

This file is to document my experience of trying to run the wordcount application in the SCI Hadoop Cluster.

In the Hadoop Cluster:

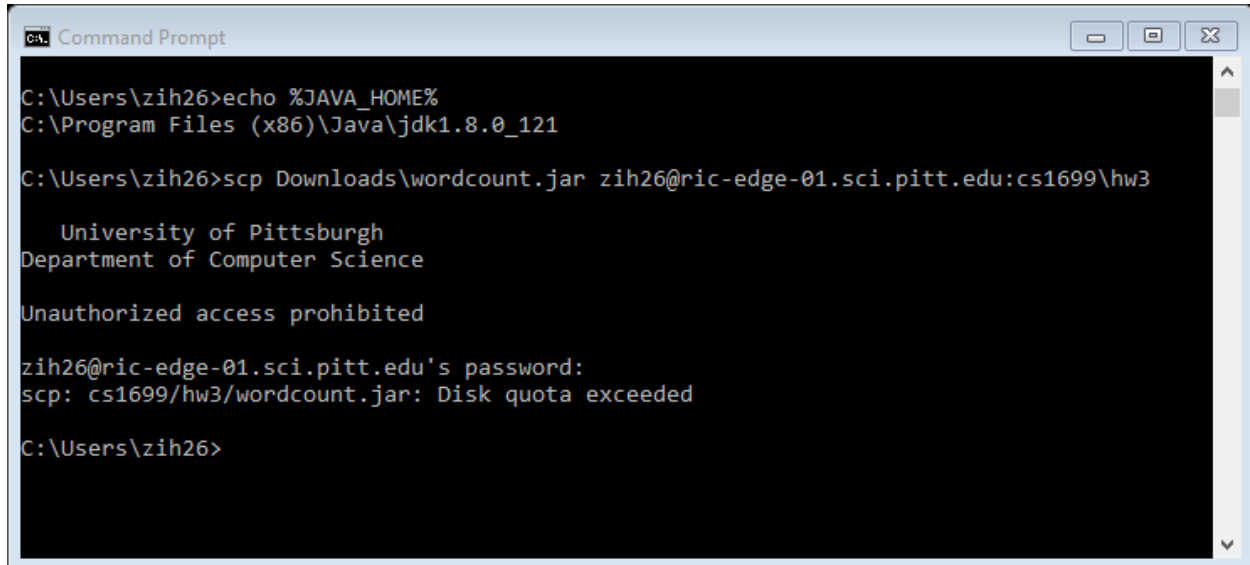
```
CA: Select OpenSSH SSH client
(1) ric-edge-01 $ ls
Backup bin cs1699 News perl5 private public
(2) ric-edge-01 $ cd cs1699
(3) ric-edge-01 $ ls
hw3 wordcount
(4) ric-edge-01 $ cd hw3
(5) ric-edge-01 $ export JAVA_HOME=/usr/local/jdk1.8.0_101
(6) ric-edge-01 $ export PATH=${JAVA_HOME}/bin:${PATH}
(7) ric-edge-01 $ export HADOOP_CLASSPATH=/opt/cloudera/parcels/CDH/lib/hadoop/hadoop-common.jar:/opt/cloudera/parcels/CDH/lib/hadoop-mapreduce/hadoop-mapreduce-client-core.jar
(8) ric-edge-01 $ ls
testData wordCount
(9) ric-edge-01 $ cd wordCount
(10) ric-edge-01 $ ls
WordCount2.class WordCount2$TokenizerMapper.class
WordCount2$IntSumReducer.class WordCount2$TokenizerMapper$CountersEnum.class
(11) ric-edge-01 $ cd ..
(12) ric-edge-01 $ cd testData
(13) ric-edge-01 $ ls
10MB.txt
(14) ric-edge-01 $ cd ..
(15) ric-edge-01 $ hadoop fs -put testData/ .
/opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/bin/./lib/hadoop/bin/hadoop: line 22: cd: .
/opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/bin/./lib/hadoop/bin
.
/usr/local/jdk1.8.0_101
/afs/pitt.edu/home/z/i/zih26
/usr/patch
/
/usr
/usr/local
/usr/pitt
/usr/contrib
/usr/afsws
/usr/andrew
/usr/bin
/usr/openwin
/opt/SUNWspr
/
/usr
/usr/ccs
/usr: No such file or directory
put: `testData/10MB.txt': File exists
```

```
CA: OpenSSH SSH client
(18) ric-edge-01 $ jar cvf wordcount.jar -C wordCount/ .
java.io.FileNotFoundException: wordcount.jar (Disk quota exceeded)
    at java.io.FileOutputStream.open0(Native Method)
    at java.io.FileOutputStream.open(FileOutputStream.java:270)
    at java.io.FileOutputStream.<init>(FileOutputStream.java:213)
    at java.io.FileOutputStream.<init>(FileOutputStream.java:101)
    at sun.tools.jar.Main.run(Main.java:195)
    at sun.tools.jar.Main.main(Main.java:1288)
(19) ric-edge-01 $
```

After I upload the 10MB file to the cluster, I got the error: Disk quota exceeded

After that, no other actions can be done

For example, when I was trying to upload the wordcount.jar to SCI Hadoop Cluster:



```
Command Prompt

C:\Users\zih26>echo %JAVA_HOME%
C:\Program Files (x86)\Java\jdk1.8.0_121

C:\Users\zih26>scp Downloads\wordcount.jar zih26@ric-edge-01.sci.pitt.edu:cs1699\hw3

University of Pittsburgh
Department of Computer Science

Unauthorized access prohibited

zih26@ric-edge-01.sci.pitt.edu's password:
scp: cs1699/hw3/wordcount.jar: Disk quota exceeded

C:\Users\zih26>
```