

## **PROBLEM STATEMENT**

Design and Implement an Apache Storm topology to count to the number of hashtags retrieved from the tweets received via Twitter Stream.

Using the Twitter API, the Apache Storm topology should consume the most recent tweets and extract the hashtags available within the tweets in real-time. Each hashtag should be stored in a MYSQL database table along with its count.

Note: In order to use the Twitter API, create a twitter developer account and register for a free APP. Once the APP is registered, you will be provided with the following credentials which will be used by the Twitter API inorder to communicate and authenticate with the Twitter Stream Hose.

1. API Key
2. API Secret Key
3. AccessToken
4. Access Token Secret

The HashTag Count topology should work towards resolving the following problems

1. Tuple Load Balancing
2. Guaranteed Message / Tuple Processing