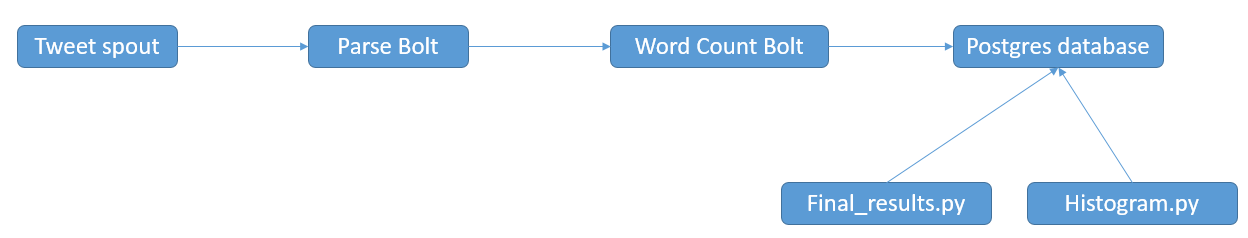
**Architecture – Jonathan Landesman**

Basic Architecture

This twitter application has:

* One spout pulling tweets from the twitter API
* Two bolts: a parsing bolt that splits the tweets into words and a wordcount bolt that counts the words and saves them into a postgres database, tcount.
* A python script final\_results.py that returns the count of the number of words in the tweet, as well as the entire database if no arguments are entered.
* A python script histogram.py that returns a list of the top 10 two integers k1,k2 and returns all the words that their total number of occurrences in the stream is more or equal than k1 and less or equal than k2.



File Structure

The files are self-contained within the folder tweetword count.

* The tweetwordcount directory holds the two files: final\_results.py and histogram.py, as well as the underlying subfolders:
* The topology is defined in the tweetwordcount/topologies folder,
* Bolts are held in tweetwordcount/src/bolts
* Spouts are held in tweetwordcount/src/bolts

To run this application:

First: the following conditions must be met:

* Storm and postgres must be installed and functional.
* **Before** running the application, the user must make a database “tcount” and a table “tweetwordcount”.

Second: from the tweetwordcount directory, type:

“sparse run”

to populate the database. This will seed the database with words collected from twitter and their counts. Note: **This program will not shut off on its own. The user must manually hit “ctrl-c” to kill script**.

Then your program is ready to run.