



Figure 1: Application Topology

Introduction:

- Using the Tweepy library, the application reads the live stream of tweets from Twitter in the tweet-spout component.
- The parse-tweet-bolt parses the tweets, extracts the words from each parsed tweet, and emits the words to the next bolt component (count-bolt) in the topology.
- The count-bolt counts the number of each word in the received tuples, and updates the counts associated with each word in the tweetwordcount table inside the tcount Postgres database.

Descriptions:

- tweet-spout - exttweetwordcount/src/spouts/tweets.py
 - pulls from twitter stream
- parse-tweet-bolt - exttweetwordcount/src/bolts/parse.py
 - parses tweets into words
- count-bolt - exttweetwordcount/src/bolts/wordcount.py
 - stores words into the tweetwordcount table