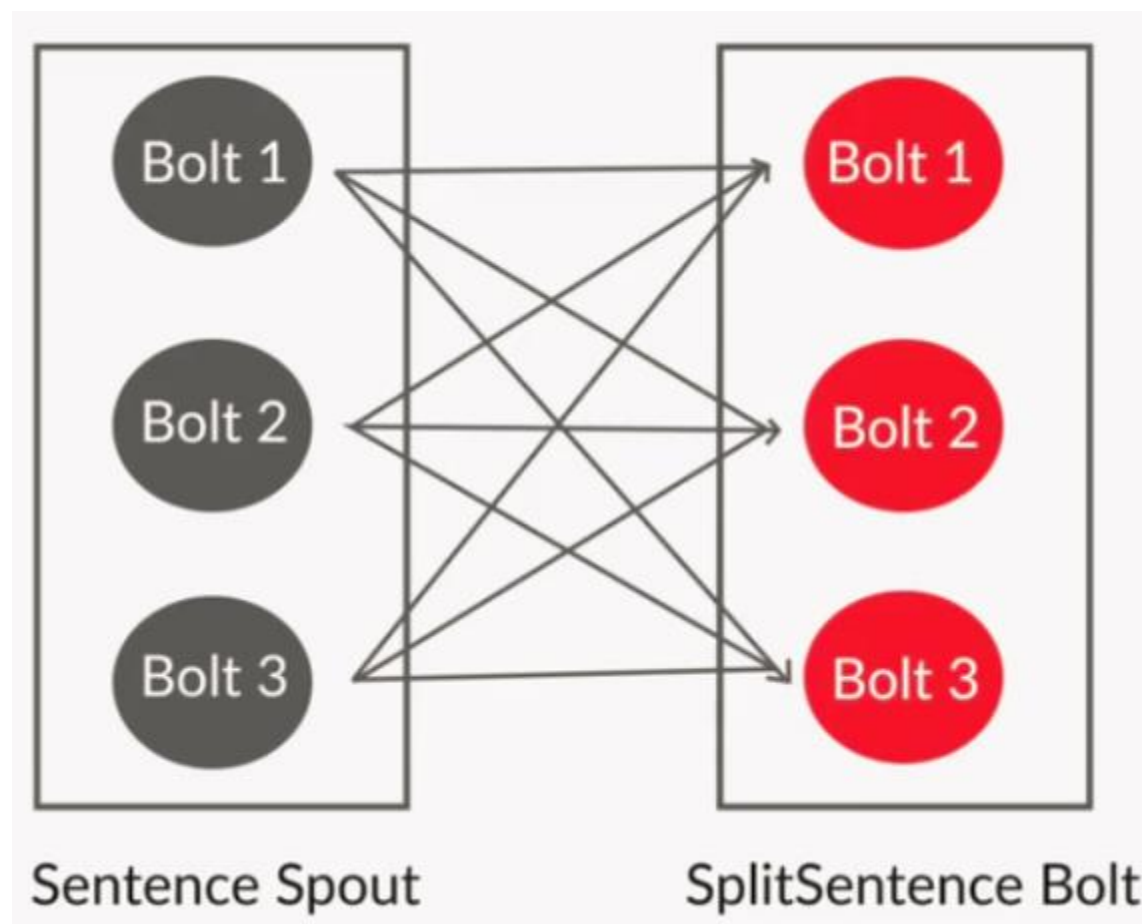


- 1) The first segment of the PDF should contain the solution to the load imbalance problem and the steps followed to ensure at least once processing of input tuples.

We are solving this by changing the grouping to shuffle groupings as in shuffle grouping storm balances the number of tuples emitted (in both the flows)

```
builder.setBolt(SPLIT_BOLT_ID, splitBolt,2)
    .shuffleGrouping(SENTENCE_SPOUT_ID);
// SplitSentenceBolt --> WordCountBolt
builder.setBolt(COUNT_BOLT_ID, countBolt,4)
    .shuffleGrouping(SPLIT_BOLT_ID);
```



The next segment of the PDF file should clearly highlight the command to run the JAR file you provided in local as well as production mode on your Cloudera instance.

Please find below the screenshots of the commands used:

~/apache-storm-1.2.1/bin/storm jar WordCount_MySQL-0.0.1-SNAPSHOT.jar WordCountTopology
"WordCountTopology_SQL2"

```
Downloads — ec2-user@ip-172-31-82-29:~ — ssh -i new_demo.pem ec2-user@ec2-18-212-181-238.compute-1.amazonaws.com — 193x30
Last login: Mon Apr 15 00:57:06 2019 from s0106bc4dfbe76b93.vc.shawcable.net
-bash: warning: setlocale: LC_CTYPE: cannot change locale (UTF-8): No such file or directory
[ec2-user@ip-172-31-82-29 ~]$ ~/apache-storm-1.2.1/bin/storm jar WordCount_MySQL-0.0.1-SNAPSHOT.jar WordCountTopology "WordCountTopology_SQL2"
Running: /usr/java/jdk1.7.0_67-cloudera/bin/java -client -Ddaemon.name= -Dstorm.options= -Dstorm.home=/home/ec2-user/apache-storm-1.2.1 -Dstorm.log.dir=/home/ec2-user/apache-storm-1.2.1/logs -D
java.library.path=/usr/local/lib:/opt/local/lib:/usr/lib -Dstorm.conf.file= -cp /home/ec2-user/apache-storm-1.2.1/lib/*:/home/ec2-user/apache-storm-1.2.1/ext
lib/*:WordCount_MySQL-0.0.1-SNAPSHOT.jar:/home/ec2-user/apache-storm-1.2.1/conf:/home/ec2-user/apache-storm-1.2.1/bin -Dstorm.jar=WordCount_MySQL-0.0.1-SNAPSHOT.jar -Dstorm.dependency.jars= -Ds
torm.dependency.artifacts={} WordCountTopology WordCountTopology_SQL2
2352 [main] WARN o.a.s.u.Utils - STORM-VERSION new 1.2.1 old null
2505 [main] INFO o.a.s.StormSubmitter - Generated ZooKeeper secret payload for MD5-digest: -5201371889635745400:-4866377062773770770
2869 [main] WARN o.a.s.u.NimbusClient - Using deprecated config nimbus.host for backward compatibility. Please update your storm.yaml so it only has config nimbus.seeds
2984 [main] INFO o.a.s.u.NimbusClient - Found leader nimbus : ip-172-31-82-29.ec2.internal:6627
3176 [main] INFO o.a.s.s.a.AuthUtils - Got AutoCreds {}
3176 [main] WARN o.a.s.u.NimbusClient - Using deprecated config nimbus.host for backward compatibility. Please update your storm.yaml so it only has config nimbus.seeds
3192 [main] INFO o.a.s.u.NimbusClient - Found leader nimbus : ip-172-31-82-29.ec2.internal:6627
3284 [main] INFO o.a.s.StormSubmitter - Uploading dependencies - jars...
3285 [main] INFO o.a.s.StormSubmitter - Uploading dependencies - artifacts...
3286 [main] INFO o.a.s.StormSubmitter - Dependency Blob keys - jars : [] / artifacts : []
3399 [main] INFO o.a.s.StormSubmitter - Uploading topology jar WordCount_MySQL-0.0.1-SNAPSHOT.jar to assigned location: /home/ec2-user/apache-storm-1.2.1/~apache-storm-1.2.1/Data/nimbus/inbox
/stormjar-f422a987-23ab-4aaf-bc42-22459d2b6b34.jar
Start uploading file 'WordCount_MySQL-0.0.1-SNAPSHOT.jar' to '/home/ec2-user/apache-storm-1.2.1/~apache-storm-1.2.1/Data/nimbus/inbox/stormjar-f422a987-23ab-4aaf-bc42-22459d2b6b34.jar' (100219
4 bytes)
[=====] 1002194 / 1002194
File 'WordCount_MySQL-0.0.1-SNAPSHOT.jar' uploaded to '/home/ec2-user/apache-storm-1.2.1/~apache-storm-1.2.1/Data/nimbus/inbox/stormjar-f422a987-23ab-4aaf-bc42-22459d2b6b34.jar' (1002194 bytes)
)
3504 [main] INFO o.a.s.StormSubmitter - Successfully uploaded topology jar to assigned location: /home/ec2-user/apache-storm-1.2.1/~apache-storm-1.2.1/Data/nimbus/inbox/stormjar-f422a987-23ab
-4aaf-bc42-22459d2b6b34.jar
3504 [main] INFO o.a.s.StormSubmitter - Submitting topology WordCountTopology_SQL2 in distributed mode with conf {"storm.zookeeper.topology.auth.scheme":"digest","storm.zookeeper.topology.auth
.payload":"-5201371889635745400:-4866377062773770770","topology.workers":3}
3504 [main] WARN o.a.s.u.Utils - STORM-VERSION new 1.2.1 old 1.2.1
3946 [main] INFO o.a.s.StormSubmitter - Finished submitting topology: WordCountTopology_SQL2
```

Mention the name of your MySQL Database and the table you used in your code to store the count of words and provide the set of steps you followed to create the same in your MySQL database.

Database used : upgrad

Table used : wordcounts

Screenshot below:

```
mysql> use upgrad;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> select * from wordcounts;
+-----+-----+
| word | count |
+-----+-----+
| a     | 1288  |
| ate   | 1310  |
| beverages | 1289  |
+-----+-----+
3 rows in set (0.00 sec)

mysql> delete from wordcounts;
Query OK, 3 rows affected (0.00 sec)

mysql> desc wordcounts;
+-----+-----+-----+-----+-----+-----+
| Field | Type | Null | Key | Default | Extra |
+-----+-----+-----+-----+-----+-----+
| word  | varchar(1024) | YES | | NULL | |
| count | bigint(20) | YES | | NULL | |
+-----+-----+-----+-----+-----+-----+
2 rows in set (0.00 sec)
```

Counts:

```
Downloads — ec2-user@ip-172-31-82-29:~ — ssh -i new_demo.pem ec2-user@ec2-18-212-181-238.compute-1.amazonaws.com — 189x27
Empty set (0.00 sec)

mysql> select * from wordcounts;
+-----+-----+
| word  | count |
+-----+-----+
| a     | 1601  |
| ate   | 1681  |
| beverages | 1704  |
| cold  | 1704  |
| cow   | 1718  |
| dog   | 3408  |
| dont  | 3211  |
| fleas | 3228  |
| has   | 1685  |
| have  | 1708  |
| homework | 1700  |
| i     | 4831  |
| like  | 3436  |
| man   | 1682  |
| my    | 3408  |
| the   | 1611  |
| think | 1684  |
+-----+-----+
17 rows in set (0.00 sec)

mysql>
```