



# Talend Open Studio for Big Data

## Release Notes

# 6.3.2

Adapted for v6.3.2. Supersedes previous releases.

Publication date September 7, 2017

## Copyright

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

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

## Notices

Talend is a trademark of Talend, Inc.

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

## License Agreement

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

This product includes software developed at AOP Alliance (Java/J2EE AOP standards), ASM, Amazon, AntLR, Apache ActiveMQ, Apache Ant, Apache Avro, Apache Axiom, Apache Axis, Apache Axis 2, Apache Batik, Apache CXF, Apache Cassandra, Apache Chemistry, Apache Common Http Client, Apache Common Http Core, Apache Commons, Apache Commons Bcel, Apache Commons JXPath, Apache Commons Lang, Apache DataFu, Apache Derby Database Engine and Embedded JDBC Driver, Apache Geronimo, Apache HCatalog, Apache Hadoop, Apache Hbase, Apache Hive, Apache HttpClient, Apache HttpComponents Client, Apache JAMES, Apache Log4j, Apache Lucene Core, Apache Neethi, Apache Oozie, Apache POI, Apache Parquet, Apache Pig, Apache PiggyBank, Apache ServiceMix, Apache Sqoop, Apache Thrift, Apache Tomcat, Apache Velocity, Apache WSS4J, Apache WebServices Common Utilities, Apache Xml-RPC, Apache Zookeeper, Box Java SDK (V2), CSV Tools, Cloudera HTrace, ConcurrentLinkedHashMap for Java, Couchbase Client, DataNucleus, DataStax Java Driver for Apache Cassandra, Ehcache, Ezmorph, Ganymed SSH-2 for Java, Google APIs Client Library for Java, Google Gson, Groovy, Guava: Google Core Libraries for Java, H2 Embedded Database and JDBC Driver, Hector: A high level Java client for Apache Cassandra, Hibernate BeanValidation API, Hibernate Validator, HighScale Lib, HsqlDB, Ini4j, JClouds, JDO-API, JLine, JSON, JSR 305: Annotations for Software Defect Detection in Java, JUnit, Jackson Java JSON-processor, Java API for RESTful Services, Java Agent for Memory Measurements, Jaxb, Jaxen, JetS3T, Jettison, Jetty, Joda-Time, Json Simple, LZ4: Extremely Fast Compression algorithm, LightCouch, MetaStuff, Metrics API, Metrics Reporter Config, Microsoft Azure SDK for Java, Mondrian, MongoDB Java Driver, Netty, Ning Compression codec for LZ4 encoding, OpenSAML, Paracel JDBC Driver, Parboiled, PostgreSQL JDBC Driver, Protocol Buffers - Google's data interchange format, Resty: A simple HTTP REST client for Java, Riak Client, Rocoto, SDSU Java Library, SL4J: Simple Logging Facade for Java, SQLite JDBC Driver, Scala Lang, Simple API for CSS, Snappy for Java a fast compressor/decompressor, SpyMemCached, SshJ, StAX API, StAXON - JSON via StAX, Super SCV, The Castor Project, The Legion of the Bouncy Castle, Twitter4J, Uuid, W3C, Windows Azure Storage libraries for Java, Woden, Woodstox: High-performance XML processor, Xalan-J, Xerces2, XmlBeans, XmlSchema Core, Xmlsec - Apache Santuario, YAML parser and emitter for Java, Zip4J, atinject, dropbox-sdk-java: Java library for the Dropbox Core API, google-guice. Licensed under their respective license.

---

## Table of Contents

System Requirements .....	1
Big Data: New Features .....	2
1. Upgraded support for Hadoop .....	2
2. MapR-DB and MapR Streams support .....	2
3. Enhanced Kafka support .....	2
4. Support of HDFS Encryption .....	2
5. AWS instance role and assume role .....	3
6. Automatic data type conversion .....	3
7. Data type mapping configurations .....	3
8. S3 multipart upload .....	3
9. Microsoft CRM .....	3
10. Other new components .....	3
11. Other version upgrades .....	4
Big Data: Deprecated Items .....	5
Big Data: Bug Fixes / Change Log .....	6
1. Bug Fixes .....	6
Big Data: Known Issues .....	7
Documentation .....	8
1. Talend Help Center .....	8
2. Videos .....	8
3. New documents .....	8
4. Revised documents .....	8
5. Open issues .....	8

# System Requirements

Users should refer to the *Talend Installation and Upgrade Guide* on the **Talend Help Center** (<http://help.talend.com>) for more information on Installation and System Requirements.

# Big Data: New Features

## 1. Upgraded support for Hadoop

1. New versions of the following Hadoop distributions are supported:

- Cloudera CDH V5.8.0
- Hortonworks Data Platform V2.5.0
- MapR V5.2.0
- EMR V5.0.0

## 2. MapR-DB and MapR Streams support

The Studio now includes MapR-DB and MapR Streams support via the following new components:

### MapR-DB

- tMapRDBConnection
- tMapRDBInput
- tMapRDBOutput
- tMapRDBCclose

### MapR Streams

- tMapRStreamsConnection
- tMapRStreamsCommit
- tMapRStreamsCreateStream
- tMapRStreamsInput
- tMapRStreamsOutput

## 3. Enhanced Kafka support

The support of Kafka 0.10 has been added with security options.

## 4. Support of HDFS Encryption

The HDFS transparent encryption is now supported.

---

## 5. AWS instance role and assume role

The instance role feature, which allows you to load credentials from the Amazon EC2 Instance Metadata Service, and the assume role feature, which allows you to use the AWS Security Token Service to assume a Role and create temporary and short-lived sessions to use for authentication, are now supported in the following components:

- tDynamoDBInput and tDynamoDBOutput
- tAmazonEMRListInstances, tAmazonEMRManage, and tAmazonEMRResize
- tAmazonRedshiftManage
- tS3BucketCreate, tS3BucketDelete, tS3BucketExist, tS3BucketList, tS3Connection, tS3Copy, tS3Delete, tS3Get, tS3List, and tS3Put

## 6. Automatic data type conversion

In standard data integration Jobs, tMap now supports automatic data type conversion at Job execution to avoid compiling errors caused by type mismatch between input and output data flows.

While it can be enabled in tMap in individual Jobs, this feature can also be activated globally in Project Settings so that any tMap component added afterwards will have this feature enabled. The default conversion behavior of tMap can be overridden by adding conversion rules in Project Settings.

## 7. Data type mapping configurations

The configurations of mappings between Talend data types and database data types can now be managed at the project level.

## 8. S3 multipart upload

The tS3Put component now supports the multipart upload feature, which allows you to upload a single object as a set of parts.

## 9. Microsoft CRM

The tMicrosoftCrmOutput component can now write the empty lookup string into Microsoft CRM as NULL value instead of empty string.

## 10. Other new components

### 10.1. Amazon SQS support

Amazon SQS (Simple Queue Service) is now supported via the following seven new components:

- tSQSConnection
- tSQSInput
- tSQSOutput
- tSQSQueueCreate
- tSQSQueueDelete
- tSQSQueueList
- tSQSQueuePurge

## 10.2. FTP connection

The new component tFTPClose, which allows you to close an active connection to an FTP server, has been created.

# 11. Other version upgrades

1. The Dropbox API has been upgraded from v1.7.5 to v2.1.2.
2. Microsoft Azure and Microsoft SQL Server 2014 and 2016 are now supported.
3. The Salesforce v37 is now supported.
4. ActiveMQ 5.14.1 is now supported.

# Big Data: Deprecated Items

The following feature is deprecated in the current release. Deprecated features are no longer recommended for use and may cease to exist in future versions of the product.

- The tAlfrescoOutput component
- The HL7 metadata wizard and the tHL7Input and tHL7Output components



# Big Data: Bug Fixes / Change Log

## 1. Bug Fixes

In addition to the above new features a number of minor improvements within the entire product and significant bug fixes have been made.

See the corresponding Change Log on our bug tracking system for more details on the individual issues:

- <https://jira.talendforge.org/secure/ReleaseNote.jspa?projectId=11430&version=22044>.

# Big Data: Known Issues

We encourage you to consult the JIRA bug tracking tool for a full list of open issues:

<https://jira.talendforge.org/issues/?filter=26283>.

Note that this list shows issues from both Talend's Community and Subscription products.

---

# Documentation

## 1. Talend Help Center

Find out more about how to get the most out of your Talend products on the Talend Help Center: <http://help.talend.com>.

## 2. Videos

From within the Studio, you can directly access the videos hosted in Talend Help Center relating to new features.

- For components, related videos are available from the F1 help view of this component.
- For other features, related videos are available in the help view of the wizard for the feature.

Note that not every feature has related videos in Talend Help Center and for this reason, not all the features have this type of access.

## 3. New documents

The following new documents have been added in this release:

- *Talend Data Stewardship User Guide* and *Talend Data Stewardship Getting Started Guide* are two new guides that accompany the new product, *Talend Data Stewardship*.

They are both available in PDF format and online at [https://help.talend.com/search/books?filters=EnrichPlatform~%2522Talend+Data+Stewardship%2522\\*EnrichVersion~%25226.4%2522](https://help.talend.com/search/books?filters=EnrichPlatform~%2522Talend+Data+Stewardship%2522*EnrichVersion~%25226.4%2522).

## 4. Revised documents

In addition to updates to the content across the documentation set, the following specific documentation changes have been made:

- In the *Talend Software Development Life Cycle Best Practices Guide*, major changes have been made to the documentation of the Talend CI Builder, including how to generate sources locally, how to manage POM files, artefact versions, and how to filter executions on specific items.

## 5. Open issues

We encourage you to consult the JIRA bug tracking tool for a full list of open issues:

<https://jira.talendforge.org/issues/?filter=18375>