

Log in to Twitter

dutta.marina@gmail.com

.....

Log in

☒ Remember me · [Forgot password?](#)

New to Twitter? [Sign up now »](#)

Already using Twitter via text message? [Activate your account »](#)

Step 2:

Go to the following link and click the 'Apply for Developer Account' button.

<https://apps.twitter.com/app>

 Application Management

Twitter Apps

As of July 2018, you must [apply for a Twitter developer account](#) and be approved before you may create new apps. Once approved, you will be able to create new apps from [developer.twitter.com](#).

For the near future, you can continue to manage existing apps here on [apps.twitter.com](#). However, we will soon retire this site and consolidate all developer tools, API access, and app management within the developer portal at [developer.twitter.com](#). You will be able to access and manage existing apps through that portal when we retire this site.

[Apply for a developer account](#)

You don't currently have any Twitter Apps.

Step 3:

Enter the necessary details.

 Application Management



Create an application

Application Details

Name *

Your application name. This is used to attribute the source of a tweet and in user-facing authorization screens. 32 characters max.

Description *

Your application description, which will be shown in user-facing authorization screens. Between 10 and 200 characters max.

Website *

Your application's publicly accessible home page, where users can go to download, make use of, or find out more information about your application. This fully-qualified URL is used in the source attribution for tweets created by your application and will be shown in user-facing authorization screens. (If you don't have a URL yet, just put a placeholder here but remember to change it later.)

Callback URL

Where should we return after successful authentication? OAuth 1.0a applications should explicitly specify their OAuth callback URL on the request token step, regardless of the value given here. To restrict your application from using callbacks, leave this field blank.

Step 4:

Accept the developer agreement and select the 'create your Twitter application' button.

Twitter, the Twitter logo, and the Twitter bird are trademarks of Twitter, Inc. or its affiliates. All other marks contained herein are the property of their respective owners.

Developer Agreement

Effective: May 18, 2015.

This Twitter Developer Agreement ("**Agreement**") is made between you (either an individual or an entity, referred to herein as "**you**") and Twitter, Inc. and Twitter International Company (collectively, "**Twitter**") and governs your access to and use of the Licensed Material (as defined below).

PLEASE READ THE TERMS AND CONDITIONS OF THIS AGREEMENT CAREFULLY, INCLUDING WITHOUT LIMITATION ANY LINKED TERMS AND CONDITIONS APPEARING OR REFERENCED BELOW, WHICH ARE HEREBY MADE PART OF THIS LICENSE AGREEMENT. BY USING THE LICENSED MATERIAL, YOU ARE AGREEING THAT YOU HAVE READ, AND THAT YOU AGREE TO COMPLY WITH AND TO BE BOUND BY THE TERMS AND CONDITIONS OF THIS AGREEMENT AND ALL APPLICABLE LAWS AND REGULATIONS IN THEIR ENTIRETY WITHOUT LIMITATION OR QUALIFICATION. IF YOU DO NOT AGREE TO BE BOUND BY THIS AGREEMENT, THEN YOU MAY NOT ACCESS OR OTHERWISE USE THE LICENSED MATERIAL. THIS AGREEMENT IS EFFECTIVE AS OF THE FIRST DATE THAT YOU USE THE LICENSED MATERIAL ("**EFFECTIVE DATE**").

IF YOU ARE AN INDIVIDUAL REPRESENTING AN ENTITY, YOU ACKNOWLEDGE THAT YOU HAVE THE APPROPRIATE AUTHORITY TO ACCEPT THIS AGREEMENT ON BEHALF OF SUCH ENTITY. YOU MAY NOT USE THE LICENSED MATERIAL AND MAY NOT ACCEPT THIS AGREEMENT IF YOU ARE NOT OF LEGAL AGE TO FORM A BINDING CONTRACT WITH TWITTER, OR YOU ARE

☐ Yes, I agree

Create your Twitter application

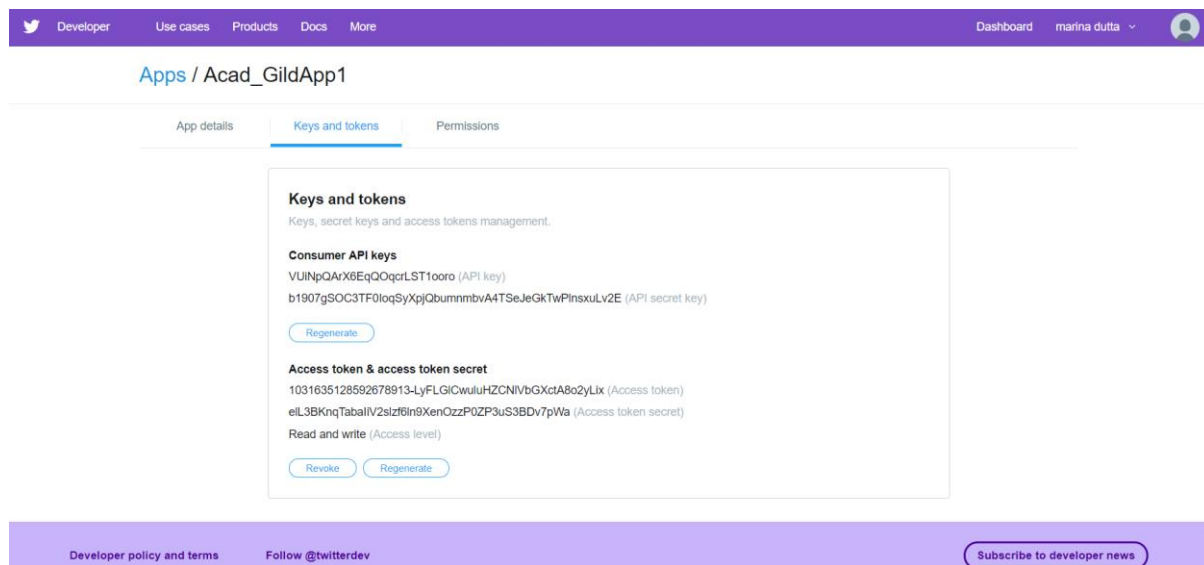
Step 5:

Create the new flume.conf file & copy the Flume configuration code from the below link and paste it in the newly created file flume.conf.

<https://drive.google.com/open?id=OB1QaXx7tpw3Sb3U4LW9SWlNidkk>

Step 6:

You would receive consumerKey, consumerSecret, accessToken, accessTokenSecret from twitter once developer account is approved.



Copy these four values within flume.config file .

Keys and tokens

Keys, secret keys and access tokens management.

Consumer API keys

VUIInpQArX6EqQOqrlST1ooro (API key)

b1907gSOC3TF0loqSyXpjQbumnmbvA4TSeJeGkTwPIInxsuLv2E (API secret key)

Regenerate

Access token & access token secret

1031635128592678913-LyFLGICwuluHZCNIVbGXctA8o2yLix (Access token)

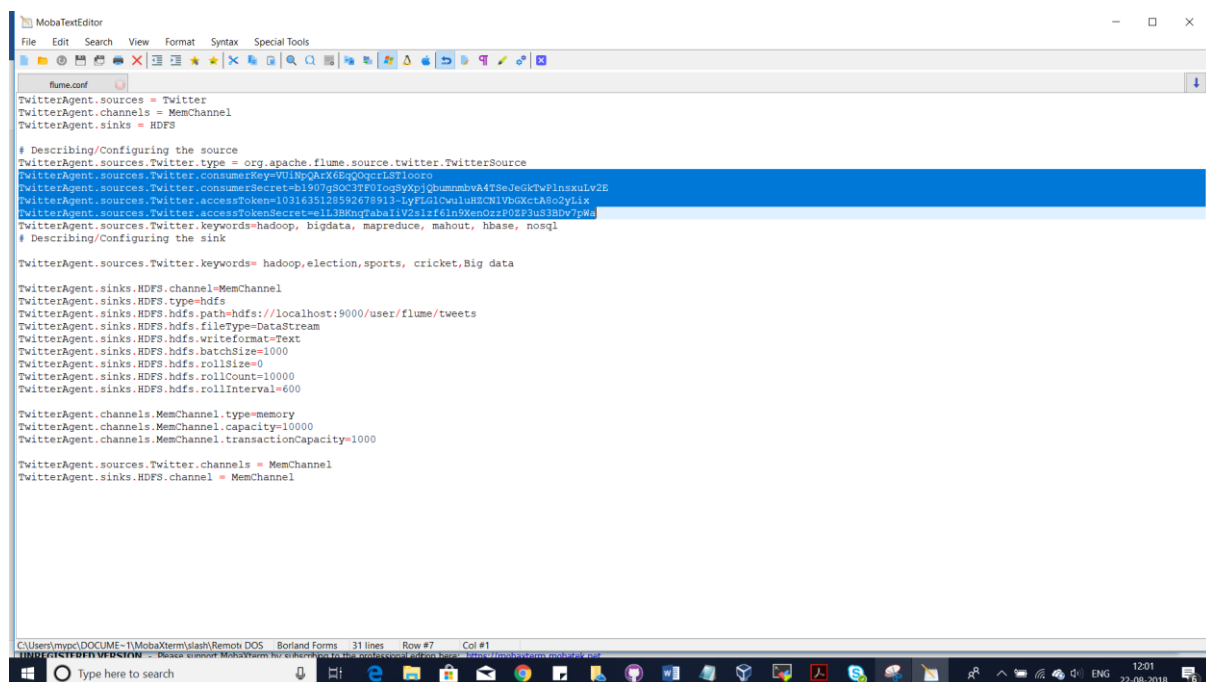
eIL3BKngTabalIV2slzf6ln9XenOzzP0ZP3uS3BDv7pWa (Access token secret)

Read and write (Access level)

Revoke

Regenerate

Copy these four values within **flume.conf** file as highlighted below.



```
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=VUIInpQArX6EqQOqrlST1ooro
TwitterAgent.sources.Twitter.consumerSecret=b1907gSOC3TF0loqSyXpjQbumnmbvA4TSeJeGkTwPIInxsuLv2E
TwitterAgent.sources.Twitter.accessToken=1031635128592678913-LyFLGICwuluHZCNIVbGXctA8o2yLix
TwitterAgent.sources.Twitter.accessTokenSecret=eIL3BKngTabalIV2slzf6ln9XenOzzP0ZP3uS3BDv7pWa
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: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
```

Step 7:

Within the same flume.conf file enter the keywords that you want to search the tweets on twitter against the key **TwitterAgent.sources.Twitter.keywords**

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

Step 8:

Create a new directory tweets which would store tweets stream by flume agent on to

HDFS

```
[acagdild@localhost ~]$ hadoop fs -mkdir -p /hadoopdata/flume/tweets
18/08/22 03:51:17 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
You have new mail in /var/spool/mail/acagdild
[acagdild@localhost ~]$ hadoop fs -ls /hadoopdata/flume/
18/08/22 03:51:46 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
drwxr-xr-x  - acagdild supergroup          0 2018-08-22 03:51 /hadoopdata/flume/tweets
[acagdild@localhost ~]$ █
```

Step 9:

Mention the newly created directory path into the flume.conf as shown

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

```
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=VUiNpQArX6EqQocrLSTlooro
TwitterAgent.sources.Twitter.consumerSecret=b1907gSOC3TF0IoqSyXpjQbummbvA4TSeJeGkTwPlnsxuLv2E
TwitterAgent.sources.Twitter.accessToken=1031635128592678913-LyFLG1CwuluHZCN1VbGXcta8o2yLix
TwitterAgent.sources.Twitter.accessTokenSecret=eLL3BKngTabaiiV2slzf6ln9XenOzzP0ZP3us3BDw7pWw
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:9000/hadoopdata/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
```

We are checking whether all the daemons are started

```
[acagdild@localhost ~]$ jps
3184 NodeManager
4449 Jps
11139 Main
2645 NameNode
23206 org.eclipse.equinox.launcher_1.4.0.v20161219-1356.jar
10840 HMaster
10953 HRegionServer
10777 HQuorumPeer
3082 ResourceManager
2938 SecondaryNameNode
2746 DataNode
8398 RunJar
You have new mail in /var/spool/mail/acagdild
[acagdild@localhost ~]$ █
```

Step 10:

For fetching data from Twitter, Use the below command to fetch the twitter tweet data into the HDFS cluster path.

```
[acadgild@localhost ~]$ flume-ng agent -n TwitterAgent -f/home/acadgild/flume.conf
Warning: No configuration directory set! Use --conf <dir> to override.
Info: Including Hadoop libraries found via (/home/acadgild/install/hadoop/hadoop-2.6.5/bin/hadoop) for HDFS access
Info: Including HBASE libraries found via (/home/acadgild/install/hbase/hbase-1.2.6/bin/hbase) for HBASE access
Info: Including Hive libraries found via (/home/acadgild/install/hive/apache-hive-2.3.2-bin) for Hive access
+ exec /usr/java/jdk1.8.0_151/bin/java -mx20m -cp '/home/acadgild/install/flume/apache-flume-1.8.0-bin/lib/*:/home/acadgild/install/hadoop/hadoop-2.6.5/etc/hadoop/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/hdfs/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/hdfs/lib/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/hdfs/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/yarn/lib/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/yarn/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/lib/*:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/mapreduce/*:/home/acadgild/install/hadoop/hadoop-2.6.5/contrib/capacity-scheduler/*:/home/acadgild/install/hbase/hbase-1.2.6/conf:/usr/java/jdk1.8.0_151/lib/tools.jar:/home/acadgild/install/hbase/hbase-1.2.6:/home/acadgild/install/hbase/hbase-1.2.6/lib/activation-1.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/aopalliance-1.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/apacheds-i18n-2.0.0-M15.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/apacheds-kerberos-codec-2.0.0-M15.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/api-asn1-api-1.0.0-M20.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/api-util-1.0.0-M20.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/asm-3.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/avro-1.7.4.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-beanutils-1.7.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-beanutils-core-1.8.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-cli-1.2.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-codec-1.9.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-collections-3.2.2.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-compress-1.4.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-configuration-1.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-daemon-1.0.13.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-digester-1.8.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-el-1.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-httpclient-3.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-io-2.4.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-lang-2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-logging-1.2.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-math-2.2.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-math3-3.1.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/commons-net-3.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/disruptor-3.3.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/findbugs-annotations-1.3.0-1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/guava-12.0.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/guice-3.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/guice-servlet-3.0.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-annotations-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-auth-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-client-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-common-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-hdfs-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-mapreduce-client-app-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-mapreduce-client-common-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-mapreduce-client-core-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-mapreduce-client-jobclient-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-mapreduce-client-shuffle-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-yarn-api-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-yarn-client-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-yarn-common-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hadoop-yarn-server-common-2.5.1.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-annotations-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-annotations-1.2.6-tests.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-client-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-common-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-common-1.2.6-tests.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-examples-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-external-blockcache-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-hadoop2-compat-1.2.6.jar:/home/acadgild/install/hbase/hbase-1.2.6/lib/hbase-hadoop-compat-1.2.6.jar:/home/acadgild/in
```

The streaming starts

To stop streaming press ctrl c.

Step 11:

To check the contents of the tweet go to the output directory at hdfs:

hadoop fs -ls /hadoopdata/flume/tweets