Talend Server Install Guide - 6.4.1

Created 2 MySQL databases for Talend Administration Center and Talend Data Quality Portal

Details of Database are:

Database hos: dev-emr-hue-metastore.cluster-cbduemg5i5w8.us-east-1.rds.amazonaws.com

Database port : 3306

Database Name :tac1

User Name :Master

Password: Adm1n!2017

Database hos: dev-emr-hue-metastore.cluster-cbduemg5i5w8.us-east-1.rds.amazonaws.com

Database port : 3306

Database Name :dqp1

User Name : Master

Password: Adm1n!2017

Download and Install Java JDK 1.8

Talend requires the Oracle Java JDK 1.8 present in order to install the applications. The

OpenJDK often found on linux distributions by default is not compatible. For non-Debian linux

installations (RedHat, Fedora, CentOS) it is recommended to download the RPM installation file.

Step 1 - Download JDK Install File

Visit the following page and download the appropriate file.

http://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html

Step 2 - Install JDK

Once downloaded run the following command to install the JDK.

**rpm -ivh jdk-8u121-linux-x64.rpm**

Step 3 - Verify Java Install

After the install completes confirm that java is installed and is reporting the proper version..

**java –version**

You should see the following output.

**java version "1.8.0\_121"**

**Java(TM) SE Runtime Environment (build 1.8.0\_121-b13)**

**Java HotSpot(TM) 64-Bit Server VM (build 25.121-b13, mixed mode)**

Download Talend Installation Files

Talend will supply an email which contains a license file attachment as well as numerous links

to download Talend installation files. Using this email you will need to download the relevant

files for your particular installation.

Step 1 - Locate Installation Email

To do a Talend server install, you will need to locate the email sent to you by Talend containing

the user and password as well as the download links for the software.

Additionally there will be a username and password in the email needed to complete start the

download.

**NOTE** : We will download the files in a later step using wget directly from the linux terminal.

Step 2 - Create Directory Structure for Install

In order to keep things organized, we will create a few directories to separate installation files

from installed files.

Create the following directories:

**/data/talend**

**/data/talend/install**

**/data/talend/Talend-6.4.1**

Step 3 - Use wget to download the Talend install files

While you can download the files directly from the email via your browser, and then SFTP them

to your linux environment, it is recommended that you use the linux utility “wget” to simply this

download process.

Provided you have the email from Talend, follow these steps:

**cd /data/talend/install**

**wget**

<http://www.opensourceetl.net/tis/tpdsbdrt_641/Talend-Installer-Starter-20170623_1246-V6.4.1-installer.zip>**--user=alpeshdesai**

**--password=\*\*\*\*\*\*\*\*\*\*\*\***

**wget** <http://www.opensourceetl.net/tis/tpdsbdrt_641/dist>

**--user=**alpeshdesai**--password==\*\*\*\*\*\*\*\*\*\*\*\***

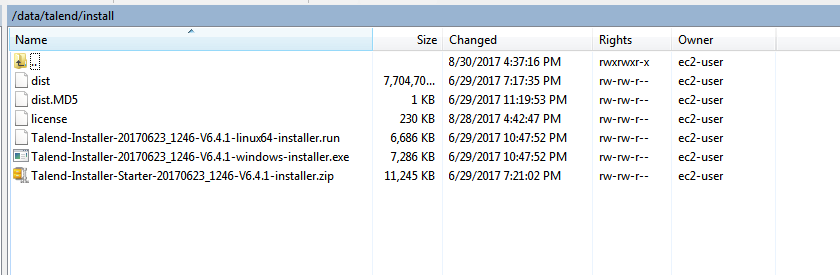
**wget** <http://www.opensourceetl.net/tis/tpdsbdrt_641/dist.MD5>**--user=**alpeshdesai**--password==\*\*\*\*\*\*\*\*\*\*\*\***

The format of the wget command is as follows:

wget <link to file> --user=<username> --pass=<password>

After running the above commands you will have the following files in the directory:

**/data/talend/install**



Step 4 - Verify dist MD5

Since the dist file is very large, we will verify that the dist file was downloaded completely.

To verify the dist file type the following commands:

**md5sum dist**

The output should read:

**18a465c7c051b2ee12c161c253115961**

Then view the contents of dist.MD5 by typing

**more dist.MD5**

The output should read:

**18a465c7c051b2ee12c161c253115961**

These two outputs should match, provided that they do, you can continue with the installation. If

they do not, you will need to re-download dist.

Step 5 - Change directory permissions

Next we will change the directory permissions to allow the script to execute.

Type the following commands

**cd /data/talend**

**chmod -R 755 install**

Step 6 - Run Talend Installer

**cd install**

**chmod+x Talend-Installer-20170623\_1246-V6.4.1-linux64-installer.run**

**./** **Talend-Installer-20170623\_1246-V6.4.1-linux64-installer.run**

----------------------------------------------------------------------------

Welcome to the Talend Installation Wizard.

----------------------------------------------------------------------------

Please read the following License Agreement. You must accept the terms of this

agreement before continuing with the installation.

---------------------

-------------------------

Press [Enter] to continue :

End User Software and Subscription Agreement

Talend EULA - June 2017

Press [Enter] to continue :

Do you accept this license? [y/n]: y

----------------------------------------------------------------------------

Talend Installer (v. 6.4.1)

Please specify the directory where Talend modules will be installed.

Installation Directory [/opt/Talend-6.4.1]: /data/talend/Talend-6.4.1

Please choose one of the installation styles below:

[1] Easy Install

[2] Advanced Install

Please choose an option [1] : 2

Please choose one of the installation types below:

[1] Server - Install all Talend server components using default configuration.

[2] Client - Install Talend Studio on your machine.

[3] Custom - Choose and configure each component to be installed individually.

Please choose an option [1] : 3

License File: []: /data/talend/install/license

----------------------------------------------------------------------------

Select the components you want to install; clear the components you do not want

to install. Click Next when you are ready to continue.

Talend Administration Center [Y/n] :Y

Talend Log Server [Y/n] :Y

Talend IAM Service [Y/n] :Y

Talend Data Stewardship [Y/n] :Y

Talend Command Line [Y/n] :Y

Talend Runtime [Y/n] :Y

Talend Job Server (Execution Server) [Y/n] :Y

Talend Data Quality Portal [Y/n] :Y

Talend Data Preparation [Y/n] :Y

Talend SAP RFC Server [Y/n] :n

Talend Studio [Y/n] :n

Talend ESB [Y/n] :Y

Talend Server Services [Y/n] :Y

Is the selection above correct? [Y/n]: Y

----------------------------------------------------------------------------

Talend Administration Center Configuration

Talend Administration Center (TAC) installs on an Apache Tomcat server. Please choose one of the options below:

[1] Install an embedded tomcat8 server (recommended).

[2] Use an existing tomcat server.

Please choose an option [1] : 1

----------------------------------------------------------------------------

Create TAC administrator user:

Admin User Name [security@company.com]: admin@capgemini.com

Admin Password [\*\*\*\*\*\*\*\*] :

----------------------------------------------------------------------------

Enable SSO (default TAC administrator user will not be created in TAC) [y/N]: N

----------------------------------------------------------------------------

Talend Administration Center Configuration

tacDB

Talend Administration Center connects to a database to store its configuration and metadata. Please choose one of the options below:

[1] Embedded H2 database (not recommended)

[2] Connect to a MySQL database

[3] Connect to an Oracle database

[4] Connect to a MS SQL Server database

[5] Connect to a PostgreSQL database

Please choose an option [1] : 2

Talend Administration Center Port (tomcat port)

[8080]: 8080

Talend's webapp (directory) name under tomcat:

[org.talend.administrator]: org.talend.administrator

Setup Email Notifications [y/N]: N

----------------------------------------------------------------------------

Talend Administration Center - Database Configuration

This page enables you to configure Talend Administration Center to use an

existing MySQL database.

Database host [localhost]: dev-emr-hue-metastore.cluster-cbduemg5i5w8.us-east-1.rds.amazonaws.com

Database port [3306]: 3306

Database Name [talend\_administrator]: tac1

User Name [talend]: Master

Password [\*\*\*\*\*\*\*\*] :

----------------------------------------------------------------------------

Talend Artifact Repository Nexus Configuration

Talend Artifact Repository Nexus Configuration

Nexus Port

[8081]: 8081

Nexus Host

[0.0.0.0]: 0.0.0.0

----------------------------------------------------------------------------

Talend Log Server Configuration

----------------------------------------------------------------------------

Cluster Name [talend-log-central]: talend-log-central

----------------------------------------------------------------------------

Talend Identity and Access Management

Select Tomcat for IAM application

[1] : Use a dedicated tomcat for IAM

[2]

Use TAC tomcat for IAM

[3] : Use an external tomcat for IAM

Please choose an option [1] : 2

Use TAC tomcat for IAM

----------------------------------------------------------------------------

TAC authentication

TAC URL: [http://ip-10-3-100-104.ec2.internal:8080/org.talend.administrator]: http://ip-10-3-100-104:8080/org.talend.administrator

TAC user: [admin@capgemini.com]: admin@capgemini.com

TAC password: [\*\*\*\*\*\*\*\*] :

----------------------------------------------------------------------------

IAM internal configuration

IAM host name: [ip-10-3-100-104.ec2.internal]: ip-10-3-100-104.ec2.internal

----------------------------------------------------------------------------

Talend Identity and Access Management

----------------------------------------------------------------------------

Talend Data Stewardship

Post-logout redirection URL:

[http://ip-10-3-100-104.ec2.internal:19999/]: http://ip-10-3-100-104.ec2.internal:19999/

----------------------------------------------------------------------------

----------------------------------------------------------------------------

Talend Data Preparation

Post-logout redirection URL:

[http://ip-10-3-100-104.ec2.internal:9999/]: http://ip-10-3-100-104.ec2.internal:9999/

----------------------------------------------------------------------------

Talend Data Stewardship Configuration

TDS installs on Apache Tomcat server. Please choose one of option below:

[1] : Embedded Tomcat 8 server (recommended)

[2] : Use an existing Tomcat server

Please choose an option [1] : 1

----------------------------------------------------------------------------

Embedded Tomcat 8 server (recommended)

UI port (Tomcat port): [19999]: 19999

----------------------------------------------------------------------------

----------------------------------------------------------------------------

Talend Data Stewardship Configuration

Choose MongoDB installation

[1] : Embedded MongoDB

[2] : External MongoDB

Please choose an option [1] : 1

----------------------------------------------------------------------------

Embedded MongoDB

MongoDB Port: [27017]: 27017

----------------------------------------------------------------------------

Embedded MongoDB

----------------------------------------------------------------------------

Talend Data Stewardship Configuration

Choose Kafka installation:

[1] : Embedded Kafka and Zookeeper

[2] : External Kafka and Zookeeper

Please choose an option [1] : 1

----------------------------------------------------------------------------

Embedded Kafka and Zookeeper

Kafka port: [9092]: 9092

Zookeeper port: [2181]: 2181

----------------------------------------------------------------------------

----------------------------------------------------------------------------

Talend Data Stewardship Configuration

Talend IAM parameters:

IAM URL: [http://ip-10-3-100-104.ec2.internal:8080]: http://ip-10-3-100-104.ec2.internal:8080

----------------------------------------------------------------------------

Talend CommandLine Configuration

This page allows you to select a port to be used by Talend CommandLine

interface. This port is used to communicate with Talend Administration Center.

CommandLine port [8002]: 8002

----------------------------------------------------------------------------

Talend Runtime parameters

Talend Runtime configuration

Select Runtime port configuration

[1] Port Config 0 (default)

[2] Port Config 1

[3] Port Config 2

Please choose an option [1] : 1

The default configuration (Port Config 0) uses the following ports:

- HTTP port: 8040

- HTTPS port: 9001

- RMI Registry port: 1099

- RMI Server port: 44444

- SSH port: 8101

- File Server Port: 8001

- Monitoring Port: 8888

Port Config 1 uses the following ports:

- HTTP port: 8041

- HTTPS port: 9002

- RMI Registry port: 1100

- RMI Server port: 44445

- SSH port: 8102

- File Server Port: 8010

- Monitoring Port: 8898

Port Config 2 uses the following ports:

- HTTP port: 8042

- HTTPS port: 9003

Press [Enter] to continue :

- RMI Registry port: 1101

- RMI Server port: 44446

- SSH port: 8103

- File Server Port: 8020

- Monitoring Port: 8908

Press [Enter] to continue :

----------------------------------------------------------------------------

Talend Job Server Configuration

This page allows you to configure Talend Remote Job Server. Please configure

ports used to communicate with the Talend Administration Center and other

setttings.

Command port [8003]: 8003

File transfer port [8004]: 8004

Monitoring port [8889]: 8889

This option enables cleaning of any cached files and logs after a given

duration. Set this value to 0 to disable log purging

Max cache duration (days) [90]: 90

----------------------------------------------------------------------------

Talend Data Quality Portal Configuration

Talend Data Quality Portal is installed on top of a Tomcat application server. Please choose one of the following options:

[1] Install with Talend Administration Center Server (Tomcat): Install with the Talend Administration Center Tomcat Server

[2] Install an embedded (standalone) Tomcat 8 server: Install an embedded Tomcat 8 server in the installation directory.

Please choose an option [2] : 1

Talend Data Quality Portal needs to connect to an existing database. Please choose one of the following options:

[1] Connect to a MySQL database

[2] Connect to an Oracle database with SID

[3] Connect to an Oracle database with service name

[4] Oracle OCI

[5] Connect to a Microsoft SQL database

[6] Connect to a PostgreSQL database

Please choose an option [1] : 1

TDQ Portal Port: (valid only with embedded tomcat install)

[8580]: 8580

----------------------------------------------------------------------------

Data Quality Portal - Database

Database Configuration:

Talend Data Quality Studio connects to a database to store the DQ reports and

analysis. Please create a database and configure the connection here:

DB Version:

[1] MYSQL 5

Please choose an option [1] : 1

Database host name (IP): [localhost]: dev-emr-hue-metastore.cluster-cbduemg5i5w8.us-east-1.rds.amazonaws.com

Database port: [3306]: 3306

Database name: [talend\_dq]: dqp1

User name: [root]: Master

Password :

----------------------------------------------------------------------------

Talend Data Quality Portal Configuration

Server Address:

[10.3.100.104]: 10.3.100.104

----------------------------------------------------------------------------

Talend Data Preparation Configuration

Enable Data Preparation on Big data [y/N]: y

----------------------------------------------------------------------------

Big data preparation port: [9060]: 9060

Spark job server port: [8090]: 8-

Hadoop conf dir: []:

Kerberos cluster [y/N]: y

----------------------------------------------------------------------------

Kerberos configuration file: [/etc/krb5.conf]:

Keytab path: []:

Principal: []:

Warning: Spark job server port value is wrong - only numbers are accepted.

Press [Enter] to continue :

Warning: Error: Hadoop configuration dir field is empty !

Please specify the path to Hadoop configuration folder.

Press [Enter] to continue :

Warning: Error: Keytab path field cannot be empty !

Press [Enter] to continue :

Warning: Error: Principal field cannot be empty !

Press [Enter] to continue :

----------------------------------------------------------------------------

Talend Data Preparation Configuration

Enable Data Preparation on Big data [Y/n]: n

----------------------------------------------------------------------------

Talend Data Preparation Configuration

Choose MongoDB installation

[1] : Embedded MongoDB

[2] : External MongoDB

Please choose an option [1] : 1

----------------------------------------------------------------------------

Embedded MongoDB

MongoDB port: 27017

----------------------------------------------------------------------------

Embedded MongoDB

----------------------------------------------------------------------------

Data Preparation Server configuration

Talend Administration Center parameters:

TAC URL [http://ip-10-3-100-104.ec2.internal:8080/org.talend.administrator/]: http://ip-10-3-100-104:8080/org.talend.administrator/

TAC user: [admin@capgemini.com]: admin@capgemini.com

TAC password: [\*\*\*\*\*\*\*\*] :

Data Preparation Server parameters:

Server Public Hostname: [ip-10-3-100-104.ec2.internal]: ip-10-3-100-104.ec2.internal

UI port: [9999]: 9999

Backend port: [8989]: 8989

Talend Identity Access Management server:

IAM URL: [http://ip-10-3-100-104.ec2.internal:8080]: http://ip-10-3-100-104.ec2.internal:8080

----------------------------------------------------------------------------

Talend Dictionary Service Configuration

----------------------------------------------------------------------------

Tomcat configuration parameters

Tomcat port: [8187]: 8187

----------------------------------------------------------------------------

----------------------------------------------------------------------------

Talend Dictionary Service Configuration

Choose MongoDB installation

[1] : Embedded MongoDB

[2] : External MongoDB

Please choose an option [1] : 1

----------------------------------------------------------------------------

Embedded MongoDB

MongoDB port: 27017

----------------------------------------------------------------------------

Embedded MongoDB

----------------------------------------------------------------------------

Talend Dictionary Service Configuration

Talend Identity Access Management server parameters

IAM URL: [http://ip-10-3-100-104.ec2.internal:8080]: http://ip-10-3-100-104.ec2.internal:8080

----------------------------------------------------------------------------

Talend Kafka and Zookeeper configuration

Apache and Zookeeper parameters

Zookeeper data dir: [/data/talend/Talend-6.4.1/kafka/temp]: /data/talend/Talend-6.4.1/kafka/temp

----------------------------------------------------------------------------

Services Installation

This page allows you to setup Talend components as system services. By

installing a component as a service the application would start automatically at

system startup.

It's recommended that you check all the boxes below.

Install Apache Kafka and Apache Zookeeper as services [Y/n]: Y

Install MongoDB as a service [Y/n]: Y

Install Talend Administration Center as a service [Y/n]: Y

Install Talend Log Server as a service [Y/n]: Y

Install Talend Command Line as a service [Y/n]: Y

Install Talend Job Server as a service [Y/n]: Y

Install Talend Data Quality Portal as a service [Y/n]: Y

Install Talend Runtime as a service [Y/n]: Y

Install Talend Data Stewardship as a service [Y/n]: Y

Install Talend Data Preparation as a service [Y/n]: Y

Install Talend Dictionary Service as a service [Y/n]: Y

Install Talend IAM server as a service [Y/n]: Y

----------------------------------------------------------------------------

Setup is now ready to begin installing Talend on your computer.

Do you want to continue? [Y/n]: Y

----------------------------------------------------------------------------

Please wait while Setup installs Talend modules on your computer.

Installing

0% \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 50% \_\_\_\_\_\_\_\_\_\_\_\_\_\_ 100%

########about to fork child process, waiting until server is ready for connections.

forked process: 27432

child process started successfully, parent exiting

#################################

----------------------------------------------------------------------------

Setup has finished installing Talend on your computer.

[root@ip-10-3-100-104 install]#

Check the TAC url: <http://ip-10-3-100-104:8080/org.talend.administrator/>

Username [admin@capgemini.com](mailto:admin@capgemini.com)

Password:Talend123

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Vrishali Mahakalkar  Senior Consultant  Capgemini  Phone:91 9870251815  [vrishali.mahakalkar@capgemini.com](mailto:vrishali.mahakalkar@capgemini.com) |  |  |