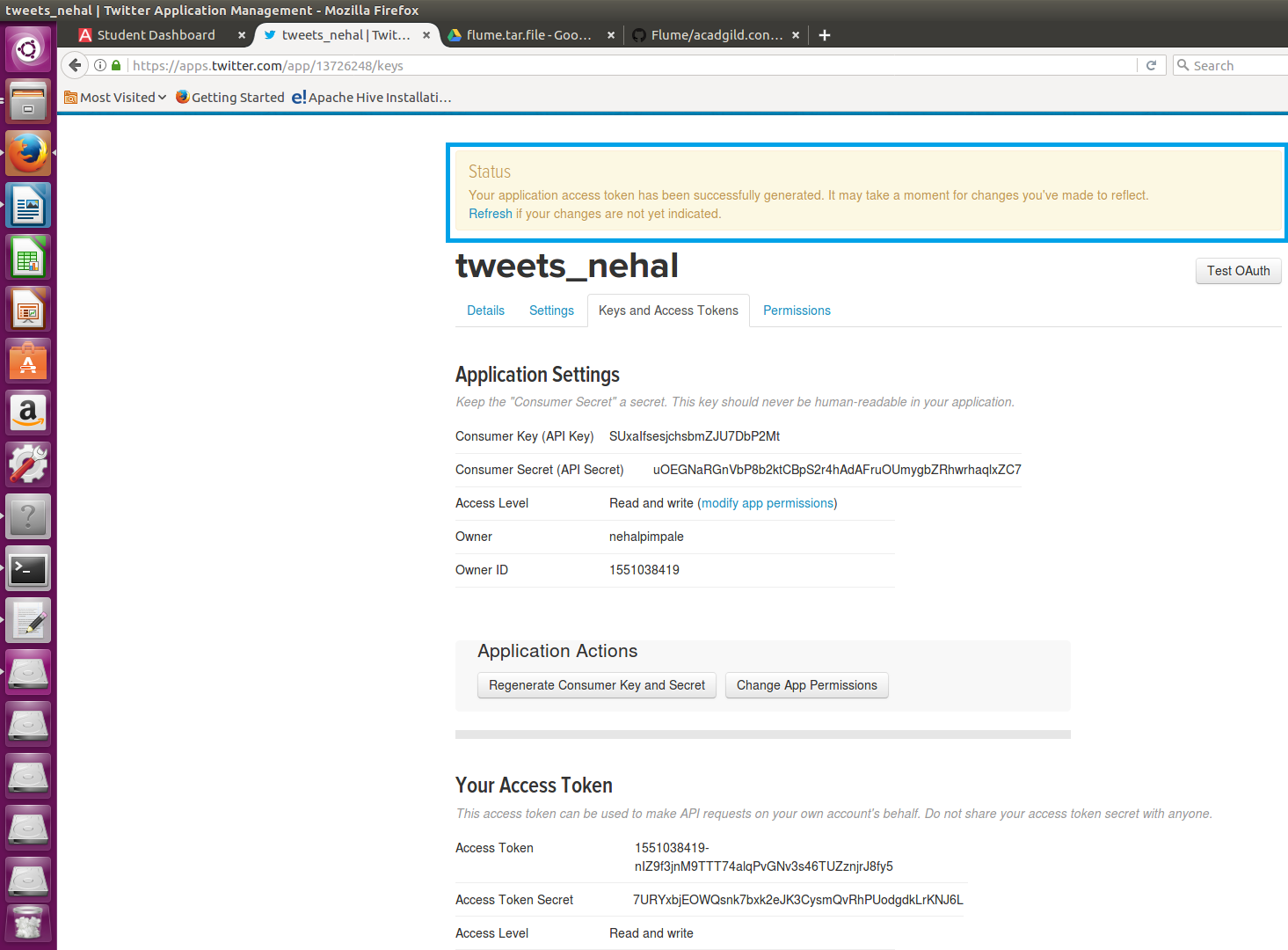
Problem Statement:

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

Step 1:

I created development application using my twitter account.

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



Step 2:

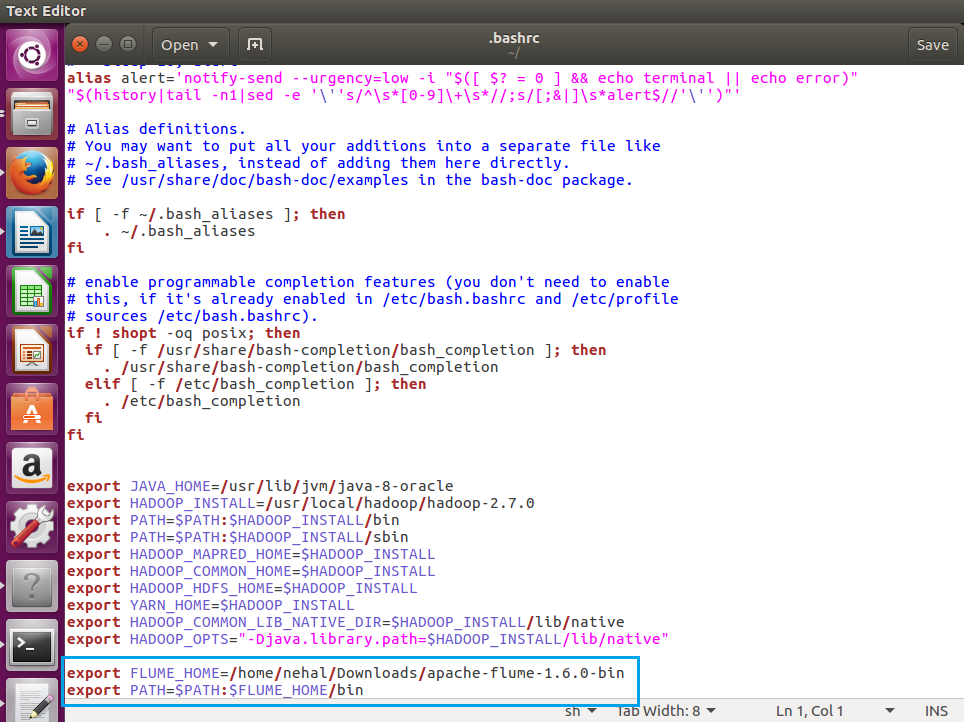
I downloaded flume tar file using below link.

<https://drive.google.com/drive/u/0/folders/0B1QaXx7tpw3SWkMwVFBkc3djNFk>

Step 3:

I extracted tar file and updated the path of extracted file in .bashrc

I updated .bashrc file with source command.

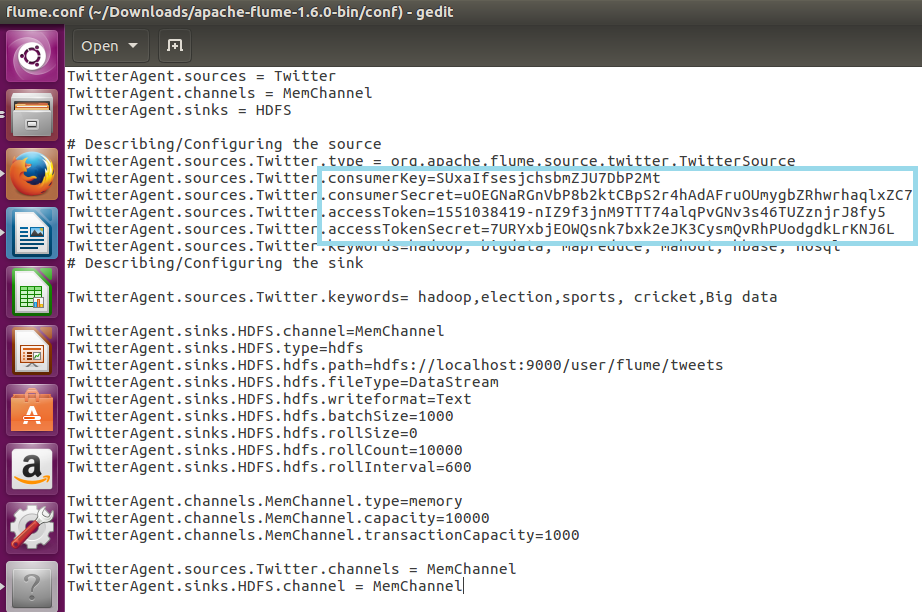


Step 4:

I added configuration file inside the conf directory inside flume extracted directory.

Step 5:

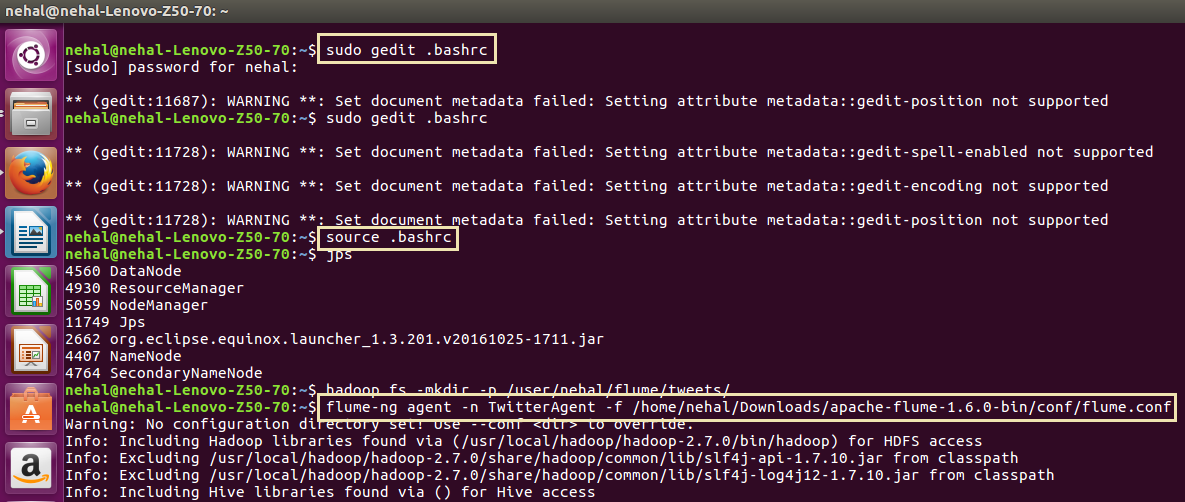
I changed the twitter api keys with the keys generated during application creation.



Step 6:

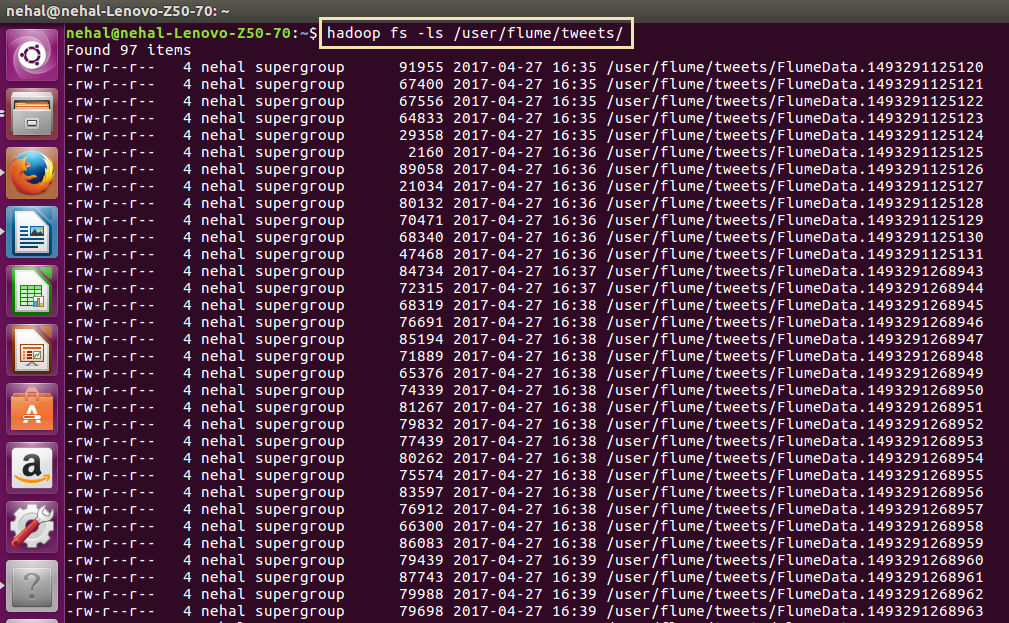
Fetched data from twitter using following command

Flume –ng agent TwitterAgent –f /home/nehal/Downloads/apache-flume-1.6.0-bin/conf/flume.conf



Step 7:

Fetched data



Step 8:

Displaying data fetched

