



Talend Open Studio for Big Data Release Notes

6.5.1

Contents

Copyright.....	3
System requirements.....	4
Talend Open Studio for Big Data: new features.....	4
Talend Open Studio for Big Data: bug fixes.....	7
Talend Open Studio for Big Data: deprecated items.....	7
Talend Open Studio for Big Data: known issues and limitations.....	7
Talend Open Studio for Big Data: hints and notes.....	8
Documentation.....	9
Reporting defects and issues.....	10

Copyright

Adapted for 6.5.1. Supersedes previous releases.

Publication date: January 18, 2018

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, ConcurrentHashMap 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.

System requirements

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

Talend Open Studio for Big Data: new features

Big Data: new features

Enhancements of Hadoop support

Feature	Description
Upgraded support for Hadoop distributions	<ul style="list-style-type: none"> Cloudera CDH V5.12 MapR 6.0
Hive application ID	The Hive components now capture the Application_ID values and write them in the Job logs.
MapR OJAI	A new component, tMapROjaiOutput, has been added to write data to a MapR Ojai database.
Hbase	Users can now read and write custom timestamps columns using the HBase components.

New NoSQL components

Feature	Description
Neo4j	<ul style="list-style-type: none"> Neo4j Batch components have been created. Neo4j V3.2 along with the Bolt protocol is now supported.

Component enhancements

Feature	Description
Upgraded support for Couchbase	New Couchbase components have been created to replace the old ones to support Couchbase V4.X and V5.X.
Enhanced Marklogic support	The Marklogic components now support Marklogic V9.

Data Integration: new features

Feature	Description
Db Connections metadata wizard enhancement	The Db Connections metadata wizard now supports specifying additional parameters and retrieving calculation view schemas for the SAP HANA database.
Enhanced Amazon S3 support	The tS3Connection component now supports connecting to Amazon S3 using a custom region endpoint.
Enhanced ELT support	<ul style="list-style-type: none"> The tELT***Map components now allow you to set the output table name and connection name to variables. The tELTJDBC*** components now support aggregating Snowflake data.
Enhanced Generic JDBC support	The generic JDBC components tJDBC*** now support the latest Amazon Redshift JDBC Driver.
Enhanced JSON support	The tExtractJSONFields and tFileInputJSON components now allow you to specify the JSONPath API version.
Enhanced Marketo support	<ul style="list-style-type: none"> The endpoint for Marketo SOAP API has been updated to the latest API version.

Feature	Description
	<ul style="list-style-type: none"> A new column <code>marketoGUID</code> of <code>String</code> type has been added to the schema of <code>tMarketoInput</code> for the <code>getLeadActivity</code> and <code>getLeadChanges</code> operations.
Enhanced Microsoft CRM support	The <code>tMicrosoftCrm***</code> components now support reading and writing data in Microsoft CRM 2016 on-premise database.
Enhanced NetSuite support	The <code>tNetsuiteInput</code> component now supports retrieving both body fields and sublist values of Netsuite records.
Enhanced SCD support	<ul style="list-style-type: none"> The <code>tJDBCSCDEL</code> component now supports tracking data changes in Vertica 7.1 and 9.0 tables. The <code>tJDBCSCDEL</code> component now supports tracking Snowflake data changes using a sequence to generate the surrogate key.
Enhanced Snowflake support	The <code>tSnowflake***</code> components now support handling data of <code>VARIANT</code> type.
Enhanced Vertica support	<ul style="list-style-type: none"> The <code>tVertica***</code> components now support Vertica 9.0. The <code>tVerticaBulkExec</code> and <code>tVerticaOutputBulkExec</code> components now support specifying more <code>COPY</code> parameters when loading data into Vertica 6.0 and later. The <code>tVerticaBulkExec</code> component can now work as an output component and you can connect a <code>tFileInputRaw</code> component to it to load data into Vertica.

Feature	Description
Other component enhancements	The tCreateTable component now supports creating Exasol, Redshift, and Vertica 7.1 database tables.

Talend Open Studio for Big Data: bug fixes

In addition to the new features a number of minor improvements within the entire products and significant bug fixes have been made. See the corresponding Change Logs on our bug tracking system for more details on the individual issues:

Module	Change Log
Big Data	https://jira.talendforge.org/issues/?filter=26980

Talend Open Studio for Big Data: deprecated items

Big Data: deprecated items

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

Deprecated item	Description
From Microsoft HD Insight 3.1 to Microsoft HD Insight 3.2	The support in the Studio to connect to these HD Insight distributions.

Data Integration: deprecated items

Feature	Description
Components	Vertica 7.1 and all previous versions

Talend Open Studio for Big Data: known issues and limitations

Big Data: known issues and known limitations

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

<https://jira.talendforge.org/issues/?filter=26981>

Limitation	Description
MapR V6.0 Limitation	MapR has removed the support of HBase from MapR V6.0 in favor of MapR-DB. For more details, see the related MapR documentation .
tMapRStreamOutput Limitation	The tMapRStreamOutput component encounters compile error when running on MapR V6.0

Data Integration: known issues and known limitations

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

<https://jira.talendforge.org/issues/?filter=26963>

Talend Open Studio for Big Data: hints and notes

Big Data: hints and notes

Item	Description
JDK Required to Build Jobs/Routes	To build Jobs or Routes, you need to install a full JDK rather than a JRE. If the system fails to initialize automatically with a JDK, you can define which JDK to use by selecting Window > Preferences > Java > Installed JREs .
HDFS components and Azure Data Lake Store (ADLS)	Although Talend does not provide reading or writing components specific to Azure Data Lake Store, you can use the existing HDFS components to easily handle this situation. For more details, you can search for information about using HDFS components to work with Azure Data Lake Store (ADLS) on Talend Help Center .

Documentation

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>.

New documents

Talend Job Script Reference Guide is a new guide that provides general information on designing Jobs using Job scripts, property settings of frequently used components in Job scripts, and scenarios of Job designs using Job scripts.

The Getting Started Guide for the subscription version of Talend Data Preparation has been translated into French.

Revised documents

The online version of the components documentation has been reorganized into domain-specific landing pages around an extensive master page.

The Talend Data Preparation documentation, which was previously published in the form of standalone articles, has been reorganized into three guides:

- Getting Started with Talend Data Preparation
- Talend Data Preparation User Guide
- Talend Data Preparation Quick Examples.

The Talend Data Stewardship documentation, which was previously published in the form of standalone articles, has been reorganized into three guides:

- Getting Started with Talend Data Stewardship
- Talend Data Stewardship Use Cases
- Talend Data Stewardship User Guide

Use cases for Big Data, Data Governance, Data Integration, Data Profiling and Cleansing, and Data Services and ESB Routing are now published separately from the Studio User Guide, as standalone guides.

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>

Reporting defects and issues

As a customer of Talend with a valid support contract, you can use our Talend Online Helpdesk or you can contact our customer support by phone and e-mail as stated in your contract.

Often the following information is needed and would help us to more easily evaluate your issue:

- Product and Version which you used (for example, Talend Data Fabric 6.5.1)
- Operating System, for example Windows Server 2008 R2 SP1 (64bit)
- Version of Java Platform JDK / JRE in use (for example, ORACLE JDK 1.8.0_60)
- Log files and/or screenshots

Also, in order to ensure that the technical support team has sufficient information to help you, describe at least the following in detail:

- your actions up to the point when the problem occurred
- the results you expected

In addition to our support helpdesk, you can also use our community support tools:

- [Talend Community](#).
- You can also look into the [Talend Bug Tracker](#).
- Just keep in mind that only cases you create via one of the Talend Support channels in our Talend Support Helpdesk are treated under the service level agreements we provided in the related contract to you.