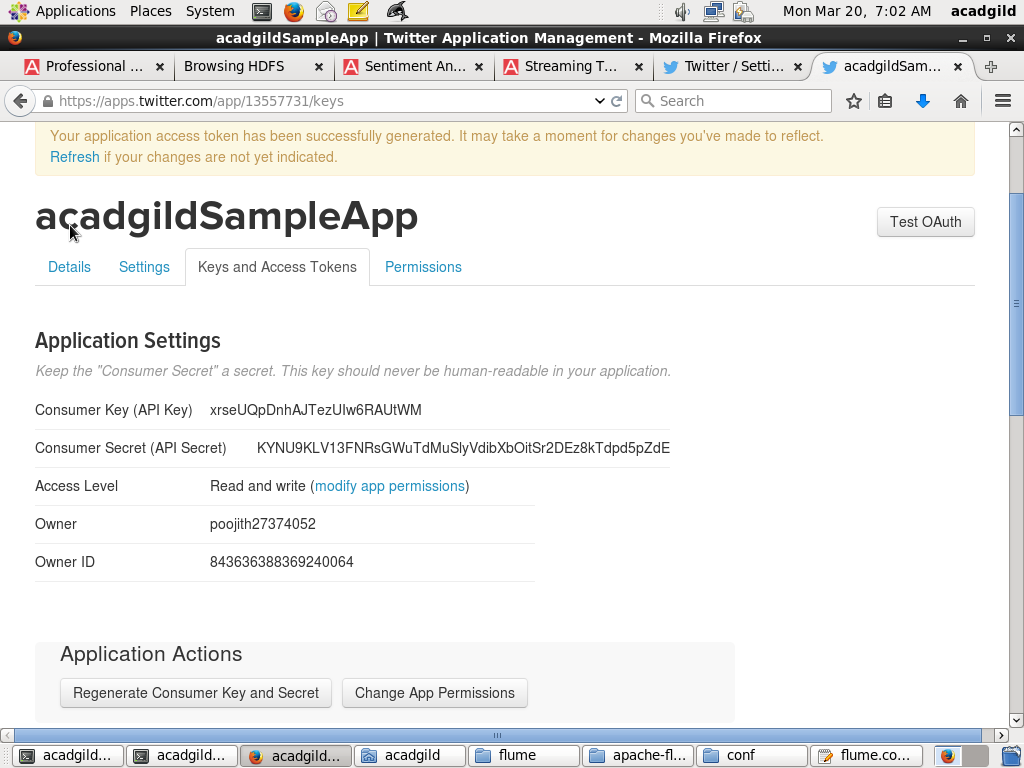
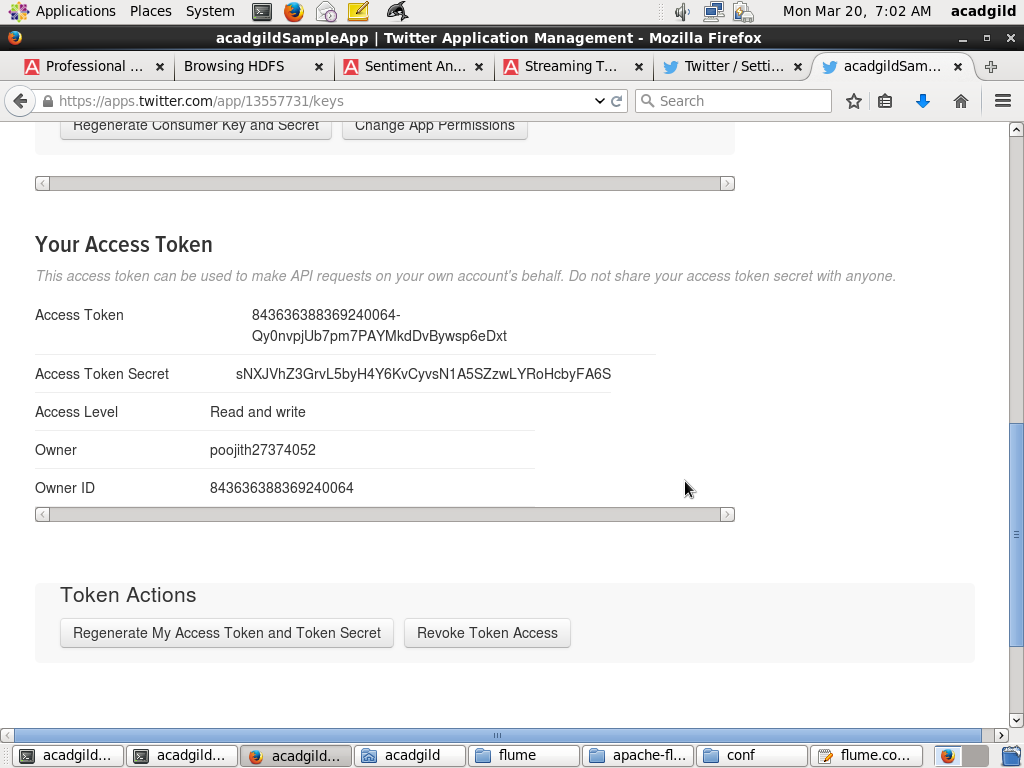
**Assignment 11 1:**

**Create a flume agent that streams data from Twitter and stores in the HDFS**

Steps involved are:

**Getting customer key , customer secret , access token and access token secret from twitter app**





**acadgild.conf: (by placing above keys)**

TwitterAgent.sources = Twitter

TwitterAgent.channels = MemChannel

TwitterAgent.sinks = HDFS

# Describing/Configuring the source

TwitterAgent.sources.Twitter.type = org.apache.flume.source.twitter.TwitterSource

TwitterAgent.sources.Twitter.consumerKey=

xrseUQpDnhAJTezUIw6RAUtWM

TwitterAgent.sources.Twitter.consumerSecret=KYNU9KLV13FNRsGWuTdMuSlyVdibXbOitSr2DEz8kTdpd5pZdE

TwitterAgent.sources.Twitter.accessToken= 843636388369240064-Qy0nvpjUb7pm7PAYMkdDvBywsp6eDxt

TwitterAgent.sources.Twitter.accessTokenSecret= sNXJVhZ3GrvL5byH4Y6KvCyvsN1A5SZzwLYRoHcbyFA6S

TwitterAgent.sources.Twitter.keywords=hadoop, bigdata, mapreduce, mahout, hbase, nosql

# Describing/Configuring the sink

TwitterAgent.sources.Twitter.keywords= hadoop,election,sports, cricket,Big data

TwitterAgent.sources.Twitter.language = en

TwitterAgent.sinks.HDFS.channel=MemChannel

TwitterAgent.sinks.HDFS.type=hdfs

TwitterAgent.sinks.HDFS.hdfs.path=hdfs://localhost:9000/user/flume/tweets

TwitterAgent.sinks.HDFS.hdfs.fileType=DataStream

TwitterAgent.sinks.HDFS.hdfs.writeformat=Text

TwitterAgent.sinks.HDFS.hdfs.batchSize=1000

TwitterAgent.sinks.HDFS.hdfs.rollSize=0

TwitterAgent.sinks.HDFS.hdfs.rollCount=10000

TwitterAgent.sinks.HDFS.hdfs.rollInterval=600

TwitterAgent.channels.MemChannel.type=memory

TwitterAgent.channels.MemChannel.capacity=10000

TwitterAgent.channels.MemChannel.transactionCapacity=1000

TwitterAgent.sources.Twitter.channels = MemChannel

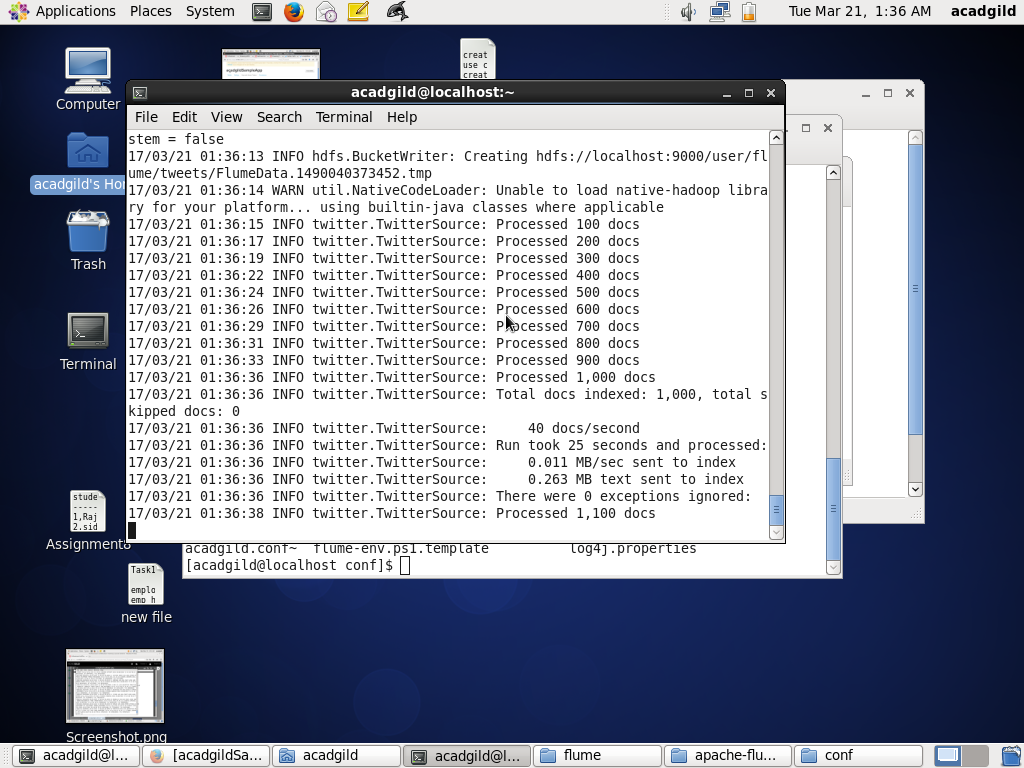
TwitterAgent.sinks.HDFS.channel = MemChannel

**Creating directory in HDFS:**

Hadoop fs -mkdir -p /user/flume/tweets

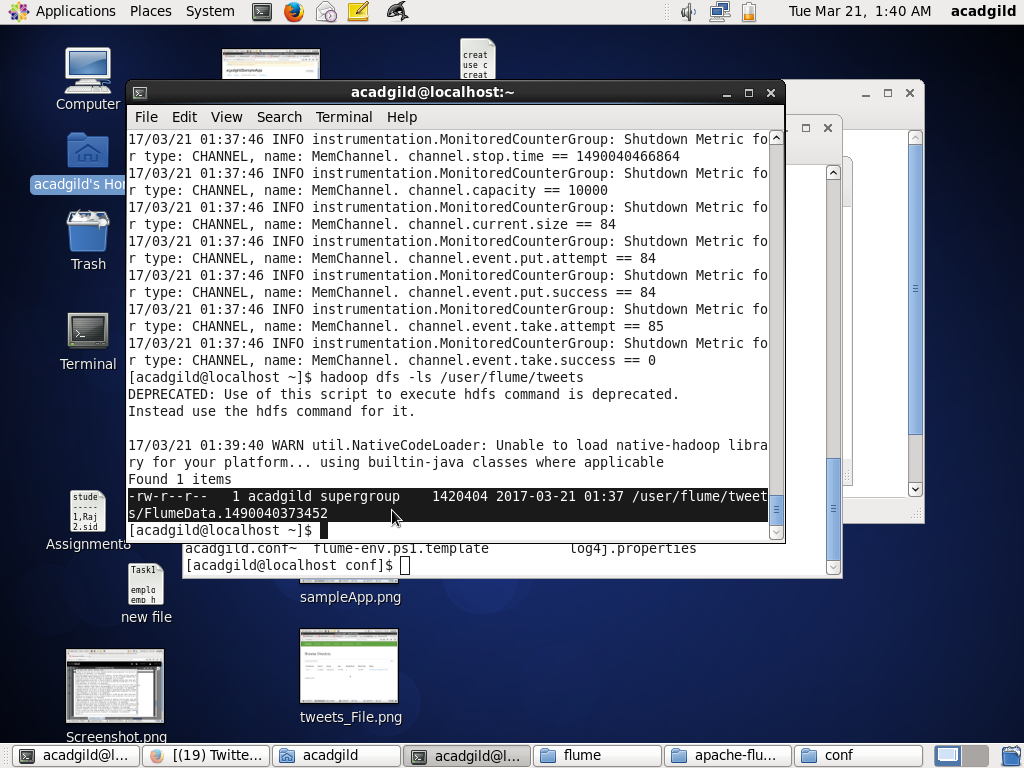
**Start fetching data:**

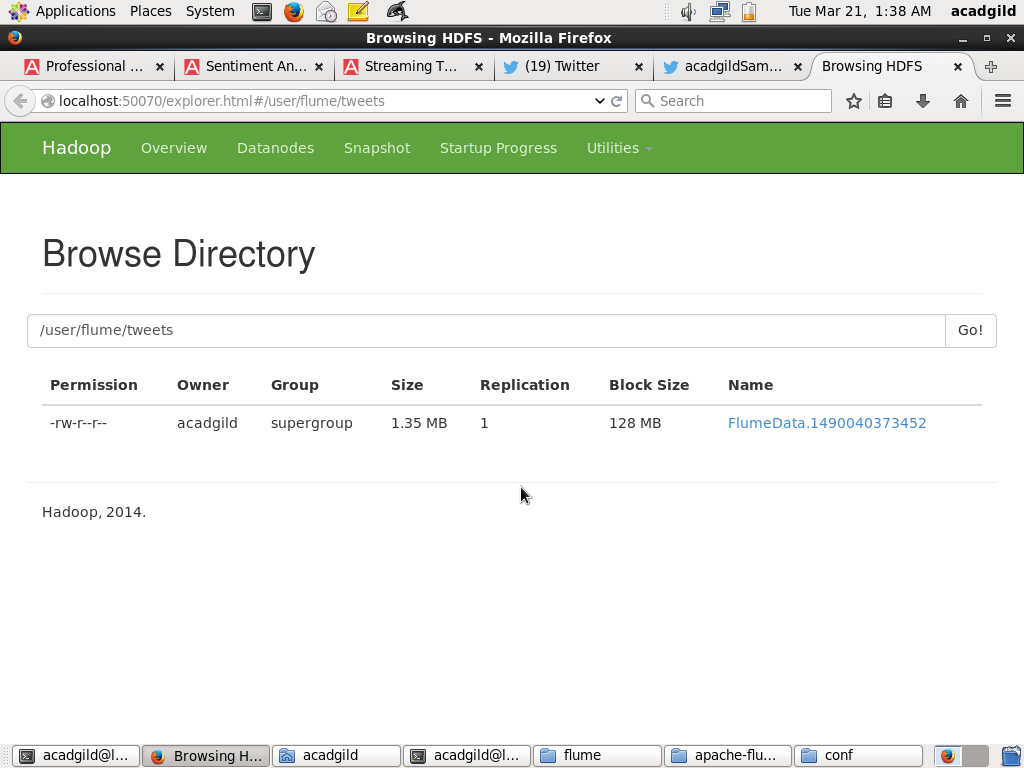
Flume-ng agent --name TwitterAgent –-conf-file /home/flume/acadgild.conf



**Checking whether file is created or not:**

Hadoop fs -ls /user/flume/tweets





**Opening the file:**

Hadoop fs -cat /user/flume/tweeys/FlumeData.1498840373452

