



Talend Open Studio for Big Data Release Notes

7.0.1

Contents

Copyright	3
System requirements	5
Talend Open Studio for Big Data: new features	6
Big Data: new features.....	6
Data Integration: new features.....	6
Talend Open Studio for Big Data: bug fixes	9
Talend Open Studio for Big Data: deprecated and removed items	10
Big Data: deprecated and removed items.....	10
Data Integration: deprecated and removed items.....	10
Talend Open Studio for Big Data: known issues and limitations	11
Big Data: known issues and known limitations.....	11
Data Integration: known issues and known limitations.....	11
Talend Open Studio for Big Data: hints and notes	12
Big Data: hints and notes.....	12
Data Integration: hints and notes.....	12
Documentation	13
Reporting defects and issues	14

Copyleft

Adapted for 7.0.1. Supersedes previous releases.

Publication date: April 13, 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, 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.

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 V5.13 EMR 5.8
Dynamic Hadoop distributions (technical preview)	<p>In the Studio, if there is no support for a Cloudera distribution to be used, users can add this distribution by themselves via a wizard to make this distribution compatible with the Studio.</p> <p>The Cloudera versions users can add vary between V5.11.0 (included) and 6.0.0 (excluded)</p> <p>The dynamic distribution added this way is not officially supported by Talend.</p>

Data Integration: new features

Feature	Description
New DMG for Talend installation on MacOS	A new DMG is provided for installation of Talend Open Studio products on MacOS to prevent macOS Sierra from setting downloaded files to "quarantine".
Enhanced Amazon EMR support	<ul style="list-style-type: none"> The tAmazonEMRManage component now allows you to specify the identifier of the Amazon EC2 security group and customize the cluster version and application information. EMR 4.9.2 and EMR 5.11.0 are now supported in the tAmazonEMRManage component.
Enhanced Amazon S3 support	<ul style="list-style-type: none"> The tS3Copy and tS3Put components now support server-side encryption with AWS KMS-Managed Keys (SSE-KMS). The tS3Put component now supports uploading data from an input stream.
Enhanced Amazon SQS support	<ul style="list-style-type: none"> A new component tSQSQueueAttributes, which allows you to get attributes of an Amazon SQS queue, has been added into Talend Studio. The tSQSOutput component now supports delivering messages to a FIFO (first-in-first-out) queue.
Enhanced DB2 support	DB2 v11 is now supported.
Enhanced FTP support	<ul style="list-style-type: none"> All FTP components now allow you to connect to an FTP server via an FTPS connection.

Feature	Description
	<ul style="list-style-type: none"> The tFTPGet and tFTPput components now support specifying the data channel protection level and the maximum size of the encoded data blocks when transferring data via an FTPS connection. The tFTPFileExists component now supports checking the existence of either a file or a folder.
Enhanced Generic JDBC support	Eight JDBC components (tJDBCClose, tJDBCCommit, tJDBCConnection, tJDBCInput, tJDBCOutput, tJDBCRollback, tJDBCRow, and tJDBCSP) and the Db Connections metadata wizard for creating and editing JDBC connections have been switched into the new framework and the user interface has been slightly changed.
Enhanced Google Drive support	<ul style="list-style-type: none"> The Google Drive components and metadata wizard have been switched into the new framework and the user interface has been slightly changed. The Google Drive components and metadata wizard support three new authentication methods, and you can now connect to Google Drive using your client ID and client secret, client secret JSON file, or service account JSON file. The tGoogleDriveList component now supports listing removed files and folders.
Enhanced JSON support	The tWriteJSONField component now supports surrounding all values including number and boolean values in the generated JSON field with double quotation marks.
Enhanced Microsoft Azure SQL Data Warehouse support	A new component tSQLDWHBulkExec, which allows you to load data into an Azure SQL Data Warehouse table from either Azure Blob Storage or Azure Data Lake Store, has been added into Talend Studio.
Enhanced Microsoft CRM support	The tMicrosoftCrmInput and tMicrosoftCrmOutput components now support connecting to Microsoft Dynamics CRM Online Web application.
Enhanced Microsoft SQL Server support	The tMSSqlBulkExec component now supports executing any insert trigger during your bulk insert operation.
Enhanced Oracle support	The tOracleConnection component and the Db Connections metadata wizard now support connecting to Oracle using the Wallet.
Enhanced positional file support	The tFileInputPositional component now allows you to specify the unit of the fixed length value for each field.
Enhanced Salesforce support	<ul style="list-style-type: none"> The Salesforce components and metadata wizard now allow you to connect to Salesforce using OAuth 2.0 authentication JSON Web Token Flow.

Feature	Description
	<ul style="list-style-type: none"> • The tSalesforceInput component now supports specifying the timeout value for your bulk query Job to complete. • The tSalesforceInput component now supports bulk querying any Salesforce field whose length is greater than 100,000 characters. • Salesforce v42 is now supported.
Enhanced Snowflake support	<ul style="list-style-type: none"> • A new component tSnowflakeRow, which allows you to execute any SQL command stated on a Snowflake database, has been added into Talend Studio. • The tSnowflakeInput and tSnowflakeOutput components now support converting lowercases in the defined table name and schema column names to uppercases.
Enhanced Sybase support	<ul style="list-style-type: none"> • The tSybaseIQBulkExec and tSybaseIQOutputBulkExec components now support loading data from the client side. • Sybase SQL Anywhere 16 is now supported.
New unified database components	Several unified database components tDB*** have been added into Talend Studio as entry points to a variety of databases.

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=27287

Talend Open Studio for Big Data: deprecated and removed items

Big Data: deprecated and removed items

Removed item	Description
The obsolete Hadoop distributions	<p>The support for the Hadoop distributions that are older or equal to the following versions has been removed:</p> <ul style="list-style-type: none"> • Amazon EMR V4.0.0 • Cloudera V5.4 • MapR V4.1.0 • Hortonworks V2.3.2 • Microsoft HDInsight V3.2 • All Pivotal versions <p>These versions were deprecated in the previous releases of the Studio and are removed in the current release.</p>

Data Integration: deprecated and removed items

Feature	Description
SAP	The tSAPBWInput component has been deprecated and hidden from the Palette.
Alfresco	The tAlfrescoOutput component has been removed.
Build Job types	The Axis WebService (WAR) and Axis WebService (ZIP) build job types have been removed.
Database components	Three database components tDBInput, tDBOutput, and tDBSQLRow have been removed.
Component Designer perspective	The Component Designer perspective has been removed.
HL7	The tHL7Input and tHL7Output components and the HL7 metadata wizard have been removed. We recommend you use Talend Data Mapper.

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=27288>

Limitation	Description
Latest Java version supported on Mac	The latest version of Java supported on Mac is 1.8.0_151 and later versions will not work correctly. See https://jira.talendforge.org/browse/TUP-19313 for details.

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=27297>

Limitation	Description
Latest Java version supported on Mac	The latest version of Java supported on Mac is 1.8.0_151 and later versions will not work correctly. See https://jira.talendforge.org/browse/TUP-19313 for details.

Talend Open Studio for Big Data: hints and notes

Big Data: hints and notes

Item	Description
MapR 6.0 connection on Windows	<p>You need to add the following properties to the Studio startup .ini file to ensure successful connection to MapR 6.0:</p> <ul style="list-style-type: none"> • -Djava.security.auth.login.config=/opt/mapr/conf/mapr.login.conf • -Dhadoop.login=hybrid • -Djava.library.path=/opt/mapr/lib <p>Related issue: https://jira.talendforge.org/browse/TBD-6411.</p>

Data Integration: hints and notes

Item	Description
Use of recursive Jobs	<p>Avoid using recursive Jobs because this forces recursive Jobs with the same dependencies build. In a normal case each Job has its own Maven project, but with the use of recursive Jobs, all the Jobs linked might be compiled inside the same project. This can easily cause some compilation issues between the Jobs.</p>

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>.

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 7.0.1)
- 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_161)
- 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.