

Supported Sources, Sinks, and Channels [\(#topic 12 9\)](#)

The following tables list the only currently-supported sources, sinks, and channels. For more information, including information on developing custom components, see the documents listed under [Viewing the Flume Documentation \(cdh ig flume doc.html#topic 12 11\)](#).

Sources [\(#topic 12 9 1\)](#)

Type	Description	Implementati
avro	Avro Netty RPC event source. Listens on Avro port and receives events from external Avro streams.	AvroSource
netcat	Netcat style TCP event source. Listens on a given port and turns each line of text into an event.	NetcatSource
seq	Monotonically incrementing sequence generator event source	SequenceGene
exec	Execute a long-lived Unix process and read from stdout.	ExecSource
syslogtcp	Reads syslog data and generates flume events. Creates a new event for a string of characters separated by carriage return (\n).	SyslogTcpSou
syslogudp	Reads syslog data and generates flume events. Treats an entire message as a single event.	SyslogUDPSou
org.apache.flume.source.avroLegacy. AvroLegacySource	Allows the Flume 1.x agent to receive events from Flume 0.9.4 agents over avro rpc.	AvroLegacySo
org.apache.flume.source.thriftLegacy. ThriftLegacySource	Allows the Flume 1.x agent to receive events from Flume 0.9.4 agents over thrift rpc.	ThriftLegacyS
org.apache.flume.source.StressSource	Mainly for testing purposes. Not meant for production use. Serves as a continuous source of events where each event has the same payload.	StressSource
org.apache.flume.source.scribe. ScribeSource	Scribe event source. Listens on Scribe port and receives events from Scribe.	ScribeSource
multiport_syslogtcp	Multi-port capable version of the SyslogTcpSource.	MultiportSysl
spooldir	Used for ingesting data by placing files to be ingested into a "spooling" directory on disk.	SpoolDirectory

Type	Description	Implementation
http	Accepts Flume events by HTTP POST and GET. GET should be used for experimentation only.	HTTPSource
org.apache.flume.source.jms.JMSSource	Reads messages from a JMS destination such as a queue or topic.	JMSSource
org.apache.flume.agent.embedded.EmbeddedSource	Used only by the Flume embedded agent. See Flume Developer Guide (https://archive.cloudera.com/cdh5/cdh/5/flume-ng/FlumeDeveloperGuide.html) for more details.	EmbeddedSource
org.apache.flume.source.kafka.KafkaSource	Streams data from Kafka to Hadoop or from any Flume source to Kafka.	KafkaSource
org.apache.flume.source.taildir.TaildirSource	<p>Watches specified files, and tails them in near real-time when it detects appends to these files.</p> <ul style="list-style-type: none">• This source is reliable and does not miss data, even when the tailing files rotate.• It periodically writes the last read position of each file in a position file using the JSON format.• If Flume is stopped or down for some reason, it can restart tailing from the position written in the existing position file.• It can add event headers to each tailing file group.	TaildirSource



Sinks ([#topic 12 9 2](#))

Type	Description	Implementation
logger	Log events at INFO level using configured logging subsystem (log4j by default)	LoggerSink
avro	Sink that invokes a pre-defined Avro protocol method for all events it receives (when paired with an avro source, forms tiered collection)	AvroSink
hdfs	Writes all events received to HDFS (with support for rolling, bucketing, HDFS-200 append, and more)	HDFS Sink
file_roll	Writes all events received to one or more files.	File Sink
org.apache.flume.hbase.HBaseSink	A simple sink that reads events from a channel and writes them synchronously to HBase. The AsyncHBaseSink is recommended. See Importing Data Into HBase (admin hbase import.html#xd_583c10bfdbd326ba-69adf108-1492ec0ce48--7f83) .	HBase Sink

Type	Description	
org.apache.flume.sink.hbase.AsyncHBaseSink	A simple sink that reads events from a channel and writes them asynchronously to HBase. This is the recommended HBase sink, but note that it does not support Kerberos. See Importing Data Into HBase (admin_hbase_import.html#xd_583c10bfdbd326ba-69adf108-1492ecoce48--7f83) .	
org.apache.flume.sink.solr.morphline.MorphlineSolrSink	Extracts and transforms data from Flume events, and loads it into Apache Solr servers. See the section on MorphlineSolrSink in the Flume User Guide listed under Viewing the Flume Documentation (cdh_ig_flume_doc.html#topic_12_11) .	
org.apache.flume.sink.kafka.KafkaSink	Used to send data to Kafka from a Flume source. You can use the Kafka sink in addition to Flume sinks such as HBase or HDFS.	



Channels [\(#topic_12_9_3\)](#)

Type	Description	Implementation Class
memory	In-memory, fast, non-durable event transport	MemoryChannel
jdbc	JDBC-based, durable event transport (Derby-based)	JDBCChannel
file	File-based, durable event transport	FileChannel
org.apache.flume.channel.kafka.KafkaChannel	Use the Kafka channel: <ul style="list-style-type: none">• To write to Hadoop directly from Kafka without using a source.• To write to Kafka directly from Flume sources without additional buffering.• As a reliable and highly available channel for any source/sink combination.	KafkaChannel

Providing for Disk Space Usage [\(#topic_12_9_3_section_qjs_h5v_ls\)](#)

It's important to provide plenty of disk space for any Flume File Channel. The largest consumers of disk space in the File Channel are the data logs. You can configure the File Channel to write these logs to multiple data directories. The following space will be consumed by default in each data directory:

- Current log file (up to 2 GB)
- Last log file (up to 2 GB)
- Pending delete log file (up to 2 GB)

Events in the queue could cause many more log files to be written, each of them up 2 GB in size by default.

You can configure both the maximum log file size (`MaxFileSize`) and the directories the logs will be written to (`DataDirs`) when you configure the File Channel; see the File Channel section of the [Flume User Guide \(http://flume.apache.org/FlumeUserGuide.html#file-channel\)](http://flume.apache.org/FlumeUserGuide.html#file-channel) for details.

Categories: [Data Analysts \(../categories/hub_data_analysts.html\)](#) | [Developers \(../categories/hub_developers.html\)](#) | [ETL \(../categories/hub_etl.html\)](#) | [Flume \(../categories/hub_flume.html\)](#) | [Ingest \(../categories/hub_ingest.html\)](#) | [All Categories \(../categories/hub.html\)](#)

- [About Cloudera \(http://www.cloudera.com/about-cloudera.html\)](http://www.cloudera.com/about-cloudera.html)
- [Resources \(http://www.cloudera.com/resources.html\)](http://www.cloudera.com/resources.html)
- [Contact \(http://www.cloudera.com/contact-us.html\)](http://www.cloudera.com/contact-us.html)
- [Careers \(http://www.cloudera.com/about-cloudera/careers.html\)](http://www.cloudera.com/about-cloudera/careers.html)
- [Press \(http://www.cloudera.com/about-cloudera/press-center.html\)](http://www.cloudera.com/about-cloudera/press-center.html)
- [Documentation \(http://www.cloudera.com/documentation.html\)](http://www.cloudera.com/documentation.html)

United States: +1 888 789 1488
Outside the US: +1 650 362 0488

© 2017 Cloudera, Inc. All rights reserved. [Apache Hadoop \(http://hadoop.apache.org\)](http://hadoop.apache.org) and associated open source project names are trademarks of the [Apache Software Foundation \(http://apache.org\)](http://apache.org). For a complete list of trademarks, [click here. \(http://www.cloudera.com/legal/terms-and-conditions.html#trademarks\)](http://www.cloudera.com/legal/terms-and-conditions.html#trademarks)

- [_ \(https://www.linkedin.com/company/cloudera\)](https://www.linkedin.com/company/cloudera)
- [_ \(https://www.facebook.com/cloudera\)](https://www.facebook.com/cloudera)
- [_ \(https://twitter.com/cloudera\)](https://twitter.com/cloudera)
- [_ \(http://www.cloudera.com/contact-us.html\)](http://www.cloudera.com/contact-us.html)

[Terms & Conditions \(http://www.cloudera.com/legal/terms-and-conditions.html\)](http://www.cloudera.com/legal/terms-and-conditions.html) | [Privacy Policy \(http://www.cloudera.com/legal/policies.html\)](http://www.cloudera.com/legal/policies.html)

Page generated June 15, 2017.