Cloudera Flow Management 4.0.0

Cloudera Flow Management Release Notes

Date published: 2019-06-26
Date modified: 2024-12-11

CLOUDERA

https://docs.cloudera.com/

Legal Notice

Contents

What's new in Cloudera Flow Management 4.0.0 [Technical Preview]

Explore the new features and improvements in Cloudera Flow Management and learn how these new additions can enhance your workflows.

Cloudera Flow Management 4.0.0 is a NiFi 2 Technical Preview based on the General Availability (GA) version of Apache NiFi 2.0.0. This release also includes several new NiFi 2 and Cloudera-exclusive features and improvements, including a powerful Python API for developing native NiFi components using Python. While this NiFi version brings significant enhancements, the transition from NiFi 1.x to NiFi 2.x comes with numerous breaking changes, potentially making the upgrade process complex and challenging.

Currently, there is no supported upgrade path from Cloudera Flow Management versions using NiFi 1 (2.1.7 and lower) to Cloudera Flow Management 4.0.0. Cloudera will provide tooling to help with upgrades and manage breaking changes in the future. The NiFi Migration Tool will offer a semi-automatic solution for migrating NiFi flows from 1.x to 2.x, focusing on compatibility with new NiFi 2 features, such as parameter contexts and new supported components. The tool will not address all breaking changes, but will aim to cover as many as possible during migration. It will reduce manual work and will ensure the best compatibility.

Important:

To upgrade to Cloudera Flow Management 4.0.0 [Technical Preview] or any future NiFi 2-based release and use the Migration Tool provided by Cloudera for migrating your flows, you must be on Cloudera Flow Management 2.1.7. For detailed guidance on preparing for the upgrade, see Preparing for upgrade to NiFi 2.

For installation instructions, see Cloudera Flow Management installation workflow.

Here are the most important new features and improvements of this release:

Flow Analysis Tool

The Flow Analysis tool introduces a rules engine that provides real-time feedback during flow design, ensuring that flows adhere to best practices and configuration guidelines. This feature allows NiFi administrators to define a set of rules to enforce best practices in flow design in NiFi. The defined rules prevent invalid flows from being deployed to production, improving performance and compliance. It is fully integrated into the NiFi UI, offering an intuitive new view that can be customized based on customer requirements.

Python API and AI processors supporting Python

There is now a first-class citizen Python API allowing users to develop NiFi processors using Python. For more information, see the NiFi Python Developer's Guide.

One of the key features introduced in Apache NiFi 2 is native support for Python processors. This capability allows you to create custom processors using Python, enabling seamless integration of Python scripts into your dataflows. With each milestone release of NiFi 2, Python integration continues to evolve, providing developers with enhanced functionality, greater flexibility, and more powerful tools for building robust dataflows.

The below list shows the Python processors that are available in Cloudera Flow Management 4.0.0.

- Bedrock
- ChunkData
- ChunkDocument
- EmbedData
- InsertToMilvus
- LexicalQueryMilvus
- ParseDocument
- PartitionCsv

- PartitionDocx
- PartitionHtml
- PartitionPdf
- PartitionText
- PromptChatGPT
- PutChroma
- PutOpenSearchVector
- PutPinecone
- PutQdrant
- QueryChroma
- QueryOpenSearchVector
- QueryPinecone
- QueryQdrant
- VectorQueryMilvus

Stateless Engine at Process Group level

It is now possible to configure a Process Group to use the Stateless Engine for running flows. This is particularly useful for transactional use cases such as Change Data Capture (CDC), or scenarios where a message broker is the source, aiming to achieve exactly-once semantics. For more information, see the Apache NiFi User Guide.

New NiFi components

Cloudera Flow Management 4.0.0 has introduced several new NiFi components. For a comprehensive list of supported NiFi components, see the Support Matrix.

Breaking changes

There are a number of breaking changes between NiFi 1 and NiFi 2. See Behavioral changes for more information about these changes. Additionally, you can anticipate further changes in the upcoming releases, particularly concerning components completely removed in favor of better and more efficient alternatives.

Support matrix

Review the support matrix before you start installing Cloudera Flow Management.

Component versions

Review the Cloudera Flow Management component versions for compatibility with other applications.

Note: NiFi is compatible with the version of NiFi Registry bundled with your Cloudera Flow Management release as well as any later version.

Cloudera Flow Management 4.0.0 [Technical Preview]

- Apache NiFi 2.0.0.4.0.0.0
- Apache NiFi Registry 2.0.0.4.0.0.0

Cloudera Flow Management 2.1.7.1000 (SP1)

- Apache NiFi 1.26.0.2.1.7.1000
- Apache NiFi Registry 1.26.0.2.1.7.1000

Cloudera Flow Management 2.1.7

- Apache NiFi 1.26.0.2.1.7.0
- Apache NiFi Registry 1.26.0.2.1.7.0

Cloudera Flow Management 2.1.6.1000 (SP1)

- Apache NiFi 1.23.1.2.1.6.1000
- Apache NiFi Registry 1.23.1.2.1.6.1000

Cloudera Flow Management 2.1.6

- Apache NiFi 1.23.1.2.1.6.0
- Apache NiFi Registry 1.23.1.2.1.6.0

Cloudera Flow Management 2.1.5.1000 (SP1)

- Apache NiFi 1.18.0.2.1.5.1000
- Apache NiFi Registry 1.18.0.2.1.5.1000

Cloudera Flow Management 2.1.5

- Apache NiFi 1.18.0.2.1.5.0
- Apache NiFi Registry 1.18.0.2.1.5.0

Cloudera Flow Management 2.1.4.1000 (SP1)

- Apache NiFi 1.16.0.2.1.4.1000
- Apache NiFi Registry 1.16.0.2.1.4.1000

Cloudera Flow Management 2.1.4

- Apache NiFi 1.16.0.2.1.4.0
- Apache NiFi Registry 1.16.0.2.1.4.0

Cloudera Flow Management 2.1.3

- Apache NiFi 1.15.2.2.1.3.0
- Apache NiFi Registry 1.15.2.2.1.3.0

Note: Apache NiFi and Apache NiFi Registry versions are unified in the 1.15.x release.

Cloudera Flow Management 2.1.2

- Apache NiFi 1.13.2.2.1.2.0
- Apache NiFi Registry 0.8.0.2.1.2.0

Cloudera Flow Management 2.1.1

- Apache NiFi 1.13.2.2.1.1.0
- Apache NiFi Registry 0.8.0.2.1.1.0

Cloudera Flow Management 2.0.4

- Apache NiFi 1.11.4
- Apache NiFi Registry 0.6.0

Cloudera Flow Management 2.0.1

- Apache NiFi 1.11.4
- Apache NiFi Registry 0.6.0

System requirements

Review the system requirements before getting started with installing Cloudera Flow Management.

Supported Cloudera versions

Cloudera Flow Management 4.0.0 supports the following versions of Cloudera Private Cloud Base:

- Cloudera 7.3.1

Supported JAVA Development Kits (JDK)

Cloudera Flow Management requires a minimum of JDK 21 for proper functionality. Ensure your environment meets this requirement before installation.

Other system requirements

ZooKeeper

You need to install the ZooKeeper service included with your Cloudera Private Cloud Base cluster.

Python

- Minimum requirement: Python 3.11
- Recommended version: Python 3.12

Number of cores

- Minimum: Four cores per NiFi node are required for Cloudera support.
- Recommended: Eight cores per NiFi node, which typically provides an optimal starting point for most common use cases.

Supported operating systems

Review the list of operating systems supported by Cloudera Flow Management.

Operating system	Versions
CentOS	• 8.2 • 8.4
RHEL	• 8.2 • 8.4 • 8.6 • 8.7 • 8.8 • 8.9 • 8.10 • 9.1 • 9.2
Oracle	• 8.8
SLES	• 12 SP5 • 15 SP4

Operating system	Versions
Ubuntu	• 20.04 • 22.04
Windows	• 10 • Server 2016 • Server 2019

Note:

NiFi on Windows is only supported in standalone mode, not managed by Cloudera Manager or as part of a Cloudera cluster, and as a single instance installation. Clustering NiFi on Windows is not supported.

NiFi Registry is not supported on Windows.

Supported NiFi Registry databases

Review the list of databases supported by NiFi Registry.

- H2
- PostgreSQL 10.x
- PostgreSQL 11.x
- PostgreSQL 12.x
- PostgreSQL 13.x
- PostgreSQL 14.x
- MySQL 8.x

Related Information

Supported NiFi processors
Supported NiFi controller services
Supported NiFi reporting tasks

Supported NiFi processors

Cloudera Flow Management is based on Apache NiFi and includes a set of processors, most of which are supported by Cloudera. To ensure optimal performance and reliable support, it is crucial to use only supported processors and avoid deploying unsupported ones in production environments.

Additional processors are developed and tested by the community but are not officially supported by Cloudera. Processors may be excluded for various reasons, including insufficient reliability, incomplete test coverage, community declaration of non-production readiness, or deviations from Cloudera best practices.

By adhering to the above guidelines, you can maintain stable and reliable workflows in your production environments.

AttributesToCSV	GetElasticsearch	PutFile
AttributesToJSON	GetFile	PutFTP1
CalculateParquetOffsets	GetFTP	PutGCSObject
CalculateParquetRowGroupOffsets	GetGcpVisionAnnotateFilesOperationStatus	PutGoogleDrive
CalculateRecordStats	GetGcpVisionAnnotateImagesOperationStatus	PutGridFS
CaptureChangeDebeziumDB2 [Technial Preview]	GetHBase	PutHBaseCell
CaptureChangeDebeziumMySQL [Technial Preview]	GetHDFS	PutHBaseJSON

CaptureChangeDebeziumOracle [Technial Preview]	GetHDFSFileInfo	PutHBaseRecord1
CaptureChangeDebeziumPostgreSQL [Technial Preview]	GetHDFSSequenceFile	PutHDFS
CaptureChangeDebeziumSQLServer [Technial Preview]	GetHubSpot	PutHive3QL
CaptureChangeMySQL	GetJiraIssue	PutHive3Streaming
ChunkDocument [Technial Preview]	GetMongoRecord	PutIceberg
CompressContent1, 2	GetSFTP	PutIcebergCDC [Technial Preview]
ConnectWebSocket	GetShopify	PutJiraIssue
ConsumeAMQP	GetSNMP	PutKinesisFirehose
ConsumeAzureEventHub	GetSnowflakeIngestStatus	PutKinesisStream
ConsumeElasticsearch	GetSolr	PutKudu
ConsumeGCPubSub	GetSplunk	PutLambda
ConsumeGCPubSubLite	GetSQS	PutMongoBulkOperations
ConsumeJMS	GetWorkdayReport	PutMongoRecord
ConsumeKafka_2_6	GetZendesk	PutORC1
ConsumeKafka2CDP	HandleHttpRequest	PutParquet
ConsumeKafka2RecordCDP	HandleHttpResponse	PutPinecone [Technial Preview]
ConsumeKafkaRecord_2_6	IdentifyMimeType	PutPLC [Technial Preview]
ConsumeKinesisStream	InvokeAWSGatewayApi	PutRecord
ConsumeMQTT1	InvokeGRPC	PutRedisHashRecord [Technial Preview]
ConsumePLC [Technial Preview]	InvokeHTTP	PutS3Object
ConsumeSlack	InvokeScriptedProcessor	PutSalesforceObject
ConsumeTwitter	JoinEnrichment	PutSFTP
ConsumeWindowsEventLog	JoltTransformJSON	PutSmbFile
ControlRate	JoltTransformRecord	PutSnowflakeInternalStage
ConvertAvroToJSON	JSLTTransformJSON	PutSNS
ConvertAvroToParquet	JsonQueryElasticsearch	PutSolrContentStream
ConvertCharacterSet	ListAzureBlobStorage_v12	PutSolrRecord
ConvertJSONToSQL	ListAzureDataLakeStorage	PutSplunk
ConvertProtobuf	ListBoxFile	PutSplunkHTTP
ConvertRecord	ListCDPObjectStore	PutSQL
CopyAzureBlobStorage_v12	ListDatabaseTables	PutSQS1
CreateHadoopSequenceFile	ListDropbox	PutSyslog
CryptographicHashContent	ListenBeats	PutTCP
DecryptContent	ListenFTP	PutUDP
DecryptContentAge	ListenGRPC	PutWebSocket
DecryptContentCompatibility	ListenHTTP	PutZendeskTicket
DecryptContentPGP	ListenNetFlow	QueryAirtableTable
DeduplicateRecord	ListenOTLP	QueryCassandra
DeleteAzureBlobStorage_v12	ListenRELP	QueryChroma [Technial Preview]

DeleteAzureDataLakeStorage	ListenSlack	QueryDatabaseTable1
DeleteByQueryElasticsearch	ListenSyslog	QueryDatabaseTableRecord
DeleteCDPObjectStore	ListenTCP	QueryPinecone [Technial Preview]
DeleteDynamoDB	ListenTCPRecord	QueryRecord
DeleteGCSObject	ListenTrapSNMP	QuerySalesforceObject
DeleteGridFS	ListenUDP	QuerySolr
DeleteHBaseCells	ListenUDPRecord	QuerySplunkIndexingStatus
DeleteHBaseRow	ListenWebSocket	QueryWhois
DeleteHDFS	ListFile	RemoveRecordField
DeleteS3Object	ListFTP	RenameRecordField
DeleteSQS	ListGCSBucket	ReplaceText
DetectDuplicate	ListGoogleDrive	ReplaceTextWithMapping
DistributeLoad	ListHDFS	ResizeImage1
DuplicateFlowFile	ListS3	RetryFlowFile
EncodeContent	ListSFTP	RouteHL7
EncryptContentAge	ListSmb	RouteOnAttribute
EncryptContentPGP	LogAttribute	RouteOnContent
EnforceOrder	LogMessage	RouteText
EvaluateJsonPath	LookupAttribute	SampleRecord
EvaluateXPath	LookupRecord	ScanAccumulo
EvaluateXQuery	MergeContent	ScanAttribute1
ExecuteGroovyScript	MergeRecord1	ScanContent
ExecuteProcess	ModifyCompression	ScanHBase
ExecuteScript	MonitorActivity	ScriptedFilterRecord
ExecuteSQL	MoveAzureDataLakeStorage	ScriptedPartitionRecord
ExecuteSQLRecord	MoveHDFS	ScriptedTransformRecord
ExecuteStateless1, 2	Notify	ScriptedValidateRecord
ExecuteStreamCommand	PackageFlowFile	SearchElasticsearch
ExtractAvroMetadata	PaginatedJsonQueryElasticsearch	SegmentContent
ExtractGrok	ParseCEF1	SelectClouderaHiveQL
ExtractHL7Attributes	ParseDocument [Technial Preview]	SelectHive3QL1
ExtractImageMetadata	ParseEvtx	SendTrapSNMP
ExtractRecordSchema	ParseSyslog	SetSNMP
ExtractText	PartitionRecord	SignContentPGP
FetchAzureBlobStorage_v12	PromptChatGPT [Technial Preview]	SplitAvro
FetchAzureDataLakeStorage	PublishAMQP	SplitContent
FetchBoxFile	PublishGCPubSub1	SplitJson1
FetchCDPObjectStore	PublishGCPubSubLite1	SplitRecord1
FetchDistributedMapCache	PublishJMS1	SplitText1
FetchDropbox	PublishKafka_2_6	SplitXml

FetchFile	PublishKafka2CDP	StartAwsPollyJob
FetchFTP	PublishKafka2RecordCDP	StartAwsTextractJob
FetchGCSObject	PublishKafkaRecord_2_6	StartAwsTranscribeJob
FetchGoogleDrive	PublishMQTT	StartAwsTranslateJob
FetchGridFS	PublishSlack	StartGcpVisionAnnotateFilesOperation
FetchHBaseRow	PutAccumuloRecord1	StartGcpVisionAnnotateImagesOperation
FetchHDFS	PutAzureBlobStorage_v12	StartSnowflakeIngest
FetchParquet	PutAzureCosmosDBRecord	TagS3Object
FetchPLC [Technial Preview]	PutAzureDataLakeStorage1	TailFile
FetchS3Object	PutAzureEventHub	TransformXml
FetchSFTP	PutAzureQueueStorage_v12	TriggerClouderaHiveMetaStoreEvent
FetchSmb	PutBigQuery	TriggerHiveMetaStoreEvent
FilterAttribute	PutBoxFile	UnpackContent
FlattenJson	PutCassandraQL1	UpdateAttribute
ForkEnrichment	PutCassandraRecord1	UpdateByQueryElasticsearch
ForkRecord	PutCDPObjectStore	UpdateClouderaHiveTable
GenerateFlowFile	PutChroma [Technial Preview]	UpdateCounter
GenerateRecord	PutClouderaHiveQL	UpdateDatabaseTable
GenerateTableFetch	PutClouderaHiveStreaming	UpdateDeltaLakeTable [Technial Preview]
GeoEnrichIP	PutClouderaORC	UpdateHive3Table
GeoEnrichIPRecord	PutCloudWatchMetric	UpdateRecord
GeohashRecord	PutCouchbaseKey	ValidateCsv
GetAsanaObject	PutDatabaseRecord1	ValidateJson
GetAwsPollyJobStatus	PutDistributedMapCache	ValidateRecord
GetAwsTextractJobStatus	PutDropbox	ValidateXml
GetAwsTranscribeJobStatus	PutDynamoDB	VerifyContentMAC
GetAwsTranslateJobStatus	PutDynamoDBRecord	VerifyContentPGP
GetAzureEventHub	PutElasticsearchJson	Wait
GetAzureQueueStorage_v12	PutElasticsearchRecord1	YandexTranslate
GetCouchbaseKey1	PutEmail	

Footnotes

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

Related Information

Supported NiFi Registry databases
Supported NiFi controller services
Supported NiFi reporting tasks

Supported NiFi controller services

Cloudera Flow Management is based on Apache NiFi and includes a set of controller services, most of which are supported by Cloudera. To ensure optimal performance and reliable support, it is crucial to use only supported controller services and avoid deploying unsupported ones in production environments.

Additional controller services are developed and tested by the community but are not officially supported by Cloudera. Controller services may be excluded for various reasons, including insufficient reliability, incomplete test coverage, community declaration of non-production readiness, or deviations from Cloudera best practices.

By adhering to the above guidelines, you can maintain stable and reliable workflows in your production environments.

AccumuloService	IPLookupService
ActiveMQJMSConnectionFactoryProvider	JASN1Reader
ADLSCredentialsControllerService	JiraRecordSink
ADLSCredentialsControllerServiceLookup	JMSConnectionFactoryProvider
ADLSIDBrokerCloudCredentialsProviderControllerService	JndiJmsConnectionFactoryProvider
AmazonGlueSchemaRegistry	JsonConfigBasedBoxClientService
ApicurioSchemaRegistry	JsonPathReader
AvroReader	JsonRecordSetWriter
AvroRecordSetWriter	JsonTreeReader
AvroSchemaRegistry	KafkaRecordSink_2_6
AWSCredentialsProviderControllerService	KerberosKeytabUserService
AWSIDBrokerCloudCredentialsProviderControllerService	KerberosPasswordUserService
AzureBlobIDBrokerCloudCredentialsProviderControllerService	KerberosTicketCacheUserService
AzureCosmosDBClientService	KeytabCredentialsService
AzureEventHubRecordSink	KuduLookupService
AzureServiceBusJMSConnectionFactoryProvider	LoggingRecordSink
AzureStorageCredentialsControllerService_v12	MongoDBControllerService
AzureStorageCredentialsControllerServiceLookup_v12	MongoDBLookupService
CassandraDistributedMapCache	ParquetReader
CassandraSessionProvider	ParquetRecordSetWriter
CdpCredentialsProviderControllerService	PostgreSQLConnectionPool
CdpOauth2AccessTokenProviderControllerService	PrometheusRecordSink
CEFReader	ProxyPLC4XConnectionPool [Technical Preview]
CiscoEmblemSyslogMessageReader	RabbitMQJMSConnectionFactoryProvider
ClouderaHiveConnectionPool	ReaderLookup
ClouderaSchemaRegistry	RecordSetWriterLookup
CMLLookupService	RecordSinkServiceLookup
ConfluentEncodedSchemaReferenceReader	RedisConnectionPoolService
ConfluentEncodedSchemaReferenceWriter	RedisDistributedMapCacheClientService
ConfluentSchemaRegistry	RedshiftConnectionPool
CouchbaseClusterService	RestLookupService

CouchbaseKeyValueLookupService	ScriptedLookupService
CouchbaseMapCacheClient	ScriptedReader
CouchbaseRecordLookupService	ScriptedRecordSetWriter
CSVReader	ScriptedRecordSink
CSVRecordLookupService	SimpleDatabaseLookupService
CSVRecordSetWriter	SimpleKeyValueLookupService
DatabaseRecordLookupService	SimpleRedisDistributedMapCacheClientService
DatabaseRecordSink	SimpleScriptedLookupService
DatabaseTableSchemaRegistry	SiteToSiteReportingRecordSink
DBCPConnectionPool	SlackRecordSink
DBCPConnectionPoolLookup	SmbjClientProviderService
DistributedMapCacheClientService	SnowflakeComputingConnectionPool
DistributedMapCacheLookupService	StandardAsanaClientProviderService
DistributedMapCacheServer	StandardAzureCredentialsControllerService
DistributedSetCacheClientService	StandardDropboxCredentialService
DistributedSetCacheServer	StandardFileResourceService
EBCDICRecordReader [Technical Preview]	StandardHashiCorpVaultClientService
ElasticSearchClientServiceImpl	StandardHttpContextMap
ElasticSearchLookupService	StandardJsonSchemaRegistry [Technical Preview]
ElasticSearchStringLookupService	StandardOauth2AccessTokenProvider
EmailRecordSink	StandardPGPPrivateKeyService
EmbeddedHazelcastCacheManager	StandardPGPPublicKeyService
ExcelReader	StandardPLC4XConnectionPool [Technical Preview]
ExternalHazelcastCacheManager	StandardPrivateKeyService
FreeFormTextRecordSetWriter	StandardProxyConfigurationService
GCPCredentialsControllerService	StandardRestrictedSSLContextService
GCSFileResourceService	StandardS3EncryptionService
GenericPLC4XConnectionPool [Technical Preview]	StandardSnowflakeIngestManagerProviderService
GrokReader	StandardSSLContextService
HadoopCatalogService	StandardWebClientServiceProvider
HadoopDBCPConnectionPool	Syslog5424Reader
HazelcastMapCacheClient	SyslogReader
HBase_2_ClientMapCacheService	UDPEventRecordSink
HBase_2_ClientService	VolatileSchemaCache
HBase_2_RecordLookupService	WindowsEventLogReader
Hive3ConnectionPool	XMLReader
HiveCatalogService	XMLRecordSetWriter
ImpalaConnectionPool	YamlTreeReader
IPFIXReader	ZendeskRecordSink

Related Information

Supported NiFi Registry databases
Supported NiFi processors
Supported NiFi reporting tasks

Supported NiFi reporting tasks

Cloudera Flow Management is based on Apache NiFi and includes a set of reporting tasks, most of which are supported by Cloudera. To ensure optimal performance and reliable support, it is crucial to use only supported reporting tasks and avoid deploying unsupported ones in production environments.

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

Additional reporting tasks are developed and tested by the community but are not officially supported by Cloudera. Reporting tasks may be excluded for various reasons, including insufficient reliability, incomplete test coverage, community declaration of non-production readiness, or deviations from Cloudera best practices.

Related Information

Supported NiFi Registry databases
Supported NiFi processors
Supported NiFi controller services

Supported NiFi parameter providers

Cloudera Flow Management is shipped with Apache NiFi and includes a set of parameter providers, most of which are supported by Cloudera. To ensure optimal performance and reliable support, it is crucial to use only supported parameter providers and avoid deploying unsupported ones in production environments.

- AwsSecretsManagerParameterProvider
- AzureKeyVaultSecretsParameterProvider
- CyberArkConjurParameterProvider
- DatabaseParameterProvider
- EnvironmentVariableParameterProvider
- FileParameterProvider
- GcpSecretManagerParameterProvider
- HashiCorpVaultParameterProvider
- OnePasswordParameterProvider

Additional parameter providers are developed and tested by the community but are not officially supported by Cloudera. Parameter providers may be excluded for various reasons, including insufficient reliability, incomplete test coverage, community declaration of non-production readiness, or deviations from Cloudera best practices.

Supported NiFi flow analysis rules

Apache NiFi 2.0 introduces flow analysis rules, a new feature designed to enhance flow validation and management by evaluating components or parts of a flow and may generate rule violations to help optimize or maintain flow design.

Currently, only the DisallowComponentType flow analysis rule is available for use.

Supported Cloudera exclusive components

Cloudera Flow Management provides a set of NiFi components available only to Cloudera customers. These components provide additional functionality and are tailored to enhance the Cloudera NiFi experience. The list of these components is provided below.

Processors

- CaptureChangeDebeziumDB2
- CaptureChangeDebeziumMySQL
- CaptureChangeDebeziumOracle
- CaptureChangeDebeziumPostgreSQL
- CaptureChangeDebeziumSQLServer
- ConsumeKafka2CDP
- ConsumeKafka2RecordCDP
- ConsumePLC
- ConvertProtobuf
- DeleteCDPObjectStore
- FetchCDPObjectStore
- FetchPLC
- GetJiraIssue
- InvokeGRPC
- ListCDPObjectStore
- ListenGRPC
- ListenNetFlow
- PublishKafka2CDP
- PublishKafka2RecordCDP
- PutCDPObjectStore
- PutClouderaHiveQL
- PutClouderaHiveStreaming
- PutClouderaORC
- PutIcebergCDC
- PutJiraIssue
- PutPLC
- SelectClouderaHiveQL
- TriggerClouderaHiveMetaStoreEvent
- UpdateClouderaHiveTable
- UpdateDeltaLakeTable

Controller services

- ActiveMQJMSConnectionFactoryProvider
- ADLSIDBrokerCloudCredentialsProviderControllerService
- AWSIDBrokerCloudCredentialsProviderControllerService

- AzureBlobIDBrokerCloudCredentialsProviderControllerService
- AzureServiceBusJMSConnectionFactoryProvider
- CdpCredentialsProviderControllerService
- CdpOauth2AccessTokenProviderControllerService
- CiscoEmblemSyslogMessageReader
- ClouderaHiveConnectionPool
- ClouderaSchemaRegistry
- CMLLookupService
- EBCDICRecordReader
- GenericPLC4XConnectionPool
- ImpalaConnectionPool
- IPFIXReader
- JiraRecordSink
- PostgreSQLConnectionPool
- ProxyPLC4XConnectionPool
- RabbitMQJMSConnectionFactoryProvider
- RedshiftConnectionPool
- StandardPLC4XConnectionPool

Parameter providers

- CyberArkConjurParameterProvider

Download locations

You can download the Cloudera Flow Management software artifacts from the Cloudera Archive. There are different artifacts for different operating systems, standalone components, and Windows files.

Use the following tables to identify the Cloudera Flow Management repository location for your operating system and operational objectives.

Note:

You must have credentials to download Cloudera Flow Management files. Your download credential is not the same as the credential you use to access the Cloudera Support Portal.

You can get download credentials in the following ways:

- Contact your Cloudera sales representative.
- Check the Welcome email you have received for your Cloudera Flow Management account.
- File a non-technical case on the Cloudera Support Portal for the Cloudera Support team to assist you.

Table 1: RHEL/CentOS 8

File	Location
Manifest	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/parcel/manifest.json
Parcel	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/parcel/CFM-4.0.0.0-383-el8.parcel
Parcel sha file	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/parcel/CFM-4.0.0.0-383-el8.parcel.sha

File	Location
CSD	**NiFi:** https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/parcel/NIFI-2.0.0.4.0.0.0-383.jar **NiFi Registry:** https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/parcel/NIFIREGISTRY-2.0.0.4.0.0.0-383.jar

Table 2: RHEL 9

File	Location
Manifest	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat9/yum/tars/parcel/manifest.json
Parcel	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat9/yum/tars/parcel/CFM-4.0.0.0-383-el9.parcel
Parcel sha file	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat9/yum/tars/parcel/CFM-4.0.0.0-383-el9.parcel.sha
CSD	**NiFi:** https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat9/yum/tars/parcel/NIFI-2.0.0.4.0.0.0-383.jar **NiFi Registry:** https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat9/yum/tars/parcel/NIFIREGISTRY-2.0.0.4.0.0.0-383.jar

Table 3: SLES 15

File	Location
Manifest	https://archive.cloudera.com/p/cfm2/4.0.0.0/sles15/yum/tars/parcel/manifest.json
Parcel	https://archive.cloudera.com/p/cfm2/4.0.0.0/sles15/yum/tars/parcel/CFM-4.0.0.0-383-sles15.parcel
Parcel sha file	https://archive.cloudera.com/p/cfm2/4.0.0.0/sles15/yum/tars/parcel/CFM-4.0.0.0-383-sles15.parcel.sha
CSD	**NiFi:** https://archive.cloudera.com/p/cfm2/4.0.0.0/sles15/yum/tars/parcel/NIFI-2.0.0.4.0.0.0-383.jar **NiFi Registry:** https://archive.cloudera.com/p/cfm2/4.0.0.0/sles15/yum/tars/parcel/NIFIREGISTRY-2.0.0.4.0.0.0-383.jar

Table 4: Ubuntu 20

File	Location
Manifest	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu20/apt/tars/parcel/manifest.json
Parcel	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu20/apt/tars/parcel/CFM-4.0.0.0-383-focal.parcel
Parcel SHA file	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu20/apt/tars/parcel/CFM-4.0.0.0-383-focal.parcel.sha
CSD	**NiFi:** https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu20/apt/tars/parcel/NIFI-2.0.0.4.0.0.0-383.jar **NiFi Registry:** https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu20/apt/tars/parcel/NIFIREGISTRY-2.0.0.4.0.0.0-383.jar

Table 5: Ubuntu 22

File	Location
Manifest	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu22/apt/tars/parcel/manifest.json

File	Location
Parcel	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu22/apt/tars/parcel/CFM-4.0.0.0-383-jammy.parcel
Parcel SHA file	https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu22/apt/tars/parcel/CFM-4.0.0.0-383-jammy.parcel.sha
CSD	**NiFi:** https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu22/apt/tars/parcel/NIFI-2.0.0.4.0.0.0-383.jar **NiFi Registry:** https://archive.cloudera.com/p/cfm2/4.0.0.0/ubuntu22/apt/tars/parcel/NIFIREGISTRY-2.0.0.4.0.0.0-383.jar

Table 6: Standalone components (OS agnostic)

File	Location
NiFi (.tar.gz)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/cdf_extensions/cdf_extensions-1.0.0.4.0.0.0-383-source.tar.gz
NiFi (.tar.gz.sha256)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/cdf_extensions/nifi-2.0.0.4.0.0.0-383-bin.tar.gz.sha256
NiFi Registry (.tar.gz)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/nifi/nifi-registry-2.0.0.4.0.0.0-383-bin.tar.gz
NiFi Registry (.tar.gz.sha256)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/nifi/nifi-registry-2.0.0.4.0.0.0-383-bin.tar.gz.sha256
NiFi Toolkit (.tar.gz)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/nifi/nifi-toolkit-2.0.0.4.0.0.0-383-bin.tar.gz
NiFi Toolkit (.tar.gz.sha256)	https://archive.cloudera.com/p/cfm2/4.0.0.0/redhat8/yum/tars/nifi/nifi-toolkit-2.0.0.4.0.0.0-383-bin.tar.gz.sha256

Unsupported features

The following features are developed and tested by the Cloudera community but are not officially supported by Cloudera. These features 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.

Unsupported NiFi components

NARs

NiFi 1 custom NARs cannot be successfully loaded into NiFi 2. If your NiFi setup includes custom NARs, it is a requirement to update your dependencies to align with NiFi 2. This entails making the necessary adjustments and rebuilding your NARs using Java 21. The below components are not supported and should not be used anymore.

- nifi-cybersecurity-nar
- nifi-email-nar
- nifi-hive-nar
- nifi-rethinkdb-nar
- nifi-influxdb-nar
- nifi-ccda-nar
- nifi-html-nar
- nifi-ignite-nar
- nifi-tcp-nar
- nifi-riemann-nar
- nifi-spring-nar
- nifi-kite-nar

- nifi-rules-action-handler-nar
- nifi-azure-nar
- nifi-easyrules-nar
- nifi-metrics-reporting-nar
- nifi-other-graph-services-nar
- nifi-hbase_1_1_2-client-service-nar
- nifi-scripting-nar
- nifi-ambari-nar
- nifi-sql-reporting-nar
- nifi-aws-nar
- nifi-accumulo-nar
- nifi-solr-nar
- nifi-accumulo-service-nar
- nifi-datadog-nar
- nifi-atlas-nar
- nifi-beats-nar
- nifi-standard-nar
- nifi-language-translation-nar
- nifi-livy-nar
- nifi-pulsar-nar

Rules engine processors

Rules engine components and handlers are removed in NiFi 2, so the below processors are not supported and should not be used anymore.

- ActionHandlerLookup
- AlertHandler
- EasyRulesEngineProvider
- EasyRulesEngineService
- ExpressionHandler
- LogHandler
- RecordSinkHandler
- ScriptedActionHandler
- ScriptedRulesEngine

Unsupported customizations

Cloudera cannot guarantee that default NiFi processors are compatible with proprietary protocol implementations or proprietary interface extensions. For example, Cloudera supports interfaces like JMS and JDBC that are built around standards, specifications, or open protocols, but does not support customizations of those interfaces, or proprietary extensions built on top of those interfaces.

Behavioral changes

Learn about behavioral changes in Cloudera Flow Management 4.0.0.

NiFi 2.0 introduces a lot of significant changes and enhancements, including some breaking changes. It is important to familiarize yourself with the following points before migrating your existing flows.

Java 21

Java 21 is the minimum Java version required with NiFi 2.0, so make sure you have Java 21 installed on your NiFi nodes before upgrading.

Templates and XML flow definitions

The concept of templates in NiFi has been deprecated, and the XML templates are stored in memory in NiFi as well as in the persisted flow definition.

Additionally, flow.xml.gz no longer exists, only flow.json.gz can be used in NiFi clusters for defining flows in the canvas.

If you have templates, export those templates as JSON definitions or version the templates into a NiFi Registry instance. The best practice is to use a NiFi Registry in combination with NiFi when it comes to version control and share / reuse flow definitions.

Custom components / NARs

Although not certain, it is very likely that a custom NAR designed for NiFi 1 will not be successfully loaded into NiFi 2. If your NiFi setup includes custom components or NARs, it is a requirement to update your dependencies to align with NiFi 2. This entails making the necessary adjustments and rebuilding your NARs using Java 21.

Variables replaced by parameters

Variables and the variable registry have been removed from NiFi. Only Parameter contexts and parameters are available for use going forward. In future Cloudera Flow Management releases, tooling will be provided to help with the conversion of variables to parameters. In the meantime, this conversion should be done manually when migrating flows to NiFi 2. Any variables left will simply be ignored when loading the flow definition.

Event driven thread pool no longer exists

The event driven scheduling strategy was an option available on some processors. This was an experimental feature in NiFi and did not prove to bring any significant performance improvements. The event driven thread pool has been removed, leaving only the time driven thread pool available. Any components previously configured using the event driven scheduling strategy should be switched to the time driven scheduling strategy.

Removed languages in scripted components

In NiFi 2.0, support for certain languages in scripted components has been removed. The affected languages are: ECMAScript, Lua, Ruby, and Python. It is recommended to switch to Groovy or to leverage the new Python API feature for developing processors.

Removed components and replacement options

The following list contains the list of the components that have been removed between clusters based on NiFi 1.26 and clusters based on NiFi 2.0, along with the recommended alternatives where available.

- Processors
  - Base64EncodeContent => EncodeContent
  - CompareFuzzyHash => no replacement
  - ConsumeEWS => no replacement
  - ConsumeKafka_1_0 => ConsumeKafka_2_6
  - ConsumeKafka_2_0 => ConsumeKafka_2_6
  - ConsumeKafkaRecord_1_0 => ConsumeKafkaRecord_2_6
  - ConsumeKafkaRecord_2_0 => ConsumeKafkaRecord_2_6
  - ConvertAvroSchema => ConvertRecord
  - ConvertAvroToORC => no replacement
  - ConvertCSVToAvro => ConvertRecord
  - ConvertExcelToCSVProcessor => ConvertRecord with ExcelReader
  - ConvertJSONToAvro => ConvertRecord
  - CryptographicHashAttribute => UpdateAttribute
  - DeleteAzureBlobStorage => DeleteAzureBlobStorage_v12
  - DeleteRethinkDB => no replacement
  - EncryptContent => EncryptContentAge or EncryptContentPGP
  - ExecuteInfluxDBQuery => use Influx Data NARs for NiFi
  - ExtractCCDAAttributes => no replacement
  - FetchAzureBlobStorage => FetchAzureBlobStorage_v12
  - FetchElasticsearchHttp => GetElasticsearch
  - FuzzyHashContent => no replacement
  - GetAzureQueueStorage => GetAzureQueueStorage_v12
  - GetHTMLElement => no replacement
  - GetHTTP => InvokeHTTP
  - GetIgniteCache => no replacement
  - GetJMSQueue => ConsumeJMS
  - GetJMSTopic => ConsumeJMS
  - GetRethinkDB => no replacement
  - GetTCP => no replacement
  - GetTwitter => ConsumeTwitter
  - HashAttribute => CryptographicHashAttribute
  - HashContent => CryptographicHashContent
  - InferAvroSchema => ExtractRecordSchema
  - ListAzureBlobStorage => ListAzureBlobStorage_v12
  - ModifyHTMLElement => no replacement
  - PostHTTP => InvokeHTTP
  - PostSlack => PublishSlack
  - PublishKafka_1_0 => PublishKafka_2_6
  - PublishKafka_2_0 => PublishKafka_2_6
  - PublishKafkaRecord_1_0 => PublishKafkaRecord_2_6
  - PublishKafkaRecord_2_0 => PublishKafkaRecord_2_6
  - PutAzureBlobStorage => PutAzureBlobStorage_v12
  - PutAzureQueueStorage => PutAzureQueueStorage_v12
  - PutBigQueryBatch => PutBigQuery
  - PutBigQueryStreaming => PutBigQuery
  - PutElasticsearchHttp => PutElasticsearchJson
  - PutElasticsearchHttpRecord => PutElasticsearchRecord
  - PutHiveQL => PutClouderaHiveQL
  - PutHiveStreaming => PutClouderaHiveStreaming

    - PutHTMLElement => no replacement
    - PutIgniteCache => no replacement
    - PutInfluxDB => use Influx Data NARs for NiFi
    - PutJMS => PublishJMS
    - PutRethinkDB => no replacement
    - PutRiemann => no replacement
    - PutSlack => PublishSlack
    - QueryElasticsearchHttp => PaginatedJsonQueryElasticsearch
    - ScrollElasticsearchHttp => SearchElasticsearch
    - SelectHiveQL => SelectClouderaHiveQL
    - SpringContextProcessor => no replacement
    - StoreInKiteDataset => no replacement
    - UpdateHiveTable => UpdateClouderaHiveTable
- Controller services
    - ActionHandlerLookup => no replacement
    - AlertHandler => no replacement
    - AzureStorageCredentialsControllerService => AzureStorageCredentialsControllerService_v12
    - AzureStorageCredentialsControllerServiceLookup => AzureStorageCredentialsControllerServiceLookup_v12
    - AzureStorageEmulatorCredentialsControllerService => no replacement
    - EasyRulesEngineProvider => no replacement
    - EasyRulesEngineService => no replacement
    - ExpressionHandler => no replacement
    - GraphiteMetricReporterService => no replacement
    - GremlinClientService => no replacement
    - HBase_1_1_2_ClientMapCacheService => HBase_2_ClientMapCacheService
    - HBase_1_1_2_ClientService => HBase_2_ClientService
    - HBase_1_1_2_ListLookupService => no replacement
    - HBase_1_1_2_RecordLookupService => HBase_2_RecordLookupService
    - HiveConnectionPool => ClouderaHiveConnectionPool
    - HortonworksSchemaRegistry => ClouderaSchemaRegistry
    - KafkaRecordSink_1_0 => KafkaRecordSink_2_6
    - KafkaRecordSink_2_0 => KafkaRecordSink_2_6
    - LogHandler => no replacement
    - OAuth2TokenProviderImpl => StandardOauth2AccessTokenProvider
    - OpenCypherClientService => no replacement
    - RecordSinkHandler => no replacement
    - ScriptedActionHandler => no replacement
    - ScriptedRulesEngine => no replacement
- Reporting tasks
    - AmbariReportingTask => no replacement
    - MetricsEventReportingTask => no replacement
    - MetricsReportingTask => no replacement

- Components with new coordinates
  - InvokeGRPC => moved into nifi-cdf-grpc-nar
  - ListenGRPC => moved into nifi-cdf-grpc-nar
  - KerberosKeytabUserService => moved into nifi-kerberos-user-service-nar
  - KerberosPasswordUserService => moved into nifi-kerberos-user-service-nar
  - KerberosTicketCacheUserService => moved into nifi-kerberos-user-service-nar

  Tooling will be provided in upcoming Cloudera Flow Management releases to automatically handle these changes. Currently, two options are available:

  - Manually edit the flow.json.gz file to update the coordinates of the impacted components.
  - Make the changes after the flow is imported in NiFi 2.0 by replacing the ghost components with the new implementations for each instance of the components listed above.
- Pulsar components

  All Pulsar components have been temporarily removed. They will be reintroduced in an upcoming release. In the meantime, you can download the NARs from a public Maven repository and deploy them as custom NARs.

  - nifi-pulsar-nar
  - nifi-pulsar-client-service-nar

Known issues in Cloudera Flow Management

Review the list of known issues in Cloudera Flow Management 4.0.0 [Technical Preview].

Known issues

The following known issues remain unresolved in Cloudera Flow Management 4.0.0 [Technical Preview].