WORD COUNT PROGRAM (YARN)

Question might come create a file in local environment, put it into hadoop and perform some operation. (We will do this in CDH)

vim good.txt (Create this file in you local environment, pwd for the path)

hdfs dfs -mkdir /user/root/medera (create a directory in hadoop)

sudo -u hdfs hadoop fs-chown root /user/root/medera (give root permission to the newly created directory)

hdfs dfs -put /root/good.txt /user/root/medera (copying file from local machine to hadoop environment)

cd /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/ (get into this directory, TAB TAB TAB if feeling lazy)

yarn jar /hadoop-mapreduce-examples.jarwordcount /user/root/medera/good.txt /user/root/good.txt\_result (This will perform word count, and it will store the result in good.txt\_result)

I don’t think any error should occur while performing this in CDH, if still some issues come, try to resolve it yourselves because you guys are really sharp :)