

StreamSets Data Collector 1.1.1 Release Notes

Fixed Issues

JIRA	Description
SDC-1776	<p>When the Directory origin processes Avro data, it applies the Buffer Limit property to the file instead of individual records.</p> <p>Workaround: You can increase the buffer limit up to 1000 KB. Files larger than the buffer limit are not processed. They are handled based on the error handling configured for the stage.</p>

Known Issues

JIRA	Description
SDC-1794	<p>You cannot use a custom delimited file type when writing delimited data to destinations at this time.</p>
SDC-1774	<p>The Data Collector command line interface fails to connect when the Data Collector uses digest authentication.</p> <p>To use the command line interface, when possible, in the Data Collector configuration file, configure the <code>http.authentication</code> property to use a different authentication type.</p>
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. 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:</p> <pre><SDCinstalldir>/data/runInfo/ <cluster pipeline name>/<revision>/pipelineState.json</pre> <p>In the file, change <code>CONNECT_ERROR</code> to <code>STOPPED</code> and save the file.</p>

StreamSets Data Collector 1.1.1 Release Notes

JIRA	Description
SDC-1556	<p>You cannot use cluster mode pipelines to read from HDP 2.3 due to a HDP integration issue with Kafka and Spark Streaming. (1567)</p> <p>For more information about the HDP issue, see http://hortonworks.com/community/forums/topic/kafka-and-spark-streaming-nosuchmethoderror-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>