Understanding MAP-REDUCE

1) Download attachment in **STAGING\_AREA**

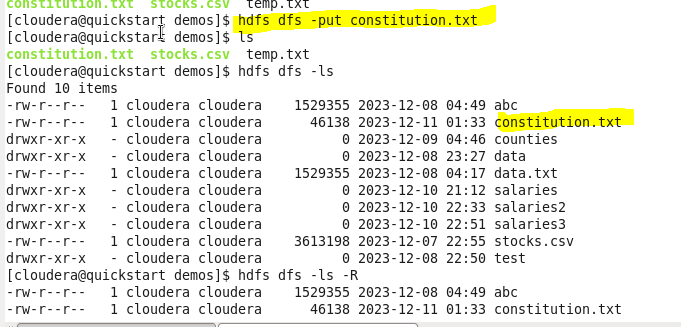


2) Pull it in **LABS\_HOME/demos**

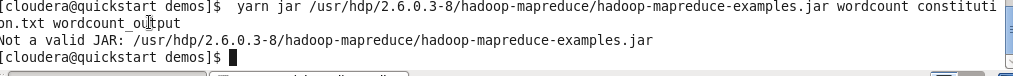


Lab Manual steps

Putting file into hdfs



Running wordcount jar

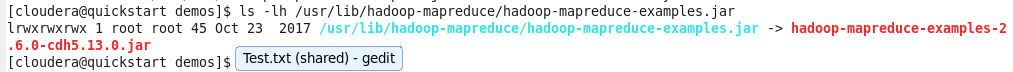


Error

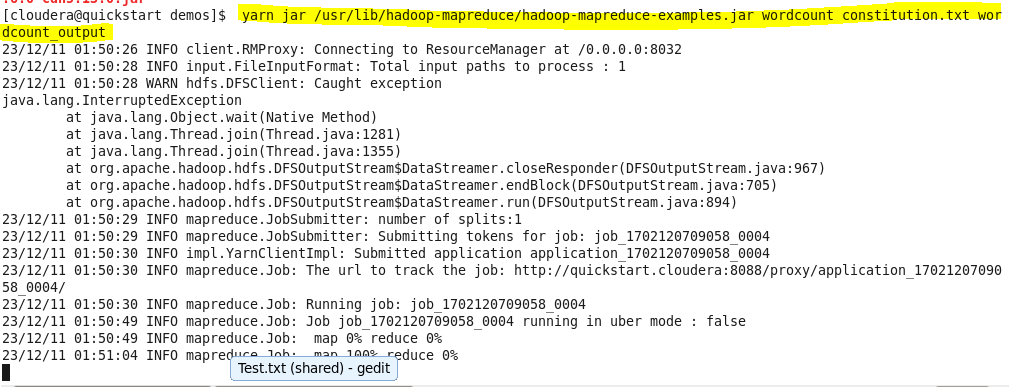
Find the file



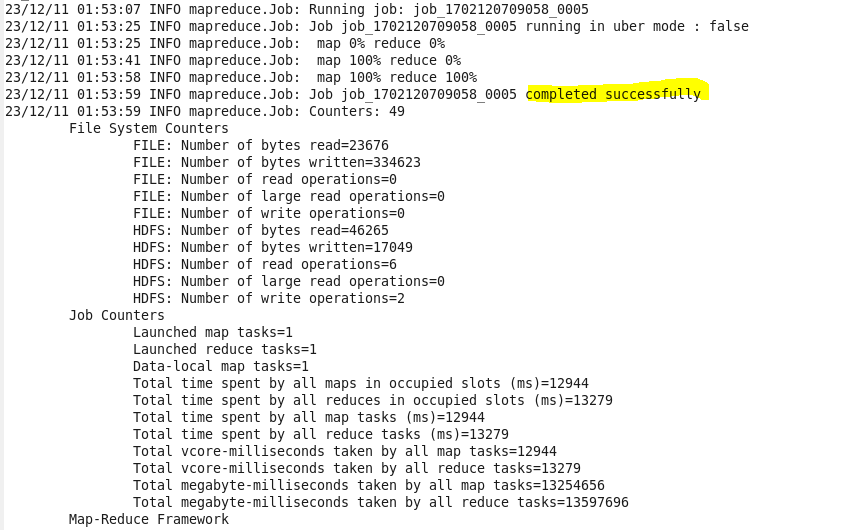
To find soft link (Blue color)



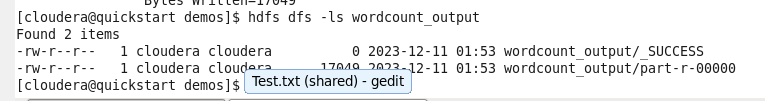
..



Delete wordcount file and then again yarn

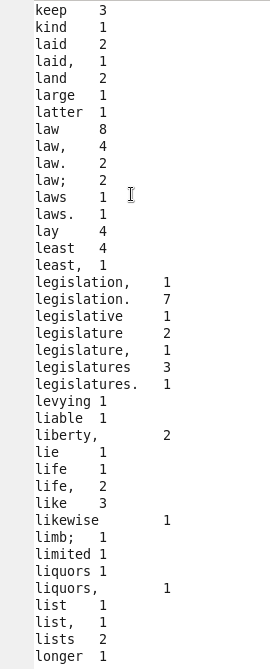


…



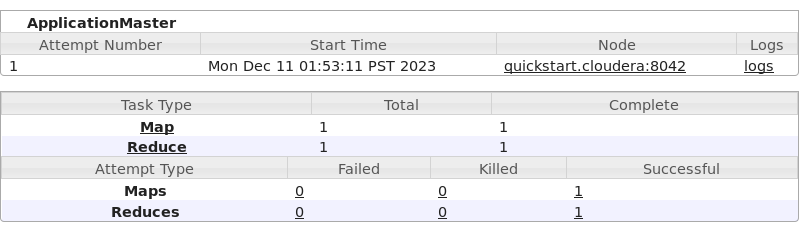
…

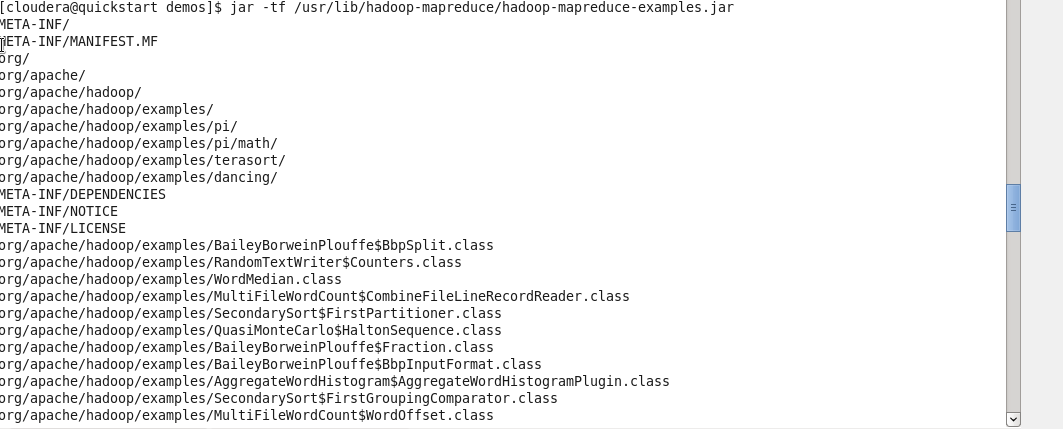


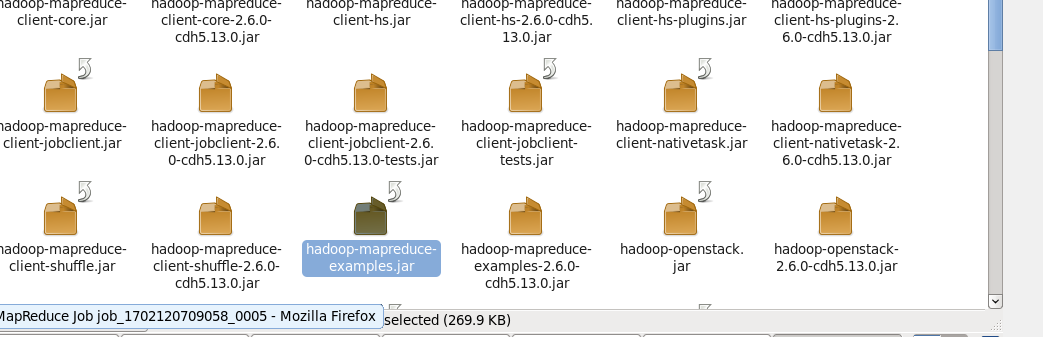


Cccc







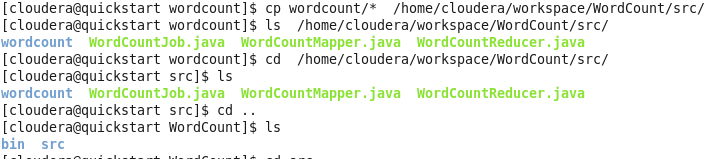


…

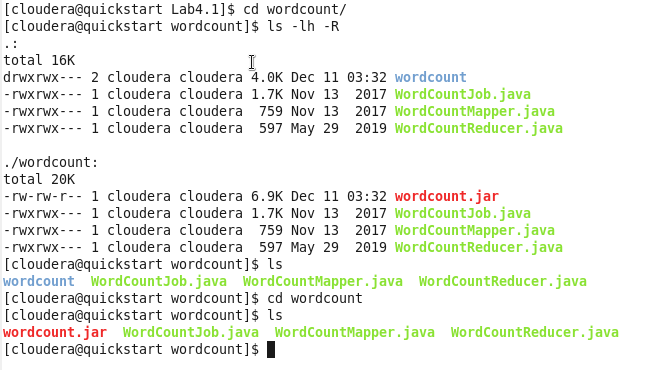


Assignment

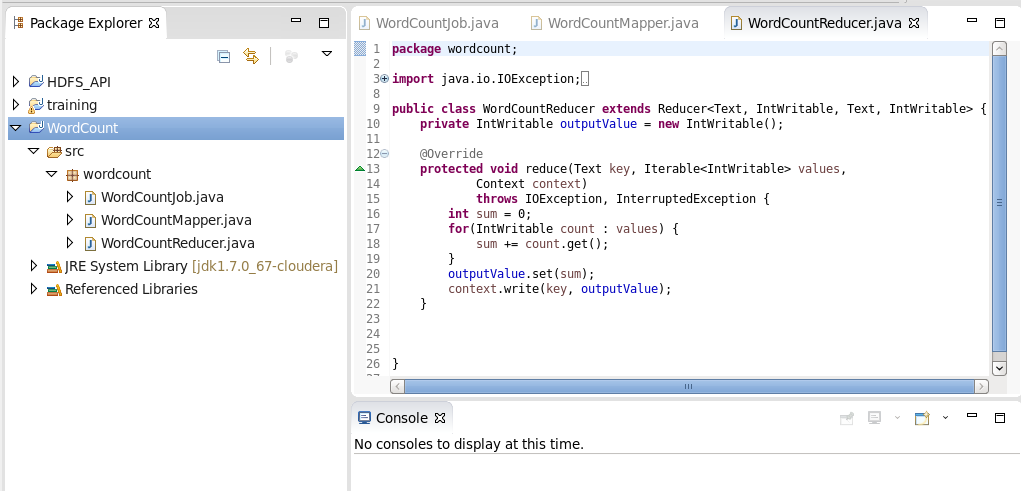
1.

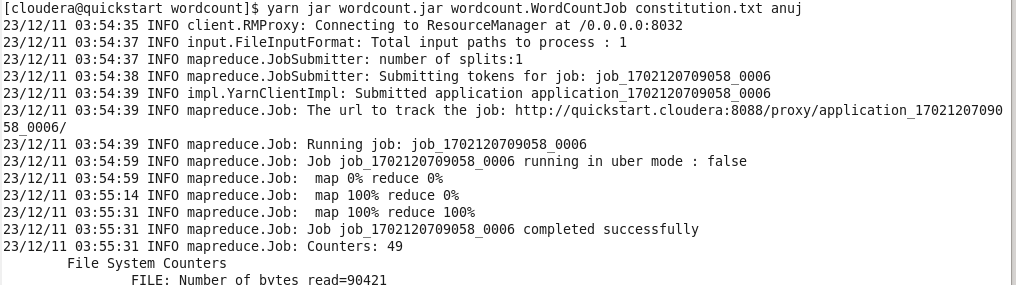


…

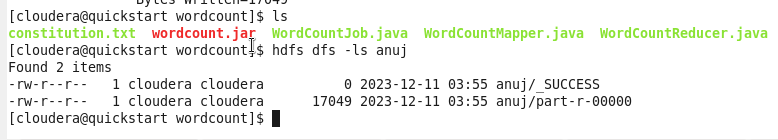


…





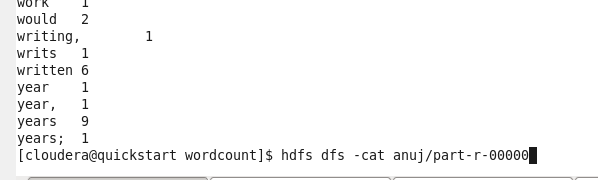
…



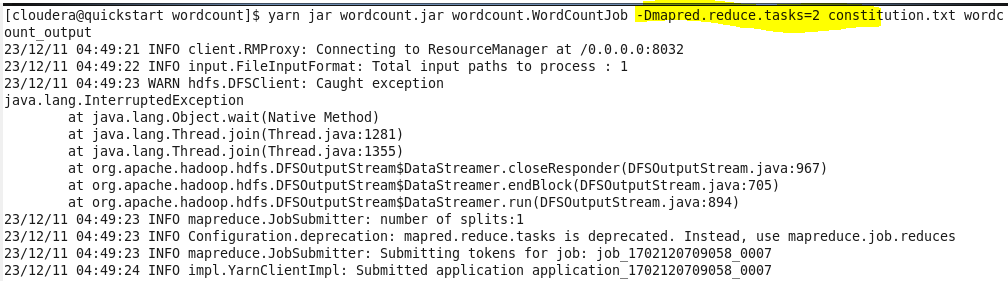
…



Here wordcount.jar is jarfile name, wordcount is package and WordCountJob is class name and constitution.txt ->input of text file name and anuj -> we put output here



Full output with map reduce like key value with words counts is shown.



Setting the configuration here, -Dmapred.reduce.tasks=2, we can increase as per our requirement and by default value is 1



