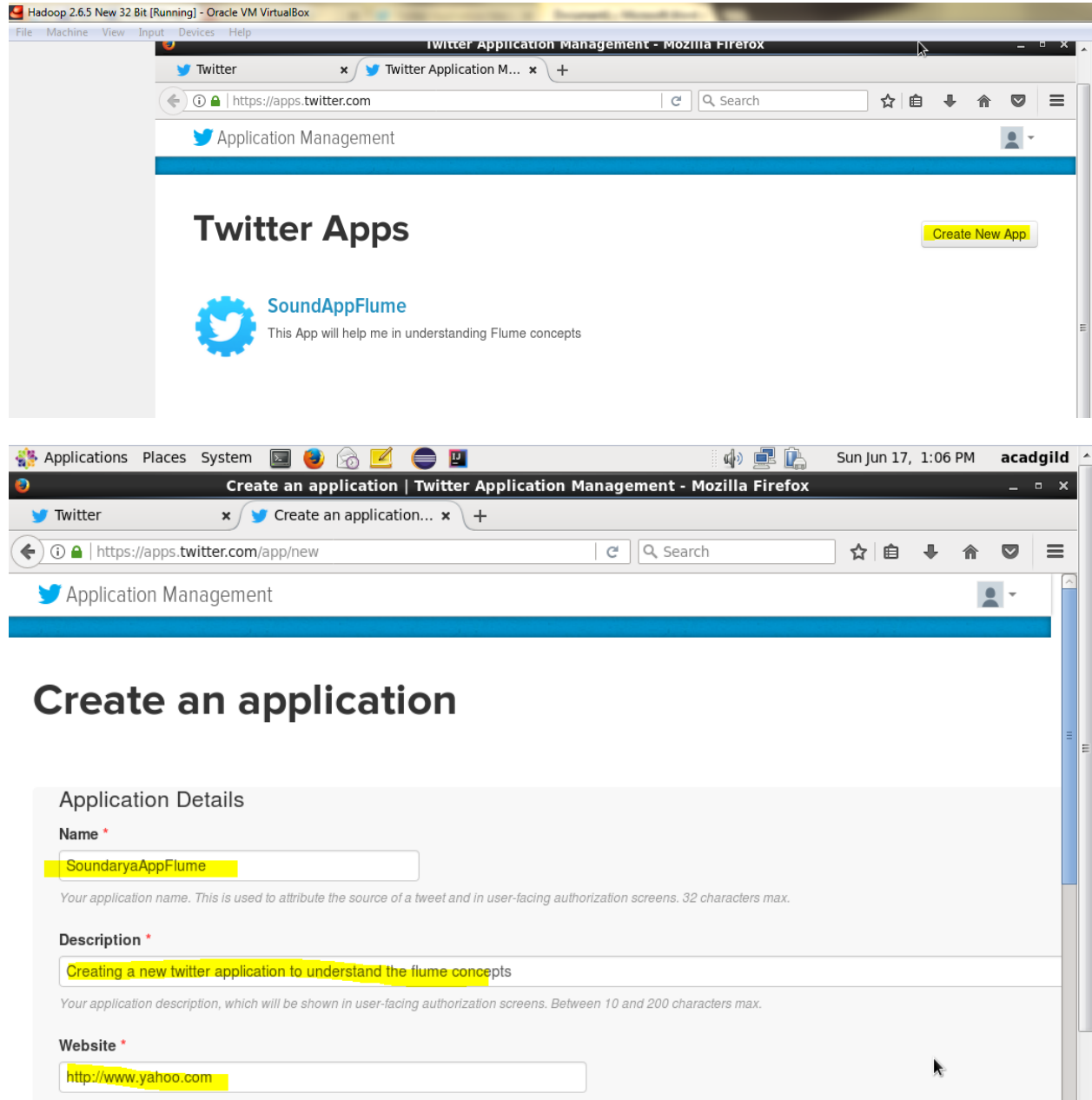**Create a flume agent that streams data from Twitter and stores in the HDFS.**

Step 1 : Logged in to my twitter account.

Step 2: Created a new application as follows

SoundaryaAppFlume | Twitter Application Management - Mozilla Firefox

Twitter   ×   SoundaryaAppFlume ...   ×   +

https://apps.twitter.com/app/15372955/keys   Search

Application Management

SoundaryaAppFlume

Test OAuth

Details   Settings   **Keys and Access Tokens**   Permissions

Application Settings

Keep the "Consumer Secret" a secret. This key should never be human-readable in your application.

Consumer Key (API Key)	kWWbKJQXPSVKnN8ShKfhEIyx2
Consumer Secret (API Secret)	qk4vPqfeJh2gDwQE9GkkmRuEeGEmwagZVOqfZ0HqKX3HSSPgqZ
Access Level	Read and write (modify app permissions)
Owner	Soundar89593389
Owner ID	1008219371171295235

Your Access Token

This access token can be used to make API requests on your own account's behalf. Do not share your access token secret with anyone.

Access Token	1008219371171295235-zhXHILy0J9Yhwv0IGTUP45EETGMgtP
Access Token Secret	xilbMzWs2IHySNofSlaLslpLtdLr56dSrEjUQJACqZX5C
Access Level	Read and write
Owner	Soundar89593389
Owner ID	1008219371171295235

STEP 3: New flume conf file with related consumer keys and access keys

```
*flume.conf (~/install/flume/apache-flume-1.8.0-bin/conf) - gedit

File   Edit   View   Search   Tools   Documents   Help

   Open  ▼  Save        Undo        ✂      📋      🔍  🔍

acadgild.conf  ✖   *flume.conf  ✖

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=kWWbKJQXPSVKnN8ShKfhEIyx2
TwitterAgent.sources.Twitter.consumerSecret=qk4vPqfeJh2gDwQE9GkkmRuEeGEmwagZVOqfZ0HqKX3HSSPgqZ
TwitterAgent.sources.Twitter.accessToken=1008219371171295235-zhXHILy0J9Yhwv0IGTUP45EETGMgtP
TwitterAgent.sources.Twitter.accessTokenSecret=xilbMzWs2IHySNofSlaLslpLtdLr56dSrEjUQJACqZX5C
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.sinks.HDFS.channel=MemChannel
TwitterAgent.sinks.HDFS.type=hdfs
TwitterAgent.sinks.HDFS.hdfs.path=hdfs://localhost:8020/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
```

Step 4:

```
out  Devices   Help

   Applications   Places   System   📺  🦊  📷  📝  ⬤  📘

                                          acadgild@localhost:~

File   Edit   View   Search   Terminal   Help

[acadgild@localhost ~]$ jps
3618 NodeManager
10019 Jps
3989 JobHistoryServer
3303 SecondaryNameNode
4215 SupportedKafka
3176 DataNode
4888 KafkaRestMain
3080 NameNode
3516 ResourceManager
4095 QuorumPeerMain
```

Created tweets directory under /user/flume path

```
c[acadgild@localhost ~]$ hadoop fs -ls /user/flume
^C[acadgild@localhost ~]$ hadoop fs -ls /user/flume
18/06/17 13:18:13 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 1 items
drwxr-xr-x   - acadgild supergroup          0 2018-06-17 12:11 /user/flume/tweets
[acadgild@localhost ~]$
```

STEP 5: Ran the flume agent to fetch the data from twitter to hdfs

```
cat. unable to write to output stream.
[acadgild@localhost ~]$ flume-ng agent -n TwitterAgent -f /home/acadgild/install/flume/apache-flume-1.8.0-bin/conf/flume.conf
```

STEP 6: HDFS OUTPUT FILE

```
[acadgild@localhost ~]$ hadoop fs -ls /user/flume/tweets
18/06/17 13:31:24 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 1 items
-rw-r--r--   1 acadgild supergroup    416262 2018-06-17 12:11 /user/flume/tweets/FlumeData.1529217653516
[acadgild@localhost ~]$
```

STEP 7:

```
[acadgild@localhost ~]$ hadoop fs -cat /user/flume/tweets/FlumeData.1529217653516 |head -20
18/06/17 13:27:49 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
{"type":"record","name":"Doc","doc":"adoc","fields":[{"name":"id","type":"string"},{"name":"user_friends_count","type":["int"
,"null"]},{"name":"user_location","type":["string","null"]},{"name":"user_description","type":["string","null"]},{"name":"use
r_statuses_count","type":["int","null"]},{"name":"user_followers_count","type":["int","null"]},{"name":"user_name","type":["s
tring","null"]},{"name":"user_screen_name","type":["string","null"]},{"name":"created_at","type":["string","null"]},{"name":"
text","type":["string","null"]},{"name":"retweet_count","type":["long","null"]},{"name":"retweeted","type":["boolean","null"]
},{"name":"in_reply_to_user_id","type":["long","null"]},{"name":"source","type":["string","null"]},{"name":"in_reply_to_statu
s_id","type":["long","null"]},{"name":"media_url_https","type":["string","null"]},{"name":"expanded_url","type":["string","nu
ll"]}]}��S��2��0���&10082380278842613760���Malaysiabmainly WINNER (esp Kang Seungyoon) & Kim Jaejoong��'�amalJae_Y8n(2018
-06-17T12:10:52Z�RT @chrissy96_: The members are going to come offstage and walk close to the audience while performing HELL
O��a href="http://twitter.com/download/android" rel="nofollow">Twitter for Android</a>��S��2��0
{"type":"record","name":"Doc","doc":"adoc","fields":[{"name":"id","type":"string"},{"name":"user_friends_count","type":["int"
,"null"]},{"name":"user_location","type":["string","null"]},{"name":"user_description","type":["string","null"]},{"name":"use
r_statuses_count","type":["int","null"]},{"name":"user_followers_count","type":["int","null"]},{"name":"user_name","type":["s
tring","null"]},{"name":"user_screen_name","type":["string","null"]},{"name":"created_at","type":["string","null"]},{"name":"
```