

StreamSets Data Collector 1.2.0.0 Release Notes

January 27, 2016

Fixed Issues

The following table lists some of known issues from 1.1.4 that are fixed with this release. We fixed over 250 issues with this release. For the full list, click [here](#).

JIRA	Description
SDC-1975	Using the JDBC Consumer to read from Netezza is not supported at this time.
SDC-1794	You cannot use a custom delimited file type when writing delimited data to destinations at this time.

Known Issues

Please note the following known issues with this release.

For a full list of known issues, check out [our JIRA](#).

JIRA	Description
SDC-2358 SDC-2359	When unable to connect to a Kafka 0.9.0.0 broker, a pipeline with a Kafka Consumer or Kafka Producer can hang during validation or display initialization errors. The Data Collector log indicates that the broker is unavailable.
SDC-2340	When processing Avro files, leave the Compression Format set to None in the origin. Typically compression is handled within Avro, and the Avro file does not need to be compressed.
SDC-2226	When the Data Collector UI is configured for HTTPS, pipelines with the Hadoop FS origin fail to start.
SDC-2190	Pipelines with Hadoop FS origins that access kerberized clusters fail to start.
SDC-2133	You cannot stop a pipeline in the middle of a long-running processor, such as a Jython Evaluator performing a complex script.

StreamSets Data Collector 1.2.0.0 Release Notes

JIRA	Description
SDC-2122	Due to the current Java keystore handling, you cannot use HTTPS with Cloudera Manager installations of the Data Collector.
SDC-1731	<p>When using the Kafka Consumer or Kafka Producer on HDP 2.3 with Kerberos enabled, set the Kafka broker configuration property <code>security.inter.broker.protocol</code> to PLAINTEXT.</p> <p>When enabling Kerberos, HDP 2.3 sets the <code>security.inter.broker.protocol</code> property to PLAINTEXTSASL, which is not supported.</p> <p>If the property is not set to PLAINTEXT, when the pipeline starts, validation errors indicate a problem connecting to Kafka.</p>
SDC-1626	<p>A cluster mode pipeline can hang with a <code>CONNECT_ERROR</code> status. This can be a temporary connection problem that resolves, returning the pipeline to the <code>RUNNING</code> status.</p> <p>If the problem is not temporary, you might need to manually edit the pipeline state file to set the pipeline to <code>STOPPED</code>. Edit the file only after you confirm that the pipeline is no longer running on the cluster or that the cluster has been decommissioned.</p> <p>To manually change the pipeline state, edit the following file: <SDCinstalldir>/data/runInfo/ <cluster pipeline name>/<revision>/pipelineState.json</p> <p>In the file, change <code>CONNECT_ERROR</code> to <code>STOPPED</code> and save the file.</p>
SDC-1567	<p>You cannot use cluster mode pipelines to read from HDP 2.3 due to a HDP integration issue with Kafka and Spark Streaming.</p> <p>For more information about the HDP issue, see http://hortonworks.com/community/forums/topic/kafka-and-spark-streaming-nosuchmet hoderror-kafka-consumer-simpleconsumer/.</p>
SDC-891	<p>At this time, writing to error records to file is not supported for cluster mode pipelines.</p> <p>Workaround: Write error records to Kafka or to an RPC pipeline.</p>
SDC-890	<p>For cluster mode pipelines configured to stop on error or to stop upon reaching a memory limit, the Data Collector cannot stop all worker pipelines as expected.</p> <p>Workaround: To stop all pipelines, use the Stop icon in the Data Collector console.</p>

StreamSets Data Collector 1.2.0.0 Release Notes

Contact Information

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

To report an issue, ask for help, or find out about our next meetup, check out our Community page: <http://streamsets.com/community/>.

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