

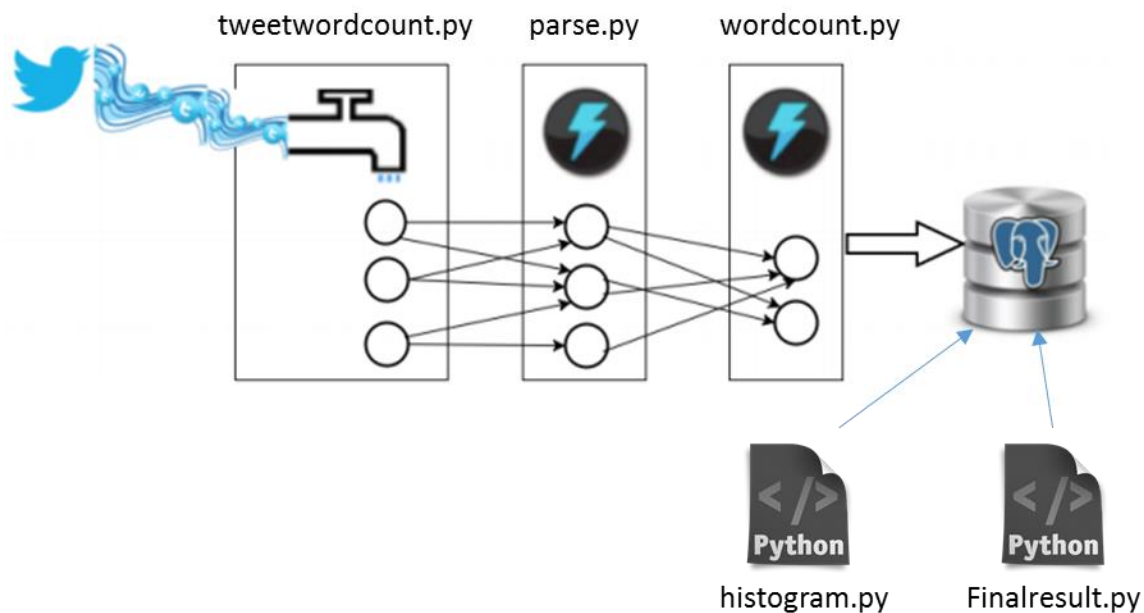
Exercise 2 Architecture:

Overview:

For exercise 2, I followed the instructions to create a twitter application that connected a twitter applications with StreamParse and a Postgres database. The application parses twitter feeds, counts the words and stores those counts in a database table.

Architecture:

The application uses a tweepy spout to get feeds from twitter, it then passes it to 2 bolts a parsing bolt and a word count bolt which completes the word counting and puts that information into a postgres database table. There are then 2 python scripts that are used to access the postgres database. Below is a picture of the overall StreamParse architecture along with the python scripts that access the postgres database.



Running the application:

You can run the application by following these steps (also available in the setup.txt that is part of the git repository).

- 1.) Clone the code from the git hub repository
root@ip #git clone https://github.com/mneuhard/exercise2 exercise2
- 2.) Create tweetwordcount StreamParse structure

```

root@ip#sparse quickstart tweetwordcount
root@ip#cd tweetwordcount
root@ip#cp -r ../exercise2.
2.) Create the database
root@ip#cd tweetwordcount
root@ip#python db_setup.py
3.) Run StreamParse
root@ip#cd tweetwordcount
root@ip#sparse run
4.) Stop StreamParse after at least 1 minute
Press Ctrl-C
5.) Run the finalresults.py
root@ip#python finalresults.py
6.) Run the finalresults.py with a word
root@ip#python finalresults.py hello
7.) Run histogram.py with 2 numbers
root@ip#python histogram.py 3 8

```

Directory and File structure:

```

.:
README.md _resources db_setup.py final_results.py histogram.py project.clj tasks.py virtualenvs
_build config.json fabfile.py finalresults.py logs src topologies

```

```

./_build:
classes stale

```

```

./_build/classes:
META-INF

```

```

./_build/classes/META-INF:
maven

```

```

./_build/classes/META-INF/maven:
tweetwordcount

```

```

./_build/classes/META-INF/maven/tweetwordcount:
tweetwordcount

```

```

./_build/classes/META-INF/maven/tweetwordcount/tweetwordcount:
pom.properties

```

```

./_build/stale:
leiningen.core.classpath.extract-native-dependencies

```

./_resources:
resources

./_resources/resources:
bolts spouts

./_resources/resources/bolts:
__init__.py parse.py wordcount.py

./_resources/resources/spouts:
__init__.py tweets.py words.py

./logs:
streamparse_tweetwordcount_count-bolt_3_11922.log streamparse_tweetwordcount_parse-
tweet-bolt_5_15464.log
streamparse_tweetwordcount_count-bolt_3_12131.log streamparse_tweetwordcount_parse-
tweet-bolt_5_15684.log
streamparse_tweetwordcount_count-bolt_3_12338.log streamparse_tweetwordcount_parse-
tweet-bolt_5_15894.log
streamparse_tweetwordcount_count-bolt_3_12577.log streamparse_tweetwordcount_parse-
tweet-bolt_5_26146.log
streamparse_tweetwordcount_count-bolt_3_12807.log streamparse_tweetwordcount_parse-
tweet-bolt_5_7503.log
streamparse_tweetwordcount_count-bolt_3_13082.log streamparse_tweetwordcount_parse-
tweet-bolt_5_7726.log
streamparse_tweetwordcount_count-bolt_3_13298.log streamparse_tweetwordcount_parse-
tweet-bolt_5_8976.log
streamparse_tweetwordcount_count-bolt_3_13507.log streamparse_tweetwordcount_parse-
tweet-bolt_5_9568.log
streamparse_tweetwordcount_count-bolt_3_13706.log streamparse_tweetwordcount_parse-
tweet-bolt_5_9756.log
streamparse_tweetwordcount_count-bolt_3_14489.log streamparse_tweetwordcount_parse-
tweet-bolt_6_11899.log
streamparse_tweetwordcount_count-bolt_3_15081.log streamparse_tweetwordcount_parse-
tweet-bolt_6_12132.log
streamparse_tweetwordcount_count-bolt_3_15465.log streamparse_tweetwordcount_parse-
tweet-bolt_6_12352.log
streamparse_tweetwordcount_count-bolt_3_15687.log streamparse_tweetwordcount_parse-
tweet-bolt_6_12578.log
streamparse_tweetwordcount_count-bolt_3_15893.log streamparse_tweetwordcount_parse-
tweet-bolt_6_12809.log
streamparse_tweetwordcount_count-bolt_3_26140.log streamparse_tweetwordcount_parse-
tweet-bolt_6_13078.log

streamparse_tweetwordcount_count-bolt_3_7502.log	streamparse_tweetwordcount_parse-tweet-
bolt_6_13300.log	
streamparse_tweetwordcount_count-bolt_3_7725.log	streamparse_tweetwordcount_parse-tweet-
bolt_6_13504.log	
streamparse_tweetwordcount_count-bolt_3_8980.log	streamparse_tweetwordcount_parse-tweet-
bolt_6_13700.log	
streamparse_tweetwordcount_count-bolt_3_9569.log	streamparse_tweetwordcount_parse-tweet-
bolt_6_14289.log	
streamparse_tweetwordcount_count-bolt_3_9759.log	streamparse_tweetwordcount_parse-tweet-
bolt_6_14460.log	
streamparse_tweetwordcount_count-bolt_4_11898.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_14912.log	
streamparse_tweetwordcount_count-bolt_4_12129.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_15097.log	
streamparse_tweetwordcount_count-bolt_4_12351.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_15466.log	
streamparse_tweetwordcount_count-bolt_4_12574.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_15657.log	
streamparse_tweetwordcount_count-bolt_4_12808.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_15897.log	
streamparse_tweetwordcount_count-bolt_4_13074.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_26138.log	
streamparse_tweetwordcount_count-bolt_4_13301.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_7505.log	
streamparse_tweetwordcount_count-bolt_4_13503.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_7723.log	
streamparse_tweetwordcount_count-bolt_4_13703.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_8966.log	
streamparse_tweetwordcount_count-bolt_4_14461.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_9575.log	
streamparse_tweetwordcount_count-bolt_4_15099.log	streamparse_tweetwordcount_parse-
tweet-bolt_6_9771.log	
streamparse_tweetwordcount_count-bolt_4_15462.log	streamparse_tweetwordcount_tweet-
spout_7_11920.log	
streamparse_tweetwordcount_count-bolt_4_15658.log	streamparse_tweetwordcount_tweet-
spout_7_12130.log	
streamparse_tweetwordcount_count-bolt_4_15895.log	streamparse_tweetwordcount_tweet-
spout_7_12337.log	
streamparse_tweetwordcount_count-bolt_4_26139.log	streamparse_tweetwordcount_tweet-
spout_7_12575.log	
streamparse_tweetwordcount_count-bolt_4_7501.log	streamparse_tweetwordcount_tweet-
spout_7_12805.log	
streamparse_tweetwordcount_count-bolt_4_7724.log	streamparse_tweetwordcount_tweet-
spout_7_13076.log	

streamparse_tweetwordcount_count-bolt_4_8968.log streamparse_tweetwordcount_tweet-spout_7_13297.log
streamparse_tweetwordcount_count-bolt_4_9573.log streamparse_tweetwordcount_tweet-spout_7_13510.log
streamparse_tweetwordcount_count-bolt_4_9770.log streamparse_tweetwordcount_tweet-spout_7_13701.log
streamparse_tweetwordcount_parse-tweet-bolt_5_11919.log streamparse_tweetwordcount_tweet-spout_7_14293.log
streamparse_tweetwordcount_parse-tweet-bolt_5_12135.log streamparse_tweetwordcount_tweet-spout_7_14491.log
streamparse_tweetwordcount_parse-tweet-bolt_5_12339.log streamparse_tweetwordcount_tweet-spout_7_14897.log
streamparse_tweetwordcount_parse-tweet-bolt_5_12576.log streamparse_tweetwordcount_tweet-spout_7_15080.log
streamparse_tweetwordcount_parse-tweet-bolt_5_12810.log streamparse_tweetwordcount_tweet-spout_7_15463.log
streamparse_tweetwordcount_parse-tweet-bolt_5_13077.log streamparse_tweetwordcount_tweet-spout_7_15686.log
streamparse_tweetwordcount_parse-tweet-bolt_5_13299.log streamparse_tweetwordcount_tweet-spout_7_15901.log
streamparse_tweetwordcount_parse-tweet-bolt_5_13509.log streamparse_tweetwordcount_tweet-spout_7_26141.log
streamparse_tweetwordcount_parse-tweet-bolt_5_13705.log streamparse_tweetwordcount_tweet-spout_7_7504.log
streamparse_tweetwordcount_parse-tweet-bolt_5_14297.log streamparse_tweetwordcount_tweet-spout_7_7728.log
streamparse_tweetwordcount_parse-tweet-bolt_5_14487.log streamparse_tweetwordcount_tweet-spout_7_8978.log
streamparse_tweetwordcount_parse-tweet-bolt_5_14898.log streamparse_tweetwordcount_tweet-spout_7_9567.log
streamparse_tweetwordcount_parse-tweet-bolt_5_15082.log streamparse_tweetwordcount_tweet-spout_7_9760.log

./src:
bolts spouts

./src/bolts:
__init__.py parse.py wordcount.py

./src/spouts:
__init__.py tweets.py words.py

./topologies:
tweetwordcount.clj

./virtualenvs:
wordcount.txt