

1) Solution to the load imbalance problem:

- Used shuffle grouping to load balanced the operation among the 4 wordcount bolt instances. By this grouping all the words are evenly distributed across the available wordcount bolt instances.
- Since the word can reach any of the wordcount bolt instance the emit value will always have the count as 1 for each word. The aggregation of count for each word is done inside Reportbolt.

2) Running the Jar:

a. Production mode:

```
./storm jar ../../storm/WordCount_Bal_MySQL-0.0.1-SNAPSHOT.jar WordCountTopology "Storm_Assign_WC_Bal"
```

b. Local mode:

```
./storm jar ../../storm/WordCount_Bal_MySQL-0.0.1-SNAPSHOT.jar  
WordCountTopology
```

3) Database Details:

Name: upgrad

Table: wordcounts

Login: root

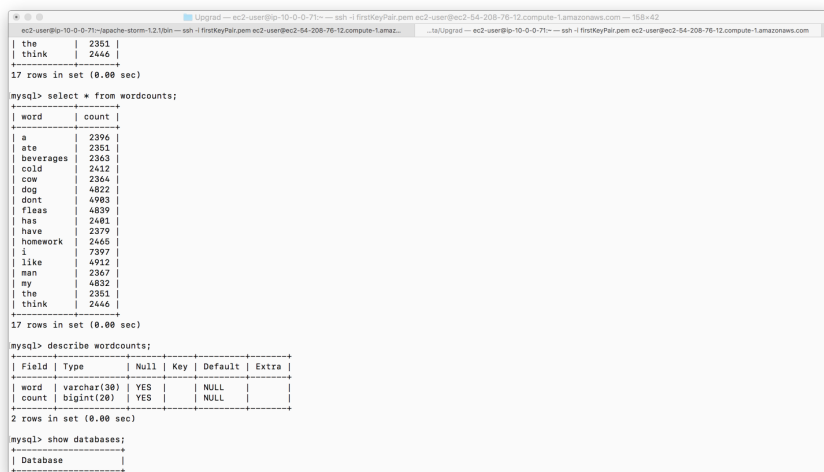
Password: 123

Steps to create db:

```
create database upgrad;  
use upgrad;  
create table wordcounts (word varchar(30), count BIGINT);
```

4) Runtime Snapshots:

- a. snapshot of the table in your database



```
mysql> select * from wordcounts;  
+-----+-----+  
| word | count |  
+-----+-----+  
| the  | 2351  |  
| think| 2446  |  
+-----+-----+  
17 rows in set (0.00 sec)  
  
mysql> select * from wordcounts;  
+-----+-----+  
| word | count |  
+-----+-----+  
| a     | 2396  |  
| ate   | 2351  |  
| beverages | 2363 |  
| cold  | 2412  |  
| cow   | 2364  |  
| dog   | 4822  |  
| dont  | 4983  |  
| flies | 4839  |  
| has   | 2481  |  
| have  | 2379  |  
| homework | 2465 |  
| i     | 7397  |  
| like  | 4912  |  
| man   | 2367  |  
| my    | 4832  |  
| the   | 2351  |  
| think | 2446  |  
+-----+-----+  
17 rows in set (0.00 sec)  
  
mysql> describe wordcounts;  
+-----+-----+  
| Field | Type          | Null | Key | Default | Extra |  
+-----+-----+  
| word  | varchar(30)   | YES  |     | NULL    |       |  
| count | bigint(20)    | YES  |     | NULL    |       |  
+-----+-----+  
2 rows in set (0.00 sec)  
  
mysql> show databases;  
+-----+  
| Database |  
+-----+
```

b. Snapshots of following tables from the Storm UI Topology Summary:

Storm UI

Search Storm_Assign_WC_Bal-2-1534906887: Search Search Archived Logs:

Topology summary

Name	Id	Owner	Status	Uptime	Num workers	Num executors	Num tasks	Replication count	Assigned Mem (MB)	Scheduler Info
Storm_Assign_WC_Bal	Storm_Assign_WC_Bal-2-1534906887	ec2-user	ACTIVE	21m 20s	3	11	11	1	2496	

Topology actions

[Activate](#) [Deactivate](#) [Rebalance](#) [Kill](#) [Debug](#) [Stop Debug](#) [Change Log Level](#)

Topology stats

Window	Emitted	Transferred	Complete latency (ms)	Acked	Failed
10m 0s	34357038	34357038	13709.817	11520	2390540
3h 0m 0s	67430060	67430060	13709.817	11520	2390540
1d 0h 0m 0s	67430060	67430060	13709.817	11520	2390540
All time	67430060	67430060	13709.817	11520	2390540

Spouts and Bolts Table:

Topology summary

Topology actions

Topology stats

Spouts (All time)

Id	Executors	Tasks	Emitted	Transferred	Complete latency (ms)	Acked	Failed	Error Host	Error Port	Last error	Error Time
SentinelSpout	1	1	1157980	1157980	13709.817	11520	372760				

Showing 1 to 1 of 1 entries

Bolts (All time)

Id	Executors	Tasks	Emitted	Transferred	Capacity (est 10s)	Execute latency (ms)	Executed	Process latency (ms)	Acked	Failed	Error Host	Error Port	Last error	Error Time
ReportBolts	1	1	0	0	0.935	1.271	75000	0.063	75000	0				
SpoutServiceBolts	2	2	29890980	29890980	0.311	0.060	6203440	0.060	6203440	0				
WordCountBolts	4	4	17709060	17709060	0.027	0.022	17697660	0.027	17697660	0				

Showing 1 to 3 of 3 entries

ETRADE Brokerage T...pdf 2018-06-19-06-01...pdf 2018-06-19-06-00...pdf 2018-06-19-07-27...pdf 2018-06-19-07-26...pdf Show All X

Topology Visualization:

