talend*
*open integration solutions

Talend Enterprise Data Quality

Installation Guide

5.0_a

Talend Enterprise Data Quality: Installation Guide

Adapted for Talend Enterprise Data Quality Studio v5.0.x. Supersedes any previous Installation Guide.
Copyright © 2008-2011

Table of Contents

Preface

1. General information

1.1. Purpose

This Installation Guide explains how to install and configure Talend Enterprise Data Quality modules and related applications. For detailed explanation on how to use and fine-tune Talend Enterprise Data Quality applications, please refer to your Talend Enterprise Data Quality User Guide.

Information presented in this document applies to *Talend Enterprise Data Quality* releases beginning with **5.0.x**.

1.2. Audience

This guide is devoted for administrators of *Talend Enterprise Data Quality Studio*, *Talend Administration Center*, and/or *Activity Monitoring Console*.

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 and 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 need to be aware of or pay special attention to.

```
Any command is highlighted with a grey background or code typeface.
```

2. History of changes

The below table lists changes made in the *Talend Enterprise Data Quality* Installation Guide.

Version	Date	History of Change
v4.2a	2011/05/19	Update of Installation Guide includes: - Moved DQ Portal specific information chapter (incl. information on memory configuration of the Tomcat server) to *Talend Installer* User guide. - In prerequisites chapter, small edits to supported Tomcat versions and to compatible OS and Browsers - In First installation chapter, changed *Talend Administration Center* configuration steps - In Upgrading chapter, added information H2 backup. - In appendix A, added section re. Perl discontinuation and re-enabling and added section regarding clustering setup issue on MSSQL server.
v4.2b	2011/07/11	Update of Installation Guide includes: - Updated documentation to reflect new product names. For further information on these changes, see the *Talend* website. - Added information regarding *Talend Enterprise Data Quality Portal* configurations. - In Section 5.4, “Updating *Talend Administration Center* Web application”, added information regarding database configuration. - Updated Section 4.3.3, “For non Linux/Windows platforms”. - Added information regarding the creation of CommandLine/ JobServer services. - Updated Section C.3, “Setting up the H2 database for access from other machines” and Section 5.1.3, “Backing up the administration database (MySQL, H2, MS SQL Server, or Oracle)”. - Added information regarding the user authentication in Section 3.7, “Installing the execution servers (JobServers)”. - Reorganized Section 6.2.1, “CommandLine” and Section 6.2.2, “JobServer”.
v5.0a		Update of Installation Guide includes: - Updated documentation to reflect new product names. For further information on these changes, see the *Talend* website. - Added information in chapter Prerequisites, especially Section 1.2, “Hardware requirements”. - Added information in the introduction chapter regarding the software update repository. - Added information in chapter First installation regarding the installation of the software update repository (Section 3.6, “Installing and configuring the software update repository”) and the installation of *Activity Monitoring Console* (Section 3.8, “Installing the *Activity Monitoring Console*”).

Version	Date	History of Change
		- Updated chapter First installation regarding the JobServer configuration. - Added information in chapter Installating services regarding the installation of Windows services. - Added information in FAQ.

3. 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*'s Forum website at:

http://talendforge.org/forum

Chapter 1. Prior to installing *Talend Enterprise Data Quality*

This chapter provides useful information on software and hardware prerequisites you should be aware of, prior to starting the installation of *Talend Enterprise Data Quality* modules.

1.1. License key

You should have received an email from **Talend**, including the following information:

- your personal license key in a file with no extension:

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

  For more information on how to use your license key, see Section 3.4.1, “Configuring the Web application access” and Section 4.2, “Configuring the *Talend Enterprise Data Quality Studio*”.

- the software parts in archive files:

  Unzip each archive file at the root or in a short path of the machine where they are to be deployed (*Talend Enterprise Data Quality Studio* on a developer’s machine and *Talend Administration Center* on the execution server for example).

  Use preferably 7-zip if you are on Windows.

- the documentation (including the one you are reading now):

  Each module of *Talend Enterprise Data Quality* has a dedicated User Guide available in pdf.

1.2. Hardware requirements

To make the most out of the *Talend* products you subscribed to, please consider the following hardware recommendations.

1.2.1. Memory usage

Memory usage heavily depends on the size and nature of your *Talend* projects. However, to make it short, 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.

Product	**Client/Server**	**Recommended alloc. memory**
Talend Administration Center	Server	4GB minimum, 8GB recommended
Commandline	Server	2GB minimum, 5 GB recommended
JobServer	Server	1GB minimum, more recommended[2]
Studio	Client	3GB minimum, 4 GB recommended

2. Memory requirements depend on the executed Jobs.

1.2.2. Disk usage

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

Product	Client/Server	Required disk space for installation	Required disk space for use
Talend Administration Center with Archiva (for Talend Software update and Talend Artifact repository)	Server	500MB	500MB minimum + project size = 20GB+ recommended
Commandline	Server	3GB	2GB minimum + project size = 20 GB+ recommended
Job Server	Server	20MB	2GB minimum + Jobs deployed = 20+ GB recommended
Studio	Client	3GB	3+ GB
Talend Runtime	Server	400MB	400+ MB

1.2.3. Compatible Web browsers

Despite our intensive tests, you might encounter some issues when accessing *Talend Administration Center* with some Web browser.

Please refer to the table below for a summary of supported Web browser. Based on reported issues, we considered that some Web browsers are not supported even though the issue can be resolved in particular conditions. A note has been added providing configuration details.

Web browser	*Talend Administration Center*
Mozilla Firefox	Working
Microsoft Internet Explorer 7 and 8	Working
Microsoft Internet Explorer 9	Working
Google Chrome	Working[1]

1. Only limited support is provided. Contact Support for details.

1.2.4. Naming conventions

In the email you received from *Talend* are listed a number of links to software modules for you to download. The file naming convention is as follows:

Zip/jar file naming convention	Example	Description
Talend-AdministrationCenter-rYYYY-**V**A.B.C	Talend-AdministrationCenter-r72851-V5.0.0.zip	*Talend Administration Center* Web-based application used to administrate Talend Enterprise Data Quality projects and users + Apache Archiva Talend.
Talend-Studio-rYYYY-**V**A.B.C	Talend-Studio-r72851-V5.0.0.zip	Commandline interface to the IDE + Studio IDE (GUI)
Talend-JobServer-rYYYY-**V**A.B.C	Talend-JobServer-r72851-V5.0.0.zip	JobServer: Standalone execution server
Talend-Soamanager-rYYYY-**V**A.B.C	Talend-Soamanager-r72851-V5.0.0.jar	SOA Manager: helps deploying Web services Jobs

Zip/jar file naming convention	Example	Description
Talend-Runtime-VA.B.C	Talend-Runtime-V5.0.1.zip	Talend Runtime: OSGI Container including JobServer

Where:

- YYYY: Revision number,
- A.B.C.: Major. Minor. Patch: revision level if relevant.

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

1.2.5. Third-party softwares

Some additional third-party applications are required for Talend Enterprise Data Quality modules to work together smoothly.

- A MySQL/Oracle/SQLServer database to hold Talend Administration Center settings: *http://www.mysql.com; http://www.oracle.com; http://www.microsoft.com* .
- A Web application server being able to deploy WAR files, e.g.:

  - Apache Tomcat version 5.5, 6.0 or 7.0 (version 6.0 or 7.0 are recommended) - http://tomcat.apache.org/or

  - JBoss Application Server version 4.2.2 - http://www.jboss.org/jbossas/downloads/
- Oracle (JDK or JRE) JVM 1.5+ (but version 1.6+ is recommended)- http://java.sun.com/javase/downloads/index.jsp
- Subversion for storing your projects http://subversion.tigris.org/ or http://www.visualsvn.com/server/download/
- JSL to launch Java applications (CommandLine, JobServer, SVN) as Windows services - http://sourceforge.net/projects/jslwin/.

1.3. Supported Platforms

Despite our intensive tests, you might encounter some issues when installing our products on some Operating Systems.

Please refer to the grids below for a summary of supported OS and Java Runtime environments.

1.3.1. Platform compatible with *Talend* Studio and Commandline

OS	Processor	JDK/JRE	Support type
MAC-OS Lion / 10.6	64 bits	Oracle Java 1.6	recommended
Microsoft Windows XP SP3	32 bits	Oracle Java 1.6	recommended
Microsoft Windows Vista SP1	32 & 64 bits	Oracle Java 1.6	supported

OS	Processor	JDK/JRE	Support type
Microsoft Windows 7	64 bits	Oracle Java 1.6	recommended
Microsoft Windows 2003 SP2	64 bits	Oracle Java 1.6	recommended[1]
Microsoft Windows 2008 R2	64 bits	Oracle Java 1.6	recommended[1]
Microsoft Windows XP SP3	64 bits	Oracle Java 1.6	supported
Microsoft Windows 7	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 10.04	64 bits	Oracle Java 1.6	recommended
Linux CentOS 5.4	32 bits	Oracle Java 1.6	recommended
Redhat Linux Enterprise Server Edition 6.1	64 bits	Oracle Java 1.6	recommended
Redhat Linux Enterprise Server Edition 5.4	64 bits	Oracle Java 1.6	recommended
Solaris 10	x86 64 bits	Oracle Java 1.6	recommended[1]
Linux Ubuntu 10.4	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 11.10	32 & 64 bits	Oracle Java 1.6	supported
Redhat Linux Enterprise Server Edition 5.3/5.4/5.6	32 bits	Oracle Java 1.6	supported
Redhat Linux Enterprise Server Edition 6.1	32 bits	Oracle Java 1.6	supported

1. Only for Commandline

1.3.2. Platform compatible with *Talend* Runtime

OS	Processor	JDK/JRE	Supported type
Linux Ubuntu 10.04	64 bits	Oracle Java 1.6	recommended
Linux CentOS 5.4	32 bits	Oracle Java 1.6	recommended
Red Hat Linux Enterprise Server Edition 5.3/5.4/5.6	64 bits	Oracle Java 1.6	supported
Red Hat Linux Enterprise Server Edition 6.1	64 bits	Oracle Java 1.6	recommended
Solaris 10	sparc 64 bits	Oracle Java 1.6	recommended
Linux CentOS 5.3/5.6	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 10.04	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 11.10	32 & 64 bits	Oracle Java 1.6	supported
SUSE SLES 10	32 & 64 bits	Oracle Java 1.6	supported
Solaris 10	x86 64 bits	Oracle Java 1.6	supported
Windows Server 2008 R2	64 bits	Oracle Java 1.6	recommended
Windows Server 2003 SP2	64 bits	Oracle Java 1.6	supported

1.3.3. Platform compatible with *Talend Administration Center*

OS	Processor	JDK/JRE	Supported type
Linux Ubuntu 10.04	64 bits	Oracle Java 1.6	recommended

OS	Processor	JDK/JRE	Supported type
Red Hat Linux Enterprise Server Edition 5.4	32 bits	Oracle Java 1.6	recommended
Red Hat Linux Enterprise Server Edition 6.1	64 bits	Oracle Java 1.6	recommended
Solaris 10	x86 64 bits	Oracle Java 1.6	recommended
Linux CentOS 5.4	32 bits	Oracle Java 1.6	supported
Red Hat Linux Enterprise Server Edition 5.4	64 bits	Oracle Java 1.6	supported
SUSE SLES 10	32 & 64 bits	Oracle Java 1.6	supported
Solaris 10	sparc 64 bits	Oracle Java 1.6	supported
Windows Server 2003 SP2	64 bits	Oracle Java 1.6	supported
Windows Server 2008 R2	64 bits	Oracle Java 1.6	recommended

1.3.4. Platform compatible with JobServer

OS	Processor	JDK/JRE	Supported type
Linux Ubuntu 10.04	64 bits	Oracle Java 1.6	recommended
Red Hat Linux Enterprise Server Edition 5.4	32 bits	Oracle Java 1.6	recommended
Red Hat Linux Enterprise Server Edition 6.1	64 bits	Oracle Java 1.6	recommended
Solaris 10	x86 64 bits	Oracle Java 1.6	recommended
Linux CentOS 5.3/5.4/5.6	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 10.04	32 bits	Oracle Java 1.6	supported
Linux Ubuntu 11.10	32 & 64 bits	Oracle Java 1.6	supported
Red Hat Linux Enterprise Server Edition 5.3/5.4/5.6	32 bits	Oracle Java 1.6	supported
SUSE SLES 10	32 & 64 bits	Oracle Java 1.6	supported
Solaris 10	x86 64 bits	Oracle Java 1.6	supported
Windows Server 2008 R2	64 bits	Oracle Java 1.6	recommended
Windows Server 2003 SP2	64 bits	Oracle Java 1.6	supported

1.4. Compatible Runtime Containers

Please refer to the grids below for a summary of supported runtime containers. Based on reported issues, we considered that some runtime containers are not supported even though the issue can be resolved in particular conditions. A note has been added providing configuration details.

1.4.1. Runtime Containers compatible with Talend Runtime

Runtime Containers	Version	Supported type
Talend ESB Container (Apache Karaf)	5.0 (2.2.4)	recommended
Apache Tomcat	6	supported
JBoss Application Server	4.2.2	supported with limitations[1]

1. ESB/CXF service participants.

1.4.2. Runtime Containers compatible with *Talend Administration Center*

Runtime Containers	Version	Supported type
Apache Tomcat	6	recommended
Apache Tomcat	7	supported
JBoss Application Server	4.2.2	supported

1.5. Compatible Databases

Please refer to the grids below for a summary of supported databases. Based on reported issues, we considered that some databases are not supported even though the issue can be resolved in particular conditions. A note has been added providing configuration details.

1.5.1. Databases compatible with *Talend Administration Center*

Databases	Version	Support type
MySQL	5.5	recommended
Oracle	10	recommended
H2	1.3	recommended
MySQL	5.0/5.1	supported
MS SQL Server	2008	supported
MS SQL Server	2005	supported
Oracle	11	supported

1.5.2. Databases compatible with *Activity Monitoring Console*

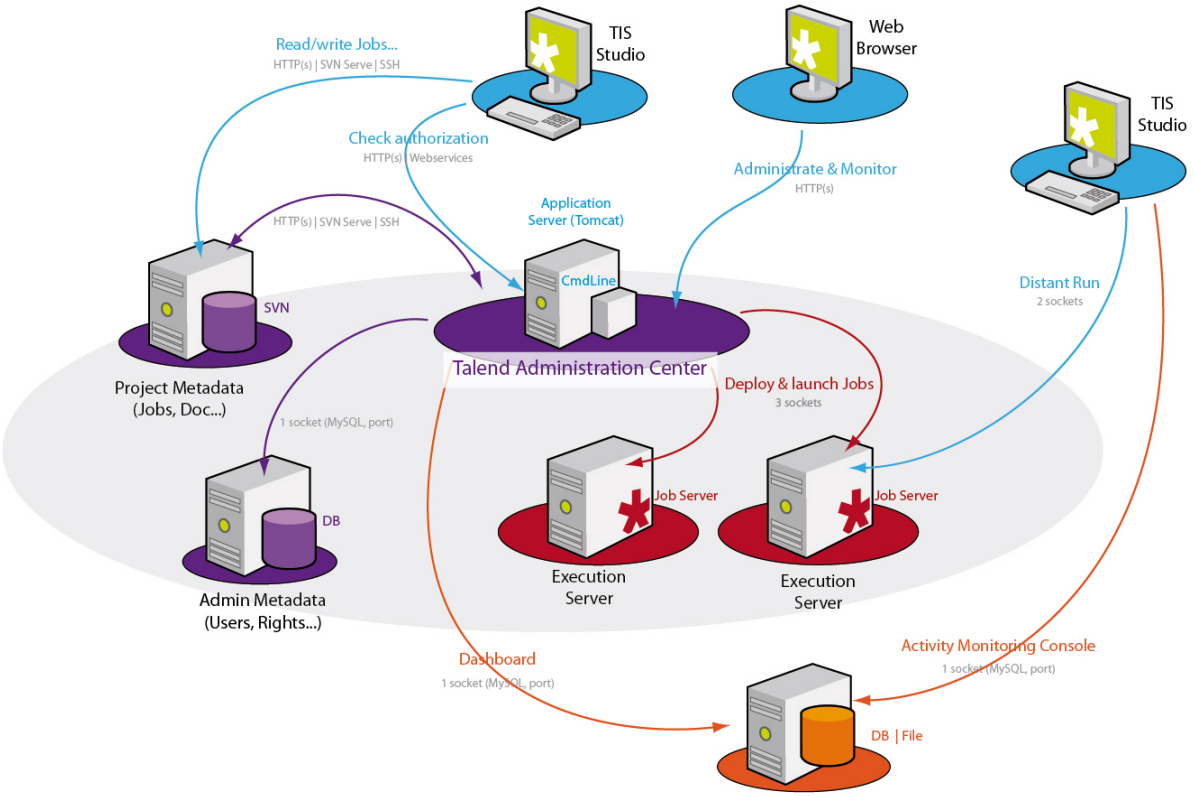
Databases	**Version**	**Support type**
MySQL	5.5	recommended
Oracle	10	recommended
H2	1.3	recommended
MySQL	5.0/5.1	supported
MS SQL Server	2008	supported
IBM DB2	9.7	supported
Informix	11	supported
PostgreSQL	9.1	supported
Sybase	15	supported
MS SQL Server	2005	supported
Firebird	2.1	supported
Oracle	11	supported

Chapter 2. Introducing *Talend Enterprise Data Quality*

This chapter introduces *Talend Enterprise Data Quality*. It provides a detailed description of the architecture and lists all modules that are required for the installation and configuration of *Talend Enterprise Data Quality Studio*, *Talend Administration Center* and/or *Activity Monitoring Console* .

2.1. General introduction to *Talend Enterprise Data Quality*

The following image shows the distributed architecture of *Talend Enterprise Data Quality*. However, we recommend you to centralize all the server modules on one single system to ease their management:



All Talend applications installed must be in the same version.

The elements that are required at a minimum for using *Talend Enterprise Data Quality* include:

- A server that hosts *Talend Administration Center* (Tomcat server + CommandLine),
- A storage server for Project metadata (SVN),
- A storage database for administration metadata (by default, an embedded H2 database is used).
- A software update repository for software updates.

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

2.1.1. Apache Tomcat Server

The Apache Tomcat server is an application server that hosts the *Talend* Web application. The Web application gives access to all management and administration functionalities for a data integration project. These functionalities include:

- the definition of projects,

- the management of user rights,
- the Job monitoring,
- the Job deployment,
- the Job execution.

Talend Web application can also be hosted by JBoss application server. For more information on how to deploy the Web application on JBoss, see Section 3.2.2, “Deploying the Web application on JBoss”.

2.1.2. Database

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

2.1.3. SVN

The SVN server is used to manage the persistence of all the data relating to the objects stored in “svn” mode in the Repository (Jobs, business models, metadata, routines, etc.).

Talend Enterprise Data Quality 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.

2.1.4. Software update repository

The software update repository is used to manage application updates. By default two repositories are embedded within this repository provided by *Talend*. These repositories allow the user to visualize the updates available.

2.2. Detailed introduction to *Talend Enterprise Data Quality* Server

This section describes in more details the architecture of each of the modules depicted in the previous section.

2.2.1. Administration database

By default, an embedded H2 database is used to store administration information. But you have the possibility to change database to MySQL, MSSQL or Oracle if you prefer so. For more information about how to set up a different database, see the FAQ section: Section A.16, “I do not want to use H2 as default database” .

The 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 (storage mode), user declaration and authorization, task list, trigger 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 Enterprise Data Quality.*

2.2.2. *Activity Monitoring Console* log database

If you want to use the *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 section *Configuring sources* of the *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.

2.2.3. Apache Tomcat server

Version 5.5 or later (version 6.0 recommended) should be used. This software is an application server that can host the Web application which allows users to (depending on their role):

- Manage user accounts and roles/rights,
- Access to the Job Conductor to run, schedule and deploy Jobs,
- Access to the Dashboard to monitor Jobs (available from Professional Edition),
- Manage projects.

Note that the best recommended versions for Tomcat are 6.0 and 7.0, however the 5.5 version is also supported.

Talend Web application can also be hosted by JBoss application server. For more information on how to deploy the Web application on JBoss, see Section 3.2.2, "Deploying the Web application on JBoss".

2.2.4. SVN server

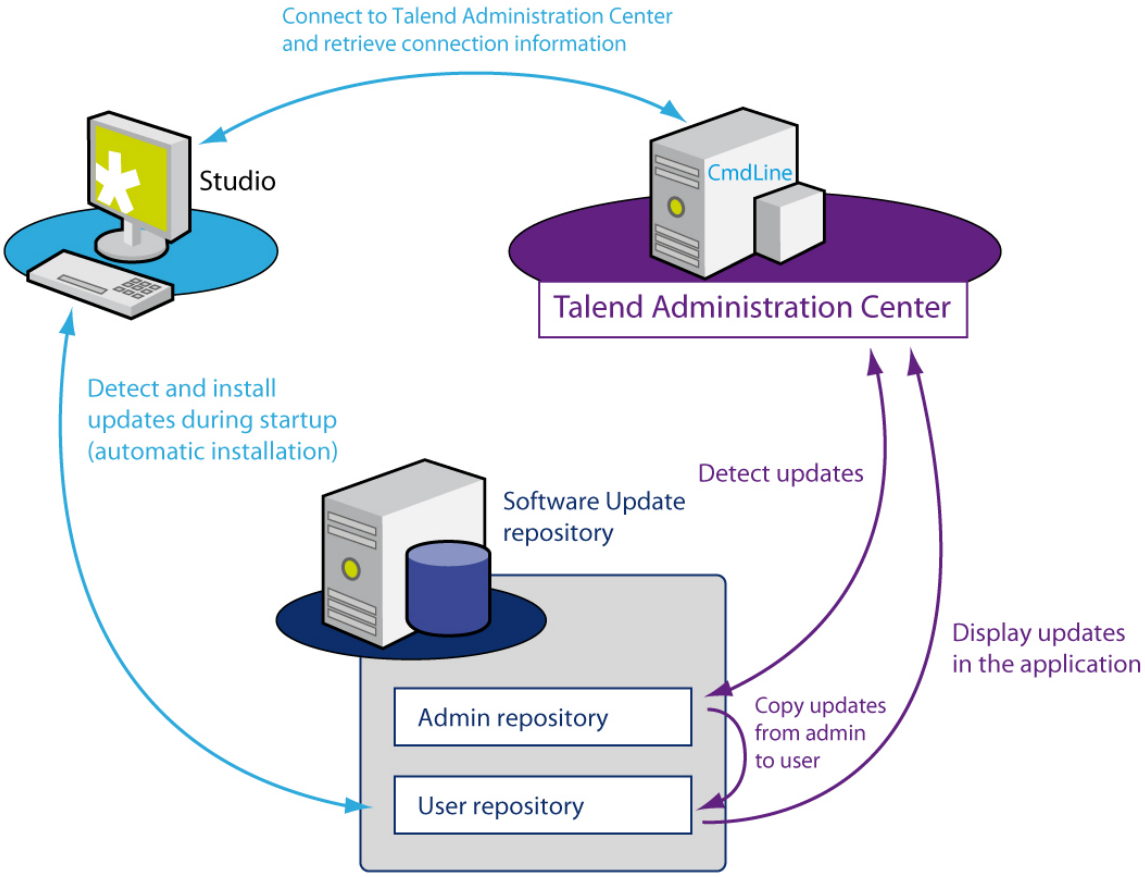
In SVN mode, project metadata (Jobs, Business Models, Documentation, etc.) are stored in an SVN repository rather than in a database. This is the recommended storage mode and the only mode that allows to store large projects.

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 SVN repository, if you prefer so.

You can also have several SVN repositories each containing several projects. For more information on how to create projects and store them in an SVN repository, see *Talend Administration Center* User Guide (formerly Administrator).

2.2.5. Software update repositories

The following image shows the architecture of the software update repository linked to *Talend Administration Center* and to the Studio.



To download and install some software updates, you need to use an Archiva repository and its two embedded repositories, i.e.:

- an *admin* repository named *talend-update-admin* used to publish the updates;
- a *user* repository named *talend-update-admin* used to retrieve the updates.

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

In Talend Administration Center, the new versions and patches are detected and the administrator can accept them.

For more information on how to install and configure this software update repository, see Section 3.6, “Installing and configuring the software update repository”.

Chapter 3. Installing *Talend Enterprise Data Quality* for the first time

We strongly encourage you to read the chapters: Chapter 1, *Prior to installing Talend Enterprise Data Quality* and Section 2.2, “Detailed introduction to *Talend Enterprise Data Quality* Server” before starting this chapter.

This chapter details the procedures required to install each part of Talend Enterprise Data Quality.

The following parts usually need to be installed in the following order:

1. SVN Server, see Section 3.1, “Setting up your project server on SVN”
2. Tomcat, see Section 3.2.1, “ Deploying the Web application on Tomcat”
3. High Availability (available from *Talend Enterprise Data Quality Cluster edition* and *Talend Enterprise Data Quality Big Data edition*, see Section 3.3, “Setting up High Availability”
4. Administration center Web application, see Section 3.4, “Configuring *Talend Administration Center* (Web application)”
5. Commandline, see Section 3.5, “Installing Commandline”
6. Software update repository, see Section 3.6, “Installing and configuring the software update repository”
7. Execution servers (Jobservers or "agents"), see Section 3.7, “Installing the execution servers (JobServers)”
8. Activity Monitoring Console, see Section 3.8, “Installing the *Activity Monitoring Console*”
9. SOA manager (available from *Talend Enterprise Data Quality RTx Edition)*, see Section 3.9, “Installing the SOA Manager”
10. JBoss Business Rules Management System (Available from *Talend Enterprise Data Quality Professional Edition)*
11. Automatic log purge (optional), see Section 3.11, “ Setting up the automatic log purge (optional)”

In the following documentation, <*TomcatPath*> designates the Tomcat's installation path and <*ApplicationPath*> designates the directory in <*TomcatPath*>*/Webapps* where *Talend Administration Center* war file has been deployed.

3.1. Setting up your project server on SVN

This section describes how to install the SVN server. This server will store all your project data (Jobs, Database connections, Routines, Joblets, etc.).

The installation instructions for Subversion will differ depending on:

- your development environment,
- the access method,
- the authentication mode.

The access methods and authentication modes to the Subversion server include:

- Apache is a Web server and can be used to access a Subversion repository. Only the authenticated access is documented here.

Note that VisualSVN has both Apache and Subversion embedded

- SSH is another method for accessing a Subversion repository. Authentication is mandatory and brings a higher level of security thanks to encrypted communication.

Choose the procedure that is most relevant to your environment among the following instruction sections. These sections are independent from each other.

3.1.1. Windows

Download the SVN server installer, for example: http://www.visualsvn.com/server/download/

1. Launch the wizard.
2. Clear the check box: Use secure connection (https://).
3. Complete the installation process without changing the other parameters.
4. Launch VisualSVN.

In the main window of VisualSVN Server Manager:

Check in the server URL that the port is *80*, if not then perform the following operation:

1. Right-click on **VisualSVN Server** on the left tree view.
2. Select **Properties.**
3. Click on the **Network** tab.
4. Change the port selection to *80*.
5. Click **Apply** then **OK**.

Then create a new repository:

1. Right-click on **Repositories** in the main window of VisualSVN.

2. Click **Create new repository...**
3. Name the repository newly created, for example: <*repotis*>.
4. Click **OK**.

Having done this, create a new user:

1. Click **Create new user...** in the main window of VisualSVN.
2. Then define a user name and his password.

Then grant read-write authorization to this user:

1. Right-click on <*repotis*>.
2. Select **Properties**.
3. Change the default user (everyone) permissions to no access.
4. Click **Add**.
5. Select your username in the list displaying and click **OK**.
6. Check that the granted permissions are read/write.

Your SVN server is now installed and the repository is ready to store all Talend Enterprise Data Quality projects.

One thing left to do is to link *Talend Administration Center* to this newly created repository.

In VisualSVN:

1. Right-click on your new repository: <*repotis*>.
2. Select **Copy URL to Clipboard**.

This copied URL is to be used along with the username and password when you configure the *Talend Administration Center* application (see section Section 3.4.1, “Configuring the Web application access”).

The SVN URL should read:

http://host_or_ip_address:port80/svn/repository_name.

such as: http://talend-esa.talend.com:80/svn/repotis/

3.1.2. Debian, apache

In the following instructions, lines starting with "#" mean that they must be executed as root. A command line starting with "$" means that it must be executed as standard user.

```
# apt-get install subversion apache2 libapache2-svn

# mkdir /svn

# chown www-data:www-data /svn
```

Switch to webserver user, which is *www-data* by default on Debian systems.

```
# su - www-data
```

```
$ svnadmin create /svn/repotis

$ htpasswd -cmb /svn/passwd tisadmin secretpassword

# vi /etc/apache2/sites-available/default
```

```
<Location /repotis> DAV svn  SVNPath /svn/repotis  AuthUserFile /svn/passwd
Require valid-user  AuthType basic  AuthName "Subversion"  </Location>
```

```
# /etc/init.d/apache2 restart
```

Your SVN server is now installed and the repository is ready to store your Talend Enterprise Data Quality projects.

One thing left to do is to link *Talend Administration Center* to this newly created repository.

You will need the URL as well as the username and password of your newly created SVN when configuring *Talend Administration Center* (see Section 3.4.1, “Configuring the Web application access”).

The URL syntax should read: *http://host_or_ip_address/repotis/project_name*.

For example:

- *http://10.42.0.10/repotis/first_project*
- *http://tisserver/repotis/another_project*

Then give the user “*tisadmin*”, password “*secretpassword*” as described in the *passwd* file.

3.1.3. Debian, ssh

You can start with the following line:

```
# apt-get install ssh subversion
```

Create a user "*tisadmin*" with password "*secretpassword*".

```
# adduser tisadmin

# mkdir /svn

# chown tisadmin:tisadmin /svn
```

Switch to “*tisadmin*” user.

```
# su - tisadmin
$ svnadmin create /svn/repotis
```

Your SVN server is now installed and the repository is ready to store your Talend Enterprise Data Quality projects.

One thing left to do is to link *Talend Administration Center* to this newly created repository.

You will need the URL as well as the username and password of your newly created SVN when configuring *Talend Administration Center* (See Section 3.4.1, “Configuring the Web application access”).

The URL syntax should read:

svn+ssh://tisadmin@host_or_ip_address/svn/repotis/project_name.

For example:

- • *svn+ssh://tisadmin@10.42.0.10/svn/repotis/first_project*
- *svn+ssh://tisadmin@tisserver/svn/repotis/another_project*

Then give the user "*tisadmin*", password "*secretpassword*".

3.1.4. Redhat (Fedora), Apache

As User: apache

-OR-

As Root

```
# yum install httpd mod_dav_svn
```

The above command installs the following:

- - *httpd-2.2.9-1.fc9.i386.rpm*
- - *mod_dav_svn.i386 0:1.4.6-7*
- - *subversion.i386 0:1.4.6-7*

```
# mkdir /var/www/svn
# svnadmin create /var/www/svn/repository
# htpasswd -cmb /var/www/svn/passwd admin@company.com admin
# chown -R apache.apache /var/www/svn
# vim /etc/httpd/conf.d/subversion.conf
```

Add:

```
<Location /svn>

DAV svn
SVNPath /var/www/svn/repository
Require valid-user
AuthType basic
AuthName "Subversion"
AuthUserFile /var/www/svn/passwd

</Location>
```

```
# /etc/init.d/httpd start
```

3.2. Deploying the Web application on an application server

This section shows how to deploy the Web application on an application server; it can be an Apache Tomcat or a JBoss server application.

3.2.1. Deploying the Web application on Tomcat

First, install the Apache Tomcat application server and stop the Tomcat service if it is automatically started.

In the rest of this documentation, <*TomcatPath*> points out the path where Tomcat has been installed.

Indeed we recommend you to stop this service prior to deploying the Web application (*org.talend.administrator.war*).

- Using 7zip, unzip the archive delivered by **Talend** into the following directory: <*TomcatPath*>*/webapps/*

On Linux:

1. Edit the following file: *etc/default/tomcat6*
2. and uncomment the tomcat security setting and change the default setting, in order to read:

```
#TOMCAT6_SECURITY=no
```

The storage of log outputs is managed by Tomcat application server, by default, but you are also able to define your own path for storing the logs. From 4.0, you can configure the path directly from Talend Administration Center . In prior versions, refer to Section 3.2.4, “Log storage mode” *for more information on manual configuration.*

Then start Tomcat using the following commands:

	Linux	**Windows**
Start the Tomcat service	sh <TomcatPath>/bin/startup.sh	net start “Apache Tomcat”
Stop the Tomcat service	sh <TomcatPath>/bin/shutdown.sh	net stop “Apache Tomcat”

For reasons of right management, make sure you launch Tomcat using the same administrator account as for the Commandline. E.g.: Create an account TISAdmin for both Tomcat and Commandline.

If you deploy a large number of applications on Tomcat, you should increase its memory to improve its performance. For more information on this process, see section about Tomcat server configuration in **Talend Installer User Guide**.

3.2.2. Deploying the Web application on JBoss

First, install the JBoss application server and stop the JBoss service if it is automatically started.

In the rest of this documentation, <*JBossPath*> points out the path where JBoss has been installed.

Indeed we recommend you to stop this service prior to deploying the Web application (*org.talend.administrator.war*).

- Using 7zip, unzip the archive delivered by **Talend**.
- Manually change the file extension from *org.talend.administrator.war* to *org.talend.administrator.zip*.
- Unzip the file to an *org.talend.administrator* folder into the following directory:

  <*JBossPath*>*/server/default/deploy/*

The storage of log outputs is managed by JBoss application server, by default, but you are also able to define your own path for storing the logs. From 4.0, you can configure the path directly from Talend Administration Center . In prior versions, refer to Section 3.2.4, “Log storage mode” *for more information on manual configuration.*

Then start JBoss using the following commands:

	Linux	**Windows**
Start the JBoss service	sh <JBossPath>/bin/run.sh	net start “JBoss”
Stop the JBoss service	sh <JBossPath>/bin/shutdown.sh	net stop “JBoss”

For reasons of right management, make sure you launch JBoss using the same administrator account as for the Commandline. E.g.: Create an account TISAdmin for both JBoss and Commandline.

3.2.3. Synchronizing Web application and server time zones

To make sure that the DST change and the time zones are correctly taken into account, check that your OS includes an environment variable set as follows:

```
On Windows: TZ=Europe/Paris

On Linux: Export TZ="Europe/Paris"
```

If you wish to implement high availability to the task execution scheduling using **Talend** clustering feature, then read Section 3.3, “Setting up High Availability”.

3.2.4. Log storage mode

The log outputs are stored by default in the server application standard log file (STDOUT) as defined in the *Log4j.xml* file. However you can store the log in a different file by setting the path to this file in the *Log4j.xml* file.

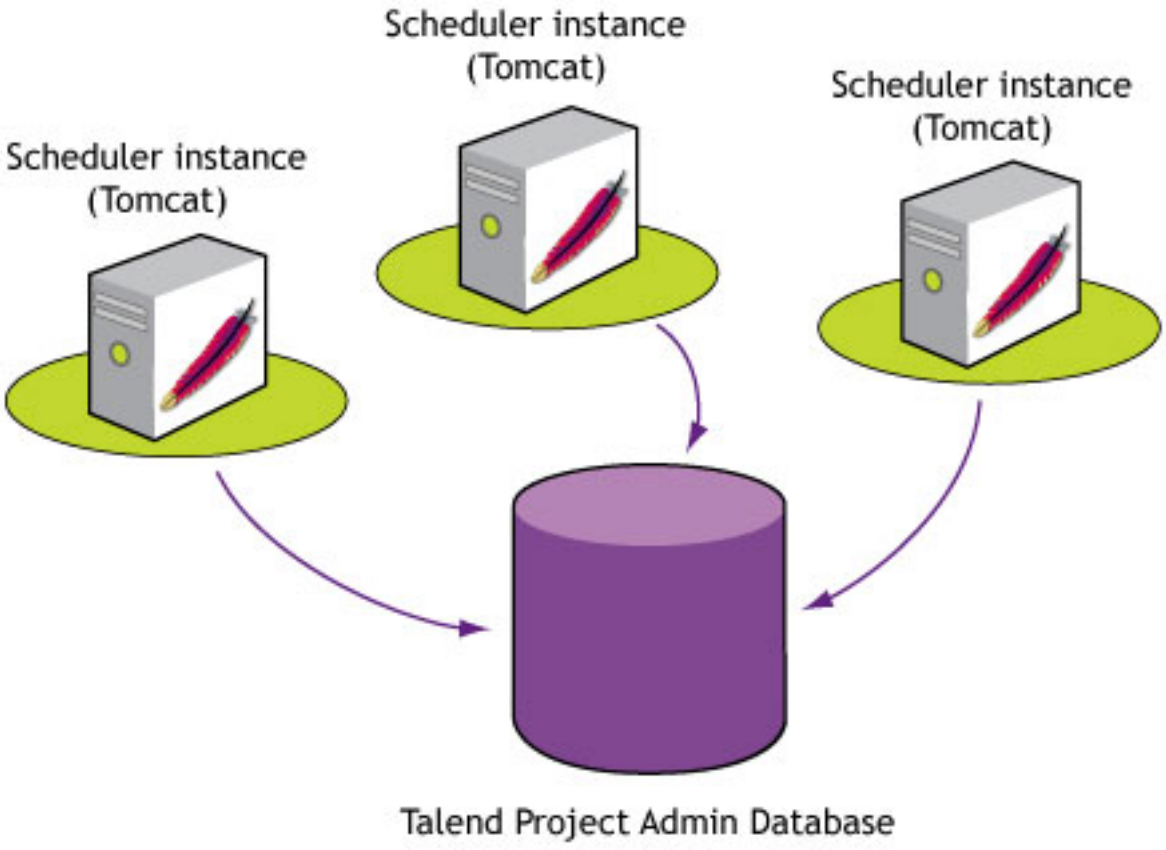
To do so, simply set the path in the **Configuration** page in *Talend Administration Center*. For more information, refer to your *Talend Administration Center* User Guide.

If you leave the **Path** field blank in the **Configuration** page, then you can also customize the *Log4j.xml* to address your custom needs.

3.3. Setting up High Availability

This feature is ONLY available in Talend Enterprise Data Quality Cluster edition and Big Data edition.

Using **Talend** clustering system, you can add high availability and failover to your task execution scheduling controlled by the **Job Conductor** of *Talend Administration Center*.



To implement this High Availability architecture, you need to install and configure several instances of the Tomcat application server and connect them all to the project administration database.

1- Install one Tomcat server as described in Section 3.2, “Deploying the Web application on an application server”.

2- Before starting Tomcat and deploying the Web application, set it into *cluster mode*. To do so:

- Edit the following file:

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

- Uncomment the following lines to enable the Quartz cluster, by removing the dash character preceding the command (in bold):

```
#org.quartz.scheduler.instanceName = MyClusteredScheduler
```

```
#org.quartz.scheduler.instanceId = AUTO
```

```
#org.quartz.jobStore.isClustered = true
```

```
#org.quartz.jobStore.clusterCheckinInterval = 20000
```

4- Then duplicate this Tomcat instance as many times as needed.

Make sure that all Tomcat instances use different port numbers.

It's not recommended to install the instances on different machines, as you need to make sure that all system clocks are synchronized using some form of time-sync service (daemon) that runs very regularly (the clocks must be within a second of each other). In addition, logs cannot be centralized unless they are stored on a different database, residing on one machine, all Tomcat can connect to.

5- Then duplicate the *org.talend.administrator* Web application to all Tomcat instances. Make sure that all Web application configurations are identical.

6- Then launch one Tomcat instance following the commands given at the end of Section 3.2, "Deploying the Web application on an application server".

7- And finally, launch the other instances of Tomcat following the same procedure.

Fail-over will occur when one of the Tomcat application servers fails while in the midst of executing one or more tasks. When a server fails, the other servers of the cluster detect the condition and identify the tasks in the database that were in progress within the failed server. Any tasks marked for recovery will be taken over by another server.

One known minor issue related to the DST change might prevent the failover to operate properly. However as a simple workaround, simply restart Tomcat after the time change. This should have no impact on executions.

You can also deploy **Talend** Web application on a JBoss application server (instead of a Tomcat). So, you can follow the same above instructions for Jboss. From more information on how to deploy the Web application on JBoss, see Section 3.2.2, "Deploying the Web application on JBoss".

3.4. Configuring *Talend Administration Center* (Web application)

Talend Administration Center is the Web administration application of Talend Enterprise Data Quality.

You can configure the way *Talend Administration Center* will operate directly from its Web interface, especially you need to link *Talend Administration Center* to the shared Repository (stored on SVN).

For more details, see Section 3.4.1, "Configuring the Web application access".

If needed, you can also configure the connection pool to be managed via Tomcat. However this advanced configuration mode cannot be done through the Web interface of *Talend Administration Center*.

For more details, see Section 3.4.2, "Managing the connection pool via Tomcat".

3.4.1. Configuring the Web application access

The recommended way to configure the connection to the database and to the shared repository (SVN) is through the Web interface of *Talend Administration Center*.

- Open a Web browser (Internet Explorer, Firefox, etc.) and type in the following URL: http://<localhost>:8080/<ApplicationPath> [http://localhost:8080/org.talend.administrator].
- Replace <*localhost*> with the IP address of the Web server if the Web browser IP is different from the machine you are on and <*ApplicationPath*> with the *Talend Administration Center* Web application path.

This login corresponds to the default user of the Web application. You can create a new one via the **Users** menu in *Talend Administration Center*, and then delete the *admin@company.com* user after you are connected using the newly created user login.

When connecting for the first time, H2 database connection parameters display and some automatic checks are performed on driver, url, connection, version information.

The administration database (storing users, rights, etc.) being an H2 embedded database, its access information is automatically set. However, if you do not want to use the embedded H2 database, you can set up a database server (MySQL, MSSQL or Oracle) in which case, set the required parameters or read Appendix: Section A.16, “I do not want to use H2 as default database”.

1. Click **Set new license**.
2. Browse your system to the License file you received from **Talend** and click **Upload**. A final License check is performed.
3. Click **Go to Login**.
4. On the **Login** page, type in the default connection login for your first access:

   Default login: *admin@company.com*,

   Default password: *admin*.

If your Web access is restricted, you may need to click **Generate Validation request** to perform the validation of your license key. Follow the instructions on screen.

Once the license is validated, the navigation bar of *Talend Administration Center* opens with all the pages accessible for the default administrator user account. For more information on which pages of *Talend Administration Center* an administrator user can access, see the *Talend Administration Center* User Guide.

You can now configure the link to the shared repository (SVN):

1. Click **Configuration** to access the setting page of *Talend Administration Center*.
2. Change the following parameters for the SVN module using the parameters you have set during the installation process of the SVN server.

   - **Server Location URL**: your <*repotis*> SVN URL.

   - **Username**: your <*repotis*> user.

   - **Password**: your <*repotis*> SVN password.

For more details, see Section 3.1, “Setting up your project server on SVN”, and Section 3.1.1, “Windows”, Section 3.1.2, “Debian, apache”, Section 3.1.3, “Debian, ssh”.

If you use several SVN repositories to store your projects, refer to the User Guide of *Talend Administration Center* and check the *Advanced settings* procedure.

The link to the SVN is now established, you can thus create a new project from the **Projects** page in order for the Talend Enterprise Data Quality clients to have at least one project in their workspace.

For more details on how to add a project in *Talend Administration Center*, see the *Talend Administration Center* User Guide.

You must now install the Commandline. The Commandline application can be installed on the same system as the Web application (but not necessarily). It allows the JAVA code to be generated for the Jobs on the basis of the XML files contained in the database. This application must be started in order to allow Jobs to be started and deployed remotely.

For all details on how to install the Commandline, see Section 3.5, “Installing Commandline”.

3.4.2. Managing the connection pool via Tomcat

By default, a third-party application (c3p0) has been embedded into the configuration file of *Talend Administration Center*, to manage the connection pool.

However if you want Tomcat to manage directly the connection pool, perform the following operations:

In the Web application installation directory, proceed as follows:

1- In the *<ApplicationPath>/WEB-INF/Classes* folder, edit the *configuration.properties* file:

- Change the default setting to:

```
database.useContext
= True
```

2- In the *WEB-INF* folder, edit the *web.xml* file:

- Uncomment the following piece of code (at the end of the file), to make it read as follows:

```
<resource-ref>

     <description>Our Datasource</description>
     <res-ref-name>jdbc/ADMINISTRATOR_CONNECTION</res-ref-name>
     <res-type>javax.sql.DataSource</res-type>
     <res-auth>Container</res-auth>

</resource-ref>
```

3- In the file *<ApplicationPath>/META-INF/context.xml*, set the fields as follows:

- Configure the parameters of connection to the database.

```
schema name= talend_administrator

server= localhost or IP of the server on which the database is installed
login= tisadmin or your login for MySQL

password= tisadmin or your password for MySQL

driverClassName="org.gjt.mm.mysql.Driver" (for MySQL DB)

url="jdbc:mysql://<host>:3306/talend_administrator" (for MySQL
DB)
```

4- Copy over the following .jar file from *<ApplicationPath>/WEB-INF/lib* to *<TomcatPath>/lib/* in case your administration data are stored in MySQL:

mysql-connector-java-5.1.3-bin.jar

In case your data are stored in a different database, copy the relevant .jar in *<TomcatPath>/lib/*.

In this case, the DB config page of the Web application will not be operating.

You can also deploy **Talend** Web application on a JBoss application server (instead of a Tomcat). So, you can follow the same above instructions for Jboss. From more information on how to deploy the Web application on JBoss, see Section 3.2.2, “Deploying the Web application on JBoss”.

3.4.3. Advanced configuration

You can fine-tune the usage parameters for the various Talend Enterprise Data Quality modules to address your particular needs.

To access the advanced parameter configuration, edit the following file:

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

For example the following parameters help you change the default credentials:

```
install.defaultaccount.login=admin@company.com
install.defaultaccount.password=admin
install.defaultaccount.firstname=admin
install.defaultaccount.lastname=admin
```

3.5. Installing Commandline

Commandline is required for generating and deploying the processes (Job) developed with *Talend Enterprise Data Quality Studio* onto the Job servers. To install the commandline, follow the procedure below:

- Copy your *Talend Enterprise Data Quality Studio* archive file onto the machine where you want to install the CommandLine. This machine can be the same as your Web application, *Talend Administration Center*, but not necessarily.
- Unzip it under a folder the name of which does NOT contain any space character.
- You can then rename the decompressed folder to *CmdLine* for more clarity.

Since version 4.0, renaming your commandline “CommandLine” is causing problems, so we recommand to rename it differently or not to rename it at all.

In this directory, the *commandline.bat* or *commandline.sh* file lets you launch the *CommandLine* program.

- Run the file.

For reasons of right management, make sure you launch the CommandLine using the same administrator account as for Tomcat. E.g.: Create an account TISAdmin for both Tomcat and Commandline.

If you need to change the port number (by default 8002), simply edit the commandline file (*.sh* or *.bat* based on your system) and change the port number.

- If you chose to install the commandline on a different machine than *Talend Administration Center*, then you need to configure the CommandLine access parameters in the **Configuration** page of *Talend Administration Center* Web application. For more information, see *Talend Administration Center* User Guide.

For more information about the CommandLine usage, see Appendix A of the *Talend Administration Center* User Guide.

After finishing the commandline installation, you still need to configure the execution servers. For more information, see Section 3.7, “Installing the execution servers (JobServers)”.

To install your commandline as a service, see Chapter 6, *Installing services* .

To install the Commandline on other platforms than Windows and Linux and/or on 64-bit platforms, see Section 4.3, “Memory configuration and launching issues”.

3.5.1. Accessing user-defined components from the Commandline

If you need to install user-defined components (that you developed locally or downloaded from Talend Exchange for example), then you need to notify the Commandline with the user component folder.

To configure the path to these components, simply use the following command:

```
setUserComponentPath -up <UserComponentPath>
```

To clear this path, type in the command:

```
setUserComponentPath -c
```

You can also configure the user component path directly from the **Configuration** page of *Talend Administration Center*. For more information, see the *Talend Administration Center* User Guide.

3.6. Installing and configuring the software update repository

The software update repository is an instance that holds two repositories, *talend-update-admin* where the updates are uploaded and *talend-update-user* where the updates are retrieved by the user. This instance is embedded in the *.zip* file of *Talend Administration Center* Web application and it allows you to display the updates on versions and patches available for download. For more information, see *Talend Administration Center* User Guide.

3.6.1. Repository installation

To install the software update repository, follow the procedure below:

1. Unzip the repository archive file delivered by *Talend* in the directory of your choice.

You have the possibility to change the default port of the repository by editing the corresponding line of the *jetty.xml* file located in the *conf* folder of this directory.

2. From the *bin* folder of this directory, run the *archiva.bat* file to launch the repository.

3.6.2. Repository configuration

Once you have launched the repository, open your Web browser and type in the URL to your repository, *http://localhost:8082/archiva* for example, then proceed as follows:

1. Log in with the following credentials:

Username	Password
admin	1q2w3e4r

2. Click **Repositories** on the menu tree view.
3. Edit the remote repository named *talend-update-admin-remote*. To do so:
   - Click **Edit** to the right of the repository.
   - Fill in the **Username** and **Password** fields with the credentials provided by Talend.
   - Save these changes by clicking **Update Repository**.

3.6.3. Configuration in *Talend Administration Center* Web application

Once you have launched and configured the repository, go to the **Configuration** page of *Talend Administration Center* and fill in the following information in the **Software Update** group:

- **Repository URL** : Type in the location URL to your software update repository, *http://localhost:8082/archiva/* for example.
- **Admin username** : Type in the name of the repository user with Manager role. By default, it is *SoftwareUpdateAdmin*.
- **Admin password** : Type in the password of the repository user with Manager role. By default, it is *SoftwareUpdateAdmin*.
- **Reader username** : Type in the name of the repository user with Observer role. By default, it is *SoftwareUpdateReader*.
- **Reader password** : Type in the password of the repository user with Observer role. By default, it is *SoftwareUpdateReader*.

In the **Software Update** page of *Talend Administration Center*, you can now see the versions and patches available and download them according to your needs.

3.7. Installing the execution servers (JobServers)

3.7.1. Installing the Jobservers

You need now to define which systems will be the execution servers.

1. First select the servers that will be used to execute the Jobs (processes) developed with *Talend Enterprise Data Quality Studio*.

2. Then, on each server, uncompress the archive file containing the JobServer application matching your release version of Talend Enterprise Data Quality.

   The archive file name for example reads: *org.talend.remote.jobserver_r29781*.

3. In the uncompressed file you need to configure the file *TalendJobServer.properties* that you can find in the directory <*root*>*/conf/* where <root> is the JobServer path.

4. Modify the installation directory of the JobServer and check that the *8000*, *8001* and *8888* ports are available.

5. To enable user authentication on the Job server, you need to define one or more lines of username and password pairs in the file *users.csv* that you can find in the directory <*root*>*/conf/* where <root> is the JobServer path.

6. In the directory you have unzipped, you will find the files *start_rs.bat* and *start_rs.sh*, which will allow you to start the Job server on Windows and Linux respectively, and the files *stop_rs.bat* and *stop_rs.sh* that will let you stop the Job server.

You may need to change the *java.library* path in order to load the correct native library for your system. In this case, adapt the variable `MY_JSYSMON_LIB_DIR` in the script *start_rs.sh*.

The JobServer is an application that allows a system installed on the same network as the Web application to declare itself as an execution server. These systems must obviously have a working JVM version 1.5 or 1.6.

For more information about the prerequisites of the JobServer, see Section 1.3, “ Supported Platforms”.

Now we simply have to declare these execution servers in the Web application and their resources (CPU, RAM, etc.) should become available. To do this:

1. Go to the **Servers** page of *Talend Administration Center*.

Only users that have Operation Manager role and rights can have a read-write access to this page. For more information on access rights, see your *Talend Administration Center User Guide. So, you have to connect to Talend Administration Center as an Operation Manager to be able to configure job servers.*

2. And define the server as follows:

Label	`TestingServer`
Description	`Type in the description of server.`
Host	`localhost`
Command port	`8000`
File transfer port	`8001`
Monitoring port	`8888`
Timeout on unknown status(s)	`120`
Username	Type in the username for user authentication to access a Job server.
Password	Type in the password for user authentication to access a Job server.

   This corresponds to the configuration of a Job server on the system that hosts the Web application. For any other system, the host field should contain the IP address of the system. Check also that the ports *8000*, *8001* and *8888* are available. These ports must be the same as defined in the *TalendJobServer.properties* defined above. Note that if no username and password pairs are defined in the file *users.csv* in the directory <*root*>*/conf/* where <root> is the JobServer path, then you do not have to set the **Username** and the **Password**.

3. Click the **Servers** page again so that the execution servers appear with their properties.

For some operating systems, the CPU information may not be available. You can test your system by setting up the following variable as `true`:

`org.talend.monitoring.jmx.api.OsInfoRetriever.FORCE_LOAD` in the file *TalendJobServer.properties*.

3.7.2. Configuring the JVM (optional)

The Jobserver provided by Talend allows you to choose another JVM than the one used by default to launch your Jobs.

To change the Job launcher path, proceed as follows:

1. Go to the directory *<root>/conf/* and open the *TalendJobServer.properties* file to edit it.

```
# Set the executable path of the launcher which will run the job, example: /usr/bin/java/MyJava1.6
org.talend.remote.jobserver.server.JobServerConfiguration.JOB_LAUNCHER_PATH="C:/Program Files/Java/jdk1.6.0_22/bin/java.exe"
```

2. In the line dedicated to the Job launcher path, add the path to your java executable after the equal sign.

The use of quotes is only necessary when your path contains spaces, as shown in the capture. Otherwise, type in the path without quotes.

3. Save your changes and close the file.

The next time you launch your Jobserver, the java executable used will be the one you have previously set in the *TalendJobServer.properties* file.

3.7.3. Configuring the SSL Keystore (optional)

The Jobserver provided by Talend allows you to encrypt data prior to transmission via an existing SSL Keystore. You are also able to choose another Keystore if needed.

To override the existing Keystore file, you have to:

- generate a new Keystore with the utility tool called Keytool (Key and Certificate Management Tool);
- set the new Keystore location;
- enable the SSL Keystore at server side.

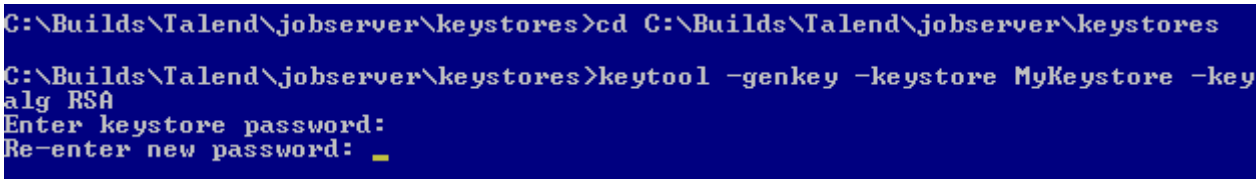
3.7.3.1. Generating Keystore

To generate a new Keystore file, complete the following.

1. Open a command prompt and change directory to *<root>/keystores* where <root> is the JobServer path.
2. Type in the following:

   `keytool -genkey -keystore <myKeystoreName> -keyalg RSA`

   where `<myKeystoreName>` refers to the name of the Keystore you are creating.



3. Enter the password for your Keystore twice, then enter the other optional information, such as your name, the name of your organization, your state etc., if needed.
4. Type in `yes` to confirm your information.
5. Type in the password you have previously defined. The new Keystore file has been created in *<root>/keystores*.

3.7.3.2. Setting the location of the new Keystore

To set the new Keystore location, you can either edit the JAVA_OPTS environment variable or edit the launching script (*start_rs.bat*) of the Jobserver.

To edit the JAVA_OPTS environment variable, complete the following:

If you haven't created the JAVA_OPTS environment variable yet, you have to create it before completing this procedure.

- Add

  ```
  -Djavax.net.ssl.keyStore=/<myDirectory>/<myKeystore>
  ```

  ```
  -Djavax.net.ssl.keyStorePassword=<myPassword>
  ```

  to your JAVA_OPTS environment variable, where <myDirectory> is the installation directory of your Keystore, <myKeystore> is the name of your Keystore and <myPassword> is the password you have previously defined for your Keystore.

Or

To edit the launching script, complete the following:

1. Open the *start_rs.bat* file to edit it.

   ```
   rem set the JVM arguments here
   set MY_JMV_ARGS=-Dcom.sun.management.jmxremote -Djava.library.path="%MY_ROOT_PATH%sigar-bin\lib>"
   -Djavax.net.ssl.keyStore=/C:/Builds/Talend/jobserver/keystores/MyKeystore
   -Djavax.net.ssl.keyStorePassword=
   ```

2. As shown in the capture, add

   ```
   -Djavax.net.ssl.keyStore=/<myDirectory>/<myKeystore> -
   Djavax.net.ssl.keyStorePassword=<myPassword>
   ```

   to the JVM arguments location, where <myDirectory> is the installation directory of your Keystore, <myKeystore> is the name of your Keystore and <myPassword> is the password you have previously defined for your Keystore.

3.7.3.3. Enabling the SSL encryption

After having defined your new Keystore, you have to enable Secure Sockets Layer (SSL) at server side in order to establish an encrypted link between the Jobserver and its clients. To do so, proceed as follows:

1. Go to the directory *<root>/conf/* and open the *TalendJobServer.properties* file to edit it.
2. Add the following line

   ```
   org.talend.remote.jobserver.server.TalendJobServer.USE_SSL=true
   ```

at the end of the file. The next time you launch your Jobserver, it will use the Keystore you have defined instead of the default one.

From *Talend Administration Center*, you have to select the **Use SSL** check box to enable the encryption.

3.8. Installing the *Activity Monitoring Console*

According to your monitoring needs, you may want to also install the *Activity Monitoring Console*.

3.8.1. Deploying the *Activity Monitoring Console* on Tomcat

To install the *Activity Monitoring Console*, follow the procedure below:

1. Unzip the archive file containing the *amc.war* file on the same machine as *Talend Administration Center* Web application.
2. Paste this *amc.war* file under the following directory: *<TomcatPath>/webapps/*

The next time you will start Tomcat, the *Activity Monitoring Console* application will be automatically deployed on the server.

3.8.2. Configuring the *Activity Monitoring Console* in *Talend Administration Center*

In the *Talend Administration Center* web application, you have to set up the link to the *Activity Monitoring Console*.

To do so, specify the following information on the **Dashboard** group of the **Configuration** page , :

- **AMC url**: type in the URL address of the *Activity Monitoring Console* application, *http://localhost:8080/amc* for example.

For more information about this application, see **Talend Activity Monitoring Console User Guide**.

3.9. Installing the SOA Manager

This feature is ONLY available from Talend Enterprise Data Quality RTx Edition.

JDK is recommended for the use of this feature.

Depending on your edition of Talend Enterprise Data Quality, and according to your needs, you may want to also install the SOA Manager.

To install the SOA Manager, follow the procedure below:

1. Unzip the archive file containing the *SOAmanager.jar* file on the same machine as *Talend Administration Center* Web application.
2. Launch the *.jar* using the following command:

```
java -jar <path_to_SOAManager>.jar <port_number>
```

On the **Configuration** page of *Talend Administration Center*, specify the following information:

- **Server port**: change the port number if needed.
- **Jobs deploy folder**: specify the path to the deployment folder.
- **Soamanager host**: specify the IP address of the SOA Manager server.
- **Default service namespace**: specify the default service namespace.

Only users that have Operation Manager role and rights can have a read-write access to this page. For more information on access rights, see your *Talend Administration Center User Guide. So, you have to connect to Talend Administration Center as an Operation Manager to be able to configure the SOA Manager.*

3.10. Installing a Business Rules Management System (Drools)

Depending on the edition you have subscribed to, you may want to install a Business Rules Management System based on JBoss Rules Governor for a centralized definition and administration of JBoss compliant business rules.

This feature is ONLY available from Talend Enterprise Data Quality Professional Edition.

A dedicated *.war* file has been delivered to you in the **Talend** archive.

- Unzip the *drools-guvnor.war* file in:

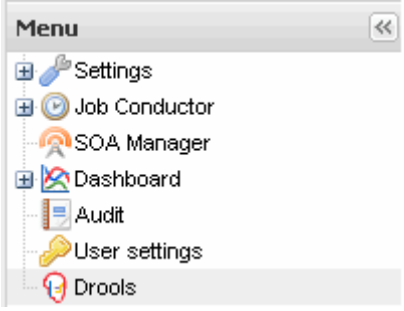
  <TomcatPath>/webapps/

After deploying the Drools BRMS, you will need to configure its URL in the *Talend Administration Center* Web application **Configuration** page. For more information about configuring the Web application, first read the Section 3.4, “Configuring *Talend Administration Center* (Web application)”.

Then after you access *Talend Administration Center*’s **Configuration** page, set Drools Guvnor’s URL to read http://localhost:8080/drools-guvnor/org.drools.guvnor.Guvnor/Guvnor.html

For more information about the actual data contained on the **Configuration** page of *Talend Administration Center*, check out the *Talend Administration Center* User Guide.

When Drools Guvnor URL is correctly set up, the application shows under the menu tree view of *Talend Administration Center*:



3.11. Setting up the automatic log purge (optional)

When deploying and executing tasks using the Job Conductor in *Talend Administration Center*, multiple files are created and stored locally.

These files include the Job archives (before deploying over job servers), execution & recovery logs, deployment logs.

You can configure *Talend Administration Center* to get rid of these legacy files when they are not used anymore:

1. Edit the configuration file of *Talend Administration Center*:

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

2. Set the following parameters:

```
### Temporary data cleaning parameters ###

# Time in seconds between each cleaning action, set with 0 to disable all
 cleanings
scheduler.conf.cleaning.frequencyCleaningAction=600

# Time in seconds before cleaning executions logs from the current date,
3 months=3*30*24*60*60=7776000, set with 0 to disable delete action with
 this method
scheduler.conf.cleaning.maxDurationBeforeCleaningOldExecutionsLogs=7776000

# Max executions logs to keep, set with 0 to disable delete action with
 this method.
scheduler.conf.cleaning.maxOldExecutionsLogs=1000
```

```
# Time in seconds before cleaning archives jobs from the current date, 3
 months=*30*24*60*60=7776000,
set with 0 to disable delete action with this method
scheduler.conf.cleaning.maxDurationBeforeCleaningOldJobs=7776000

# Max archive and deployed jobs to keep, set with 0 to disable delete
 action with this method.
scheduler.conf.cleaning.maxOldJobs=200
```

Take into account your available disk space when setting these parameters.

1. In the Job Server configuration file, the same parameters can be adapted:

   root/conf/configuration.properties

The following files are impacted by the auto-cleaning operation:

- Job archives in:

  <ApplicationPath>/archiveJobs

- Job execution & recovery logs in:

  <ApplicationPath>/executionLogs

- Job Server files in:

<ApplicationPath>/TalendJobServersFiles

Chapter 4. Configuring *Talend Enterprise Data Quality Studio* and AMC

This Chapter describes the configuration steps for *Talend Enterprise Data Quality Studio* and *Activity Monitoring Console*. It also provides useful information related to memory and launching issues you might encounter when working with *Talend Enterprise Data Quality Studio* and *Talend Administration Center*.

4.1. Prerequisites of *Talend Enterprise Data Quality Studio*

To use *Talend Enterprise Data Quality Studio* properly, you first need to install the following programs:

- Java Virtual Machine version 1.5+.
- External programs specific to bulk components (if you want to use Oracle, Sybase, Informix or Ingres bulk functionality). For more information, see the below sub-section.

On Windows XP and Windows Server 2003, the GDI is already installed. However, on Windows 2000, this installation is required. The GDI can be downloaded from Microsoft's Website [http://www.microsoft.com/downloads/en/details.aspx?FamilyID=6a63ab9c-df12-4d41-933c-be590feaa05a&DisplayLang=en]. For further information, visit *Eclipse's FAQ*.

On some Linux distributions, the xulrunner package is required for the Welcome page to display properly.

4.1.1. Installing database client software (for bulk mode)

Some bulk components, like Oracle, Sybase, Informix or Ingres, require database client software to run properly:

- OracleBulkExec uses the *sqlldr* external utility. This utility is available in Oracle clients that must be installed on the computer.
- Informix uses the *dbload* external utility.
- Ingres uses the *sql* external utility.
- Sybase uses the *bcp.exe* external utility. This utility is asked for in the Sybase bulk components' Basic Settings view. For more information, see *tSybaseBulkExec*, *tSybaseOutputBulk* and *tSybaseOutputBulkExec* components on *Talend Enterprise Data Quality Components* Reference Guide.

4.1.2. Installing the FileScale program (available for *Talend Enterprise Data Quality Big Data edition*)

If you subscribed to *Talend Enterprise Data Quality Big Data edition*, you will be able to use the FileScale technology which leverages the execution server hardware architecture and maximizes the performance of low-level sort algorithms. With FileScale components, you can transform and sort large quantities of data in large scale files.

The FileScale program can only be run on the following Windows and Linux operating system architecture:

- *win32-i386* for windows 32/64 bits (Windows 2000 or greater),
- *linux_elf-i386* for Linux 32 bits (Linux 2.6.18 or greater),
- *linux-amd64* for Linux 64 bits (Linux 2.6.18 or greater).

To install the FileScale executable:

- Unzip the archive file on the same client machine(s) as the one you installed the Studio on.
- Unzip the .tar file.

The DOC directory contains user manual and other data (example, benchmark, etc.).

- Open the directory corresponding to the architecture of your server.
- In this directory, create a file (without extension) named "*license*" containing your license key and only your license key (found in your email - For more information, see Section 1.1, "License key").

When installing the executable for windows, also copy subdirectory containing manifest files (in win32-i386).

FileScale is now installed on the client machine(s), you can use the FileScale components to handle large data files. For more information about the various components, see the *Talend Enterprise Data Quality Components* Reference Guide.

4.2. Configuring the *Talend Enterprise Data Quality Studio*

Installing *Talend Enterprise Data Quality Studio* on your client machines is as easy as these three steps:

- Unzip the archive under a directory. If possible, keep the path length at a minimum.
- Create a file (without extension) named "*license*" containing your license key and only your license key (found in your email - For more information, see Section 1.1, "License key").
- Paste the file at the root of the extracted directory at the same level as the *.exe* or *.sh* files.

Make sure this file is present at the root of the extracted directory and has no extension.

- Run the software according to your architecture (if you are running under a 64-bit capable OS then use x86_64 executable).

This license key is valid for ALL the modules and is particularly required for the Studio to start. For more information on the license key request and management, see sections Section 1.1, "License key" *and* Section 3.4.1, "Configuring the Web application access".

When launching *Talend Enterprise Data Quality Studio* for the first time, you'll need to set up the connection to the *Talend Administration Center* Web application in order to access your allocated projects.

You will thus be asked to fill in a **URL** field with the *Talend Administration Center*'s **URL**.

For more details about how to connect to the *Talend Administration Center*, see section Section 3.4.1, "Configuring the Web application access" and read the *Talend Administration Center* User Guide.

When filling out the URL field in the Studio, users should never use Localhost, but the actual IP address or the hostname of your database server. (e.g.: 192.168.1.1:8080/org.talend.administrator).

You may need to install the *Activity Monitoring Console* now.

For more information on *Activity Monitoring Console*, see section Section 4.4, "Configuring *Activity Monitoring Console*".

4.2.1. Installing external modules

Some specific components need specific java library classes to be installed to run properly. For more information on how to install those external libraries, see section *How to install external modules* of **Talend** Studio User Guide.

4.3. Memory configuration and launching issues

4.3.1. Configuring the *.ini* file on Linux/Windows

To gain in performance at run time and when launching the Studio, you can edit the memory settings in the *.ini* file corresponding to your OS.

Examples:

- Talend-linux-gtk-x86 -> Talend-linux-gtk-x86.ini
- Talend-win32-x86.exe -> Talend-win32-x86.ini

The default values are:

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

If your system resources allow you to change it, edit the default values to:

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

For more details: http://java.sun.com/docs/hotspot/HotSpotFAQ.html

If your Studio is installed on Unix-like system, add execution rights on the relevant Studio binary file. On a standard Linux distribution, type in the following command:

```
$ chmod +x TalendOpenStudio-linux-gtk-x86
```

```
$ ./TalendOpenStudio-linux-gtk-x86
```

4.3.2. For 64-bit platforms on Linux/Windows

You need to add the following JVM attributes to the "*-vmargs*":

```
-XX:CompileCommand=exclude,org/eclipse/core/internal/dtree/
DataTreeNode,forwardDeltaWith
```

in the *.ini* file corresponding to the binary file called in the *commandline.sh*.

4.3.3. For non Linux/Windows platforms

To be able to use a different JVM than the default one (for platforms other than Windows and Linux for example), edit the *commandline.sh* file and replace the following content (in **bold**):

```
./Talend-linux-gtk-x86
-nosplash -application org.talend.commandline.CommandLine -consoleLog
-data commandline-workspace startServer -p 8002
```

with:

```
/My_Jvm_Path/java -Xmx1024m -XX:MaxPermSize=256m -jar
plugins/org.eclipse.equinox.launcher_1.0.201.R35x_v20100507.jar
-nosplash -application org.talend.commandline.CommandLine -consoleLog
-data commandline-workspace startServer -p 8002 -vmargs
```

Don't forget the *-vmargs* information as the corresponding *.ini* file will not be taken into account any longer.

4.3.4. For 64-bit platform on non Linux/Windows OS

For non Linux/Windows 64-bit platforms, change the following information (in **bold**):

```
./Talend-linux-gtk-x86 -nosplash -application
org.talend.commandline.CommandLine -consoleLog -data
commandline-workspace startServer -p 8002
```

to

```
/My_Jvm_Path/java -jar
plugins/org.eclipse.equinox.launcher_1.0.201.R35x_v20100507.jar
-nosplash -application org.talend.commandline.CommandLine
-consoleLog -data commandline-workspace startServer -p 8002
-vmargs -Xmx500m -XX:MaxPermSize=256m
-XX:CompileCommand=exclude,org/eclipse/core/internal/dtree/
DataTreeNode,forwardDeltaWith
```

4.4. Configuring *Activity Monitoring Console*

Activity Monitoring Console is an application that allows you to monitor Job executions. The Job executions are monitored using three database tables that relate to the following data:

- collection of logs,
- component statistics,
- data flow volumes.

To store this data, you need to create three database tables respectively using the schema of the **tLogcatcher**, **tStatcatcher**, **tFlowMeterCatcher** connectors (present in the **Palette** of your *Talend Enterprise Data Quality Studio*).

1. Create a database that you name *amc*, for example.
2. Then, create a Job to create these tables in the *amc* database.
3. Launch a *Talend Enterprise Data Quality Studio* and connect to the *WORK* project that you just created when configuring the Talend Enterprise Data Quality client.

4. Create a Job that contains three **tCreateTable** components.
5. Define the connection parameters to the *amc* in the three **tCreateTable** components.
6. Define the data structure by taking the schema provided in the components: **tLogCatcher**, **tStatCatcher** and **tFlowMeterCatcher**.

In the tStatsCatcher schema, make sure the "Moment" data type is set to "datetime" type and not "date". To change this data type, click Edit Schema in the corresponding tCreateTable.

Now that the *Activity Monitoring Console* database is created, you can monitor the Job executions either from *Talend Enterprise Data Quality Studio* or from the **Dashboard** module of *Talend Administration Center* (available from *Talend Enterprise Data Quality Professional Edition*).

1. To do so, you simply need to tell *Talend Enterprise Data Quality Studio* and *Talend Administration Center* about the connection parameters to these tables.

From the *Talend Enterprise Data Quality Studio*, configure the *Activity Monitoring Console*:

1. Click **Edit>Project Settings>Job Settings>Stats&Logs.**
2. Then define the connection parameters to the tables.

The *Activity Monitoring Console* builds charts on the basis of the data collected during execution of various Jobs.

For more details on how to use the *Activity Monitoring Console*, refer to the *Activity Monitoring Console* User Guide.

From the Web application (*Talend Administration Center*):

1. Select the **Dashboard** menu, then click **Manage Connections.**
2. Set the connection parameters to the *LogCatcher* and *StatCatcher* tables.

Only users that have Operation Manager role and rights can have a read-write access to this page. For more information on access rights, see the Talend Administration Center User Guide. So, you have to connect to Talend Administration Center as an Operation Manager to be able to configure the amc monitoring database connection information.

For more details on how to use the **Dashboard**, refer to the *Talend Administration Center* User Guide.

Chapter 5. Upgrading Talend Enterprise Data Quality

This chapter describes the various operations required to migrate version of Talend Enterprise Data Quality .

In this chapter we assume that you have installed and configured Talend Enterprise Data Quality as described in Chapter 3, *Installing Talend Enterprise Data Quality for the first time*.

The migration and upgrade process includes the following mandatory steps:

1. Section 5.1, “Backing up the environment”
2. Section 5.2, “Migrating the administration database ”
3. Section 5.3, “Re-installing Commandline and JobServer”
4. Section 5.4, “Updating *Talend Administration Center* Web application”
5. Section 5.5, “Getting and managing the License key”
6. Section 5.6, “(specific to DB projects) Switching from database to SVN”

5.1. Backing up the environment

Before you start migrating your *Talend* solution, make sure your environment is correctly backed up and your *Talend Administration Center* configuration settings are saved.

Follow the procedure described in the subsections below.

5.1.1. Saving the configuration parameters

As part of the environment backup operation, we encourage you to export your current configuration parameters for *Talend Administration Center*:

1. Log on to *Talend Administration Center* as Administrator
2. Click **Configuration** on the menu tree view to the left. The **Configuration** page opens on the workspace.
3. Click the **Export parameters** button located at the top of the **Configuration** page to export your current setting into a file.

This file will be used when your new instance of *Talend Administration Center* will be set up in order to easily restore your configuration parameters.

5.1.2. Stopping the services

All users should be disconnected from Talend Enterprise Data Quality Studio.

Before you start the backup operations, make sure the following modules/services are stopped:

- Tomcat,
- Commandline,
- Job Server.

For more information about the various services, refer to Chapter 6, *Installing services* .

5.1.3. Backing up the administration database (MySQL, H2, MS SQL Server, or Oracle)

As a start of the migration operation, you need to back up the project administration database (*talend_administrator*).

5.1.3.1. In case your database is MySQL

Execute the following dump command:

On Windows:

```
mysqldump -u tisadmin -ptisadmin talend_administrator >
c:/<DB_Backuppath>/DumpTalendAdministrator.sql
```

On Linux:

```
mysqldump -u tisadmin -ptisadmin talend_administrator >
/home/<DB_Backuppath>/DumpTalendAdministrator.sql
```

If you want to load this back-up file to the *talend_administrator* database in case of error:

```
mysql -u tisadmin -ptisadmin talend_administrator</home/<DB_Backuppath>/
DumpTalendAdministrator.sql
```

where *tisadmin*, *ptisadmin* and *talend_administrator* are respectively your login, password and database name that you are required to use to connect to the administration database for *Talend* projects.

5.1.3.2. In case your database is H2

To backup your current H2 database, simply perform the following operation:

1. Copy the following files

   <Applicationpath>/WEB-INF/Database/talend_administrator.h2.db *<Applicationpath>/WEB-INF/Database/talend_administrator.trace.db*

2. Save these files in a convenient and safe place for backup.

Note that an automatic backup is performed every day and at each startup and shutdown of the application server. Oldest backups are deleted when their number is above 30. All these parameters are set by default in the *configuration.properties* file, you can customize them according to your need. For more information, see Section C.2.2, “Backing up the H2 database”.

5.1.4. Backing up the SVN repository

If your projects are currently stored in a database and you want to switch to SVN, skip this section and read directly Section 5.2, “Migrating the administration database ”.

After you backed up the administration database, you need to back up the SVN repository where all projects items and metadata are stored.

To do so, execute locally the following command on the machine hosting the SVN repository, using the

```
svnadmin
dump
```

command:

```
$> svnadmin dump /talend/svn/repotis > svn_repotis.dump
```

where */talend/svn/repotis* is the path on the hard drive where the repository is stored and > is the copy command into a dump file

If you need to restore your SVN repository, then perform the following operations:

1. Shut down the subversion daemon.
2. Delete the */talend/svn/repotis* repository
3. Create again the repository from the backup file, using this command:

```
$> mkdir /talend/svn
$> svnadmin create /talend/svn/repotis
$> svnadmin load /talend/svn/repotis

< svn_repotis.dump
```

where */talend/svn/repotis* and *svn_repotis.dump* are the respective path and name used to store the dump file you created.

Beware that only the repository is restored, not the users.

5.2. Migrating the administration database

After the environment is properly backed-up, you are required to launch migration script described in the following sections to upgrade the *Talend Administration Center* parameters in the database. For further information about database backup, see Section 5.1.3, “Backing up the administration database (MySQL, H2, MS SQL Server, or Oracle)”.

5.2.1. For MySQL, SQL Server, or Oracle

Depending on your OS and architecture, perform one of the following commands:

On 32-bit Linux:

```
./Talend-linux-gtk-x86 -nosplash -application
 org.talend.repository.TalendMigration
-consoleLog -data migration-workspace jdbc:mysql://10.42.0.10:3306/
talend_administrator
talend pwd
```

On 32-bit Windows:

```
Talend-win32-x86.exe -nosplash -application
 org.talend.repository.TalendMigration
-consoleLog -data migration-workspace jdbc:mysql://10.42.0.10:3306/
talend_administrator
talend pwd
```

“Talend-linux-gtk-x86” as well as “Talend-win32-x86.exe” are only given as examples. Use the actual .exe name of the NEW edition.

This script performs some checks while upgrading the parameters and will prevent you to upgrade if some projects are still stored in Database mode, in which case, you need to switch from database to SVN mode. For more information about SVN storage mode for your projects, see Section 5.6, “(specific to DB projects) Switching from database to SVN”.

5.2.2. For H2

Before migrating your H2 database, it is preferable to copy the existing database where the new web application is installed.

Depending on your OS and architecture, perform one of the following commands:

On 32-bit Linux:

```
./Talend-linux-gtk-x86 -nosplash -application
 org.talend.repository.TalendMigration
-consoleLog -data migration-workspace “jdbc:h2://<tomcat or
jboss installation path>/webapps/org.talend.administrator/WEB-INF/database/
talend_administrator;AUTO_SERVER=TRUE;MVCC=TRUE"
tisadmin tisadmin
```

On 32-bit Windows:

```
Talend-win32-x86.exe -nosplash -application
 org.talend.repository.TalendMigration
-consoleLog -data migration-workspace “jdbc:h2://<tomcat or
jboss installation path>/webapps/org.talend.administrator/WEB-INF/database/
talend_administrator;AUTO_SERVER=TRUE;MVCC=TRUE"
tisadmin tisadmin
```

“Talend-linux-gtk-x86” as well as “Talend-win32-x86.exe” are only given as examples. Use the actual .exe name of the NEW edition.

5.3. Re-installing Commandline and JobServer

Extract the new Talend Enterprise Data Quality clients and reinstall CommandLine and JobServer on the relevant systems as described in Section 3.5, “Installing Commandline” and Section 3.7, “Installing the execution servers (JobServers)”.

In case you are migrating your projects from database to SVN repository, read now the Section 5.6, “(specific to DB projects) Switching from database to SVN” before you update the Web application.

5.4. Updating *Talend Administration Center* Web application

When upgrading Talend Enterprise Data Quality, you need to update also the *Talend Administration Center*. To do so, follow the procedure described below:

1- Clear your previous *Talend Administration Center* files

- Via the Tomcat manager, undeploy the previous version of the *Talend Administration Center* in Tomcat.
- Delete all the remnants as well as the *Talend Administration Center* folder (highlighted in **bold**) from the Tomcat Web application folder:

*<TomcatPath>/work/Catalina/localhost/<**ApplicationPath**>*

- Delete the content of the *temp* folder from the Tomcat Web application folder:

  apache-tomcat-6.0.18/temp

- Flush the browser cache.

- Reinstall the *Talend Administration Center .war* file (*org.talend.administrator.war*).

  To do this, follow the procedure described in detail in Section 3.2, “Deploying the Web application on an application server”.

- Restart Tomcat.

2- Set up your *Talend Administration Center* configuration back:

- In your Internet browser address bar, type in the *Talend Administration Center* URL to open the **Login** page.

  http://<IPAddress>:<Port>/<TalendAdministrationCenterWebApplicationName>/

where <TalendAdministrationCenterApplicationName> is the name for *org.talend.administrator*, if it was changed when deploying the .war.

e.g.:

http://10.42.10.56:8080/MyTAC/

- On the **Login** page, click the **Go to db config page** link.

If the **Go to db config page** *link of the* **Login** *page is hidden, you need to activate it from the configuration.properties file. For more information on how to get the DB Config page link back, see* Section A.17, “Where is the “Go to DB Config” button gone?”.

- You will be prompted to enter a password to unlock the page. By default the password is “*admin*”.

- On the **Database Configuration** page, click the **Import parameters** button.

- In the **Parameters Import** dialog box, click **Browse** to browse to the *administrator_config.txt* file that contains the parameters exported from *Talend Administration Center*, and then click the **Upload** button.

Your configuration settings are now ported over.

Beware that the migration from a version prior to 4.0 involves a change of project architecture within the SVN repository. The project architecture now includes various levels such as trunk and branches. This migration is carried out automatically during the migration. No action is required by the user. However if your project is big, this automatic operation can be rather long. For more information, see Section A.15, “Following a migration to 4.x, I get an error when I try to connect to my project in *Talend Administration Center*”.

5.5. Getting and managing the License key

In the case of migration from one version to another, you might be required to manage a license key at various application levels: *Talend Administration Center*, *Talend Enterprise Data Quality Studio* and possibly the *Activity Monitoring Console*.

Before starting to use *Talend Administration Center*, you need to enter your license key:

1. Click **License** on the navigation bar.

2. Then click **Edit License key**.

3. Copy-paste the license key.

If your Web access is restricted, you may need to click **Generate Validation request** to perform the validation of your license key. Follow the instructions on screen.

For more information about the Talend Enterprise Data Quality license key request and management, see Section 3.4.1, “Configuring the Web application access”.

Then to be able to launch the Studio, copy and paste the License key in a file (without extension) that you will place **at the root of the application’s installation directory.**

This license key is required for the Studio to start.

5.6. (specific to DB projects) Switching from database to SVN

If you want to migrate your projects from a database to an SVN repository, follow the procedure described below before starting any migration to a newer version of Talend Enterprise Data Quality. The version of *Talend Enterprise Data Quality Studio* and *Talend Administration Center* should be the same and prior to 4.1.

1- Prior to migrating your projects from the DB to svn, perform the following operations:

- Back up the database. For more information see Section 5.1, “Backing up the environment”.
- Then, in the Studio, export all items into an archive file for each DB project you have. For more information about how to export items from the Studio, see *Talend Enterprise Data Quality Studio* User Guide.
- In your *Talend Administration Center* Web application, delete each DB project one by one.

  For each project, wait until the message **project <projectname> deleted** displays on the **Client Log** page of *Talend Administration Center*. The deletion process can take some time.

Make sure you are logged in with sufficient privileges to delete the DB projects from your *Talend Administration Center*.

In case of error during the project deletion process from *Talend Administration Center*, you can use the following command from the Commandline:

For Linux:

```
./TalendDataQuality-linux-gtk-x86 -nosplash -application
org.talend.dbrepocleaner.TalendDBRepoCleaner -consoleLog -data temp
 org.gjt.mm.mysql.Driver
<DatabaseJDBC_URL> <DB_UserName> <DB_Password>
```

For Windows:

```
TalendDataQuality-win32-x86.exe -nosplash -application
 org.talend.dbrepocleaner.TalendDBRepoCleaner
-consoleLog -data temp org.gjt.mm.mysql.Driver <DatabaseJDBC_URL>
<DB_UserName> <DB_Password>
```

where <*DatabaseJDBC_URL*>, <*DB_UserName*> and <*DB_Password*> are the login details to connect to the old Project database.

- Then, install SVN. For more information, see Section 3.1, “Setting up your project server on SVN”.

All operation steps described in this procedure are to be carried out from your Talend Enterprise Data Quality Studio and Talend Administration Center versions older than 4.1.

2- For EACH project to be moved from the DB to SVN repository, follow the procedure:

- Create an SVN project from the *Talend Administration Center* Web application. The project url should be of type: `<repository url>/<project name>` (e.g.: http:/miro/reposvn/Project1).
- Set the user authorizations for this new project as described in the corresponding section of *Talend Administration Center* User Guide.
- Connect to this project via the *Talend Enterprise Data Quality Studio*.
- Import items from the relevant project created in the SVN repository.

Now that you have transformed your DB projects into SVN projects, you can now update *Talend Enterprise Data Quality Studio* and *Talend Administration Center* to a newer version. For details, see Section 5.4, “Updating *Talend Administration Center* Web application”.

Chapter 6. Installing services

This chapter explains primarily the procedure of installing Commandline and JobServer as a service.

6.1. Configuring JSL and *.ini* file

A *jsl.ini* file is integrated in the JobServer and ready for configuration. However, before starting to install CommandLine as a service, you must have previously downloaded JSL.

Java Service Launcher allows you to launch Java applications as Windows services. To download it, you can go to http://sourceforge.net/projects/jslwin/.

Once you have downloaded JSL, proceed as follows:

1. Unzip the file *jsl.zip* into a directory called *jsl*.

Once JSL is configured, you will be able to copy the content of this new directory into the directory of the application you want to install as a service. However, note that this operation is not necessary for the JobServer that already holds an integrated JSL. For more information, see Section 6.2, “On Windows” and Section 6.3, “On Linux”.

2. Open the file *jsl_static.ini* that is inside the directory in order to configure it.

The *jsl_static.ini* file may not exists in the JSL directory. In that case, copy the *jsl.ini* file and rename it as follows: *jsl_static.ini*.

If you have already installed your JobServer, a *jsl.ini* file is directly available from the directory *<JobserverPath>\conf*.

```
[service]
appname = TalendJobServer
servicename = TalendJobServer
displayname = Talend JobServer
```

At the beginning of the file, you can see different names: `appname` corresponds to the name of your application, `servicename` corresponds to the name you want to give to the service and `displayname` corresponds to the name that will be displayed when creating the service.

If you are not satisfied with those default names, edit the part that is after the = sign according to your needs.

```
;environment variable substitution is possible in any value
;e.g.
;jrepath=%JAVA_HOME%
;% is escaped by %%
```

Under this area, you can see that, by default, the JAVA_HOME environnement variable is pointing towards the JRE path.

```
[java]
;Path to the java runtime used
;If this option is not used the default from the registry will be used
;jrepath=c:\java\jdk15
jrepath=C:\Program Files\Java\jdk1.5.0_14
```

This path is set in the **[java]** part of the *.ini* file. If the default path is not the path of the installation directory of your application, you can edit the part that is after the = sign and write this path.

6.1.1. Configuring the CommandLine parameters

To configure the CommandLine parameters that are passed to the Java application, you have to edit the last lines of the *jsl_static.ini* file which indicate the path to different files (such as configuration files, *.jar* files and/or java libraries). To do so, edit the parameters as shown below:

```
params = 15
param00 = -Xms64m
param01 = -Xmx1024m
param02 = -XX:MaxPermSize=256m
param03 = -jar
param04 = <CommandLinePath>\plugins
\org.eclipse.equinox.launcher_1.1.0.v20100507.jar
param05 = -nosplash
param06 = -application
param07 = org.talend.commandline.CommandLine
param08 = -consoleLog
param09 = -data
param10 = <CommandLinePath>\commandline-workspace
param11 = startServer
param12 = -p
param13 = 8002
```

where <*CommandLinePath*> refers to the installation directory of the CommandLine application.

Note that the first parameter indicated is named "param00" and not "param01", and that the use of quotes is not required around strings holding spaces. Note also that every parameter of the command line must be indicated in a seperate parameter.

6.1.2. Configuring the JobServer parameters

To configure the JobServer parameters that are passed to the Java application, you have to edit the last lines of the *jsl.ini* file which indicate the path to different files (such as configuration files, *.jar* files and/or java libraries). To do so, edit the parameters as shown below:

```
params = 6
param00 = -cp
param01 = <JobServerPath>;<JobServerPath>\lib\jmxremote_optional.jar;
<JobServerPath>\lib\log4j-1.2.15.jar;<JobServerPath>\lib\commons-
io-1.3.2.jar;
<JobServerPath>\lib\jsl.jar
param02 = -Djava.library.path=<JobServerPath>\jsysmon-20061229\native
\windows\
param03 = -Dcom.sun.management.jmxremote
param04 = -Dtalend_props=<JobServerPath>\conf\TalendJobServer.properties
param05 = org.talend.remote.jobserver.service.WindowsJobServer
```

where <*JobServerPath*> refers to the installation directory of the JobServer application.

Note that the first parameter indicated is named "param00" and not "param01", and that the use of quotes is not required around strings holding spaces. Note also that every parameter of the command line must be indicated in a seperate parameter.

6.2. On Windows

6.2.1. CommandLine

We strongly encourage you to read Section 6.1, « Configuring JSL and *.ini* file » before starting this section.

To install CommandLine as a service on 32-bit Windows, follow the procedure below:

For a CommandLine installed on 64-bit Windows, use jsl_static64.exe in place of jsl_static.exe in the relevant below procedure steps.

1. Unzip the *Talend-All<subscription>-V<version_number>.zip* file where *<subscription>* corresponds to your subscription and *<version_number>* corresponds to the product version you have. For example, *Talend-All-r59885-V4.2.0.zip*.

2. To facilitate the installation, rename the *Talend-All-<subscription>-V<version_number>* directory into *CMDLine*.

   To reproduce the following steps, you have to put the downloaded JSL in the *CMDLine* directory.

3. Open a CMD window in Administrator mode.

4. Change the directory to the CommandLine installation directory.

5. Type in the following command:

   ```
   jsl_static.exe -debug
   ```

   ```
   Env Entry: export = CLASSPATH,PATH
   Env Entry: CLASSPATH =
   Env Entry: PATH = %PATH%
   Env afterlookup: PATH 0
   Env: PATH = C:\WINDOWS\system32;C:\WINDOWS;C:\WINDOWS\System32\Wbem;C:\Program F
   iles\Intel\Wireless\Bin\;C:\Program Files\MySQL\MySQL Server 5.0\bin;C:\Program
   Files\TortoiseSVN\bin;C:\Program Files\TortoiseGit\bin
   Debugging CMDLine.
   Now starting JVM
   Path for JVM: C:\WINDOWS\system32;C:\WINDOWS;C:\WINDOWS\System32\Wbem;C:\Program
    Files\Intel\Wireless\Bin\;C:\Program Files\MySQL\MySQL Server 5.0\bin;C:\Progra
   m Files\TortoiseSVN\bin;C:\Program Files\TortoiseGit\bin;C:\Program Files\Java\j
   re6\bin
   ControlHandler registered after JVM start
   ```

   This command creates a CommandLine service. It starts, you can now check that the *<CMDLINE_DATA_DIR>\commandline-workspace* directory has been created as well.

6. Stop execution by pressing **Ctrl**+**C**.

   ```
   Stopping CMDLine.
   Trying to stop service via JNI call to java/lang/System.exit
   Done
   Exiting

   C:\Documents and Settings\lgaudens>
   ```

   The service stops, and the message `Exiting` is returned.

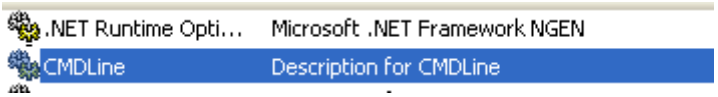
7. In the CMD window, type in the following command:

   ```
   jsl_static.exe -install
   ```

```
C:\Documents and Settings\lgaudens>D:\Builds\CMDLine\jsl_static.exe -install
Installing service with command: D:\Builds\CMDLine\jsl_static.exe
CMDLine  installed as a Windows service.

C:\Documents and Settings\lgaudens>_
```

The **Talend** CommandLine service is created and can be viewed by selecting **Control Panel**>**Administrative Tools**>**Services**.



6.2.2. JobServer

We strongly encourage you to read Section 6.1, « Configuring JSL and *.ini* file » before starting this section.

The procedure for enabling JobServer to run as a service under Windows is slightly different. To make it, follow the steps below:

1. Edit the file *jsl.ini* in `<JobServerPath>/conf`:

   Replace all the occurrences of the string *c:\jobserver* with your own JobServer installation directory `<JobServerPath>`.

2. Change the path for your JVM by entering, for example: `jrepath=<JREPath>`.

3. Open a command window in administrator mode and change the directory to `<JobServerPath>`.

4. Execute the command:

   ```
   jsl.exe -debug
   ```

```
Env Entry: param02 = -Djava.library.path=D:\Builds\Talend\JobServer\jsysmon-2006
1229\native\windows\
Env Entry: param03 = -Dcom.sun.management.jmxremote
Env Entry: param04 = -Dtalend_props=D:\Builds\Talend\JobServer\conf\TalendJobSer
ver.properties
Env Entry: param05 = org.talend.remote.jobserver.service.WindowsJobServer
Debugging Talend JobServer.
Now starting JVM
Path for JVM: C:\WINDOWS\system32;C:\WINDOWS;C:\WINDOWS\System32\Wbem;C:\Program
 Files\Intel\Wireless\Bin\;C:\Program Files\MySQL\MySQL Server 5.0\bin;C:\Progra
m Files\TortoiseSVN\bin;C:\Program Files\TortoiseGit\bin;C:\Program Files\Java\j
re6\bin
ControlHandler registered after JVM start
```

5. Stop execution by pressing **Ctrl**+**C**.

6. In the CMD window, execute the following command

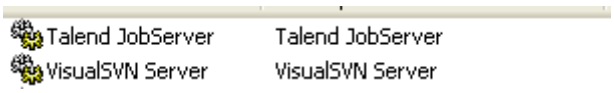
   ```
   jsl.exe -install
   ```

```
C:\Documents and Settings\lgaudens>D:\Builds\Talend\JobServer\jsl.exe -install

Installing service with command: D:\Builds\Talend\JobServer\jsl.exe
Talend JobServer installed as a Windows service.

C:\Documents and Settings\lgaudens>_
```

The **Talend** JobServer service is created and can be viewed by selecting **Control Panel**>**Administrative Tools**>**Services**.



Talend JobServer is the default name of the service, which can be configured in the file *<JobServerPath>/conf/jsl.ini*.

To remove the service you can enter the following command in the same command window:

```
jsl.exe -remove
```

6.2.3. SVN

VisualSVN should launch automatically as a service.

But if it does not automatically run as a service, the procedure on Windows is similar to the JobServer as a service procedure.

6.3. On Linux

The method is the same for CommandLine and JobServer.

Note that the following information is provided for Debian/Ubuntu distributions. For information regarding Redhat, Solaris or OpenSuse, see Appendix B, *Installing services on non-Debian distributions*

6.3.1. CommandLine

To install CommandLine as a service on Linux, carry out the following procedure:

1. Create an executable from which commandline can be run in the directory */usr/bin*.

   To do this, you need to create two files, *commandline_start* and *commandline_stop* containing the following:

   For *commandline_start*:

   ```
   #!/bin/sh

   cd <CommandLinePath>

   nohup sh <CommandLinePath>&
   ```

   For *commandline_stop*:

   ```
   #!/bin/sh

   (echo stopServer ; sleep 2) | telnet localhost
   ```

   The path *<CommandLinePath> should be absolute.*

2. Ensure that the two files above are executable. To do this, you can execute the two commands below in the */usr/bin* directory:

```
chmod +x /usr/bin/commandline_start

chmod +x /usr/bin/commandline_stop
```

3. Paste the *sh* file: */<TISPath>/addons/scripts/tis_commandline* into the directory: */etc/init.d* in order to create the service related to the two executables above.
4. Make the file executable using:

   ```
   chmod +x /etc/init.d/tis_commandline
   ```

5. Execute the following command:

   ```
   sudo update-rc.d tis_commandline defaults 60
   ```

Now you have created the service, related to the two executable files *commandline_start* and *commandline_stop*.

6.3.2. JobServer

To create a service for the JobServer, carry out the following procedure:

1. Create an executable from which the JobServer can be run in the directory */usr/bin*.

   To do this, you need to create two files, *jobserver_start* and *jobserver_stop* containing the following:

   For *jobserver_start*:

   ```
   #!/bin/sh

   cd <JobServerPath>

   nohup sh <JobServerPath>&
   ```

   For *jobserver_stop*:

   ```
   #!/bin/sh

   stop_rs.sh
   ```

   The path <*JobServerPath*> should be absolute. Note also that the file stop_rs.sh can be found in the directory named *org.talend.remoteserver*.

2. Ensure that the two files above are executable. To do this, you can execute the two commands below in the */usr/bin* directory:

   ```
   chmod +x /usr/bin/jobserver_start

   chmod +x /usr/bin/jobserver_stop
   ```

3. Paste the *sh* file: */<TISPath>/addons/scripts/tis_jobserver* into the directory: */etc/init.d* in order to create the service related to the two executables above.
4. Make the file executable using:

   ```
   chmod +x /etc/init.d/tis_jobserver
   ```

5. Execute the following command:

```
sudo update-rc.d tis_jobserver defaults 60
```

Now you have created the service, related to the two executable files *jobserver_start* and *jobserver_stop*.

6.3.3. SVN

Because users access SVN via Apache and Apache is automatically installed as a service, then SVN does not need installing as a service.

SVN works automatically as a service.

Appendix A. FAQ

This Chapter addresses some frequently asked questions related to the installation of Talend Enterprise Data Quality.

A.1. I cannot launch *Talend Enterprise Data Quality Studio* although I have a valid License Key

To be able to launch the *Talend Enterprise Data Quality Studio*, you need to paste the License Key you received by email in a file (without extension) that you place at the root of the extraction folder (where you unzipped *Talend Enterprise Data Quality Studio* application).

A.2. I cannot connect to *Talend Administration Center*

When installing or updating Talend Enterprise Data Quality, you cannot access the web application, *Talend Administration Center*. Check the database URL and make sure you replaced the default URL with the relevant DB server URL.

For more details on how to configure *Talend Administration Center*, see Section 3.4.1, “Configuring the Web application access”

A.3. A “Malformed token” message is displayed on the top of the License page of *Talend Administration Center*, what should I do?

Validate your license or update your license token by generating a validation request. For more information, see *Talend Administration Center* User Guide.

A.4. I cannot install the JVM (Sun) on my Debian distribution

You may be missing an entry in the **source.list** file.

Edit the file:

`/etc/apt/source.list`

Then add:

```
deb http://www.backports.org/debian etch-backports main
contrib non-free
```

Then type in the following commands to install the JVM (Sun):

```
$> apt-get update

$> apt-get install sun-java5-jdk
```

A.5. I cannot install Apache tomcat 5.5 on a Debian distribution

To install the Tomcat server, type in the following command:

```
$> apt-get install tomcat5.5
```

When getting the error at launching, edit the file: `/etc/init.d/tomcat5.5` and add the JDK 1.5

To do so, edit the following line:

```
JDK_DIRS="/usr/lib/jvm/java-6-sun [...]
```

```
$> apt-get install tomcat5.5-admin
```

Add in `/etc/tomcat5.5/policy.d/04webapps.policy` the following lines:

```
permission java.util.logging.LoggingPermission "control",
"";

permission java.io.FilePermission "./*", "read,write,execute,delete";

permission java.util.PropertyPermission "user.dir", "read";

permission java.util.PropertyPermission "disableLuceneLocks", "read";

permission java.util.PropertyPermission "java.io.tmpdir", "read";

permission java.util.PropertyPermission "org.apache.*", "read";

permission java.io.FilePermission "/-", "read,write,execute,delete";

permission java.lang.RuntimePermission "createClassLoader", "";

permission java.net.SocketPermission "*", "resolve, connect";
```

And add in `/etc/tomcat5.5/policy.d/50user.policy` the following lines:

```
// These permissions apply to all JARs from Debian packages

grant codeBase "file:/var/lib/tomcat5.5/webapps/-" {

? permission java.security.AllPermission;};

grant codeBase "file:/var/lib/tomcat5.5/webapps/-" { permission
 java.security.AllPermission;
```

}

A.6. Tomcat is installed as a service on Windows and *Talend Administration Center* returns several errors. In the log file, I get some “Java heap space” exceptions.

By default, Tomcat service limits the max memory of the JVM to 64Mb. However 256Mb seems to be the minimum threshold for *Talend Administration Center* to operate correctly.

Go to Tomcat **Status** page (ex: [http://localhost:8080/manager/status]*http://localhost:8080/manager/status*).

In the **JVM** area, check the **Free memory**, **Total memory** and **Max memory** values.

When using *Talend Administration Center*, if **Free memory** < 32Mb and **Max memory** = **Total memory**, you must increase the JVM max memory.

To increase the JVM max memory, run `<ApacheInstallPath>\bin\tomcatw.exe`.

On the **Java** tab, fill in the **Maximum memory pool** field with *256* (or more if other applications are installed on Tomcat).

A.7. When I access a page of *Talend Administration Center*, the memory usage jumps and never goes back down. I get some “Java heap space” exceptions.

This issue, related to Talend Enterprise Data Quality v3.0.0 up to 3.0.5, is linked to a Tomcat memory leak as well as SVN libraries used.

To fix this issue, the following updates are recommended:

- Update the SVN libraries, in:

```
<ApplicationPath>/WEB-INF/lib/
```

using the *.jar* files provided (*svnkit.jar svnkit-javahl.jar trilead.jar jna.jar*) in

http://www.svnkit.com/org.tmatesoft.svn_1.2.3.standalone.zip

- Upgrade to Tomcat 6.0.

A.8. *Talend Enterprise Data Quality Studio* fails to start on Linux environments containing an unexpected version of XULRunner.

Be aware that up to 3.2 (included) Talend Enterprise Data Quality requires Mozilla 1.4 GTK2 - 1.7.x GTK2 and XULRunner 1.8.x - 1.9.0.x to function properly.

If you run Talend Enterprise Data Quality with the latest distributions of Linux you may have only XULRunner 1.9.1 installed. *Talend Enterprise Data Quality Studio* is not yet compliant with this XULRunner version as it doesn't operate yet on Eclipse 3.5.

To resolve this incompatibility issue, proceed as follows.

- Set up an older version of XULRunner than the default one.
- To automate the launching of XULRunner, add the following commandline to the *.ini* file:

  ```
  Dorg.eclipse.swt.browser.XULRunnerPath=/usr/lib/xulrunner-1.9.0.10
  ```

This should resolve the issue.

A.9. When I click on a task in the Task table or when Job Conductor refreshes Jobs, the Job, Version and Context fields in the Task parameter view may display "Data unavailable" or null values.

The problem may be caused in one of the two following occasions:

- You are using the versions 3.1.2 or 3.1.3 of Talend Enterprise Data Quality.

  In this case, download the patch from this address: *http://www.talendforge.org/bugs/file_download.php?file_id=4547&type=bug* [http://www.talendforge.org/bugs/file_download.php?file_id=4547&type=bug],

  and install it using the file *install.txt* provided with the patch.

Or

- You are using version 3.1.4 of Talend Enterprise Data Quality, but the network between the Web application and SVN repository is slow.

  In this case, increase the **SVN timeout** value in the *configuration.properties* file of *Talend Administration Center*.

A.10. I get the error: java.lang.UnsatisfiedLinkError: no jsysmon in java.library.path

This error can occur when information on CPU is not available for the system in use. Nevertheless, the JobServer agent can still operate.

OS	CPU info. avail.
Linux i386 2.6	yes
Linux i386 2.4	yes
Linux amd64 2.6	yes
windows xp x86 5.1	yes
windows xp	yes
windows	yes
windows 2003 x86 5.2	no
windows 2003	no

This list is not exhaustive. Other systems could be supported too.

If you want to test your system, set the following variable in the *TalendJobServer.properties* file:

```
org.talend.monitoring.jmx.api.OsInfoRetriever.FORCE_LOAD
= true
```

If you are on Linux amd64, then you need to manually replace *jsysmon-20061229/native/linux_2.6/libjsysmon.so*

by

jsysmon-20061229/native/linux_2.6/amd64/libjsysmon.so

A.11. I have set up various triggers which do not run consistently nor on the times specified.

To resolve this issue, you need to increase the value of the quartz property in order for this value to always be greater than the concurrent tasks' max running time:

- In *quartz.properties* file, increase the value of the following property: `org.quartz.threadPool.threadCount`.

For example, for 20 max concurrent executions, you can set:

```
org.quartz.threadPool.threadCount = 20
```

A.12. *Talend Administration Center* does not automatically validate the license when a http proxy is in use.

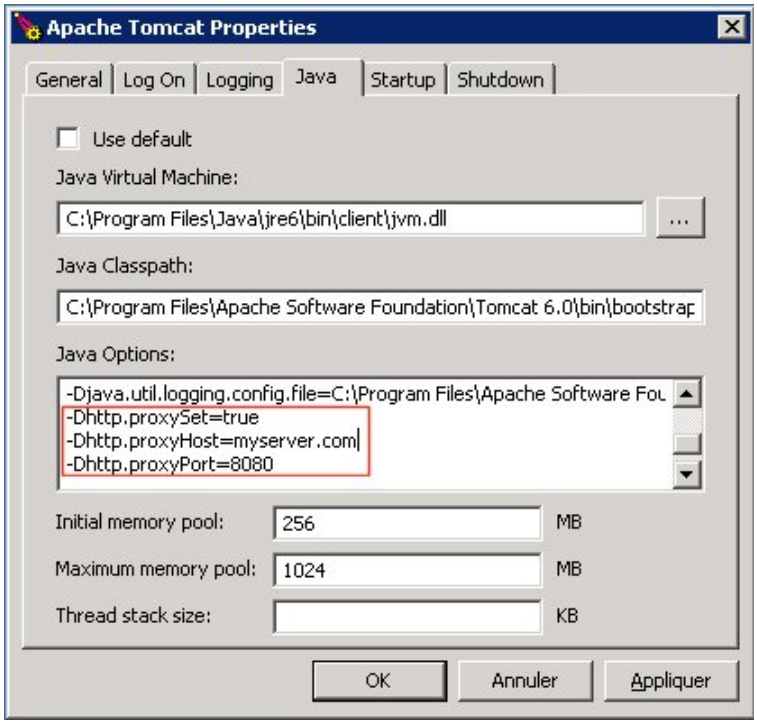
You need to set the http proxy properties with the Java options in the Tomcat manager.

To do so, proceed as follows. The example given below is for Windows:

- Click the **[Java]** tab in the **[Apache Tomcat Properties]** dialog box.
- In the **Java Options** area, add the following parameters:

```
-Dhttp.proxySet=true
-Dhttp.proxyHost=<yourproxyserver.com>
-Dhttp.proxyPort=8080
```

The window reads as follows



- Click **OK**.

A.13. I cannot create a project or I have graphical issues on Ubuntu 9.10

There is a known issue between Ubuntu 9.10 and *Talend Enterprise Data Quality Studio* version 3.2 or earlier, due to version compatibility with Xulrunner 1.9.1.

This issue should be solved starting from Talend Enterprise Data Quality 4.0 as Xulrunner is now embedded (with Eclipse 3.5).

If you encounter this issue, you might want to try the following operation:

- Install Xulrunner 1.9.0.x using the following command:

```
sudo apt-get install xulrunner-1.9
```

- Edit the .ini file residing at the root of your *Talend Enterprise Data Quality Studio* installation folder, and add the following line:

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

For more complete information, please read the following wiki page that provides up-to-date information regarding installations on Ubuntu 9.10, OpenSuse 11.2, etc: http://www.talendforge.org/wiki/doku.php?id=doc:installation_on_ubuntu. [http://www.talendforge.org/wiki/doku.php?id=doc:installation_on_ubuntu]

A.14. I get a 500 error ‘java.lang.ClassNotFoundException: javax.faces.convert.ConverterException’ after deploying JBoss Guvnor

In order for your JBoss Guvnor web application to operate, you need to paste some *.jar* files into the *tomcat lib* directory.

- Add the following *.jar* files in the *tomcat lib* directory, such as, for Linux:

usr/share/tomcat6/lib/jsf-impl.jar
usr/share/tomcat6/lib/jsf-api.jar
usr/share/tomcat6/lib/jboss-faces.jar

The .jar files can be downloaded from http://download.jboss.org/drools/dependencies/jboss_jsf_libs.zip.

A.15. Following a migration to 4.x, I get an error when I try to connect to my project in *Talend Administration Center*

Prior to v4.x, all **Talend** projects were stored at the root of the SVN repository (“trunk” level). With v4.x, **Talend** introduced the capability to manage various branches of a project within the SVN repository.

During a migration to v4.x of *Talend Administration Center*, the SVN repository project architecture has thus been changed to match the SVN typical architecture. The projects now include various levels such as: trunk, branches, tags.

- The architecture migration operation can take quite a long time, depending on the size of your project. Therefore, in case of connection error, wait a moment after you restart Tomcat before trying to connect to your project.

or

- If during the migration, the project URL could not be accessed, the architecture migration within the SVN cannot be performed. In this case, you will be required to enter the correct URL in order to connect to your project.

A.16. I do not want to use H2 as default database

By default an H2 embedded database is used to store cross-project data such as users, projects, authorization, license, tasks, triggers, dashboard, etc. But if you prefer, you can set up a different database, (MySQL, SQL Server or Oracle) to store these data and metadata. To set up this database and link it to your *Talend Administration Center*, follow the procedure hereafter.

A.16.1. Setting up the Database server

The project administration information (user accounts, rights...) can be stored and managed in a MySQL, SQL Server or Oracle database.

Take care to have at hand the java database driver (.jar) corresponding to the type and version of the database you are using. Contact your database vendor for more information about java database driver.

Before you get started, make sure to paste your database in the path, corresponding to the way you intend to administrate your users and project in *Talend Administration Center*:

/<ApplicationPath>/WEB-INF/lib/

1. If you want to administrate projects and users via the *Talend Administration Center* Web application, then paste the .jar file into:

   Note that the MySQL java database driver (.jar) file is generally provided and thus is already present in the above path. However for SQL Server and Oracle, the relevant file is NOT provided.Contact your database vendor for more information about java database driver.

2. If you want to administrate projects and users connection pool via Tomcat, then paste the .jar file directly into:

   <TomcatPath>/webapps/

   Note that the database engine should support transactions and foreign keys, therefore you need to make sure that the database you choose includes a transactional engine, such as Oracle, MySQL InnoDB, MySQL Falcon (not MyIsam), for example.

A.16.1.1. Setting up the DB server on MySQL

In the rest of this document, all procedures provided by default are relative to MySQL use.

1. On Windows, start the installer.
2. Select the custom installation.
3. Select "**Server**" as Type of installation.

Lets assume that the user login is "*tisadmin*" and that the password is "*tisadmin*". For more MySQL commands, see Section A.16.1.2, "Examples of MySQL Commands".

After you have created the **talend_administrator** database, make sure that the "*tisadmin*" user has, for example, read-write authorization to it from any IP address. For details on how to grant read-write rights, see Section A.16.1.3, "Granting access authorizations".

A.16.1.2. Examples of MySQL Commands

Here are some commands that are required to work with MySQL in order to launch the service by hand:

	Linux	**Windows**
Start the MySQL service	sudo /etc/init.d/mysql start	net start "MySQL"
Stop the MySQL service	sudo /etc/init.d/mysql stop	net stop "MySQL"

To run a MySQL command window connected to the "**MyDatabase**" database using the user "*tisadmin*" and the password "*tisadmin*", you can enter:

```
mysql –u tisadmin –p tisadmin MyDatabase
```

Now, create an empty database using the following command in MySQL server:

```
create database talend_administrator
```

Here are some more MySQL commands:

	Command
Create/delete a database	create/drop database [Name];
Connect to the DB database	connect DB;
Display the tables in a database	show tables;
Display the structure of the table Table	desc Table;
Exit MySQL	quit;

A.16.1.3. Granting access authorizations

After you have created the *talend_administrator* database, make sure that the "*tisadmin*" user has, for example, read-write authorization to it from any IP address. Do this using:

```
GRANT ALL PRIVILEGES ON talend_administrator.* TO 'tisadmin'@'%'
IDENTIFIED BY 'tisadmin'
```

All the commands used to start the services can be wrapped in a **.bat* file in Windows or a **.sh* file in Linux, to allow them to be started later as though they were programs.

In the case where the MySQL instance does not run on the same server as Tomcat, make sure to authorize connections from external applications.

To authorize connections from external applications, edit the *my.cnf* file and add a comment as follows:

```
bind_address = 127.0.0.1
```

After completing the installation of the database server which stores the administration data, you need to set up the server that will store the actual **Talend** project data. For more information on how to install your project server, see Section 3.1, “Setting up your project server on SVN”

A.16.2. Configuring the database access from *Talend Administration Center*

After completing the database server setup, you need to configure *Talend Administration Center* so that it can access the database:

- In your Internet browser address bar, type in the **DB config page** URL .

  http://<IPAddress>:<Port>/<TalendAdministrationCenterWebApplicationName>/DbConfigModule.html

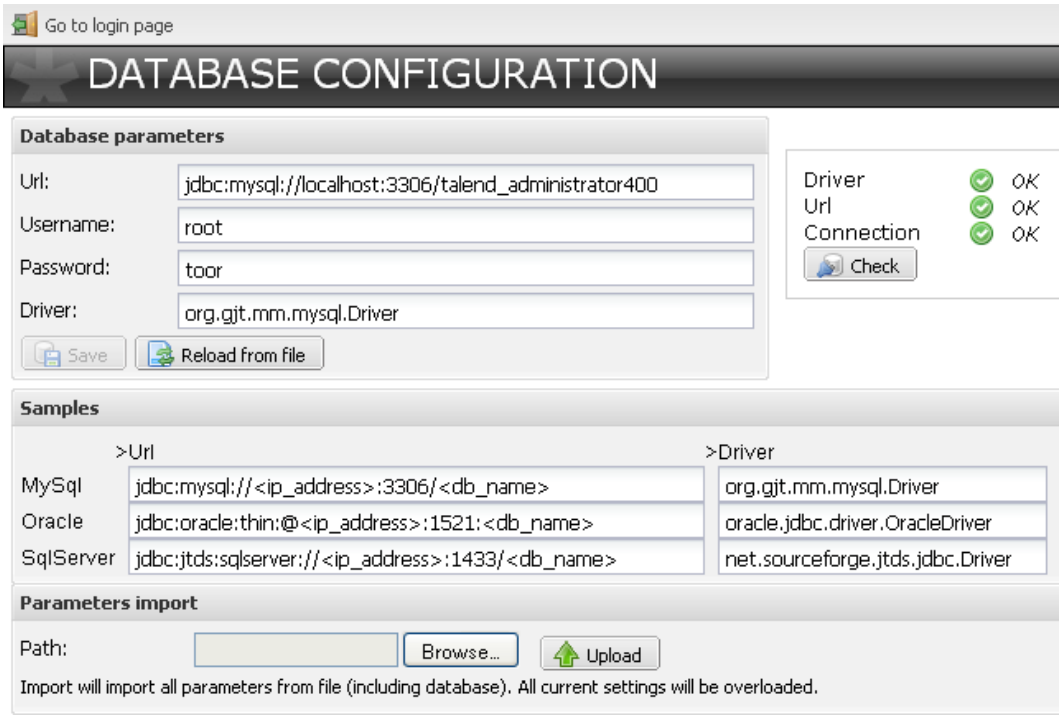
where <TalendAdministrationCenterApplicationName> is the name for *org.talend.administrator*, if it was changed when deploying the .war.

e.g.:

http://10.42.10.56:8080/org.talend.administrator/DbConfigModule.html

By default, the **Go to db config page** *link of the* **Login** *page is hidden. To display it, you have to activate it from the configuration.properties file. For more information on how to get the DB Config page link back, see Appendix* Section A.17, “Where is the “Go to DB Config” button gone?”.

- You will be prompted to enter a password to unlock the page. By default the password is “*admin*”.



- Complete the **URL** and **Driver** information (Replace “*localhost*” with the IP address of the server used to host the database). Some samples of URLs are provided for your convenience in the **Samples** area or in the section .

- Change the username and password to those you configured during the Db server configuration (see section Section 6.2, “On Windows”). By default: username=“*tisadmin*” and password=“*tisadmin*”.

- Once the DB connection parameters are set, go back to **Login** page via the **Go to Login page** link.

A.16.2.1. Samples of URL and Driver

Depending on the DB type, driver and URL information may vary:

MySQL
`driverClassName="org.gjt.mm.mysql.Driver"`
`url="jdbc:mysql://<host>:<port>/<dbname>"`

SQL server
`driverClassName="net.sourceforge.jtds.jdbc.Driver"`
`url="jdbc:jtds:sqlserver://<host>:<port>/<dbname>"`

Oracle
`driverClassName="oracle.jdbc.driver.OracleDriver"`
`url="jdbc:oracle:thin:@<host>:<port>:<dbname>"`

A.17. Where is the “Go to DB Config” button gone?

By default, the **Go to db config** link is shown on the **Login** page of *Talend Administration Center*.

However if, if it has been hidden for any reason or if you want to use a different database than the default H2 one, you may want bring this **Go to db config** link back to *Talend Administration Center*’s **Login** page again in order to access the database access configuration settings.

To do so, simply open the *configuration.properties* file and edit the following line:

```
database.showDbConfigButton=true
```

A.18. When I try to connect to *Talend Administration Center*, I get an error “Incompatible database schema version”.

The database schema does not match the schema expected by *Talend Administration Center*. The reason for it is that you try to connect to a newer version *Talend Administration Center* with an older version of the database. The schema being different, it cannot connect.

The reason for it is that you haven’t updated your administration database prior to install the newer version of *Talend Administration Center*. For more information about the administration database update and migration process, please refer to Chapter 5, *Upgrading Talend Enterprise Data Quality* and in particular Section 5.2, “Migrating the administration database ” regarding database update.

A.19. Librairies are not deployed when I deploy my Job onto the Job Servers.

From 4.1, for performance reasons, a cache is used automatically at deployment time, to avoid libraries to be redeployed every time a Job is deployed to a Job Server. This helps improve the deployment time performance as only new or updated libraries get deployed.

If you encounter issues and want to revert back to the automatic redeployment of all librairies each time a Job is deployed to a Job Server, open the *configuration.properties* file and edit the following parameter such as to read:

```
jobServer.useCache=false
```

A.20. I want to create a Perl project, but I can’t find the Perl option

From 4.2 onwards, Perl is deprecated and its support will be discontinued at the next major release

In v4.2, however, Perl is only disabled by default and can easily be enabled again.

To create Perl Projects and Jobs, you need to re-activate the Perl option by carrying out the following operation in :

- In the following file where <release_revision> corresponds to your product version.

  org.talend.resources.perl_<release_revision>/resources/config.properties

- Set the following parameter to `True`:

  ```
  enablePerl=false
  ```

A.21. I am trying to set up high availability but I get the following error on MS SQL Server: '... FOR UPDATE clause allowed only for DECLARE CURSOR ...'

In case of error while trying to set up a High availability clustering, follow the short procedure herafer:

1. Open the following file:

   <Applicationpath>/WEB-INF/classes/quartz.properties

2. Uncomment the following parameter

```
### May be required using MS SQL Server if the following error occurs:
 '... FOR UPDATE clause allowed
only for DECLARE CURSOR
...' ###
```

```
#org.quartz.jobStore.selectWithLockSQL = SELECT * FROM {0}LOCKS UPDLOCK
WHERE LOCK_NAME = ?
```

A.22. When I run a Job dealing with a .csv file with accented characters in its path, I get the error ‘Exception on component xxx‘

By default, accented characters are not supported in the Studio. To insure that accented characters are recognized, you have to change the encoding to UTF-8 in your .ini file.

To do so, proceed as follows:

The example given below is for Windows.

- Open your .ini file to edit it. It reads as follows.

```
1  -vmargs
2  -Xms64m
3  -Xmx768m
4  -XX:MaxPermSize=256m
5
```

- Edit the .ini file and type in:

```
-Dfile.encoding=UTF-8
```

This row must be added under the last row of the .ini file. In the example above, the row you have to edit is the fifth one.

- Save these settings and close the .ini file.

Close your Studio and relaunch it. The accented characters are now supported.

A.23. In the Data Quality Portal, when I try to open a report via the Analytical Processing tab or the QBE tab, I get an error message.

This is due to the fact that the resource path is not correctly configured for Windows. In order to resolve this issue, you will have to edit the server.xml file.

To do so, proceed as follows:

The example given below is for Windows.

1. Open the following file.

   <DQPortalInstallationpath>/apache-tomcat/conf/server.xml

2. Look at the line `<Environment name="spagobi_resource_path" type="java.lang.String" value="../resources"/>`. It reads as follows.

```
<Environment name="spagobi_resource_path" type="java.lang.String" value="D:/tdqp/apache-tomcat/resources"/>
```

3. Edit the line to remove the part between `value="` and `resources"/>` as shown in the following capture.

```
<Environment name="spagobi_resource_path" type="java.lang.String" value="resources"/>
```

If you still get the same error message, you should replace this path by `"<absolute_Tomcat_path>/resources/"`.

Appendix B. Installing services on non-Debian distributions

This Chapter provides examples of init files that could be used to install Talend Enterprise Data Quality on a non-Debian Linux distributions. The examples are given for Redhat, Solaris and OpenSuse distributions.

B.1. Example of CommandLine daemon file

B.1.1. On Redhat

Below are some commands that can be used to launch the CommandLine on a Redhat distribution:

```
#!/bin/sh
### BEGIN INIT INFO
# Default-Start:     2 3 4 5
# Default-Stop:      S 0 1 6
# Short-Description: Example initscript
# Description:       This file should be used to construct scripts to
be
#                    placed in /etc/init.d.
### END INIT INFO

# Author: jsd03

# Do NOT "set -e"

# PATH should only include /usr/* if it runs after the mountnfs.sh script
PATH=/usr/sbin:/usr/bin:/sbin:/bin
COMMANDLINE=/oa/TIS/commandline303
CMD_WORKSPACE=$COMMANDLINE"/commandline-workspace"
STARTUP=commandline.sh
USER=cxp
NAME=commandline
PORT=10004


# Read configuration variable file if present
[ -r /etc/default/$NAME ] && . /etc/default/$NAME

# Load the VERBOSE setting and other rcS variables
[ -f /etc/default/rcS ] && . /etc/default/rcS

# Define LSB log_* functions.
# Depend on lsb-base (>= 3.0-6) to ensure that this file is present.
. /lib/lsb/init-functions

#
# Function that starts the daemon/service
#
do_start()
{
echo -n $"Starting commandline service: "
 su - $USER -c "rm -Rf $CMD_WORKSPACE"
 su - $USER -c "cd $COMMANDLINE && screen -dmS cmdLine ./$STARTUP"
#screen -dmS cmdLine $DAEMON_START
 RETVAL=$?
 echo

}
```

```
#
```

```
# Function that stops the daemon/service
#
do_stop()
{
 echo -n $"Stopping commandline service: "
 su - $USER -c "( echo 'stopServer' ; sleep 2 ) | telnet localhost $PORT"
 RETVAL=$?
 echo

}

case "$1" in
start)
do_start
 ;;
stop)
 do_stop
 ;;
*)
 echo $"Usage: $0 {start|stop}"
 exit 1
esac

exit 0
```

B.1.2. On Solaris

Same method as to run Job Server as a service applies for the Commandline. For more information read the Section B.2, “Example of Job Server init file”, Section B.2.2, “On Solaris”.

B.2. Example of Job Server init file

Below are some init script commands in order to start up the Job Server.

B.2.1. On Redhat

```
# chkconfig: 345 91 10
# description: Starts and stops the jobserver daemon.
#

# Source function library.
. /etc/rc.d/init.d/functions

# Get config.
. /etc/sysconfig/network
```

```
# Check that networking is up.
[ "${NETWORKING}" = "no" ] && exit 0

user=cxp
```

```
jobserver=/u/bin/Talend/jobserver_3.0.1
startup=start_rs.sh
shutdown=stop_rs.sh


start(){
 echo -n $"Starting joserver service: "
 su - $user -c "cd $jobserver && sh $startup &"
 RETVAL=$?
 echo
}

stop(){
 echo -n $ "Stopping jobserver service: "
 su - $user -c  "cd $jobserver && sh $shutdown"
 RETVAL=$?
 echo
}

restart() {
 stop
 start
}

# See how we were called.
case "$1" in
start)
 start
 ;;
stop)
 stop
 ;;
restart)
 restart
 ;;
*)
 echo $"Usage: $0 {start|stop|restart}"
 exit 1
esac

exit 0
```

B.2.2. On Solaris

The method provided below may not be the only method to run the Job Server as a service. More information can be found on the Sun documentation website:

1. Create an xml file named *tisagent.xml*,
2. Place this file under the following path

   /var/svc/manifest/application/

3. Include the following commands to the file you've just created:

   ```
   <?xml version="1.0"?>
   ```

```
<!DOCTYPE service_bundle SYSTEM "/usr/share/lib/xml/dtd/
service_bundle.dtd.1">

<service_bundle type='manifest' name='tisagent'>

<service
        name='application/tisagent'
        type='service'
        version='1'>

        <create_default_instance enabled='true' />

        <single_instance/>

        <exec_method
                type='method'
                name='start'
                exec='/lib/svc/method/tisagent start'
                timeout_seconds='150'
                 />

        <exec_method
                type='method'
                name='stop'
                exec='/lib/svc/method/tisagent stop'
                timeout_seconds='150' />

        <template>
                <common_name>
                     <loctext xml:lang='C'>
                          JobServer Agent: Tis component
                     </loctext>
                </common_name>
                <documentation>
                          <manpage title='TIS JobServer agent'
 section='5'
/>
                     <doc_link name='talend.com'
                          uri='http://talend.com'/>
                </documentation>
        </template>

</service>
</service_bundle>
<>
```

4. Then create a file named *tisagent*.
5. Place it under

   /lib/svc/method

6. In this file, type in the script that will be called when enabling/disabling the service:

```
#!/bin/sh

case "$1" in
'start')
        cd <path to talend jobserver>
```

```
        sh start_rs.sh &
        ;;

'stop')
        cd <path to talend jobserver>
        sh stop_rs.sh
        ;;

*)
        echo "Usage: $0 {start|stop}"
        exit 1
        ;;
esac
exit 0
```

You can now import the configuration file using the following command:

```
svccfg import /var/svc/manifest/application/tisagent.xml
```

You can also enable the service using the following command:

```
svcadm enable tisagent
```

And you can also make sure that the service is running using the following command:

```
svcs |grep tis
```

And checking that the output is:

```
online Apr_20 svc:/application/tisagent:default
```

B.2.3. On OpenSuse

The following operation needs to be performed with root privileges.

This procedure was tested on OpenSuse 11.2 x64 bits.

1. Make sure that the three scripts *joberserver_start*, *joberserver_stop* and *joberserver* are executable.
2. Copy *usr/bin/jobserver_start* and *usr/bin/jobserver_stop* into:

   /usr/bin/
3. Copy *etc/ini.d/jobserver* in:

   /etc/init.d/
4. Edit the configuration file *etc/sysconfig/jobserver* and set the path to your installation directory
5. Copy this file into:

   /etc/sysconfig/
6. Execute the following command to create a link called *rcjobserver*:

   ```
   ln -s /etc/init.d/jobserver /usr/sbin/rcjobserver
   ```
7. To start or stop the Job Server manually, use:

```
rcjobserver start

rcjobserver stop
```

8. Install the service using:

```
Yast > System > System Services
```

9. Then type in:

```
chkconfig -e jobserver
```

10. And set the variable to *ON*

11. Run SuSEconfig.

The Job Server installation path can be edited through *Yast > /etc/sysconfig Editor* in *Applications/ Talend*.

B.3. Example of Tomcat init file

Below is an example of init script in order to launch Tomcat server as daemon.

```
# chkconfig: 345 91 10
# description: Starts and stops the Tomcat daemon.
#

# Source function library.
. /etc/rc.d/init.d/functions

# Get config.
. /etc/sysconfig/network

# Check that networking is up.
[ "${NETWORKING}" = "no" ] && exit 0

user=cxp
tomcat=/u/bin/Tomcat/apache-tomcat-5.5.26/
startup=$tomcat/bin/startup.sh
shutdown=$tomcat/bin/shutdown.sh
#export JAVA_HOME=/usr/local/jdk

status(){
ps ax --width=1000 | grep "[o]rg.apache.catalina.startup.Bootstrap start"
| awk '{printf $1 " "}' | wc | awk
'{print $2}' > /tmp/tomcat_process_count.txt
read line < /tmp/tomcat_process_count.txt
if [ $line -gt 0 ]; then
echo -n "tomcat ( pid "
ps ax --width=1000 | grep "[o]rg.apache.catalina.startup.Bootstrap start"
| awk '{printf $1 " "}'
echo -n ") is running..."
echo
else
echo "Tomcat is stopped"
fi
```

```
}

start(){
 echo -n $"Starting Tomcat service: "
 #daemon -c
 su - $user -c "$startup"
 RETVAL=$?
 echo
}

stop(){
 action $"Stopping Tomcat service: " su - $user "$shutdown"
 RETVAL=$?
 echo
}

restart(){
  stop
  start}
```

```
# See how we were called.
case "$1" in
start)
 start
 ;;
stop)
 stop
 ;;
status)
      # This doesn't work ;)
 status tomcat
 ;;
restart)
 restart
 ;;
*)
 echo $"Usage: $0 {start|stop|status|restart}"
 exit 1
esac

exit 0
```

Appendix C. 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.

C.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). For more information about the setup procedure, see FAQ Appendix Section A.16, “I do not want to use H2 as default database”.

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

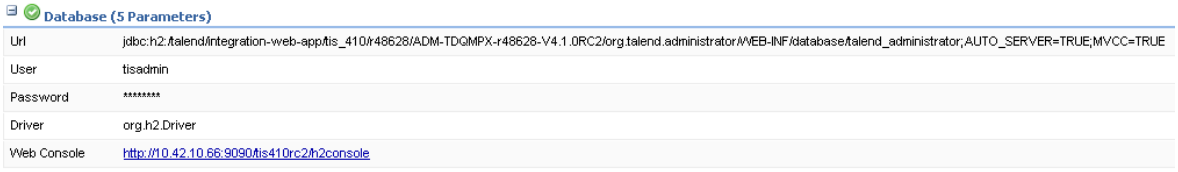
C.2. Administrating the H2 database through the Web console

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

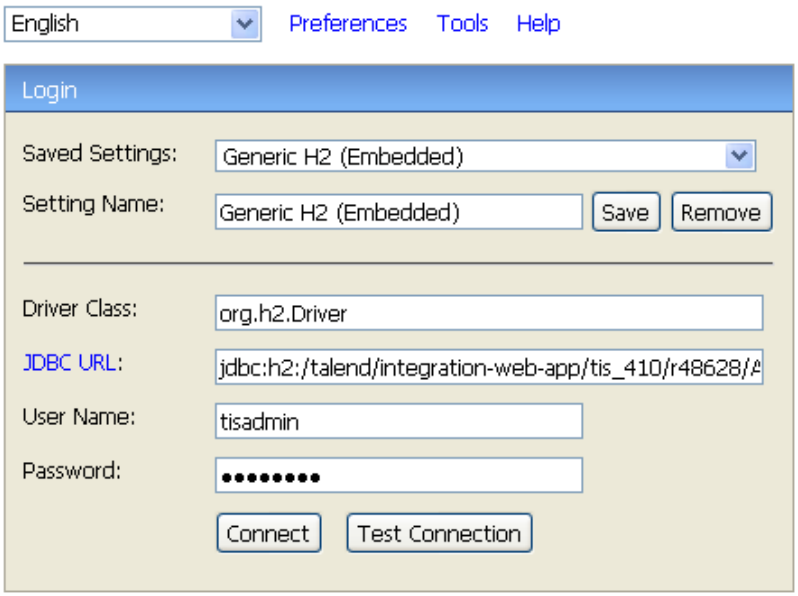
C.2.1. Connecting to the H2 Web Console

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, develop the **Database** node to display the parameters.



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



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

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

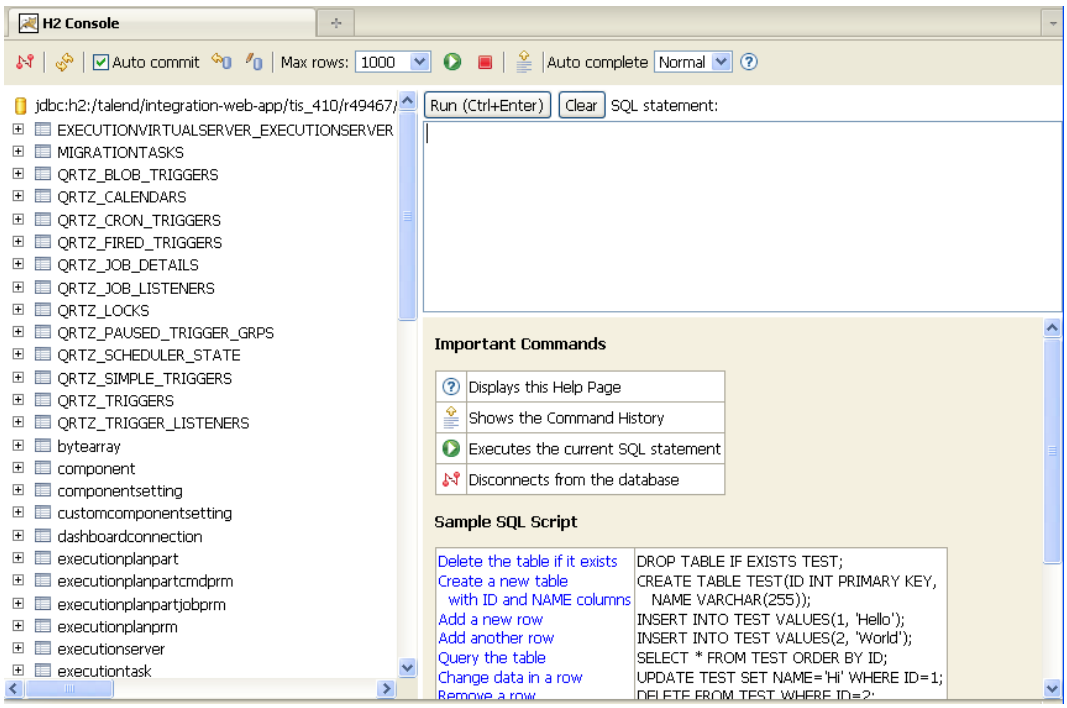
   jdbc:h2:/<ApplicationPath>/WEB-INF/database/
   talend_administrator;AUTO_SERVER=TRUE;MVCC=TRUE;LOCK_TIMEOUT=15000

   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. Then click the Test Connection button prior to clicking Connect 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.

C.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 everyday-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/docs/tutorials/crontrigger.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

C.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 www.h2database.com/html/download.html [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.