

Cloudera Runtime 7.1.3

Kafka Connect Connector Reference

Date published: 2020-05-21

Date modified: 2020-08-10

CLOUdera

<https://docs.cloudera.com/>

Legal Notice

© Cloudera Inc. 2024. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

HDFS Sink Connector Properties Reference.....	4
Amazon S3 Sink Connector Properties Reference.....	5

HDFS Sink Connector Properties Reference

HDFS Sink Connector Properties Reference.

The following table collects connector properties that are specific for the HDFS Sink Connector. For properties common to all sink connectors, see the upstream Apache Kafka documentation.

Property Name	Description	Type	Default Value	Accepted Values	Recommended Values
hdfs.uri	The Hadoop file system URI to connect to on the destination HDFS cluster.	String	None		
hdfs.output	The root directory on the HDFS cluster where all the output files will reside. The sub path has the following pattern: {topic}/{topic}_{partition}_{endoffset}. {file extension}	String	/tmp		Any path on the HDFS file system where the role has read write permission.
hdfs.kerberos.authentication	Enables or disables secure access to the HDFS cluster by authenticating with Kerberos.	Boolean	false	true or false	
hdfs.kerberos.user.principal	The kerberos user principal.	String	null	The host-dependent Kerberos principal assigned to the Kafka Connect role.	
hdfs.kerberos.keytab.path	The path to the Kerberos keytab file.	String	null		In a Cloudera Manager provisioned environment, it's recommended to use the Cloudera Manager Config Provider to automatically provision the path.
hdfs.kerberos.namenode.principal	The kerberos name node principal. Required when the HDFS cluster has data encryption on.	String	null		
hadoop.conf.path	The path to the site specific Hadoop configuration XML files. Required when the HDFS cluster has data encryption on.	String	null		

Property Name	Description	Type	Default Value	Accepted Values	Recommended Values
output.writer	The output file writer which determines the type of file to be written to the HDFS cluster. The value of this property should be the FQCN of a class that implements the PartitionWriter interface.	String	null	<ul style="list-style-type: none"> com.cloudera.dim.kafka.connect.partition.writers.avro.AvroPartitionWriter com.cloudera.dim.kafka.connect.partition.writers.json.JsonPartitionWriter com.cloudera.dim.kafka.connect.hdfs.parquet.ParquetPartitionWriter com.cloudera.dim.kafka.connect.partition.writers.txt.TxtPartitionWriter 	
output.avro.passthrough.enabled	Configures whether the output writer expects an Avro encoded Kafka Connect data record. Must match the configuration of value.converter.passthrough.enabled.	Boolean	true	true or false	True if input and output are both Avro.
value.converter	The converter to be used to translate the value field of the source Kafka record into Kafka Connect Data format.	String	Inherited from Kafka Connect worker properties.	<ul style="list-style-type: none"> org.apache.kafka.connect.json.JsonConverter org.apache.kafka.connect.storage.StringConverter com.cloudera.dim.kafka.connect.converters.AvroConverter 	
value.converter.schema.registry.url	The URL to the Schema Registry server.	String	null	true or false	
value.converter.passthrough.enabled	Configures whether the AvroConverter translates an Avro record into Kafka Connect Data or transparently passes the Avro encoded bytes as payload.	Boolean	true	true or false	True if input and output are both Avro.

Amazon S3 Sink Connector Properties Reference

Amazon S3 Sink Connector Properties Reference.

The following table collects connector properties that are specific for the Amazon S3 Sink Connector. For properties common to all sink connectors, see the upstream Apache Kafka documentation.

Property Name	Description	Type	Default Value	Accepted Values	Recommended Value
aws.s3.bucket	The target S3 bucket name.	String	none	Any valid S3 bucket name.	
aws.s3.service_endpoint	The target S3 host and port.	String	none	Any valid S3 endpoint.	
aws.access_key_id	The AWS secret key ID to authenticate.	String	none	Any valid secret key issued by AWS.	
aws.secret_access_key	The AWS secret access key to authenticate.	String	none	Any valid access key issued by AWS.	

Property Name	Description	Type	Default Value	Accepted Values	Recommended Value
value.converter	Value conversion class.	String	none		com.cloudera.dim.kafka.connect.converts
value.converter.passthrough.enabled	Configures whether the AvroConverter translates an Avro record into Kafka Connect Data or transparently passes the Avro encoded bytes as payload.	Boolean	true	true, false	True if input and output are both Avro.
value.converter.schema.registry.url	The URL to the Schema Registry server.	String	none		
output.storage	The S3 storage implementation class.	String	none		com.cloudera.dim.kafka.connect.s3.S3Pa
output.writer	The output file writer which determines the type of file to be written. The value of this property should be the FQCN of a class that implements the PartitionWriter interface.	String	none	<ul style="list-style-type: none"> com.cloudera.dim.kafka.connect.partition.writers.avro.AvroP com.cloudera.dim.kafka.connect.partition.writers.json.JsonPa com.cloudera.dim.kafka.connect.hdfs.parquet.ParquetPartitio com.cloudera.dim.kafka.connect.partition.writers.txt.TxtParti 	com.cloudera.dim.kafka.connect.partition
output.avro.passthrough.enabled	Configures whether the output writer expects an Avro encoded Kafka Connect data record. Must match the configuration of value.converter.passthrough.enabled.	Boolean	none	true, false	True if input and output are both Avro.