**ANALYZING THE SOCIAL MEDIA TRENDS USING THE TWITTER HASHTAGS**

**[Project – Mapreduce and Flume combined]**

* Prepared by Vignesh.R (15CSE107)

**Step 1:**

Download the data from twitter. The procedures are already explained in the earlier flume documentation. There the keywords are entered before compiling the .jar file and then the mentioned tweet data is downloaded to the hadoop file system.

**Step 2:**

Then a new jar file named TopTrendsFinder is created with the necessary external jar files added to it. The jar files contains the following java programs.

* ReverseComparator.java
* TopTrendsFinder.java
* TrendMapper1.java
* TrendMapper2.java
* TrendReducer1.java
* TrendReducer2.java

The configuration is set in such a way that only the hashtags are taken into account.

Two mapper classes and two reducer classes are written in which the first mapper phase collects the hashtags and its corresponding reducer phase combines it. On the other hand, the second mapper phase re-orders it to shows the top trending hashtag in its reducer phase.

The TopTrendsFinder.jar can be downloaded from the following link:-

<https://drive.google.com/file/d/0B_ip9omU6d3yeVdUTnlWdF9INXM/view?usp=sharing>

**Step 3:**

Command to execute the jar file:

hadoop jar /home/cloudera/Desktop/TopTrendsFinder.jar TopTrendsFinder.TopTrendsFinder /user/flume/tweets /user/flume/trending

If the input path is to be changed, then the input,output and temporary path are also needed to be changed in the main TopTrendsFinder.java program.

**Result:**



