1. **Loading files into HDFS using flumes Spool directory**

Link: https://acadgild.com/blog/loading-files-into-hdfs-using-flumes-spool-directory/

Fetch the data from local file system to HDFS using the steps given in the above blog and send us the screenshot for the same.

#Flume Configuration Starts

# Define a file channel called fileChannel on agent1

agent1.channels.fileChannel1\_1.type = file

# on linux FS

agent1.channels.fileChannel1\_1.capacity = 200000

agent1.channels.fileChannel1\_1.transactionCapacity = 1000

# Define a source for agent1

agent1.sources.source1\_1.type = spooldir

# on linux FS

#Spooldir in my case is /home/hadoop/Desktop/flume\_sink

agent1.sources.source1\_1.spoolDir = /home/Acadgild/Desktop/flume\_sink

agent1.sources.source1\_1.fileHeader = false

agent1.sources.source1\_1.fileSuffix = .COMPLETED

agent1.sinks.hdfs-sink1\_1.type = hdfs

#Sink is /flume\_import under hdfs

agent1.sinks.hdfs-sink1\_1.hdfs.path = hdfs://localhost.localdomain:9000/flume\_sink

agent1.sinks.hdfs-sink1\_1.hdfs.batchSize = 1000

agent1.sinks.hdfs-sink1\_1.hdfs.rollSize = 268435456

agent1.sinks.hdfs-sink1\_1.hdfs.rollInterval = 0

agent1.sinks.hdfs-sink1\_1.hdfs.rollCount = 50000000

agent1.sinks.hdfs-sink1\_1.hdfs.writeFormat=Text

agent1.sinks.hdfs-sink1\_1.hdfs.fileType = DataStream

agent1.sources.source1\_1.channels = fileChannel1\_1

agent1.sinks.hdfs-sink1\_1.channel = fileChannel1\_1

agent1.sinks = hdfs-sink1\_1

agent1.sources = source1\_1

agent1.channels = fileChannel1\_1



