

Introduction

Overview

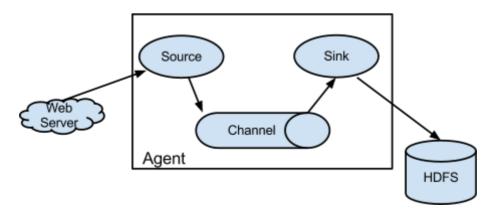
Apache Flume is a distributed, reliable, and available system for efficiently collecting, aggregating and moving large amounts of log data from many different sources to a centralized data store.

Apache Flume is a top-level project at the Apache Software Foundation. There are currently two release code lines available, versions 0.9.x and 1.x. This documentation applies to the 1.x codeline. For the 0.9.x codeline, please see the Flume 0.9.x Developer Guide.

Architecture

Data flow model

An Event is a unit of data that flows through a Flume agent. The Event flows from source to Channel to sink, and is represented by an implementation of the Event interface. An Event carries a payload (byte array) that is accompanied by an optional set of headers (string attributes). A Flume agent is a process (JVM) that hosts the components that allow Events to flow from an external source to a external destination.



A source consumes Event's having a specific format, and those Event's are delivered to the source by an external source like a web server. For example, an Avrosource can be used to receive Avro Event's from clients or from other Flume agents in the flow. When a source receives an Event, it stores it into one or more channels. The channel is a passive store that holds the Event until that Event is consumed by a

sink. One type of Channel available in Flume is the FileChannel which uses the local filesystem as its backing store. A sink is responsible for removing an Event from the Channel and putting it into an external repository like HDFS (in the case of an HDFSEventsink) or forwarding it to the source at the next hop of the flow. The source and sink within the given agent run asynchronously with the Events staged in the Channel.

Reliability

An Event is staged in a Flume agent's channel. Then it's the sink's responsibility to deliver the Event to the next agent or terminal repository (like HDFS) in the flow. The sink removes an Event from the Channel only after the Event is stored into the channel of the next agent or stored in the terminal repository. This is how the single-hop message delivery semantics in Flume provide end-to-end reliability of the flow. Flume uses a transactional approach to guarantee the reliable delivery of the Events. The sources and sinks encapsulate the storage/retrieval of the Events in a Transaction provided by the Channel. This ensures that the set of Events are reliably passed from point to point in the flow. In the case of a multi-hop flow, the sink from the previous hop and the source of the next hop both have their Transactions open to ensure that the Event data is safely stored in the Channel of the next hop.

Building Flume

Getting the source

Check-out the code using Git. Click here for the git repository root.

The Flume 1.x development happens under the branch "trunk" so this command line can be used:

git clone https://git-wip-us.apache.org/repos/asf/flume.git

Compile/test Flume

The Flume build is mavenized. You can compile Flume using the standard Maven commands:

- 1. Compile only: mvn clean compile
- 2. Compile and run unit tests: mvn clean test
- 3. Run individual test(s): mvn clean test -Dtest=<Test1>,<Test2>,... -DfailIfNoTests=false
- 4. Create tarball package: mvn clean install
- 5. Create tarball package (skip unit tests): mvn clean install -DskipTests

Please note that Flume builds requires that the Google Protocol Buffers compiler be in the path. You can download and install it by following the instructions here.

Developing custom components

Client

The client operates at the point of origin of events and delivers them to a Flume agent. Clients typically operate in the process space of the application they are consuming data from. Flume currently supports Avro, log4j, syslog, and Http POST (with a JSON body) as ways to transfer data from a external source. Additionally, there's an Execsource that can consume the output of a local process as input to Flume.

It's quite possible to have a use case where these existing options are not sufficient. In this case you can build a custom mechanism to send data to Flume. There are two ways of achieving this. The first option is to create a custom client that communicates with one of Flume's existing sources like Avrosource or SyslogTcpSource. Here the client should convert its data into messages understood by these Flume sources. The other option is to write a custom Flume source that directly talks with your existing client application using some IPC or RPC protocol, and then converts the client data into Flume Events to be sent downstream. Note that all events stored within the channel of a Flume agent must exist as Flume Events.

Client SDK

Though Flume contains a number of built-in mechanisms (i.e. sources) to ingest data, often one wants the ability to communicate with Flume directly from a custom application. The Flume Client SDK is a library that enables applications to connect to Flume and send data into Flume's data flow over RPC.

RPC client interface

An implementation of Flume's RpcClient interface encapsulates the RPC mechanism supported by Flume. The user's application can simply call the Flume Client SDK's append(Event) or appendBatch(List<Event>) to send data and not worry about the underlying message exchange details. The user can provide the required Event arg by either directly implementing the Event interface, by using a convenience implementation such as the SimpleEvent class, or by using EventBuilder's overloaded withBody() static helper methods.

RPC clients - Avro and Thrift

As of Flume 1.4.0, Avro is the default RPC protocol. The NettyAvroRpcClient and ThriftRpcClient implement the RpcClient interface. The client needs to create this object with the host and port of the target Flume agent, and can then use the RpcClient to send data into the agent. The following example shows how to use the Flume Client SDK API within a user's data-generating application:

```
import org.apache.flume.Event;
import org.apache.flume.EventDeliveryException;
import org.apache.flume.api.RpcClient;
import org.apache.flume.api.RpcClientFactory;
import org.apache.flume.event.EventBuilder;
import java.nio.charset.Charset;
public class MyApp {
  public static void main(String[] args) {
    MyRpcClientFacade client = new MyRpcClientFacade();
    // Initialize client with the remote Flume agent's host and port
    client.init("host.example.org", 41414);
    // Send 10 events to the remote Flume agent. That agent should be
    // configured to listen with an AvroSource.
    String sampleData = "Hello Flume!";
    for (int i = 0; i < 10; i++) {
      client.sendDataToFlume(sampleData);
    client.cleanUp();
}
class MyRpcClientFacade {
  private RpcClient client;
```

```
private String hostname;
  private int port;
  public void init(String hostname, int port) {
    // Setup the RPC connection
    this.hostname = hostname;
    this.port = port;
    this.client = RpcClientFactory.getDefaultInstance(hostname, port);
    // Use the following method to create a thrift client (instead of the above line):
    // this.client = RpcClientFactory.getThriftInstance(hostname, port);
  }
  public void sendDataToFlume(String data) {
    // Create a Flume Event object that encapsulates the sample data
    Event event = EventBuilder.withBody(data, Charset.forName("UTF-8"));
    // Send the event
    try {
      client.append(event);
    } catch (EventDeliveryException e) {
      // clean up and recreate the client
      client.close();
      client = null;
      client = RpcClientFactory.getDefaultInstance(hostname, port);
      // Use the following method to create a thrift client (instead of the above line
      // this.client = RpcClientFactory.getThriftInstance(hostname, port);
    }
  }
  public void cleanUp() {
    // Close the RPC connection
    client.close();
  }
}
```

The remote Flume agent needs to have an AvroSource (or a ThriftSource if you are using a Thrift client) listening on some port. Below is an example Flume agent configuration that's waiting for a connection from MyApp:

```
al.channels = c1
al.sources = r1
al.sinks = k1

al.channels.cl.type = memory

al.sources.rl.channels = c1
al.sources.rl.type = avro
# For using a thrift source set the following instead of the above line.
# al.source.rl.type = thrift
al.sources.rl.bind = 0.0.0.0
al.sources.rl.port = 41414

al.sinks.kl.channel = c1
al.sinks.kl.type = logger
```

For more flexibility, the default Flume client implementations (NettyAvroRpcClient and ThriftRpcClient) can be configured with these properties:

```
client.type = default (for avro) or thrift (for thrift)
hosts = h1  # default client accepts only 1 host
# (additional hosts will be ignored)
hosts.h1 = host1.example.org:41414 # host and port must both be specified
```

```
# (neither has a default)
batch-size = 100  # Must be >=1 (default: 100)

connect-timeout = 20000  # Must be >=1000 (default: 20000)

request-timeout = 20000  # Must be >=1000 (default: 20000)
```

Secure RPC client - Thrift

As of Flume 1.6.0, Thrift source and sink supports kerberos based authentication. The client needs to use the getThriftInstance method of secureRpcClientFactory to get hold of a secureThriftRpcClient. SecureThriftRpcClient extends ThriftRpcClient which implements the RpcClient interface. The kerberos authentication module resides in flume-ng-auth module which is required in classpath, when using the secureRpcClientFactory. Both the client principal and the client keytab should be passed in as parameters through the properties and they reflect the credentials of the client to authenticate against the kerberos KDC. In addition, the server principal of the destination Thrift source to which this client is connecting to, should also be provided. The following example shows how to use the SecureRpcClientFactory within a user's data-generating application:

```
import org.apache.flume.Event;
import org.apache.flume.EventDeliveryException;
import org.apache.flume.event.EventBuilder;
import org.apache.flume.api.SecureRpcClientFactory;
import org.apache.flume.api.RpcClientConfigurationConstants;
import org.apache.flume.api.RpcClient;
import java.nio.charset.Charset;
import java.util.Properties;
public class MyApp {
  public static void main(String[] args) {
    MySecureRpcClientFacade client = new MySecureRpcClientFacade();
    // Initialize client with the remote Flume agent's host, port
    Properties props = new Properties();
    props.setProperty(RpcClientConfigurationConstants.CONFIG_CLIENT_TYPE, "thrift");
    props.setProperty("hosts", "h1");
    props.setProperty("hosts.h1", "client.example.org"+":"+ String.valueOf(41414));
    // Initialize client with the kerberos authentication related properties
    props.setProperty("kerberos", "true");
    props.setProperty("client-principal", "flumeclient/client.example.org@EXAMPLE.ORG"
props.setProperty("client-keytab", "/tmp/flumeclient.keytab");
    props.setProperty("server-principal", "flume/server.example.org@EXAMPLE.ORG");
    client.init(props);
    // Send 10 events to the remote Flume agent. That agent should be
    // configured to listen with an AvroSource.
    String sampleData = "Hello Flume!";
    for (int i = 0; i < 10; i++) {
      client.sendDataToFlume(sampleData);
    client.cleanUp();
}
class MySecureRpcClientFacade {
  private RpcClient client;
  private Properties properties;
  public void init(Properties properties) {
    // Setup the RPC connection
    this.properties = properties;
    // Create the ThriftSecureRpcClient instance by using SecureRpcClientFactory
```

```
this.client = SecureRpcClientFactory.getThriftInstance(properties);
  }
  public void sendDataToFlume(String data) {
    // Create a Flume Event object that encapsulates the sample data
    Event event = EventBuilder.withBody(data, Charset.forName("UTF-8"));
    // Send the event
    try {
      client.append(event);
    } catch (EventDeliveryException e) {
      // clean up and recreate the client
      client.close();
      client = null;
      client = SecureRpcClientFactory.getThriftInstance(properties);
    }
  }
  public void cleanUp() {
    // Close the RPC connection
    client.close();
  }
}
```

The remote Thriftsource should be started in kerberos mode. Below is an example Flume agent configuration that's waiting for a connection from MyApp:

```
al.channels = c1
al.sources = r1
al.sinks = k1

al.channels.cl.type = memory

al.sources.rl.channels = c1
al.sources.rl.type = thrift
al.sources.rl.bind = 0.0.0.0
al.sources.rl.port = 41414
al.sources.rl.kerberos = true
al.sources.rl.agent-principal = flume/server.example.org@EXAMPLE.ORG
al.sources.rl.agent-keytab = /tmp/flume.keytab
al.sinks.kl.channel = c1
al.sinks.kl.type = logger
```

Failover Client

This class wraps the default Avro RPC client to provide failover handling capability to clients. This takes a whitespace-separated list of <host>:<port> representing the Flume agents that make-up a failover group. The Failover RPC Client currently does not support thrift. If there's a communication error with the currently selected host (i.e. agent) agent, then the failover client automatically fails-over to the next host in the list. For example:

```
// Setup properties for the failover
Properties props = new Properties();
props.put("client.type", "default_failover");

// List of hosts (space-separated list of user-chosen host aliases)
props.put("hosts", "h1 h2 h3");

// host/port pair for each host alias
String host1 = "host1.example.org:41414";
String host2 = "host2.example.org:41414";
String host3 = "host3.example.org:41414";
```

```
props.put("hosts.h1", host1);
props.put("hosts.h2", host2);
props.put("hosts.h3", host3);

// create the client with failover properties
RpcClient client = RpcClientFactory.getInstance(props);
```

For more flexibility, the failover Flume client implementation (FailoverRpcClient) can be configured with these properties:

```
client.type = default failover
hosts = h1 h2 h3
                                     # at least one is required, but 2 or
                                     # more makes better sense
hosts.h1 = host1.example.org:41414
hosts.h2 = host2.example.org:41414
hosts.h3 = host3.example.org:41414
max-attempts = 3
                                     # Must be >=0 (default: number of hosts
                                     # specified, 3 in this case). A '0'
                                     # value doesn't make much sense because
                                     # it will just cause an append call to
                                     # immmediately fail. A '1' value means
                                     # that the failover client will try only
                                     # once to send the Event, and if it
                                     # fails then there will be no failover
                                     # to a second client, so this value
                                     # causes the failover client to
                                     # degenerate into just a default client.
                                     # It makes sense to set this value to at
                                     # least the number of hosts that you
                                     # specified.
batch-size = 100
                                     # Must be >=1 (default: 100)
                                     # Must be >=1000 (default: 20000)
connect-timeout = 20000
                                     # Must be >=1000 (default: 20000)
request-timeout = 20000
```

LoadBalancing RPC client

The Flume Client SDK also supports an RpcClient which load-balances among multiple hosts. This type of client takes a whitespace-separated list of <host>:<port> representing the Flume agents that make-up a load-balancing group. This client can be configured with a load balancing strategy that either randomly selects one of the configured hosts, or selects a host in a round-robin fashion. You can also specify your own custom class that implements the LoadBalancingRpcClient\$HostSelector interface so that a custom selection order is used. In that case, the FQCN of the custom class needs to be specified as the value of the host-selector property. The LoadBalancing RPC Client currently does not support thrift.

If backoff is enabled then the client will temporarily blacklist hosts that fail, causing them to be excluded from being selected as a failover host until a given timeout. When the timeout elapses, if the host is still unresponsive then this is considered a sequential failure, and the timeout is increased exponentially to avoid potentially getting stuck in long waits on unresponsive hosts.

The maximum backoff time can be configured by setting maxBackoff (in milliseconds). The maxBackoff default is 30 seconds (specified in the orderSelector class that's the superclass of both load balancing strategies). The backoff timeout will increase exponentially with each sequential failure up to the

maximum possible backoff timeout. The maximum possible backoff is limited to 65536 seconds (about 18.2 hours). For example:

```
// Setup properties for the load balancing
Properties props = new Properties();
props.put("client.type", "default loadbalance");
// List of hosts (space-separated list of user-chosen host aliases)
props.put("hosts", "h1 h2 h3");
// host/port pair for each host alias
String host1 = "host1.example.org:41414";
String host2 = "host2.example.org:41414";
String host3 = "host3.example.org:41414";
props.put("hosts.h1", host1);
props.put("hosts.h2", host2);
props.put("hosts.h3", host3);
props.put("host-selector", "random"); // For random host selection
// props.put("host-selector", "round_robin"); // For round-robin host
                                              // selection
props.put("backoff", "true"); // Disabled by default.
props.put("maxBackoff", "10000"); // Defaults 0, which effectively
                                  // becomes 30000 ms
// Create the client with load balancing properties
RpcClient client = RpcClientFactory.getInstance(props);
```

For more flexibility, the load-balancing Flume client implementation (LoadBalancingRpcClient) can be configured with these properties:

```
client.type = default_loadbalance
hosts = h1 h2 h3
                                    # At least 2 hosts are required
hosts.h1 = host1.example.org:41414
hosts.h2 = host2.example.org:41414
hosts.h3 = host3.example.org:41414
                                     # Specifies whether the client should
backoff = false
                                     # back-off from (i.e. temporarily
                                     # blacklist) a failed host
                                     # (default: false).
maxBackoff = 0
                                     # Max timeout in millis that a will
                                     # remain inactive due to a previous
                                     # failure with that host (default: 0,
                                     # which effectively becomes 30000)
host-selector = round_robin
                                    # The host selection strategy used
                                     # when load-balancing among hosts
                                     # (default: round_robin).
                                     # Other values are include "random"
                                     # or the FQCN of a custom class
                                     # that implements
                                     # LoadBalancingRpcClient$HostSelector
batch-size = 100
                                    # Must be >=1 (default: 100)
connect-timeout = 20000
                                    # Must be >=1000 (default: 20000)
request-timeout = 20000
                                     # Must be >=1000 (default: 20000)
```

Flume has an embedded agent api which allows users to embed an agent in their application. This agent is meant to be lightweight and as such not all sources, sinks, and channels are allowed. Specifically the source used is a special embedded source and events should be send to the source via the put, putAll methods on the EmbeddedAgent object. Only File Channel and Memory Channel are allowed as channels while Avro Sink is the only supported sink. Interceptors are also supported by the embedded agent.

Note: The embedded agent has a dependency on hadoop-core.jar.

Configuration of an Embedded Agent is similar to configuration of a full Agent. The following is an exhaustive list of configration options:

Required properties are in **bold**.

Property Name	Default	Description
source.type	embedded	The only available source is the embedded source.
channel.type	_	Either memory or file which correspond to MemoryChannel and FileChannel respectively.
channel.*	-	Configuration options for the channel type requested, see MemoryChannel or FileChannel user guide for an exhaustive list.
sinks	_	List of sink names
sink.type	_	Property name must match a name in the list of sinks. Value must be avro
sink.*	_	Configuration options for the sink. See AvroSink user guide for an exhaustive list, however note AvroSink requires at least hostname and port.
processor.type	-	Either failover or load_balance which correspond to FailoverSinksProcessor and LoadBalancingSinkProcessor respectively.
processor.*	-	Configuration options for the sink processor selected. See FailoverSinksProcessor and LoadBalancingSinkProcessor user guide for an exhaustive list.
source.interceptors	_	Space-separated list of interceptors
source.interceptors.*	_	Configuration options for individual interceptors specified in the source.interceptors property

Below is an example of how to use the agent:

```
Map<String, String> properties = new HashMap<String, String>();
properties.put("channel.type", "memory");
properties.put("channel.capacity", "200");
properties.put("sinks", "sink1 sink2");
properties.put("sink1.type", "avro");
properties.put("sink2.type", "avro");
properties.put("sink1.hostname", "collector1.apache.org");
properties.put("sink1.port", "5564");
properties.put("sink2.hostname", "collector2.apache.org");
properties.put("sink2.port", "5565");
properties.put("processor.type", "load_balance");
properties.put("source.interceptors", "i1");
properties.put("source.interceptors.il.type", "static
properties.put("source.interceptors.il.key", "key1");
properties.put("source.interceptors.il.value", "value1");
EmbeddedAgent agent = new EmbeddedAgent("myagent");
```

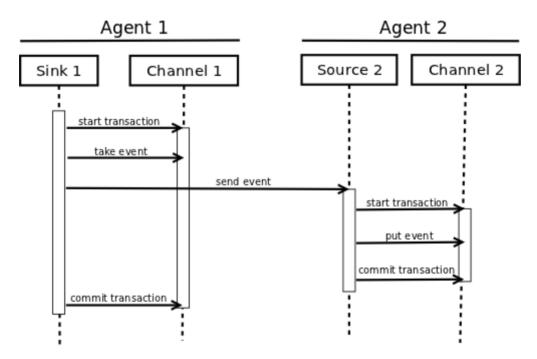
```
agent.configure(properties);
agent.start();

List<Event> events = Lists.newArrayList();

events.add(event);
events.add(event);
events.add(event);
events.add(event);
agent.putAll(events);
...
agent.stop();
```

Transaction interface

The Transaction interface is the basis of reliability for Flume. All the major components (ie. sources, sinks and channels) must use a Flume Transaction.



A Transaction is implemented within a Channel implementation. Each source and Sink that is connected to a Channel must obtain a Transaction object. The sources use a ChannelProcessor to manage the Transactions, the Sinks manage them explicitly via their configured Channel. The operation to stage an Event (put it into a Channel) or extract an Event (take it out of a Channel) is done inside an active Transaction. For example:

```
txn.rollback();

// Log exception, handle individual exceptions as needed

// re-throw all Errors
if (t instanceof Error) {
   throw (Error)t;
}

finally {
   txn.close();
}
```

Here we get hold of a Transaction from a Channel. After begin() returns, the Transaction is now active/open and the Event is then put into the Channel. If the put is successful, then the Transaction is committed and closed.

Sink

The purpose of a sink to extract Events from the Channel and forward them to the next Flume Agent in the flow or store them in an external repository. A sink is associated with exactly one Channels, as configured in the Flume properties file. There's one sinkRunner instance associated with every configured sink, and when the Flume framework calls sinkRunner.start(), a new thread is created to drive the sink (using sinkRunner.PollingRunner as the thread's Runnable). This thread manages the sink's lifecycle. The sink needs to implement the start() and stop() methods that are part of the LifecycleAware interface. The sink.start() method should initialize the sink and bring it to a state where it can forward the Events to its next destination. The sink.process() method should do the core processing of extracting the Event from the Channel and forwarding it. The sink.stop() method should do the necessary cleanup (e.g. releasing resources). The sink implementation also needs to implement the configurable interface for processing its own configuration settings. For example:

```
public class MySink extends AbstractSink implements Configurable {
  private String myProp;
  @Override
  public void configure(Context context) {
    String myProp = context.getString("myProp", "defaultValue");
    // Process the myProp value (e.g. validation)
    // Store myProp for later retrieval by process() method
    this.myProp = myProp;
  }
  @Override
  public void start() {
    // Initialize the connection to the external repository (e.g. HDFS) that
    // this Sink will forward Events to ...
  }
  @Override
  public void stop () {
    // Disconnect from the external respository and do any
    // additional cleanup (e.g. releasing resources or nulling-out
    // field values) ..
  }
  @Override
  public Status process() throws EventDeliveryException {
    Status status = null;
    // Start transaction
```

```
Channel ch = getChannel();
    Transaction txn = ch.getTransaction();
    txn.begin();
    try {
      // This try clause includes whatever Channel operations you want to do
      Event event = ch.take();
      // Send the Event to the external repository.
      // storeSomeData(e);
      txn.commit();
      status = Status.READY;
    } catch (Throwable t) {
      txn.rollback();
      // Log exception, handle individual exceptions as needed
      status = Status.BACKOFF;
      // re-throw all Errors
      if (t instanceof Error) {
        throw (Error)t;
    }
   return status;
}
```

Source

The purpose of a source is to receive data from an external client and store it into the configured channels. A source can get an instance of its own ChannelProcessor to process an Event, committed within a Channel local transaction, in serial. In the case of an exception, required Channels will propagate the exception, all Channels will rollback their transaction, but events processed previously on other Channels will remain committed.

Similar to the SinkRunner.PollingRunner Runnable, there's a PollingRunner Runnable that executes on a thread created when the Flume framework calls PollableSourceRunner.start(). Each configured PollableSource is associated with its own thread that runs a PollingRunner. This thread manages the PollableSource's lifecycle, such as starting and stopping. A PollableSource implementation must implement the start() and stop() methods that are declared in the LifecycleAware interface. The runner of a PollableSource invokes that source's process() method. The process() method should check for new data and store it into the channel as Flume Events.

Note that there are actually two types of sources. The Pollablesource was already mentioned. The other is the EventDrivenSource. The EventDrivenSource, unlike the PollableSource, must have its own callback mechanism that captures the new data and stores it into the Channel. The EventDrivenSources are not each driven by their own thread like the PollableSources are. Below is an example of a custom PollableSource:

```
public class MySource extends AbstractSource implements Configurable, PollableSource {
   private String myProp;

@Override
   public void configure(Context context) {
      String myProp = context.getString("myProp", "defaultValue");

      // Process the myProp value (e.g. validation, convert to another type, ...)
```

```
// Store myProp for later retrieval by process() method
   this.myProp = myProp;
 }
  @Override
 public void start() {
   // Initialize the connection to the external client
  @Override
  public void stop () {
   // Disconnect from external client and do any additional cleanup
    // (e.g. releasing resources or nulling-out field values) ..
  }
  @Override
  public Status process() throws EventDeliveryException {
   Status status = null;
    try {
     // This try clause includes whatever Channel/Event operations you want to do
      // Receive new data
     Event e = getSomeData();
      // Store the Event into this Source's associated Channel(s)
      getChannelProcessor().processEvent(e);
     status = Status.READY;
    } catch (Throwable t) {
     // Log exception, handle individual exceptions as needed
      status = Status.BACKOFF;
      // re-throw all Errors
      if (t instanceof Error) {
       throw (Error)t;
    } finally {
     txn.close();
   return status;
 }
}
```

Channel

TBD