

Cloudera DataFlow for Data Hub 7.2.12

Cloudera DataFlow for Data Hub Release Notes

Date published: 2019-12-16

Date modified: 2021-10-25

The Cloudera logo is displayed in a bold, orange, sans-serif font. The word "CLOUDERA" is written in all caps, with a stylized 'E' that has a horizontal bar extending to the right.

<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

What's New in Cloudera DataFlow for Data Hub 7.2.12.....	4
What's New in Flow Management.....	4
What's New in Streams Messaging.....	4
What's New in Streaming Analytics.....	5
Component Support in Cloudera DataFlow for Data Hub 7.2.12.....	6
Supported NiFi Extensions.....	6
Supported NiFi Processors.....	6
Supported NiFi Controller Services.....	8
Supported NiFi Reporting Tasks.....	10
Unsupported Features in Cloudera DataFlow for Data Hub 7.2.12.....	10
Unsupported Flow Management features.....	10
Unsupported Streams Messaging features.....	11
Unsupported Streaming Analytics features.....	11
Known Issues In Cloudera DataFlow for Data Hub 7.2.12.....	12
Known Issues in Flow Management.....	12
Known Issues in Streams Messaging.....	13
Known Issues in Streaming Analytics.....	18
Fixed Issues in Cloudera DataFlow for Data Hub 7.2.12.....	18
Fixed Issues in Flow Management.....	19
Fixed Issues in Streams Messaging.....	24
Fixed Issues in Streaming Analytics.....	25

What's New in Cloudera DataFlow for Data Hub 7.2.12

Cloudera DataFlow for Data Hub 7.2.12 includes components for Flow Management, Streaming Analytics, and Streams Messaging. Learn about the new features and improvements in each of these components.

What's New in Flow Management

Learn about the new Flow Management features in Cloudera DataFlow for Data Hub 7.2.12.

There are no new features for Flow Management in this release.

What's New in Streams Messaging

Learn about the new Streams Messaging features in Cloudera DataFlow for Data Hub 7.2.12.

Kafka

There are no new features for Kafka in this release.

Schema Registry

Schema Registry integration with Atlas

You can now enable the integration between Schema Registry and Atlas to view a version of a schema in Atlas.

Atlas serves as a metadata catalog that shows relationships between entities. Additionally, when Kafka is also integrated with Atlas you can connect Kafka topics and schemas.

For more information, see *Integrating with Atlas*.

Streams Messaging Manager

There are no new features for Streams Messaging Manager in this release.

Streams Replication Manager

SRM Service Multi Target Cluster Support

The SRM Service can now target and monitor multiple Kafka clusters. Previously, the Service could only target a single cluster. Targeting multiple clusters enables you to monitor multiple remote Kafka clusters using a single SRM deployment. Configuration is done in Cloudera Manager by adding multiple target Kafka cluster aliases to the Streams Replication Manager Service Target Cluster property. For more information, see [Configuring the service role target cluster](#).

SRM Service Remote Querying [Technical Preview]



Important: Remote Querying is available in this release of Cloudera Runtime but is not ready for production deployment. Cloudera encourages you to explore technical preview features in non-production environments and provide feedback on your experiences through the [Cloudera Community Forums](#).

A new feature, Remote Querying, is introduced for the SRM Service. Remote Querying enables the SRM Service to query other, remote SRM Services and fetch the metrics gathered by the remote SRM Services. This allows users to monitor all replications of a deployment that has multiple instances of SRM with a single SRM Service. For more information, see [Remote Querying](#) and [Configuring Remote Querying](#).

Auto topic creation is disabled by default on source clusters

A new Connector property, `disable.source.topic.auto.creation`, is introduced. This property controls whether auto topic creation is enabled for SRM's internal consumers that fetch data from source clusters. This property is enabled by default meaning that auto topic creation is disabled by default. This property is not directly available for configuration in Cloudera Manager, and does not have a dedicated configuration entry on the UI. Instead, it can be configured through the Streams Replication Manager's Replication Configs Cloudera Manager property. Configuration is possible on a global or replication level.

Cruise Control

Cruise Control service is available in CDP Public Cloud

Cruise Control is supported in CDP Public Cloud and Streams Messaging cluster definitions. Beside its basic functionalities, Cruise Control in CDP Public Cloud enables you to scale your Kafka clusters based on your requirements.

Related Information

[Integrating with Atlas](#)

What's New in Streaming Analytics

Learn about the new Streaming Analytics features in Cloudera DataFlow for Data Hub 7.2.12.

The following new features are introduced in Streaming Analytics CDF for Data Hub 7.2.12:

Apache Flink upgrade

Apache Flink 1.13 is supported in Streaming Analytics 7.2.12 cluster templates.

Default integration for SSB user in Ranger

The SQL Stream Builder (SSB) service user is by default added for Apache Kafka and YARN in Apache Ranger. When preparing your environment, you do not have to add the SSB user to the Kafka and YARN policies.

SQL Session Cluster support

Session mode is supported for SQL Stream Builder which means that the Flink jobs are running in a session cluster by default. You are also able to set the properties for the session mode directly from the SQL window on the Streaming SQL Console.

For more information, see the [Managing session for SQL jobs](#) section.

Connector support

The following new connectors are added as Templates to SQL Stream Builder:

- Change Data Capture (CDC) connectors: Oracle, MySQL, PostgreSQL
- Blackhole connector
- Faker connector

You are also able to use more types of Kafka connectors as there are individual templates for Upsert Kafka, Kafka in CDP and Kafka outside of CDP.

For more information, see the [Supported tables in SQL Stream Builder](#) section.

Streaming SQL Console updates

Expanded Flink SQL support

More SQL statements are available that can be used directly from the SQL window on the Streaming SQL Console.

INSERT INTO support

Instead of using a Sink Table on the Streaming SQL Console, you are able to use the INSERT INTO statement directly in the SQL window to output your query results.

Tables > Flink DDL option is no longer supported

Instead of using the Flink DDL wizard, you are able to use Templates under the SQL window to select and import the CREATE TABLE templates.

For more information, see the [Concept of tables in SQL Stream Builder](#) section.

Component Support in Cloudera DataFlow for Data Hub 7.2.12

Cloudera DataFlow for Data Hub 7.2.12 includes the following components.

Flow Management clusters

- Apache NiFi 1.13.2
- Apache NiFi Registry 0.8.0

Streams Messaging clusters

- Apache Kafka 2.5.0
- Schema Registry 0.10.0
- Streams Messaging Manager 2.2.0
- Streams Replication Manager 1.1.0
- Cruise Control 2.0.100

Streaming Analytics clusters

- Apache Flink 1.13

Supported NiFi Extensions

Apache NiFi 1.13.2 ships with a set of Processors, Controller Services, and Reporting Tasks, most of which are supported by Cloudera Support. Review the supported extensions and avoid using any unsupported extensions in your production environments.

Supported NiFi Processors

This release ships with Apache NiFi 1.13.2 and includes a set of Processors, most of which are supported by Cloudera Support. You should be familiar with the available supported Processors, and avoid using any unsupported Processors in production environments.

Additional Processors are developed and tested by the Cloudera community but are not officially supported by Cloudera. Processors are excluded for a variety of reasons, including insufficient reliability or incomplete test case coverage, declaration of non-production readiness by the community at large, and feature deviation from Cloudera best practices.

AttributesToCSV	GetAzureQueueStorage	PutCouchbaseKey 1
AttributesToJSON	GetCouchbaseKey 1	PutDatabaseRecord
Base64EncodeContent	GetFile	PutDistributedMapCache
CalculateRecordStats	GetFTP	PutDynamoDB 1
CaptureChangeMySQL	GetHBase	PutElasticsearch 1
CompressContent12	GetHDFS	PutElasticsearchHttp 1
ConnectWebSocket	GetHDFSFileInfo	PutElasticsearchHttpRecord
ConsumeAMQP	GetHDFSSequenceFile	PutElasticsearchRecord

ConsumeAzureEventHub	GetHTMLElement	PutEmail 1
ConsumeEWS	GetHTTP	PutFile
ConsumeGCPubSub	GetIgniteCache	PutFTP
ConsumeJMS	GetJMSQueue	PutGCSObject
ConsumeKafka	GetJMSTopic	PutGridFS
ConsumeKafka_0_10	GetKafka	PutHBaseCell 1
ConsumeKafka_1_0	GetMongoRecord	PutHBaseJSON
ConsumeKafka_2_0	GetSFTP	PutHBaseRecord
ConsumeKafka_2_6	GetSolr	PutHDFS
ConsumeKafka2CDP	GetSplunk	PutHive3QL
ConsumeKafka2RecordCDP	GetSQS	PutHive3Streaming
ConsumeKafkaRecord_0_10	GetTCP	PutHiveQL
ConsumeKafkaRecord_1_0	GetTwitter	PutHiveStreaming
ConsumeKafkaRecord_2_0	HandleHttpRequest	PutHTMLElement
ConsumeKafkaRecord_2_6	HandleHttpResponse	PutInfluxDB
ConsumeKinesisStream	HashAttribute	PutJMS
ConsumeMQTT 1	HashContent	PutKafka
ConsumeWindowsEventLog	IdentifyMimeType	PutKinesisFirehose
ControlRate	InvokeAWSGatewayApi	PutKinesisStream
ConvertAvroSchema	InvokeGRPC	PutKudu
ConvertAvroToJSON	InvokeHTTP	PutLambda
ConvertAvroToORC	InvokeScriptedProcessor	PutMongoRecord
ConvertAvroToParquet	JoltTransformJSON	PutORC
ConvertCharacterSet	JoltTransformRecord	PutParquet
ConvertCSVToAvro	JsonQueryElasticsearch	PutRecord
ConvertJSONToAvro	ListAzureBlobStorage	PutRiemann
ConvertJSONToSQL	ListAzureDataLakeStorage	PutS3Object
ConvertRecord	ListCDPObjectStore	PutSFTP
CreateHadoopSequenceFile	ListDatabaseTables	PutSNS
CryptographicHashAttribute	ListenFTP	PutSolrContentStream
CryptographicHashContent	ListenGRPC	PutSolrRecord
DecryptContentPGP	ListenHTTP	PutSplunk
DeleteAzureBlobStorage	ListenRELPL	PutSplunkHTTP 1
DeleteAzureDataLakeStorage	ListenSyslog	PutSQL
DeleteByQueryElasticsearch	ListenTCP	PutSQS
DeleteCDPObjectStore	ListenTCPRecord	PutSyslog
DeleteDynamoDB	ListenUDP	PutTCP
DeleteGCSObject	ListenUDPRecord	PutUDP
DeleteGridFS	ListenWebSocket	PutWebSocket 1
DeleteHBaseCells	ListFile	QueryCassandra
DeleteHBaseRow	ListFTP	QueryDatabaseTable
DeleteHDFS	ListGCSBucket	QueryDatabaseTableRecord
DeleteS3Object	ListHDFS	QueryElasticsearchHttp
DeleteSQS	ListS3	QueryRecord
DetectDuplicate	ListSFTP	QuerySolr
DistributeLoad	LogAttribute	QuerySplunkIndexingStatus
DuplicateFlowFile	LogMessage	QueryWhois
EncryptContent 2	LookupAttribute	ReplaceText 1
EncryptContentPGP	LookupRecord	ReplaceTextWithMapping
EnforceOrder	MergeContent 1	ResizeImage

EvaluateJsonPath	MergeRecord	RetryFlowFile
EvaluateXPath	ModifyHTMLERelement	RouteHL7
EvaluateXQuery	MonitorActivity	RouteOnAttribute
ExecuteGroovyScript	Notify	RouteOnContent
ExecuteInfluxDBQuery	ParseCEF	RouteText
ExecuteProcess	ParseEvtx	SampleRecord 1
ExecuteScript	ParseSyslog	ScanAccumulo
ExecuteSQL	PartitionRecord	ScanAttribute
ExecuteSQLRecord	PostHTTP	ScanContent
ExecuteStreamCommand	PublishAMQP 1	ScanHBase
ExtractAvroMetadata	PublishGCPubSub 1	ScriptedTransformRecord
ExtractGrok	PublishJMS 1	ScrollElasticsearchHttp
ExtractHL7Attributes	PublishKafka	SegmentContent
ExtractImageMetadata	PublishKafka_0_10	SelectHive3QL
ExtractText	PublishKafka_1_0	SelectHiveQL
FetchAzureBlobStorage	PublishKafka_2_0	SplitAvro 1
FetchAzureDataLakeStorage	PublishKafka_2_6	SplitContent 1
FetchCDPObjectStore	PublishKafka2CDP	SplitJson 1
FetchDistributedMapCache	PublishKafka2RecordCDP	SplitRecord
FetchElasticsearch	PublishKafkaRecord_0_10	SplitText 1
FetchElasticsearchHttp	PublishKafkaRecord_1_0	SplitXml 1
FetchFile	PublishKafkaRecord_2_0	TagS3Object
FetchFTP	PublishKafkaRecord_2_6	TailFile
FetchGCSObject	PublishMQTT 1	TransformXml
FetchGridFS	PutAccumuloRecord	UnpackContent
FetchHBaseRow	PutAzureBlobStorage	UpdateAttribute
FetchHDFS	PutAzureCosmosDBRecord 1	UpdateCounter
FetchParquet	PutAzureDataLakeStorage	UpdateHive3Table
FetchS3Object	PutAzureEventHub 1	UpdateHiveTable
FetchSFTP	PutAzureQueueStorage	UpdateRecord
FlattenJson	PutBigQueryBatch	ValidateCsv
ForkRecord	PutBigQueryStreaming 1	ValidateRecord
GenerateFlowFile	PutCassandraQL 1	ValidateXml
GenerateTableFetch	PutCassandraRecord	Wait
GeoEnrichIP	PutCDPObjectStore	YandexTranslate
GeoEnrichIPRecord	PutCloudWatchMetric	
GetAzureEventHub		

Footnotes

- 1 – indicates a memory intensive processor
- 2 – indicates a CPU intensive processor

Supported NiFi Controller Services

This release ships with Apache NiFi 1.13.2 and includes a set of Controller Services, most of which are supported by Cloudera Support. You should be familiar with the available supported Controller Services, and avoid using any unsupported Controller Services in production environments.

Additional Controller Services are developed and tested by the Cloudera community but are not officially supported by Cloudera. Controller Services are excluded for a variety of reasons, including insufficient reliability or incomplete

test case coverage, declaration of non-production readiness by the community at large, and feature deviation from Cloudera best practices.

AccumuloService	HiveConnectionPool
ActionHandlerLookup	HortonworksSchemaRegistry
ADLSCredentialsControllerService	IPFIXReader
ADLSIDBrokerCloudCredentialsProviderControllerService	IPLookupService
AlertHandler	JASN1Reader
AvroReader	JMSConnectionFactoryProvider
AvroRecordSetWriter	JndiJmsConnectionFactoryProvider
AvroSchemaRegistry	JsonPathReader
AWSCredentialsProviderControllerService	JsonRecordSetWriter
AWSIDBrokerCloudCredentialsProviderControllerService	JsonTreeReader
AzureBlobIDBrokerCloudCredentialsProviderControllerService	KafkaRecordSink_1_0
AzureCosmosDBClientService	KafkaRecordSink_2_0
AzureStorageCredentialsControllerService	KafkaRecordSink_2_6
AzureStorageCredentialsControllerServiceLookup	KeytabCredentialsService
CassandraDistributedMapCache	KuduLookupService
CassandraSessionProvider	LoggingRecordSink
CouchbaseClusterService	LogHandler
CouchbaseKeyValueLookupService	MongoDBControllerService
CouchbaseMapCacheClient	MongoDBLookupService
CouchbaseRecordLookupService	ParquetReader
CSVReader	ParquetRecordSetWriter
CSVRecordLookupService	PrometheusRecordSink
CSVRecordSetWriter	ReaderLookup
DatabaseRecordLookupService	RecordSetWriterLookup
DatabaseRecordSink	RecordSinkHandler
DBCPCConnectionPool	RecordSinkServiceLookup
DBCPCConnectionPoolLookup	RedisConnectionPoolService
DistributedMapCacheClientService	RedisDistributedMapCacheClientService
DistributedMapCacheLookupService	RestLookupService
DistributedMapCacheServer	ScriptedActionHandler
DistributedSetCacheClientService	ScriptedLookupService
DistributedSetCacheServer	ScriptedReader
EasyRulesEngineProvider	ScriptedRecordSetWriter
EasyRulesEngineService	ScriptedRecordSink
ElasticSearchClientServiceImpl	ScriptedRulesEngine
ElasticSearchLookupService	SimpleDatabaseLookupService
ElasticSearchStringLookupService	SimpleKeyValueLookupService
EmbeddedHazelcastCacheManager	SimpleScriptedLookupService
ExpressionHandler	SiteToSiteReportingRecordSink
ExternalHazelcastCacheManager	StandardHttpContextMap
FreeFormTextRecordSetWriter	StandardPGPPrivateKeyService
GCPCredentialsControllerService	StandardPGPPublicKeyService
GrokReader	StandardProxyConfigurationService
HadoopDBCPCConnectionPool	StandardRestrictedSSLContextService
HazelcastMapCacheClient	StandardS3EncryptionService
HBase_1_1_2_ClientMapCacheService	StandardSSLContextService
HBase_1_1_2_ClientService	Syslog5424Reader

HBase_1_1_2_ListLookupService	SyslogReader
HBase_1_1_2_RecordLookupService	VolatileSchemaCache
HBase_2_ClientMapCacheService	WindowsEventLogReader
HBase_2_ClientService	XMLReader
HBase_2_RecordLookupService	XMLRecordSetWriter
Hive3ConnectionPool	

Supported NiFi Reporting Tasks

This release ships with Apache NiFi 1.13.2 and includes a set of Reporting Tasks, most of which are supported by Cloudera Support. You should be familiar with the available supported Reporting Tasks, and avoid using any unsupported Reporting Tasks in production environments.

Additional Reporting Tasks are developed and tested by the Cloudera community but are not officially supported by Cloudera. Reporting Tasks are excluded for a variety of reasons, including insufficient reliability or incomplete test case coverage, declaration of non-production readiness by the community at large, and feature deviation from Cloudera best practices. Do not use these features in your production environments.

- AmbariReportingTask
- ControllerStatusReportingTask
- MetricsEventReportingTask
- MonitorDiskUsage
- MonitorMemory
- PrometheusReportingTask
- QueryNiFiReportingTask
- ReportLineageToAtlas
- ScriptedReportingTask
- SiteToSiteBulletinReportingTask
- SiteToSiteMetricsReportingTask
- SiteToSiteProvenanceReportingTask
- SiteToSiteStatusReportingTask

Unsupported Features in Cloudera DataFlow for Data Hub 7.2.12

Some features exist within Cloudera DataFlow for Data Hub 7.2.12 components, but are not supported by Cloudera.

Unsupported Flow Management features

There are no unsupported Flow Management features in Cloudera DataFlow for Data Hub 7.2.12

NiFi

There are no updates for this release.

NiFi Registry

There are no updates for this release.

Related Information

[Cloudera Community Forum](#)

Unsupported Streams Messaging features

Some Streams Messaging features exist in Cloudera DataFlow for Data Hub 7.2.12, but are not supported by Cloudera.

Kafka

The following Kafka features are not ready for production deployment. Cloudera encourages you to explore these features in non-production environments and provide feedback on your experiences through the *Cloudera Community Forums*.

- Only Java and .Net based clients are supported. Clients developed with C, C++, Python, and other languages are currently not supported.
- While Kafka Connect is available as part of Runtime, it is currently not supported in CDP Public Cloud. NiFi is a proven solution for batch and real time data loading that complement Kafka's message broker capability. For more information, see [Creating your first Flow Management cluster](#).
- The Kafka default authorizer is not supported. This includes setting ACLs and all related APIs, broker functionality, and command-line tools.

Schema Registry

There are no updates for this release.

Streams Messaging Manager

There are no updates for this release.

Streams Replication Manager

There are no updates for this release.

Related Information

[Cloudera Community Forum](#)

Unsupported Streaming Analytics features

Some Streaming Analytic features exist in Cloudera DataFlow for Data Hub 7.2.12, but are not supported by Cloudera.

The following features are not ready for production deployment. Cloudera encourages you to explore these features in non-production environments and provide feedback on your experiences through the *Cloudera Community Forums*.

SQL Stream Builder

- SQL Stream Builder REST API

Flink

- Apache Flink batch (DataSet) API
- GPU Resource Plugin
- Application Mode deployment
- SQL Client
- Python API

- The following features are not supported in SQL and Table API:
 - HBase Table Connector
 - Old Planner
 - Non-windowed (unbounded) joins, distinct

Related Information

[Cloudera Community Forum](#)

Known Issues In Cloudera DataFlow for Data Hub 7.2.12

You must be aware of the known issues and limitations, the areas of impact, and workaround in Cloudera DataFlow for Data Hub 7.2.12.

Known Issues in Flow Management

Learn about the known issues in Flow Management clusters, the impact or changes to the functionality, and the workaround.

NIFI-9054: Calling Nifi Registry's createExtensionBundleVersion REST endpoint will cause a NullPointerException

The `/buckets/{bucketId}/bundles/nifi-nar` API in NiFi Registry may throw a `NullPointerException`.

If you are using this API, contact Cloudera for a Hotfix.

JDK versions mismatch

If doing a software only upgrade for your Flow Management DataHub clusters and if repairing one of the NiFi nodes after the upgrade, you may be in a situation where the JDK used by NiFi is not the same across the nodes. In such a case, this may cause issues in the NiFi UI and you may get an "Unexpected error" message.

Ensure that the same JDK is used across the NiFi nodes and if there is a JDK versions mismatch, manually upgrade the JDK to match the JDK version being installed on the node that has been repaired.

Failed to import XML templates through the NiFi UI

When you try to import an XML template through the NiFi UI, you get an Invalid CORS request error.

Export or import flow definitions as JSON files. To import a JSON flow definition, drag and drop a process group on the canvas and upload the JSON file. For more information, see *Downloading a flow definition from NiFi*.

Technical Service Bulletins

TSB 2022-580: NiFi Processors cannot write to content repository

If the content repository disk is filled more than 50% (or any other value that is set in `nifi.properties` for `nifi.content.repository.archive.max.usage.percentage`), and if there is no data in the content repository archive, the following warning message can be found in the logs: "Unable to write flowfile content to content repository container default due to archive file size constraints; waiting for archive cleanup". This would block the processors and no more data is processed.

This appears to only happen if there is already data in the content repository on startup that needs to be archived, or if the following message is logged: "Found unknown file XYZ in the File System Repository; archiving file".

Upstream JIRA

- [NIFI-10023](#)

- [NIFI-9993](#)

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-580: NiFi Processors cannot write to content repository](#)

TSB 2022-589: CVE-2022-33140 Apache NiFi ShellUserGroupProvider Vulnerability

The optional ShellUserGroupProvider in Apache NiFi 1.10.0 to 1.16.2 and Apache NiFi Registry 0.6.0 to 1.16.2 does not neutralize arguments for group resolution commands, allowing injection of operating system commands on Linux and macOS platforms. The ShellUserGroupProvider is not included in the default configuration. Command injection requires ShellUserGroupProvider to be one of the enabled User Group Providers (UGP) in the Authorizers configuration. Command injection also requires an authenticated user with elevated privileges. Apache NiFi requires an authenticated user with authorization to modify access policies in order to execute the command. Apache NiFi Registry requires an authenticated user with authorization to read user groups in order to execute the command. The resolution removes command formatting based on user-provided arguments.

Knowledge article

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-589: CVE-2022-33140 Apache NiFi ShellUserGroupProvider Vulnerability](#)

Related Information

[Downloading a flow definition from NiFi](#)

Known Issues in Streams Messaging

Learn about the known issues in Streams Messaging clusters, the impact or changes to the functionality, and the workaround.

Kafka

Learn about the known issues and limitations in Kafka in this release:

Known Issues

Topics created with the kafka-topics tool are only accessible by the user who created them when the deprecated --zookeeper option is used

By default all created topics are secured. However, when topic creation and deletion is done with the kafka-topics tool using the `--zookeeper` option, the tool talks directly to Zookeeper. Because security is the responsibility of ZooKeeper authorization and authentication, Kafka cannot prevent users from making ZooKeeper changes. As a result, if the `--zookeeper` option is used, only the user who created the topic will be able to carry out administrative actions on it. In this scenario Kafka will not have permissions to perform tasks on topics created this way.

Use kafka-topics with the `--bootstrap-server` option that does not require direct access to Zookeeper.

Certain Kafka command line tools require direct access to Zookeeper

The following command line tools talk directly to ZooKeeper and therefore are not secured via Kafka:

- `kafka-reassign-partitions`

None

The `offsets.topic.replication.factor` property must be less than or equal to the number of live brokers

The `offsets.topic.replication.factor` broker configuration is now enforced upon auto topic creation. Internal auto topic creation will fail with a `GROUP_COORDINATOR_NOT_AVAILABLE` error until the cluster size meets this replication factor requirement.

None

Requests fail when sending to a nonexistent topic with auto.create.topics.enable set to true

The first few produce requests fail when sending to a nonexistent topic with auto.create.topics.enable set to true.

Increase the number of retries in the producer configuration setting retries.

KAFKA-2561: Performance degradation when SSL Is enabled

In some configuration scenarios, significant performance degradation can occur when SSL is enabled. The impact varies depending on your CPU, JVM version, Kafka configuration, and message size. Consumers are typically more affected than producers.

Configure brokers and clients with ssl.secure.random.implementation = SHA1PRNG. It often reduces this degradation drastically, but its effect is CPU and JVM dependent.

OPSAPS-43236: Kafka garbage collection logs are written to the process directory

By default Kafka garbage collection logs are written to the agent process directory. Changing the default path for these log files is currently unsupported.

None

CDPD-49304: AvroConverter does not support composite default values

AvroConverter cannot handle schemas containing a STRUCT type default value.

None.

Limitations

Collection of Partition Level Metrics May Cause Cloudera Manager's Performance to Degrade

If the Kafka service operates with a large number of partitions, collection of partition level metrics may cause Cloudera Manager's performance to degrade.

If you are observing performance degradation and your cluster is operating with a high number of partitions, you can choose to disable the collection of partition level metrics.



Important: If you are using SMM to monitor Kafka or Cruise Control for rebalancing Kafka partitions, be aware that both SMM and Cruise Control rely on partition level metrics. If partition level metric collection is disabled, SMM will not be able to display information about partitions. In addition, Cruise Control will not operate properly.

Complete the following steps to turn off the collection of partition level metrics:

1. Obtain the Kafka service name:
 - a. In Cloudera Manager, Select the Kafka service.
 - b. Select any available chart, and select Open in Chart Builder from the configuration icon drop-down.
 - c. Find \$SERVICENAME= near the top of the display.

The Kafka service name is the value of \$SERVICENAME.

2. Turn off the collection of partition level metrics:
 - a. Go to Hosts Configuration.
 - b. Find and configure the Cloudera Manager Agent Monitoring Advanced Configuration Snippet (Safety Valve) configuration property.

Enter the following to turn off the collection of partition level metrics:

```
[KAFKA_SERVICE_NAME]_feature_send_broker_topic_partition_entity_update_enabled=false
```

Replace [KAFKA_SERVICE_NAME] with the service name of Kafka obtained in step 1. The service name should always be in lower case.

- c. Click Save Changes.

Schema Registry

CDPD-49304: AvroConverter does not support composite default values

AvroConverter cannot handle schemas containing a STRUCT type default value.

None.

CDPD-49217 and CDPD-50309: Schema Registry caches user group membership indefinitely

Schema Registry caches the Kerberos user and group information indefinitely and does not catch up on group membership changes.

Restart Schema Registry after group membership changes.

CDPD-60160: Schema Registry Atlas integration does not work with Oracle databases

Schema Registry is unable to create entities in Atlas if Schema Registry uses an Oracle database. The following will be present in the Schema Registry log if you are affected by this issue:

```
ERROR com.cloudera.dim.atlas.events.AtlasEventsProcessor: An error occurred while processing Atlas events.
java.lang.IllegalArgumentException: Cannot invoke com.hortonworks.registries.schemaregistry.AtlasEventStorable.setType on bean class 'class com.hortonworks.registries.schemaregistry.AtlasEventStorable' - argument type mismatch - had objects of type "java.lang.Long" but expected signature "java.lang.Integer"
```

This issue causes the loss of audit data on Oracle environments.

None.

Streams Messaging Manager

Learn about the known issues in Streams Messaging Manager in this release.

OPSAPS-59553: SMM's bootstrap server config should be updated based on Kafka's listeners

SMM does not show any metrics for Kafka or Kafka Connect when multiple listeners are set in Kafka.

Workaround: SMM cannot identify multiple listeners and still points to bootstrap server using the default broker port (9093 for SASL_SSL). You would have to override bootstrap server URL (hostname:port as set in the listeners for broker) in the following path:

Cloudera Manager > SMM > Configuration > Streams Messaging Manager Rest Admin Server Advanced Configuration Snippet (Safety Valve) for streams-messaging-manager.yaml > Save Changes > Restart SMM.

OPSAPS-59597: SMM UI logs are not supported by Cloudera Manager

Cloudera Manager does not support the log type used by SMM UI.

Workaround: View the SMM UI logs on the host.

OPSAPS-59828: SMM cannot connect to Schema Registry when TLS is enabled

When TLS is enabled, SMM by default cannot properly connect to Schema Registry.

As a result, when viewing topics in the SMM Data Explorer with the deserializer key or value set to Avro, the following error messages are shown:

- Error deserializing key/value for partition [***PARTITION***] at offset [***OFFSET***]. If needed, please seek past the record to continue consumption.
- Failed to fetch value schema versions for topic : '[***TOPIC***]'.

In addition, the following certificate error will also be present in the SMM log:

- javax.net.ssl.SSLHandshakeException: PKIX path building failed:...

Workaround: Additional security properties must be set for SMM.

1. In Cloudera Manager, select the SMM service.
2. Go to Configuration.
3. Find and configure the SMM_JMX_OPTS property. Add the following JVM SSL properties:
 - Djavax.net.ssl.trustStore=[***SMM TRUSTSTORE LOCATION***]
 - Djavax.net.ssl.trustStorePassword=[***PASSWORD***]

Streams Replication Manager

Learn about the known issues in Streams Replication Manager in this release:

CDPD-22089: SRM does not sync re-created source topics until the offsets have caught up with target topic

Messages written to topics that were deleted and re-created are not replicated until the source topic reaches the same offset as the target topic. For example, if at the time of deletion and re-creation there are a 100 messages on the source and target clusters, new messages will only get replicated once the re-created source topic has 100 messages. This leads to messages being lost.

None

CDPD-14019: SRM may automatically re-create deleted topics

If auto.create.topics.enable is enabled, deleted topics are automatically recreated on source clusters.

Prior to deletion, remove the topic from the topic whitelist with the srm-control tool. This prevents topics from being re-created.

```
srm-control topics --source [SOURCE_CLUSTER] --target [TARGET_CLUSTER] --remove [TOPIC1][TOPIC2]
```

CDPD-13864 and CDPD-15327: Replication stops after the network configuration of a source or target cluster is changed

If the network configuration of a cluster which is taking part in a replication flow is changed, for example, port numbers are changed as a result of enabling or disabling TLS, SRM will not update its internal configuration even if SRM is reconfigured and restarted. From SRM's perspective, it is the cluster identity that has changed. SRM cannot determine whether the new identity corresponds to the same cluster or not, only the owner or administrator of that cluster can know. In this case, SRM tries to use the last known configuration of that cluster which might not be valid, resulting in the halt of replication.

The internal topic storing the configuration of SRM can be deleted. After a restart SRM will re-create and re-populate it with the configuration data loaded from its property file. The topic is hosted on the target cluster of the replication flow. The topic name is: mm2-configs.[SOURCE_ALIAS].internal. However, changing a replicated cluster's identity is generally not recommended.

CDPD-11079: Blacklisted topics appear in the list of replicated topics

If a topic was originally replicated but was later disallowed (blacklisted), it will still appear as a replicated topic under the /remote-topics REST API endpoint. As a result, if a call is made to this endpoint, the disallowed topic will be included in the response. Additionally, the disallowed topic will also be visible in the SMM UI. However, its Partitions and Consumer Groups will be 0, its Throughput, Replication Latency and Checkpoint Latency will show N/A.

None

OPSAPS-62546: Kafka External Account SSL keypassword configuration is used incorrectly by SRM

When a Kafka External Account specifies a keystore that uses an SSL key password, SRM uses it as the ssl.keystore.key configuration. Due to using the incorrect ssl.keystore.key configuration, SRM will fail to load the keystore in certain cases.

Workaround: For the keystores used by the Kafka External Accounts, the SSL key password should match the SSL keystore password, and the SSL keystore key password should not be provided. Alternatively, you can use the legacy connection configurations based on the streams.replication.manager.configs to specify the SSL key password.

Limitations

SRM cannot replicate Ranger authorization policies to or from Kafka clusters

Due to a limitation in the Kafka-Ranger plugin, SRM cannot replicate Ranger policies to or from clusters that are configured to use Ranger for authorization. If you are using SRM to replicate data to or from a cluster that uses Ranger, disable authorization policy synchronization in SRM. This can be achieved by clearing the Sync Topic Acls Enabled (sync.topic.acls.enabled) checkbox.

SRM cannot ensure the exactly-once semantics of transactional source topics

SRM data replication uses at-least-once guarantees, and as a result cannot ensure the exactly-once semantics (EOS) of transactional topics in the backup/target cluster.



Note: Even though EOS is not guaranteed, you can still replicate the data of a transactional source, but you must set isolation.level to read_committed for SRM's internal consumers. This can be done by adding [***SOURCE CLUSTER ALIAS***]->[***TARGET CLUSTER ALIAS***].consumer.isolation.level=read_committed to the Streams Replication Manager's Replication Configs SRM service property in Cloudera Manger.

SRM checkpointing is not supported for transactional source topics

SRM does not correctly translate checkpoints (committed consumer group offsets) for transactional topics. Checkpointing assumes that the offset mapping function is always increasing, but with transactional source topics this is violated. Transactional topics have control messages in them, which take up an offset in the log, but they are never returned on the consumer API. This causes the mappings to decrease, causing issues in the checkpointing feature. As a result of this limitation, consumer failover operations for transactional topics is not possible.

Cruise Control

Learn about the known issues and limitations in Cruise Control in this release:

CDPD-47616: Unable to initiate rebalance, number of valid windows (NumValidWindows) is zero

If a Cruise Control rebalance is initiated with the rebalance_disk parameter and Cruise Control is configured to fetch metrics from Cloudera Manager, Cruise Control stops collecting metrics from the partitions that are moved. This is because Cloudera Manager does not collect metrics from moved partitions due to an issue in Kafka (KAFKA-10320).

If the metrics are not available, the partition is considered invalid by Cruise Control. This results in Cruise Control blocking rebalance operations and proposal generation.

Rolling restart Kafka broker roles.

Known Issues in Streaming Analytics

Learn about the known issues in Streaming Analytics clusters, the impact or changes to the functionality, and the workaround.

SQL Stream Builder

CSA-1232: Big numbers are incorrectly represented on the Streaming SQL Console UI

The issue impacts the following scenarios in Streaming SQL Console:

- When having integers bigger than 253-1 among your values, the Input transformations and User Defined Functions are considered unsafe and produce incorrect results as these numbers will lose precision during parsing.
- When having integers bigger than 253-1 among your values, sampling to the Streaming SQL Console UI produces incorrect results as these numbers will lose precision during parsing.

None

CSA-1454: Timezone settings can cause unexpected behavior in Kafka tables

You must consider the timezone settings of your environment when using timestamps in a Kafka table as it can affect the results of your query. When the timestamp in a query is identified with `from_unixtime`, it returns the results based on the timezone of the system. If the timezone is not set in UTC+0, the timestamp of the query results will shift in time and will not be correct.

Change your local timezone settings to UTC+0.

CSA-1985: DROP TABLE limitation when using Webhook table

DROP TABLE cannot be executed against Webhook type tables. The following error message is displayed when trying to delete a Webhook table using the SQL window: Table with identifier 'xyz' does not exist.

Use the Delete button on the Streaming SQL Console.

CSA-2016: Deleting table from other teams

There is a limitation when using the Streaming SQL Console for deleting tables. It is not possible to delete a table that belongs to another team using the Delete button on the User Interface.

Use DROP TABLE statement from the SQL window.

CSA-3742: Catalogs are not working due to expired Kerberos TGT

When SSB is running for a longer period of time than the lifetime of the Kerberos Ticket Granting Ticket (TGT), authentication with the catalog services will fail and the catalogs stop working.

None

Flink

CSA-1318: Flink job submission fails in RAZ enabled environment

Flink applications will not run in environments where Ranger Authorization Service (RAZ) is enabled as the Flink job submission fails.

Workaround: None

Fixed Issues in Cloudera DataFlow for Data Hub 7.2.12

Fixed issues represent selected issues that were previously logged through Cloudera Support, but are addressed in the current release. These issues may have been reported in previous versions within the Known Issues section; meaning they were reported by customers or identified by Cloudera Quality Engineering team.

Review the list of issues that are resolved in Cloudera DataFlow for Data Hub 7.2.12.

Fixed Issues in Flow Management

Review the list of Flow Management issues that are resolved in Cloudera DataFlow for Data Hub 7.2.12.

7.2.12.7

Technical Service Bulletins

TSB 2022-580: NiFi Processors cannot write to content repository

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-580: NiFi Processors cannot write to content repository](#)

TSB 2022-589: CVE-2022-33140 Apache NiFi ShellUserGroupProvider Vulnerability

For the latest update on this issue see the corresponding Knowledge article: [TSB 2022-589: CVE-2022-33140 Apache NiFi ShellUserGroupProvider Vulnerability](#)

7.2.12

NIFI-3862

ListenHTTPServlet should set the issuerDN as well.

NIFI-5936

MockProcessSession remove() does not report a "DROP" provenance event.

NIFI-6714

Unable to start NiFi due to Illegal reflective access warning.

NIFI-7198

Support TLS in RedisConnectionPoolService.

NIFI-7361

UI - policy icon not showing up correctly.

NIFI-7443

SFTP KeepAlive not working.

NIFI-7737

Add support for String[] to PutCassandraRecord.

NIFI-7788

Start and Stop actions on a Process Group should enable/disable transmission of RPG's.

NIFI-7870

Fix anonymous access control for advanced UI resources.

NIFI-7886

FetchAzureBlobStorage.

NIFI-7947

DeleteAzureDataLakeStorage Processor cannot delete non empty folder.

NIFI-8126

Include Total Queued Duration in metrics reported via ConnectionStatus.

NIFI-8133

PutElasticsearchRecord/ElasticSearchClientService should allow null values to optionally be suppressed/output.

NIFI-8137

Record Path functions to un/escape JSON.

NIFI-8298

Refactor nifi-security-utils to reduce dependence on Bouncy Castle.

NIFI-8317

Prometheus metrics returning incorrect heap and non-heap usage values when max values are undefined.

NIFI-8325

Rework SNMP processors.

NIFI-8341

ConsumeAMQP and PublishAMQP Should support multiple hosts.

NIFI-8347

1.13.1/2 Listen Websocket Processor not working.

NIFI-8348

Advanced UI not available when building with Java 11.

NIFI-8361

UnpackContent failing for Deflate:Maximum.

NIFI-8365

JSON record reader mishandles deep CHOICE types.

NIFI-8376

Gracefully handle SQL exceptions in ResultSetRecordSet.

NIFI-8389

Grok Reader - raw data strategy when no match.

NIFI-8394

Fix NPE error when NiFi connects to external zookeeper insecurely.

NIFI-8397

Update to latest Simple-Syslog-5424 to fix BOM issue.

NIFI-8417

emoji in processor group name cause nifi to not start up.

NIFI-8428

UI - Access Policy icon not displaying properly for Controller Services.

NIFI-8431

Redundant validation of Dynamic Properties.

NIFI-8437

RecordReader 'Infer Schema' for large records with JsonPathReader may throw IOException.

NIFI-8442

Can not write Timestamp.

NIFI-8454

InvokeHTTP does not report final URL after following redirects.

NIFI-8457

When Load-balanced connections are used, offloading a node may never complete.

NIFI-8463

Custom SASL protocol name support in PutKudu processor.

NIFI-8466

Offloading the Single Node load balancing target causes ignored backpressure.

NIFI-8470

PutDatabaseRecord truncates strings to 1 character when inserting into PostgreSQL CHAR(n) field.

NIFI-8471

Parameter Contexts - show referencing process groups.

NIFI-8473

REST API Documentation Missing Parameter Contexts.

NIFI-8475

EL query parser can produce wrong output when input contains multiple escaped expressions.

NIFI-8477

If a node completely dies, can not delete it from the cluster; AKA Zombie Node.

NIFI-8496

NiFi returns 500 error when uploading templates.

NIFI-8501

Add support for Azure Storage Client-Side Encryption.

NIFI-8510

CSRF filter blocking requests that contain unrelated cookies.

NIFI-8512

NiFi is not efficient when handling nullable types with Avro Reader/Writer.

NIFI-8520

Parameter Contexts - Show the wrong information of referencing components.

NIFI-8530

PutDatabaseRecord DELETE issue with Postgres UUID.

NIFI-8535

PutDatabaseRecord should give a better message when the table cannot be found.

NIFI-8536

PrometheusReportingTask and PrometheusRecordSink do not support variable registry for port.

NIFI-8537

Components become invalid if referencing a disabled controller service.

NIFI-8542

The FlowFile content returned by Stateless TriggerResult.readContent may contain too much data.

NIFI-8609

Improve efficiency of converting Record object to Avro GenericRecord object.

NIFI-8610

Invalid bigdecimal values in ORC when multiple avro records provided.

NIFI-8620

NullPointerException on commit with multiple ProcessSession.

NIFI-8622

AvroReader not reading Decimal values correctly when avro file with no embedded schema is parsed.

NIFI-8628

Variable Registry - Variable count doesn't reset when opening the variable dialog.

NIFI-8631

ConsumeGCPubSub acknowledges messages without committing the session.

NIFI-8633

Content Repository can be improved to make fewer disks accesses on read.

NIFI-8639

Add custom headers and dynamic URLs to ConnectWebsocket.

NIFI-8643

Nifi 1.13.2 Cluster Load balancer Does not Work Properly.

NIFI-8645

Disable LongRunningTaskMonitor by default.

NIFI-8664

Upgrade json-path (jayway) dependency.

NIFI-8666

Parameters cannot be referenced from Expression Language if they have a space in the name.

NIFI-8667

Deadlock when enabling Controller Service and all Referencing Components.

NIFI-8668

ConsumeAzureEventHub NiFi processors need to support storage SAS token authentication.

NIFI-8670

Parameter Context update may fail if Controller Service referencing a parameter is Enabling.

NIFI-8674

Update Apache Calcite to latest 1.27.0 for fixing query bug.

NIFI-8689

Site-to-Site client is constantly flushing the socket's OutputStream.

NIFI-8690

NarBundleExtractor fails for latest extension docs.

NIFI-8697

If JWT Token no longer valid.

NIFI-8698

PutKudu fails if 'Handle Schema Drift' is true and using Data Record Path.

NIFI-8707

Upgrade SSHJ to 0.31.0.

NIFI-8711

Support keyboard-interactive authentication for SFTP.

NIFI-8719

ProcessSession may leave open references to files in content repo until committed/rolled back.

NIFI-8720

SplitRecord can enter infinite loop of Expression Language used to provide value of 0 Records per FlowFile.

NIFI-8721

Upgrade Spring Boot to 2.5.1.

NIFI-8722

EmbeddedQuestDbStatusHistoryRepository does not record new stats for Connections.

NIFI-8731

When updating Parameter Context, sometimes see failure: processor cannot be started because it is not stopped. Current state is STOPPING.

NIFI-8749

PutDatabaseRecord Insert Datetime Type Time Zone Issue.

NIFI-8761

Enable not setting a value for Escape Character in CSVReader controller service.

NIFI-8771

NiFi fails to start if clustered and flow consists only of 1+ Controller Services.

NIFI-8773

Allow TailFile to hold off on ingesting lines of text if the full (multi-line) message is not available.

NIFI-8785

AvroReader controller can't get confluent schema IDs except only the latest.

NIFI-8790

PutElasticsearchRecord incorrectly validates Index Operation for Expression Language input.

NIFI-8933

Jersey ObjectMapper should ignore unknown fields.

NIFI-8936

Adding Dynamic HTTP Headers to Confluent Schema Registry Requests.

NIFI-8939

Issue updating Parameter Values when referencing Controller Service is Disabling.

NIFI-8942

NiFi Registry: flow description cannot be selected and copied on the UI.

NIFI-8951

Precision is used instead of scale in QueryDatabase processors.

NIFI-8955

Add Max Connection Lifetime property to Hive(_1_1)ConnectionPool CS.

NIFI-8957

NiFi Registry - Possibility to set a description when creating a bucket.

NIFI-8969

Fix Maximum Polling Time in CuratorLeaderElectionManager.

NIFI-8973

Implement KerberosUserService API and implementations.

NIFI-8986

Upgrade Apache Commons Compress to 1.21.

NIFI-8987

Upgrade Apache PDFBox to 2.0.24.

NIFI-8990

Upgrade Groovy to 2.5.14 and 2.4.21.

NIFI-8996

PutHive*QL processors do not close JDBC statements.

NIFI-9017

If a node's load balance host/port is updated, other nodes stop sending data.

NIFI-9018

When connection points to a moved port version change with NiFi Registry may throw exception.

NIFI-9025

Build is failing because TestTailFile is references the wrong property.

NIFI-9035

Keystore Valid Check Throws NPE on null Keystore Password.

NIFI-9038

Fix fingerprinting group access control policies for Remote Port.

NIFI-9054

Calling Nifi Registry's createExtensionBundleVersion REST endpoint will cause NullPointerException.

NIFI-9055

Added handling for 0- read range to FetchS3Object.

NIFI-9061

Improve responsiveness of the UI.

NIFI-9066

PutSplunkHTTP - FF attributes support.

NIFI-9076

In MoveHDFS processor, hdfs-related operations should be wrapped in UGI.doAs().

NIFI-9079

Set log level to WARN for Apache Atlas client logs.

NIFI-9082

ZooKeeperStateProvider should use jute.maxbuffer property.

NIFI-9200

AbstractCSVLookupService cache remains on the heap after disabling the service.

NIFI-9201

NullPointerException in AbstractKerberosUser if the tgt is not renewable.

NIFI-9210

Upgrade jsoup to 1.14.2.

NIFI-9217

Possible deadlock when node is disconnected.

Fixed Issues in Streams Messaging

Review the list of Streams Messaging issues that are resolved in Cloudera DataFlow for Data Hub 7.2.12.

Kafka

CDPD-26944: Kafka client entities with null or empty client ID are created in Atlas

Atlas entities with null or empty client ID are no longer created.

CDPD-24428: Topics created with the default replication factor and partition count are incorrectly propagated into Atlas

Topics created with the default replication factor and partition count are now propagated into Atlas with the correct replication factor and partition number.

Schema Registry

CDPD-28714: Schema should connect to all Kafka topics with the same name

When having multiple Kafka clusters, with the Schema Registry-Atlas plugin, schemas are connected to all topics with the same name as the schema.

CDPD-26574: SR - Remove the spring security dependency due to CVE-2021-22112, was used by kerberos basic authentication

Removed the dependency on spring-security due to CVE-2021-22112.

CDPD-25052: No active transaction is associated with the thread

Multiple entries of No active transaction in the logs made the logs difficult to read. The logs no longer contain multiple entries.

CDPD-17737: Fix misused exception handling of Schema Registry resources

Endpoint `/api/v1/schemaregistry/schemas/versionsById/{versionId}/branch` responds with 404 Not Found when a `SchemaNotFoundException` is thrown (and not 400 Bad Request).

Streams Messaging Manager

CDPD-27197: Handle requests when KafkaConnect client not defined

When `KafkaConnect` is not configured, all the associated endpoints (except `/is-configured`) returns 404.

Streams Replication Manager

CDPD-27662: The `/remote-topics` endpoint reports null source for topics

The SRM Service now properly fills the source field of the topic info provided on the `/remote-topics` endpoint.

CDPD-14019: SRM may automatically re-create deleted topics on source clusters

Deleted topics are no longer automatically re-created on source clusters.

OPSAPS-60823: Configuring the SRM Client's secure storage is mandatory for unsecured environments

The `srm-control` tool failed to run in an unsecured cluster. Unnecessary additional configs have been removed from the `srm-control` tool's configuration when it is used in an unsecured cluster.

OPSAPS-61001: Saving configuration changes for SRM is not possible

The SRM Client's Secure Storage Password property is no longer a mandatory property.

Cruise Control

OPSAPS-60810: Cruise Control script overrides security related properties added to Safety Valve

The issue regarding the Cruise Control script is resolved, the properties added to the Safety Valve are applied for the Cruise Control service.

Fixed Issues in Streaming Analytics

Review the list of Streaming Analytics issues that are resolved in Cloudera DataFlow for Data Hub 7.2.12.

7.2.12.10

CSA-3742: Catalogs are not working due to expired Kerberos TGT

The issue regarding the expired Kerberos Ticket Granting Ticket (TGT) and catalog authentication has been fixed.

7.2.12**There is an authentication issue when accessing Streaming SQL Console**

The authentication issue when accessing Streaming SQL Console is resolved. You can access the Streaming SQL Console when using environments with assigned public IP addresses.

CSA-2093: SQL Jobs are not listed with Knox

The issue regarding the Knox authentication and SQL job list is resolved. The stopped or running jobs are listed under the SQL Jobs tab in Streaming SQL Console when using Knox authentication.

CSA-2155: Dependency version mismatch in Streaming SQL Engine logs

The issue regarding the errors of the Streaming SQL Engine logs has been resolved. Errors are not generated due to not compatible dependency versions in the Streaming SQL Engine.

CSA-2156: SASL password is displayed for Kafka Data Provider

The password you set for the Kafka Data Provider when using the wizard in Streaming SQL Console is now hidden on the UI.

CSA-2302: SQL Stream Builder ignores restart from savepoint

The issue regarding SQL Stream Builder ignoring the restart from savepoint option has been resolved.