

Step 1 : Download the Latest Version of Flume.

<https://www.apache.org/dyn/closer.lua/flume/1.9.0/apache-flume-1.9.0-bin.tar.gz>

Step 2 : `tar -xzvf apache-flume-1.9.0-bin.tar.gz`

Step 3 : `sudo mv apache-flume-1.9.0-bin /usr/local/flume`

Step 4 : `sudo nano ~/.bashrc`

Step 5 : `export FLUME_HOME=/usr/local/flume`

`export FLUME_CONF_DIR=$FLUME_HOME/conf`

`export FLUME_CLASS_PATH=$FLUME_CONF_DIR`

`export PATH=$FLUME_HOME/bin:$PATH`

Step 6 :

`cp conf/flume-env.sh.template conf/flume-env.sh`

Step 7 :

`sudo nano conf/flume-env.sh`

`JAVA_HOME=/usr/lib/jvm/jdk1.8.0`

`JAVA_OPTS="-Xms100m -Xmx200m -Dcom.sun.management.jmxremote"`

Step 8:

`sudo nano flume-conf.properties.template`

`agent.channels.memoryChannel.type = memory`

`agent.channels.memoryChannel.capacity = 100`

`# Define a source on agent and connect to channel memoryChannel.`

`agent.sources.tail-source.type = exec`

`agent.sources.tail-source.command = tail -F /opt/hadoop-2.6.0/logs/hadoop-hadoop-datanode-node1.log`

`agent.sources.tail-source.channels = memoryChannel`

`# Define a sink that outputs to logger.`

`agent.sinks.log-sink.channel = memoryChannel`

`agent.sinks.log-sink.type = logger`

`agent.sinks.hdfs-sink.channel = memoryChannel`

`agent.sinks.hdfs-sink.type = hdfs`

`agent.sinks.hdfs-sink.hdfs.path = hdfs://node1:8020/flumedata/`

`agent.sinks.hdfs-sink.hdfs.fileType = DataStream`

`# Activate channel, source and sinks`

`agent.channels = memoryChannel`

`agent.sources = tail-source`

`agent.sinks = log-sink hdfs-sink`

`Start Flume-ng agent`

`source ~/.bashrc`

`flume-ng agent -n agent -f conf/flume.conf -`

`Dflume.root.logger=DEBUG,console`

`flume-ng --help`

`sudo yum install flume-ng`