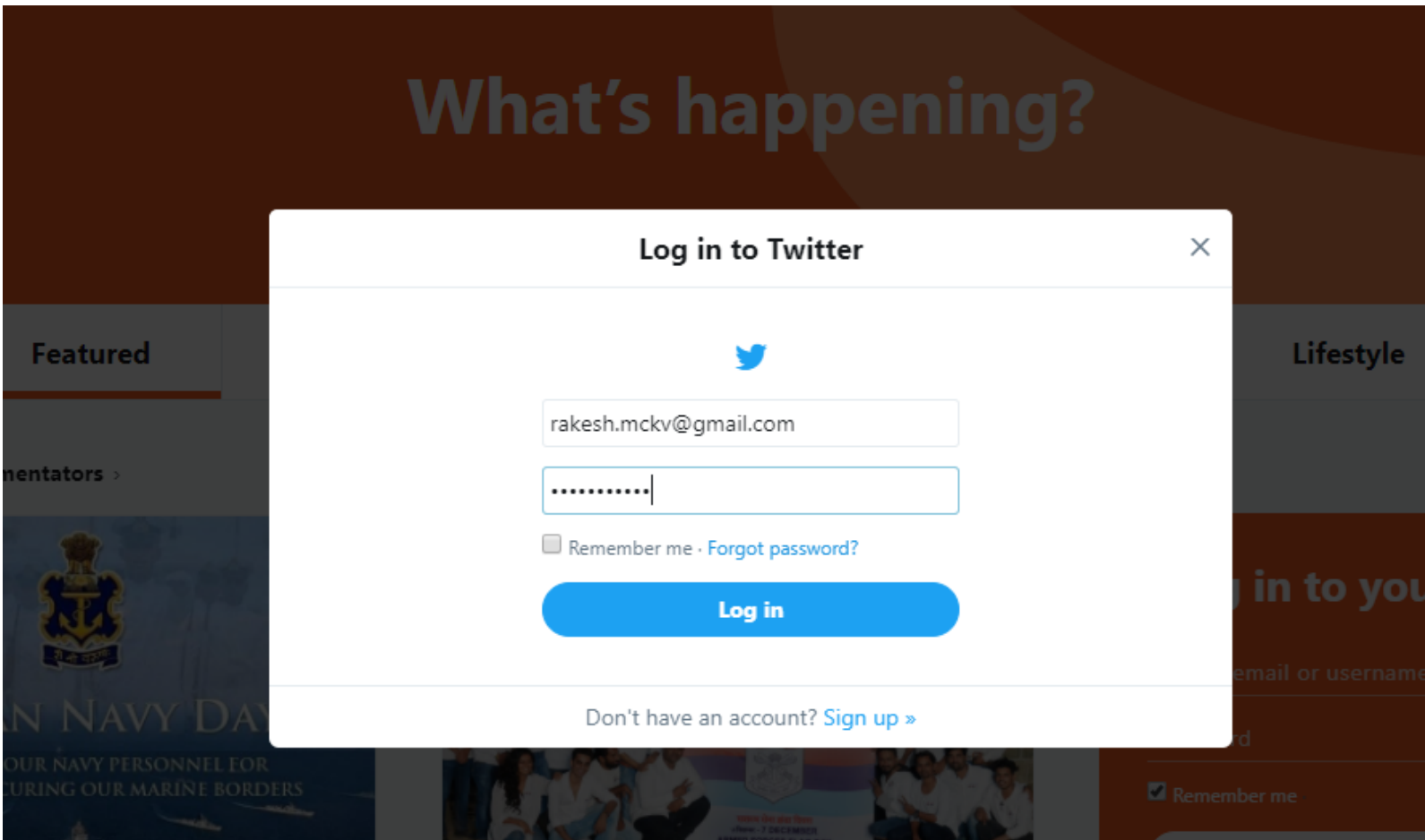


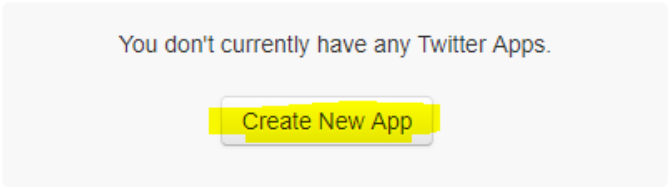
**Problem Statement -Create a flume agent that streams data from Twitter and stores in the HDFS.**

**Step-1** In order to do this we will login to Twitter account



**Step-2** Then we will go to <https://apps.twitter.com/app> and click the ‘create new app’ button.

## Twitter Apps



**Step-3-** Then we will fill up the required details as shown below-



## Create an application

Application Details

Name \*

acadgilApp

Your application name. This is used to attribute the source of a tweet and in user-facing

Description \*

This is for assignment-11.3

Your application description, which will be shown in user-facing authorization screens. B

Website \*

http://www.yahoo.com

Your application's publicly accessible home page, where users can go to download, mak  
source attribution for tweets created by your application and will be shown in user-facing  
(If you don't have a URL yet, just put a placeholder here but remember to change it later

Callback URL

Then we will agree to the developer agreement-

Developer Agreement

☒ Yes, I have read and agree to the [Twitter Developer Agreement](#).

Create your Twitter application

**Step-4** Now we will click on Keys and Access Tokens tab and make a note of the Consumer Key, Consumer Secret-

## acadgilApp

DetailsSettingsKeys and Access TokensPermissions

### Application Settings

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

Consumer Key (API Key)	uae2h893ujuuNW3GdhMmJrVFW
Consumer Secret (API Secret)	LbDoQOTjAn9Qa3MvMmrzQSzdWP34pnEjUKeu6Gx4ZWuFIONTsz
Access Level	Read and write (modify app permissions)
Owner	rakeshhon
Owner ID	704476863385640960

Application Actions

Regenerate Consumer Key and SecretChange App Permissions

**Step-5-** Now click on Create my access token to generate the access token-

## Your Access Token

You haven't authorized this application for your own account yet.

By creating your access token here, you will have everything you need to make your application's current permission level.

Token Actions

Create my access token

Status

Your application access token has been successfully generated. It may take a moment for changes you've made to reflect. Refresh if your changes are not yet indicated.

**Step-6-** Now in order to run the flume agent all HDFS services must be running-

```
[acadgild@localhost ~]$ jps
3378 SecondaryNameNode
4066 Jps
3666 NodeManager
4028 JobHistoryServer
3165 DataNode
3053 NameNode
3551 ResourceManager
[acadgild@localhost ~]$
```

**Step-7-** Set the class path as shown below for flume-

```
[acadgild@localhost ~]$ gedit .bashrc

. bashrc (~) - gedit@localhost.localdomain
File Edit View Search Tools Documents Help
New Open Save Print Undo Redo Cut Copy Paste Find Replace

.bashrc x
export PIG_INSTALL=$HOME/pig-0.16.0
export PATH=$PATH:$PIG_INSTALL/bin
export PIG_CLASSPATH=$HOME/hadoop-2.7.2/sbin

export HIVE_HOME=$HOME/apache-hive-2.1.0-bin
export PATH=$HIVE_HOME/bin:$PATH
export HADOOP_USER_CLASSPATH_FIRST=true

export HBASE_HOME=$HOME/hbase-1.0.3
export PATH=$PATH:HBASE_HOME/bin

export SQOOP_HOME=$HOME/sqoop-1.4.6.bin__hadoop-2.0.4-alpha
export PATH=$PATH:$SQOOP_HOME/bin

export FLUME_HOME=$HOME/apache-flume-1.6.0-bin
export PATH=$PATH:$FLUME_HOME/bin

export PATH=$PATH:$FLUME_HOME/bin:$PATH:$SQOOP_HOME/bin:$PIG_INSTALL/bin:
$HIVE_HOME/bin:$HBASE_HOME/bin:$HADOOP_HOME/sbin
```

**Step-8-** Now we will make a file named as **twitter.conf** inside below mentioned directory-

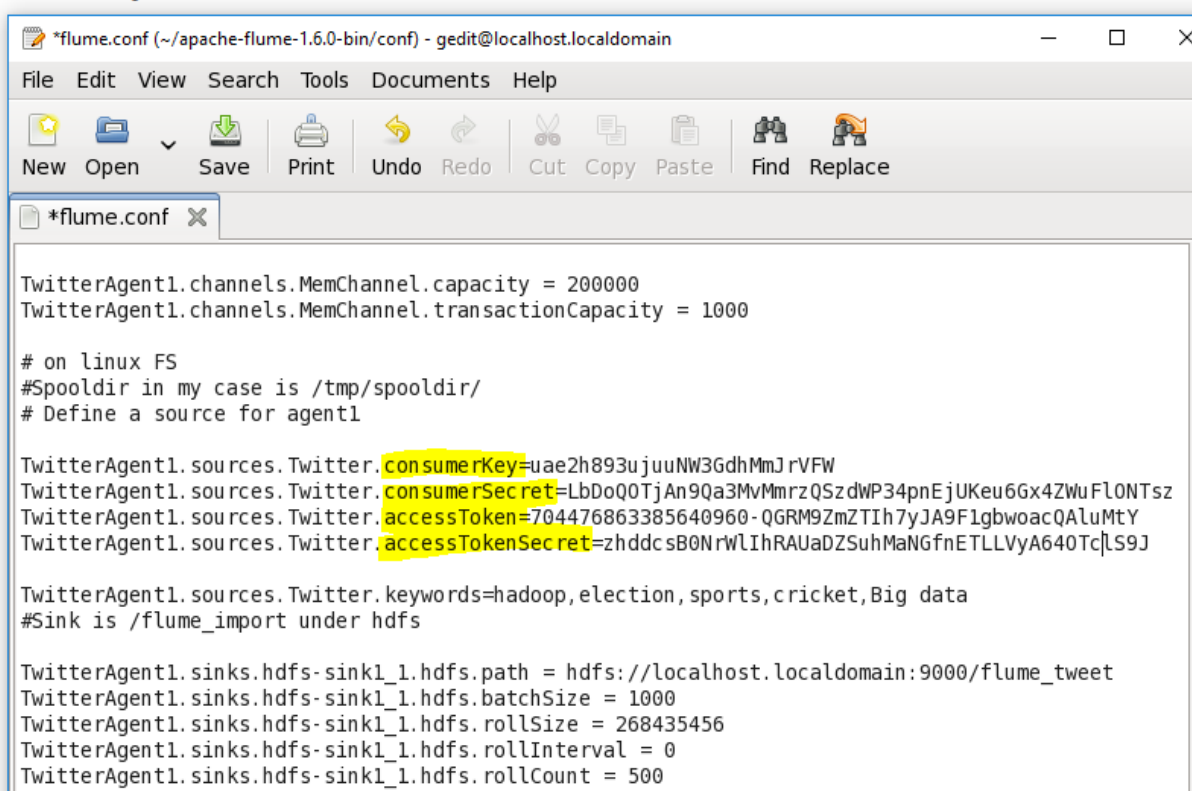
```
[acadgild@localhost conf]$ pwd
/home/acadgild/apache-flume-1.6.0-bin/conf
[acadgild@localhost conf]$ cat > twitter.conf
[acadgild@localhost conf]$ ls -l
total 16
-rw-r--r--. 1 acadgild acadgild 1661 May  9  2015 flume-conf.properties.template
-rw-r--r--. 1 acadgild acadgild 1110 May  9  2015 flume-env.ps1.template
-rw-r--r--. 1 acadgild acadgild 1214 May  9  2015 flume-env.sh.template
-rw-r--r--. 1 acadgild acadgild 3107 May  9  2015 log4j.properties
-rw-rw-r--. 1 acadgild acadgild   0 Dec  5 15:27 twitter.conf
[acadgild@localhost conf]$
```

**Step-9-** In order to execute flume agent below mentioned jars must be present-

```
-rw-r--r--. 1 acadgild acadgild 284077 Aug 23 2014 twitter4j-core-3.0.3.jar
-rw-r--r--. 1 acadgild acadgild 27698 Aug 26 2014 twitter4j-media-support-3.0.3.jar
-rw-r--r--. 1 acadgild acadgild 56307 Aug 23 2014 twitter4j-stream-3.0.3.jar
-rw-r--r--. 1 acadgild acadgild 449505 Aug 23 2014 velocity-1.7.jar
-rw-r--r--. 1 acadgild acadgild 3154938 Apr 18 2015 xalan-2.7.2.jar
-rw-r--r--. 1 acadgild acadgild 1229125 Aug 23 2014 xercesImpl-2.9.1.jar
-rw-r--r--. 1 acadgild acadgild 194354 Aug 23 2014 xml-apis-1.3.04.jar
-rw-r--r--. 1 acadgild acadgild 94672 Aug 23 2014 xz-1.0.jar
-rw-r--r--. 1 acadgild acadgild 64009 Aug 23 2014 zkclient-0.3.jar
[acadgild@localhost lib]$ pwd
/home/acadgild/apache-flume-1.6.0-bin/lib
[acadgild@localhost lib]$
```

**Step-10-** Now we will set the API keys such as consumer key, consumer secret key, access token, access token secret in the twitter.conf file created earlier-

```
calhost conf]$ gedit flume.conf
```



**Step-11-** Now we will run below agent to generate the tweets mentioned as keywords in above conf file-

```
[acadgild@localhost ~]$ flume-ng agent -n TwitterAgent -f /home/acadgild/apache-flume-1.6.0-bin/conf/twitter.conf
```

Relevant discussions can be found on the Internet at:  
<http://www.google.co.jp/search?q=d0031b0b> or  
<http://www.google.co.jp/search?q=1db75513>  
TwitterException(exceptionCode=[d0031b0b-1db75513], statusCode=401, message=null, code=-1, retryAfter=-1, rateLimitStatus=null, version=3.0.3)  
at twitter4j.internal.http.HttpClientImpl.request(HttpClientImpl.java:177)  
at twitter4j.internal.http.HttpClientWrapper.request(HttpClientWrapper.java:61)  
at twitter4j.internal.http.HttpClientWrapper.get(HttpClientWrapper.java:89)  
at twitter4j.TwitterStreamImpl.getSampleStream(TwitterStreamImpl.java:176)  
at twitter4j.TwitterStreamImpl\$.getStream(TwitterStreamImpl.java:164)  
at twitter4j.TwitterStreamImpl\$.TwitterStreamConsumer.run(TwitterStreamImpl.java:462)  
17/12/05 16:08:16 INFO twitter4j.TwitterStreamImpl: Waiting for 40000 milliseconds

**Step-12**-Once the data is loaded same can be found on HDFS path-

[illegible]

P.S- Display Settings of MobaXTerm were changed in between from white to black, that is why above screenshot is showing black.