

ejemplo-flume.conf

config Flume Agente1

TwitterAgente1.sources = Twitter
TwitterAgente1.channels = MemChannel
TwitterAgente1.sinks = HDFS

Describe/configura source

TwitterAgente1.sources.Twitter.type = source.TwitterSource
TwitterAgente1.sources.Twitter.channels = MemChannel
TwitterAgente1.sources.Twitter.consumerKey = <consumerKey>
TwitterAgente1.sources.Twitter.consumerSecret = <consumerSecret>
TwitterAgente1.sources.Twitter.accessToken = <accessToken>
TwitterAgente1.sources.Twitter.accessTokenSecret = <accessTokenSecret>
TwitterAgente1.sources.Twitter.swLngLat = -9.299269, 35.999882
TwitterAgente1.sources.Twitter.neLngLat = 4.327812, 43.79142
TwitterAgente1.sources.Twitter.keywords = @el_pais, @elmundoes, @abc_es, @larazon_es,
@voz_populi, @uah, investigacion

Describe sink

TwitterAgente1.sinks.HDFS.channel = MemChannel
TwitterAgente1.sinks.HDFS.type = hdfs
TwitterAgente1.sinks.HDFS.hdfs.path =
hdfs://localhost:9000/user/hduser/flume/twitter/agente1/%Y/%m/%d/%H
TwitterAgente1.sinks.HDFS.hdfs.fileType = DataStream
TwitterAgente1.sinks.HDFS.hdfs.writeFormat = Text
TwitterAgente1.sinks.HDFS.hdfs.batchSize = 10000
TwitterAgente1.sinks.HDFS.hdfs.rollSize = 0
TwitterAgente1.sinks.HDFS.hdfs.rollCount = 1000

Use a channel which buffers events in memory

TwitterAgente1.channels.MemChannel.type = memory
TwitterAgente1.channels.MemChannel.capacity = 10000
TwitterAgente1.channels.MemChannel.transactionCapacity = 10000