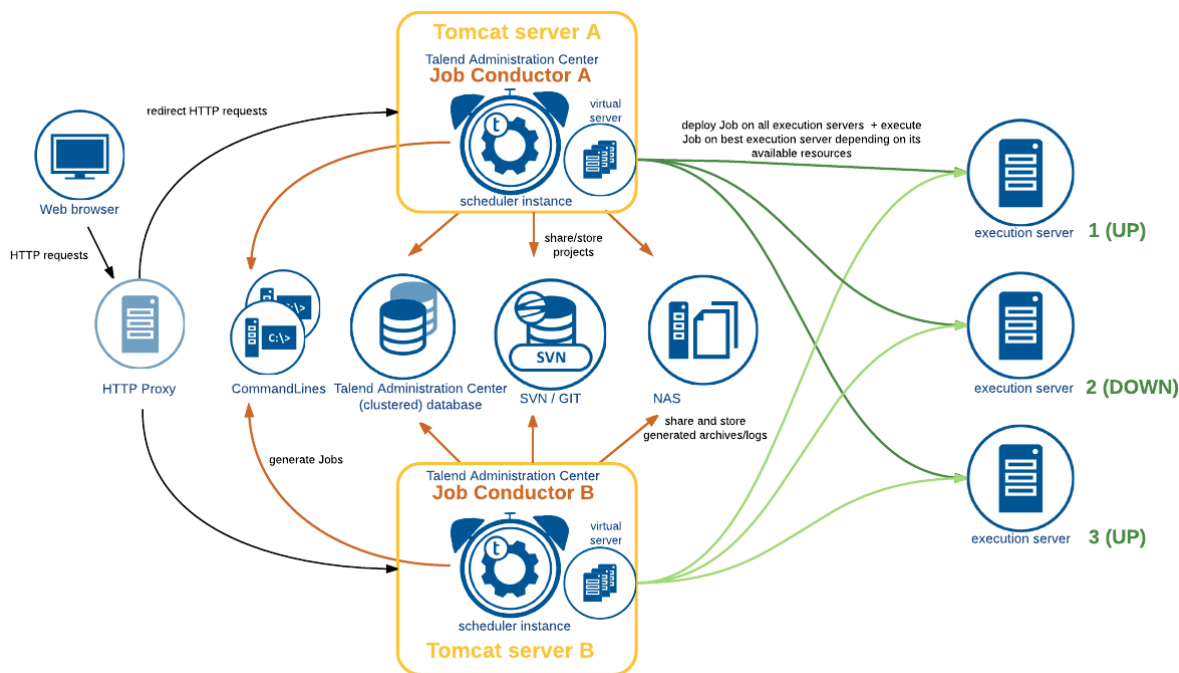


The availability of this feature depends on your license.

You can set up a cluster in your *Talend* system to provide high availability and failover features for task execution scheduling in *Talend Administration Center*. You do this by deploying multiple Job Conductors and Job execution servers on different machines.



High availability in this context refers only to the scheduling of task executions.



To summarize:

- Two application servers (Tomcat or JBoss) holding the *Talend Administration Center* Job Conductors and Virtual Servers, as well as two CommandLine applications are installed (on different machines) and point to the same SVN/GIT shared project.
- All instances of the application server are connected to the project administration database, which is presumably clustered.
- (optional) *Talend Administration Center* users are routed to the same active application instance, for example through an HTTP Proxy (switch). This feature is not provided by Talend and thus needs to be implemented separately.
- The first CommandLine generates the artifacts to be deployed. The second CommandLine is only used when the first one is down.
- When an execution server fails, the other execution servers can recover the interrupted tasks.
- A shared storage is implemented to store and share between active instances all archives and logs generated during each Job execution, for example through a Network-Attached Storage (NAS). This feature is not provided by Talend and thus needs to be implemented separately.

For more information about the failover and the various actions you can perform on a task when a server is unavailable, see the *Talend Administration Center User Guide*.

A.4. H2 Database Administration & Maintenance

This Chapter provides information about how to manage and back up the H2 embedded database.

For more information about how to use the H2 database and web console, refer to the H2 database documentation at <http://www.h2database.com>.

A.4.1. About H2 embedded database

H2 is a relational database management system written in Java. It can be embedded in Java applications or run in the client-server mode.

This database is the default solution embedded in *Talend Administration Center* to store all cross-project information such as users, authorizations, projects...

If you do not want to use this default database, you can set up a database server (MySQL, MSSQL or Oracle).

The benefits of using this H2 embedded database is that it simplifies the installation process.

A.4.2. Administrating the H2 database through the Web console

To help you administrate the H2 embedded database, a dedicated Web console is available directly from *Talend Administration Center*.

A.4.2.1. Connecting to the H2 Web Console

From *Talend Administration Center*, you can access the H2 administration console:

1. From the main **Menu**, click **Configuration** to access the **Configuration** page.
2. On the **Configuration** page, expand the **Database** node to display the parameters.

Database (5 Parameters)	
Url	jdbc:h2:talend/integration-web-app/tis_410/r48628/ADM-TDOMPX-r48628-V4.1.0RC2/org.talend.administrator/WEB-INF/database/talend_administrator;AUTO_SERVER=TRUE;MVCC=TRUE
User	tisadmin
Password	*****
Driver	org.h2.Driver
Web Console	http://10.42.10.66:9090/tis410rc2/h2console

3. In the **Web Console** field, click the link to access the H2 Web Console.
4. The H2 Web Console's **Login** page displays:

English

Login

Saved Settings: Generic H2 (Embedded)

Setting Name: Generic H2 (Embedded)

Driver Class: org.h2.Driver

JDBC URL: jdbc:h2:/talend/integration-web-app/tis_410/r48628/A

User Name: tisadmin

Password: *****

5. In the **User Name** and **Password** fields, type in the connection login and password to the database, by default talend_admin and talend_admin.

6. The **JDBC URL** field reads by default:

```
jdbc:h2:./<ApplicationPath>/WEB-INF/database/
talend_administrator;AUTO_SERVER=TRUE;MVCC=TRUE;LOCK_TIMEOUT=15000
```

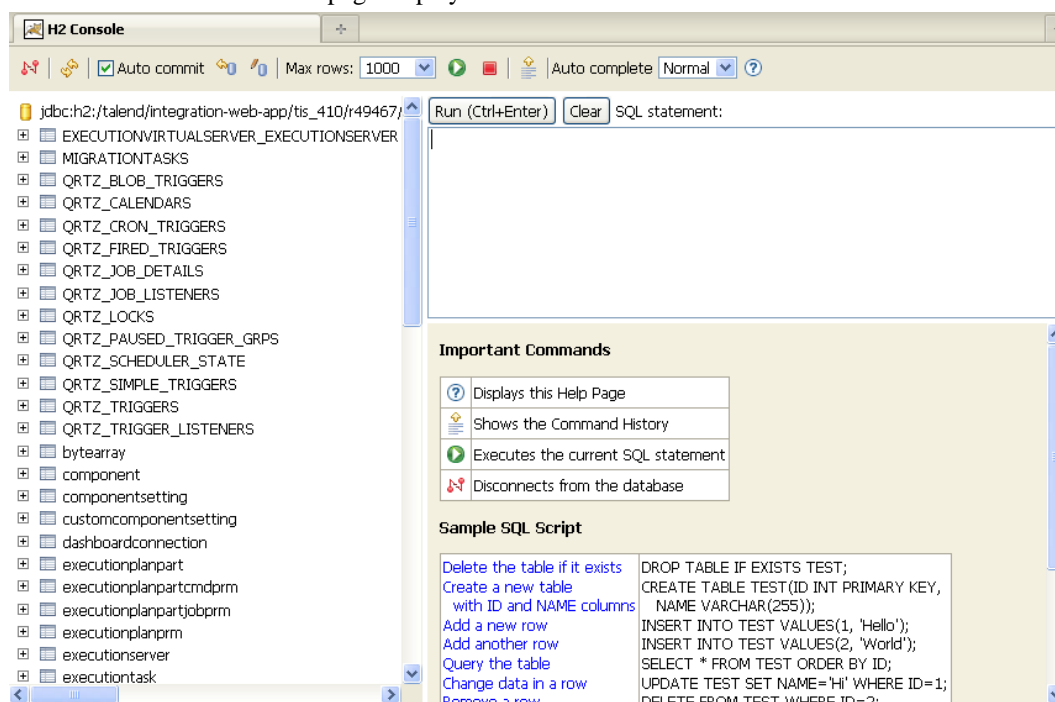
where <ApplicationPath> is the location where org.talend.administrator was deployed.



If you have moved the H2 embedded database location, then fill out the JDBC URL field with the updated URL information. Prior to clicking Connect, click the Test Connection button in order to check the new URL. In case of a mistyped URL, the JDBC URL will revert back to the original URL information.

7. Click **Connect**.

The Web database administration page displays.



For more information about H2 use and troubleshooting, please refer to the H2 online documentation on <http://www.h2database.com>.

A.4.2.2. Backing up the H2 database

The configuration parameters of the H2 database backup is already set by default so that the backup occurs on an daily basis.

If you need or want to make edits to this setting, edit the configuration file:

<ApplicationPath>/WEB-INF/classes/configuration.properties

The cron-based backup of the embedded database triggers everyday at 3.45am all year round. The syntax reads as follows "Seconds Minutes Hours Day-of-month Month Day-of-week Year", such as for example:

```
0 45 3 ? * * * (default setting - trigger every day at 3.45am)
```

```
0 45 5 ? * MON-FRI (every Monday, Tuesday, Wednesday, Thursday and Friday at 5.45 am)
```

More examples are available on <http://www.quartz-scheduler.org/documentation/quartz-2.2.x/tutorials/tutorial-lesson-06.html>.

Other automatic backups are performed at startup and shutdown of the application server:

```
database.embedded.backup.doBackupAtStartup=true
database.embedded.backup.doBackupAtShutdown=true
```

The backup files are stored at the following location, up to the 30 latest backups:

<ApplicationPath>/WEB-INF/database/backups

A.4.3. Setting up the H2 database for access from other machines

To allow other users to access the H2 database for centralized storage of cross-project information, do the following:

1. Stop Tomcat service if it is running.
2. Unzip your H2 database server package to any of your local drives. The latest H2 database server package is available at <http://www.h2database.com/html/download.html>.
3. Open a CMD window, navigate to the drive where the H2 database server package was unzipped, and change directory to *h2\bin*, which contains the *h2*.jar* file.
4. Start the H2 server as a service using the following command:

```
java -cp h2*.jar org.h2.tools.Server -tcp -tcpAllowOthers
-tcpPort <port_number>
```

Now other users can access the H2 database, but you still need to edit the database URL to make *Talend Administration Center* work.

To do so, proceed as follows:

1. Open the *configuration.properties* file in the <ApplicationPath>/WEB-INF/Classes folder, and edit the H2 database URL setting as follows:

```
database.url=jdbc:h2:tcp://<IP_address>:<port_number>/file:<ApplicationPath>/WEB-
INF/database/talend_administrator;AUTO_SERVER=TRUE;IFEXISTS=TRUE;MVCC=TRUE;
LOCK_TIMEOUT=15000
```

where <IP_address> is your IP address, <port_number> is the TCP port number specified in the command used to start the H2 server, and <ApplicationPath> is the location where *org.talend.administrator* was deployed.

2. Start the Tomcat service.
3. Start your *Talend Administration Center* Web application.

Now others can access and use the H2 database through the URL address.

A.5. 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.5.1. 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.5.2. 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 ⁵		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 ⁵		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 ²	Hive2 ²			Standalone	YARN		
	V2.1.0 (Deprecated) ⁵	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✗	✓ ■■	✓ ■■	✓ ■■
	V2.2.0 (Deprecated) ⁵	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✗	✗ ■■	✓ ■■	✓ ■■
	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 ⁵		Oozie	Pig	Spark		Sqoop	Talend MapReduce
					Hive1 ²	Hive2 ²			Standalone	YARN		
MapR	CDH 5.10 (YARN mode)	✓ ■■	✓ ■■	✓ ■■	✗	✓ ■■	✓ ■■	✓ ■■	✓	✓ ■■	✓ ■■	✓ ■■
	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	✓ ■■	✓ ■■	✓ ■■	✗	✓ ■■	✓ ■■	✓ ■■	✓	✓ ■■	✓ ■■	✓ ■■
Amazon EMR	5.2.0 (YARN mode) ⁴	✓ ■■	✓ ■■	✓ ■■	✗	✓ ■■	✓ ■■	✓ ■■	✓	✓ ■■	✓ ■■	✓ ■■
	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)	✓	✓■■	✓	✗	✓	✓	✓	✗	✓	✓	✓
Pivotal HD	EMR 5.4/5.5	✓	✓■■	✓	✗	✓	✓	✓	✗	✓	✓	✓
	1.0.1 (deprecated)	✓	✗	✓	✓	✗	✗	✓	✗	✗	✓	✓
	2.0 (deprecated)	✓ ■■	✗	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✓ ■■	✗	✗	✓ ■■	✓ ■■

		HBase	HCatalog	HDFS	Hive ⁵		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.

5. 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.

For further information about what Hadoop distributions are supported from the **Profiling** perspective, see [Supported Hive distributions for profiling data](#).

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

A.5.3. Supported Hadoop distribution versions for Talend Data Preparation with Big Data

The following table lists the supported Hadoop distributions for Talend Data Preparation with Big Data.

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](#).

Distribution	Supported version
HDP	2.4 and above
Cloudera	5.7 and above
EMR	4.5 and above
Hadoop	2.7 and above

A.5.4. Supported databases for profiling data

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

Database name	Database version
Amazon Redshift	Initial release of Amazon Redshift
AS/400	V5R2 to V7R2
Hive	For a complete list of the supported Hive distributions, see Supported Hive distributions for profiling data .
IBM DB2	10.5
IBM DB2 ZOS	2000/2005/2008
Impala (a sub-module of Cloudera)	CDH5 5.0/5.1
Informix	11.50
Ingres	9.2
Microsoft SQL Server	2000/2003/2005/2008/2012
MySQL	5.1/5.5/5.6
Netezza	6
Oracle with SID	9i to 11g
Oracle with service name	9i to 11g
PostgreSQL	8.3/9.1+
SQLite	3.6.7
Sybase (ASE and IQ)QLite	12.5/12.7/15.2
Teradata	12/13/14/15
Vertica	6.0/7.0

A.5.5. Supported Hive distributions for profiling data

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](#).

Below is a list of the supported distributions for Hive profiling.



For the time being, the embedded mode on Hive distributions is available mainly for test purposes done by Hadoop developers. The studio may not be able to run correctly with the embedded mode.

Hive distribution		version	
		Hive 1	Hive 2
HortonWorks	HDP 1.0.0 (deprecated)	Embedded and Standalone	No
	HDP 1.2	Embedded and Standalone	Embedded and Standalone
	HDP 1.3	Embedded and Standalone	Embedded and Standalone
	HDP 2.0	Embedded (only Linux) and Standalone	Embedded (only Linux) and Standalone
	HDP 2.1	Embedded (only Linux) and Standalone	Embedded (only Linux) and Standalone
	HDP 2.2	Embedded (only Linux) and Standalone	Embedded (only Linux) and Standalone
Cloudera	CDH4 (Kerberos authentication is supported)	Embedded and Standalone	Embedded and Standalone
	CDH5 (Kerberos authentication is supported)	Embedded and Standalone	Embedded and Standalone

Hive distribution		version	
		Hive 1	Hive 2
	CDH5.1 MR1 (Kerberos authentication is supported)	No	Standalone
	CDH5.4 YARN (Kerberos authentication is supported)	No	Standalone
MapR	MapR 1.2 (deprecated)	Standalone	No
	MapR 2.0	Embedded and Standalone	No
	MapR 2.1.2	Embedded and Standalone	No
	MapR 3.0.1	Embedded and Standalone	Embedded and Standalone
	MapR 3.1	Embedded and Standalone	Embedded and Standalone
Apache	Apache 1.0.0 (Hive 0.9.0)	Embedded and Standalone	No
	Apache 0.20.23 (Hive 0.7.1)	Standalone	No
Pivotal HD	Pivotal HD 1.0.1	Standalone	No
	Pivotal HD 2.0.1	Embedded (only Linux) and Standalone (Linux and Windows)	Embedded and Standalone (only Linux)