talend*
*open integration solutions

Talend Open Studio for Big Data
Release Notes

5.2.1

Copyleft

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

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

Table of Contents

Introduction

Talend Open Studio for Big Data is an Apache based product that offers Big Data capabilities for the Masses, simplifies Big Data technologies, and thus makes it easier for users to leverage Big Data clusters, through an intuitive tool, making them more productive.

System Requirements

Talend Open Studio for Big Data users should refer to the *http://www.talend.com/docs/community/prerequisites.html* and *http://talendforge.org/wiki/doku.php?id=doc:installation_guide* for more information on Installation and System Requirements.

Big Data: New Features

1. Wizards

- A wizard has been added to help install third-party libraries.

- The HDFS browser enables HDFS file navigation locally in the Studio.

2. Hadoop releases

- With the addition of support for Cloudera CDH4 and Mapr V2, users can now connect to these Hadoop distributions from Big Data components.

3. Components

- The Google BigQuery Components have been created.

- New components (MongoDB and Cassandra) are now available to perform transactions with these NoSQL databases.

- ELT components for Hive have been introduced to enable faster throughput for data transmissions.

- Additional HDFS components have been added to perform operations directly in HDFS, such as listing, copying or renaming files. These components are tHDFSList, tHDFSCopy, tHDFSExist, tHDFSRename, tHDFSProperties, tHDFSCompare and tHDFSRowCount.

- Sqoop components support "free form query".

- Kerberos security has been enabled for the components for Hortonworks HCatalog and as well for CDH4 in the Pig and the HDFS components.

4. Other features

- Enhancements of the Salesforce components:

  - The Salesforce.com connector has been upgraded to version 25.

  - The Salesforce components now support the Bulk Query and the Delete actions

  - Support for external IDs has been improved.

  - The "Ignore NULL fields values" check box has been added to tSalesforceOutput.

  - In Salesforce bulk components, the "Concurrency mode" drop-down list is now available to enable the selection between the Parallel and the Serial modes for data transmissions.

- Vertica 6.0 is now supported.

- The Exasol JDBC Driver v4.02 has been integrated.

- The MOM components have been enhanced to support local JMS transactions for ActiveMQ, and the MOM and the JMS components to support ActiveMQ failover.

- ELT mode has been enabled for the Greenplum and the Netezza components.

- Some databases components have been improved to support more functions:

  - The tSybaseSP component for FUNCTION stored procedures.

  - The tMSSqlOutput component for MS SQL 2008 Date types.

  - tPostGreSql for SAVE_POINT with Transaction.

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 done.

Please 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=10237&version=12958*

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/secure/IssueNavigator.jspa?requestId=13953

Documentation

1. New documents

The following new documentation is available for this release.

- *Talend Big Data Studio User Guide*: this guide addresses only big data related features and capabilities offered in the big data studios.

2. Revised documents

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

- *Talend ESB Getting Started User Guide* has been renamed *Talend ESB Getting Started Guide*, and *Talend ESB Service Factory User Guide* has been renamed *Talend ESB Service Developer Guide*.

- Updates have been made to several *Talend ESB* guides to reflect changes in certain Apache products.

  - *Talend ESB Development Guide*: Apache CXF 2.7

  - *Talend ESB Mediation Developer Guide*: Apache Camel 2.10

  - *Talend ESB Service Developer Guide*: Apache CXF 2.7

  - *Talend ESB STS User Guide*: Apache CXF 2.7 - STS

3. Known issues

The documentation of certain technical fixes that arrived too late to be included in the current release will be dealt with in the next release.

https://jira.talendforge.org/secure/IssueNavigator.jspa?requestId=13952