

## Architecture

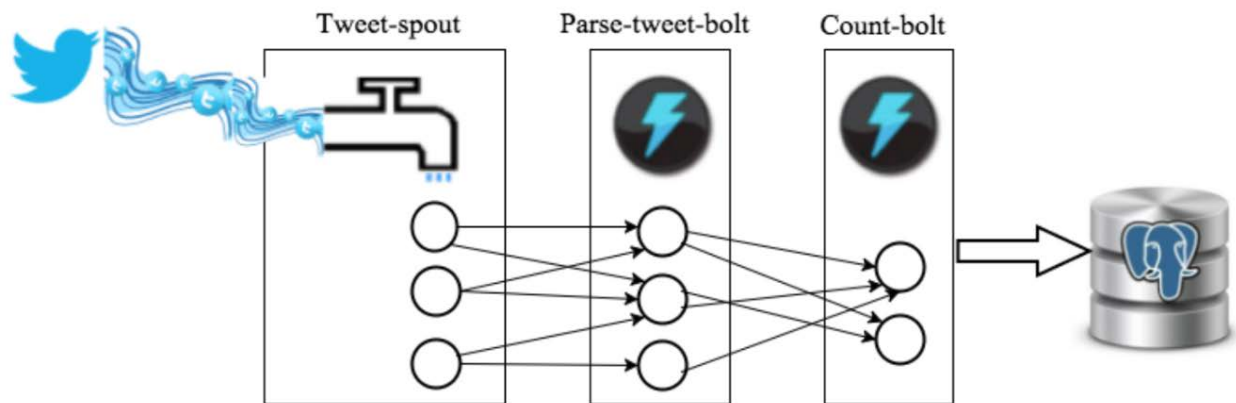
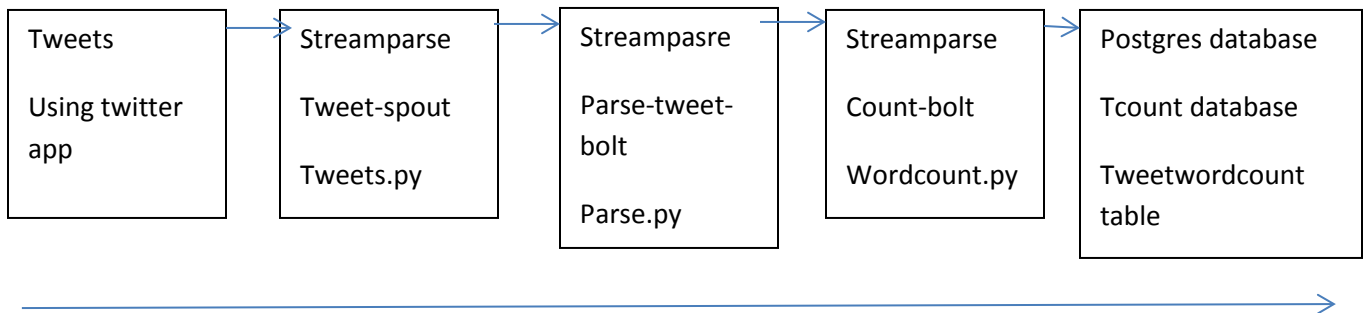


Figure 1: Application Topology

The overall project architecture was created using Streamparse and then modifying it.



In Exercise\_2\_new there is a directory called extweetwordcount

- Extweetwordcount
  - Topologies
    - Tweetwordcount.clj #defines the architecture
  - Src
    - Spouts
      - Tweets.py # gets the tweets from twitter
    - Bolts
      - Parse.py # parses the tweets
      - Wordcount.py #counts the words and writes them to the postgres Tcount database in the Tweetwordcount table

Once this has been run, you can query the results using the finalresults.py and histogram.py files