

StreamSets Transformer 3.12.0 Release Notes

January 6, 2020

We're happy to announce a new version of StreamSets Transformer. This version contains several new features and some bug fixes. We recommend that you try this latest release.

This document contains the following information:

- [New Features and Enhancements](#)
- [Upgrade Information](#)
- [Known Issues](#)
- [Documentation](#)
- [Contact Information](#)

New Features and Enhancements

This release includes the following new features and enhancements:

- **SQL Server 2019 Big Data Cluster support** - You can run pipelines using SQL Server 2019 Big Data Cluster as a cluster manager.
- **Delta Lake additional storage support** - You can now use ADLS Gen2 as a storage system for Delta Lake stages.
- **Enhanced file path validation** - File path validation has been enhanced for HDFS-based stages, such as the Amazon S3, ADLS, Delta Lake, and File origins and destinations.
- **Hive external table creation** - You can configure the Hive destination to create an external table at a specified location.
- **HTTPS self-signed certificates** - You can more easily use self-signed certificates when enabling Transformer to use HTTPS.
- **JDBC stage quote character** - You can specify a quote character for the JDBC origin, JDBC destination, and JDBC Lookup processor to use.
- **Kafka schema registry authentication** - When a Kafka origin or destination uses Avro schemas in Confluent Schema Registry, you can specify basic authentication user information when required.
- **Scala processor empty batch handling** - You can now configure the Scala processor to skip processing for empty batches.
- **Configuration property rename** - In the Transformer configuration file, `transformer.properties`, the `sdc.base.http.url` property has been renamed to `transformer.base.http.url`.

If you configured the `sdc.base.http.url` property in a previous version of Transformer, configure the new `transformer.base.http.url` property to use the same value.

- **Configuration property removal** - In the Transformer configuration file, `transformer.properties`, the `kerberos.client.enabled` property has been removed since it is no longer used.

If upgrading from a previous version of Transformer, do not add this property to the Transformer 3.12.0 configuration file. If using a previous version of the file, remove this property or set it to `false`.

Upgrade Information

To upgrade existing versions of StreamSets Transformer registered with Control Hub, complete the following steps for each registered Transformer.

We recommend upgrading one Transformer at a time and verifying that jobs continue to run as expected.

1. In Control Hub, stop all jobs running on the Transformer to be upgraded.
2. Stop the Transformer.
3. Download Transformer 3.12.0 from the [StreamSets website](#).
4. [Install and configure Transformer 3.12.0](#).
5. Compare the previous and new versions of the `$(TRANSFORMER_CONF)/transformer.properties` file, and update the new file as needed with the same customized property values.
6. [Start the 3.12.0 version of Transformer](#).
7. [Register the 3.12.0 version of Transformer with Control Hub](#).
8. To have Transformer 3.12.0 run the same jobs as the earlier version of Transformer, [assign the same label to the 3.12.0 version of Transformer](#).
9. Restart Transformer jobs.
10. If jobs run as expected, [unregister the old version of Transformer](#).

Known Issues

Please note the following known issues with this release:

Issue	Description
TRANSFORM-1401	You cannot use a PySpark processor in a pipeline that runs on a Databricks cluster at this time.
TRANSFORM-1358	Scala compilation fails for pipelines that run in Hadoop YARN cluster deployments. This affects pipelines with a preprocessing script or Scala processor that run on Hadoop YARN with the Deployment pipeline property set to Cluster. Workaround: Set the Deployment Mode pipeline property to Client.
TRANSFORM-495	The File origin processes files with mixed schemas.

Documentation

You can view the latest Transformer documentation in the Control Hub online help. Transformer details appear in the following chapters: [Transformer](#) and [Pipeline Design for Transformer](#).

Contact Information

For more information about StreamSets, visit our website: <https://streamsets.com/>.

Check out our Documentation page for doc highlights and tutorials: <https://streamsets.com/docs>

To report an issue, to get help from our Google group, Slack channel, or Ask site, or to find out about our next meetup, check out our Community page: <https://streamsets.com/community/>.

For general inquiries, email us at info@streamsets.com.