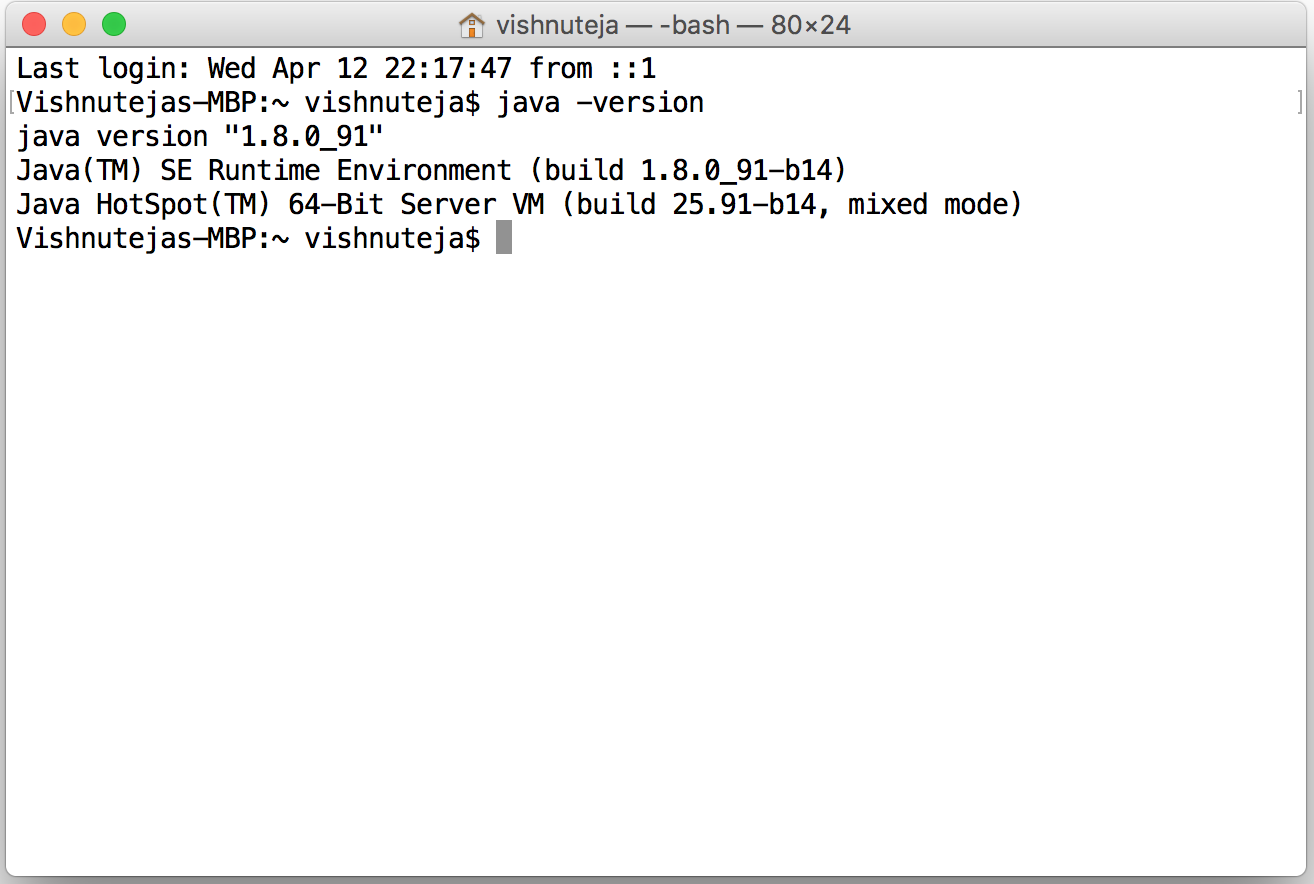
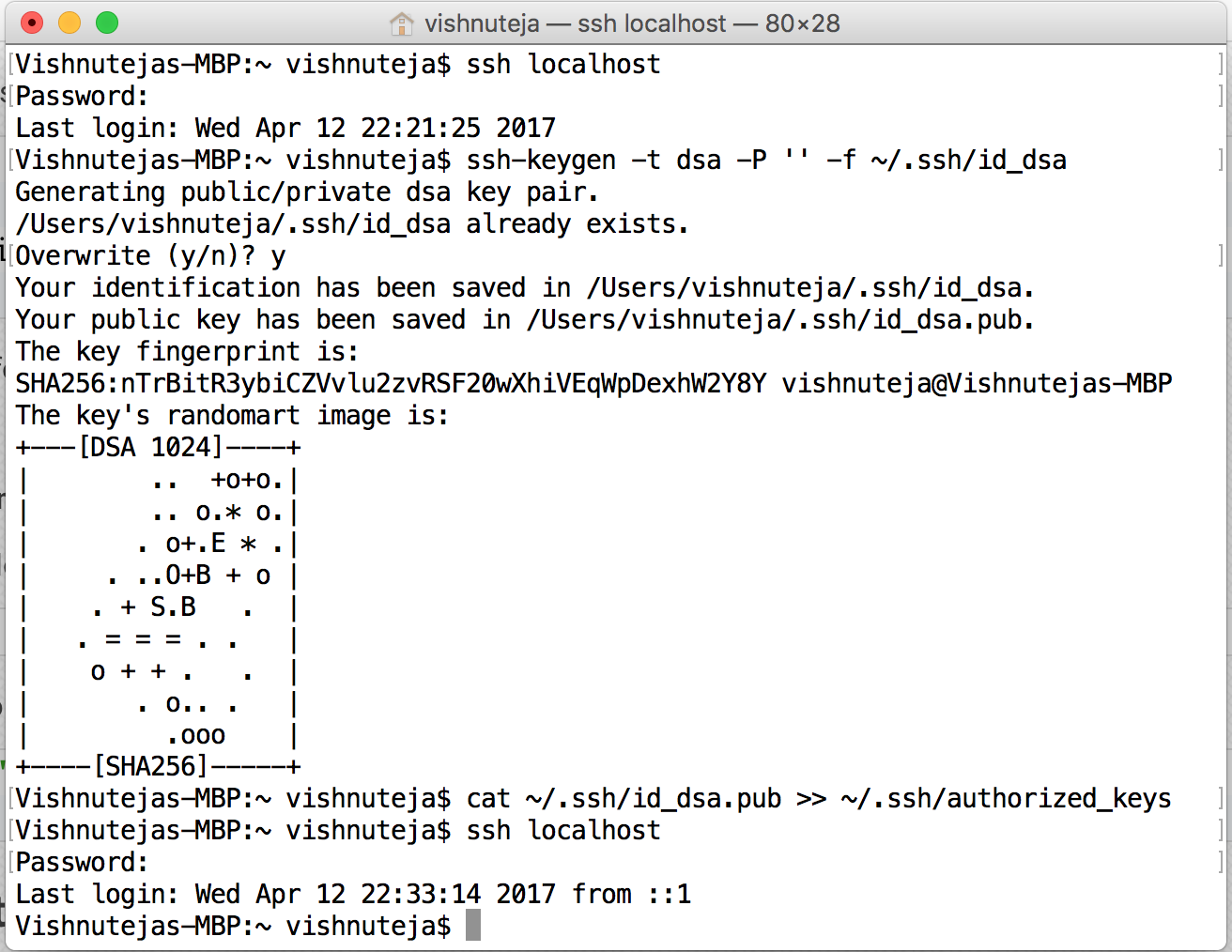
1. Checking whether Java is installed and identifying the version.



1. Enabling SSH to connect to LOCALHOST

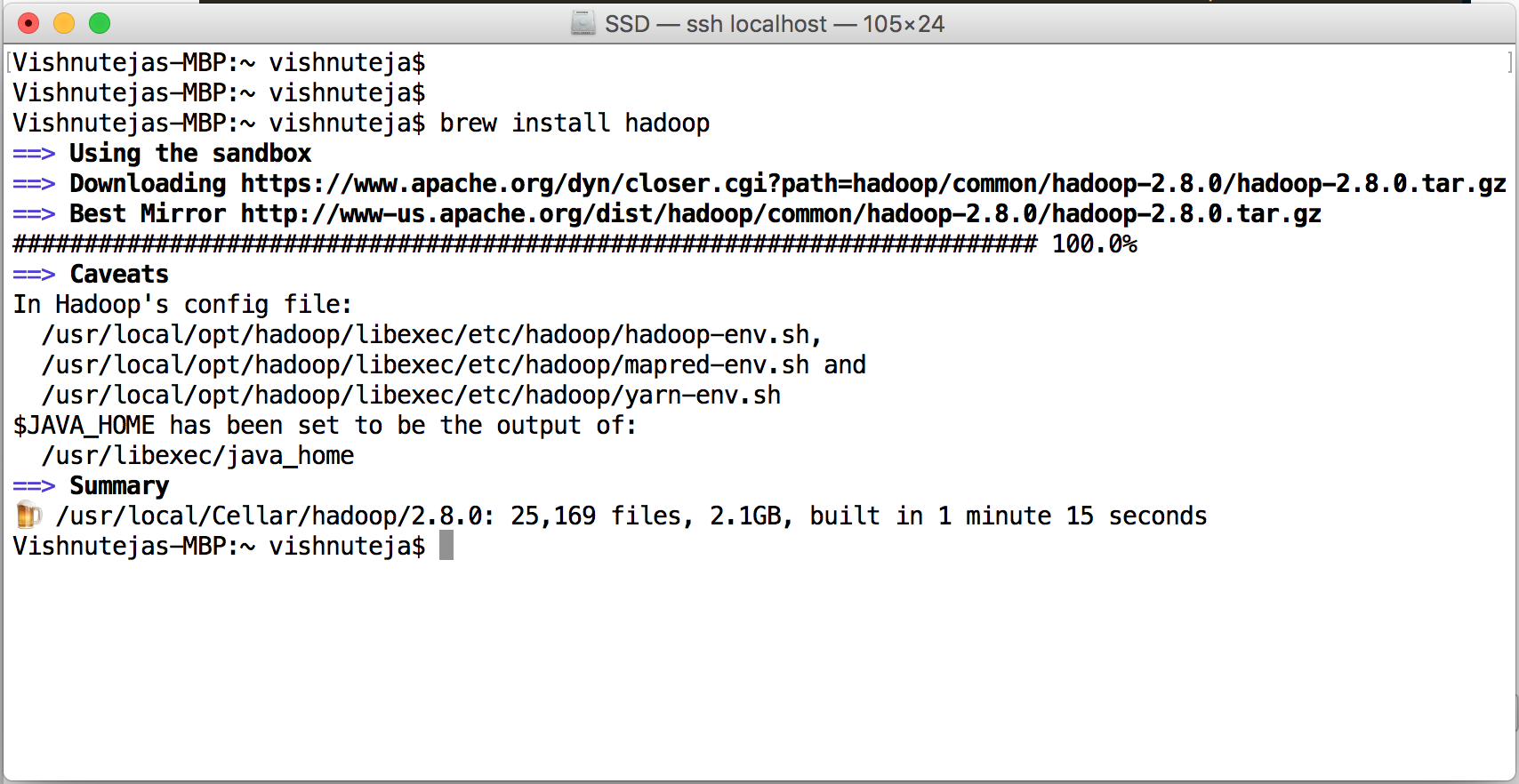


1. Install Homebrew.

$ ruby -e "$(curl –fsSL <https://raw.github.com/mxcl/homebrew/go/install>)"

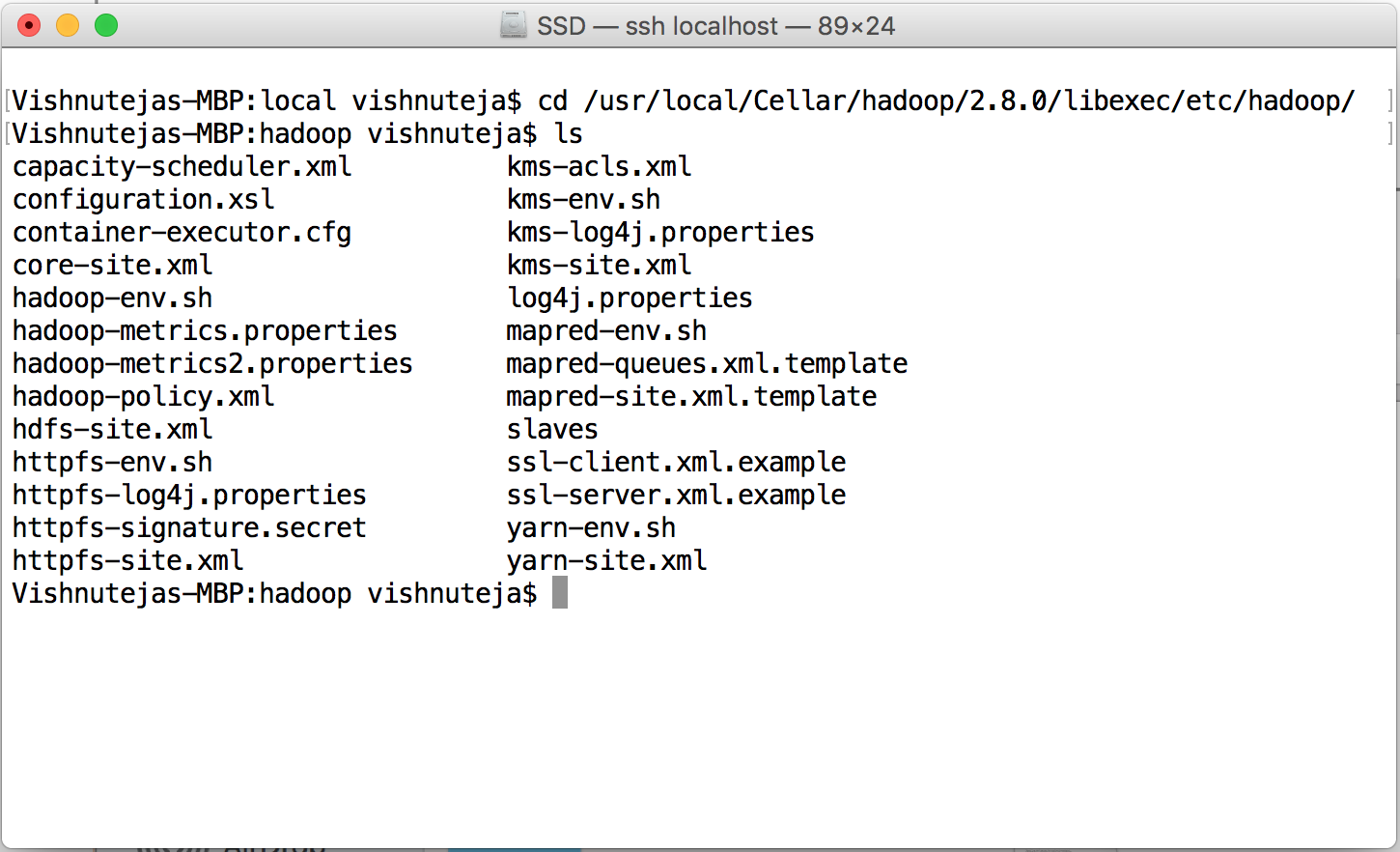


1. Installing Hadoop

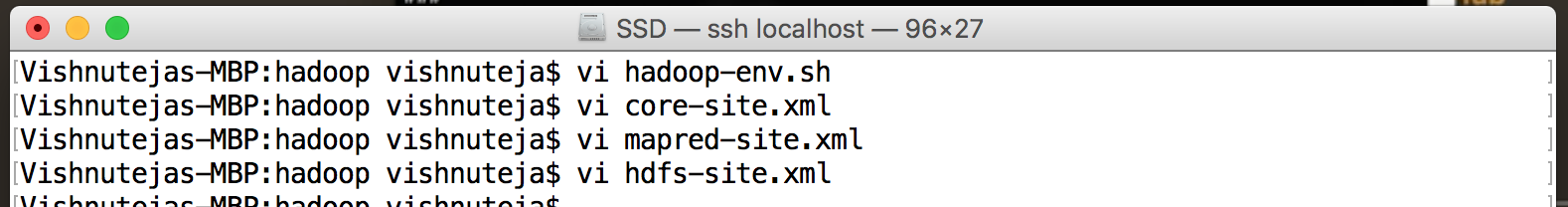


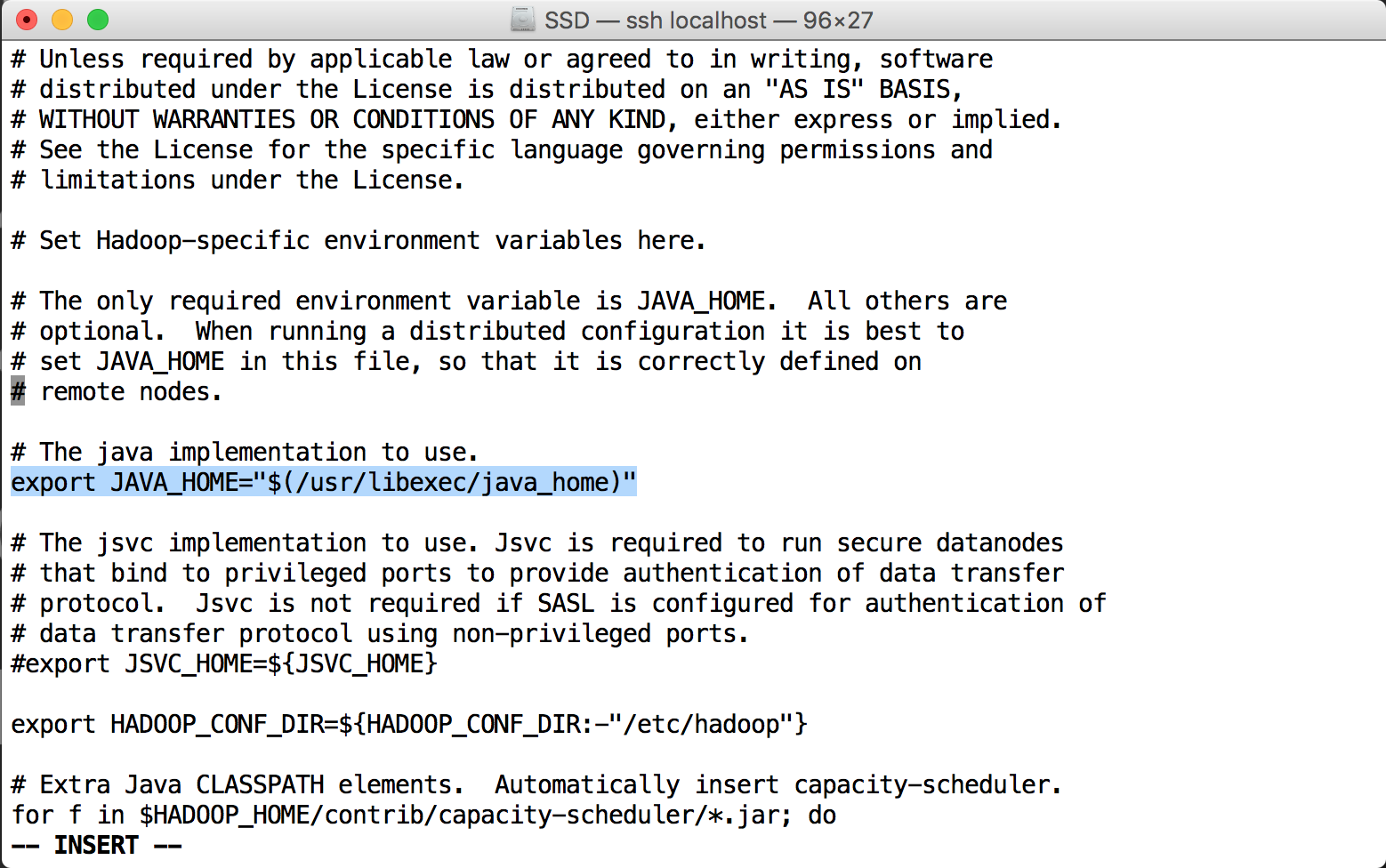
Hadoop version 2.8.0 had been installed.

Listing out the files inside the Hadoop directory



Editing the configuration files using Vi editor





**core-site.xml**

<configuration>

<property>

<name>hadoop.tmp.dir</name>

<value>/usr/local/Cellar/hadoop/hdfs/tmp</value>

<description>A base for other temporary directories.</description>

</property>

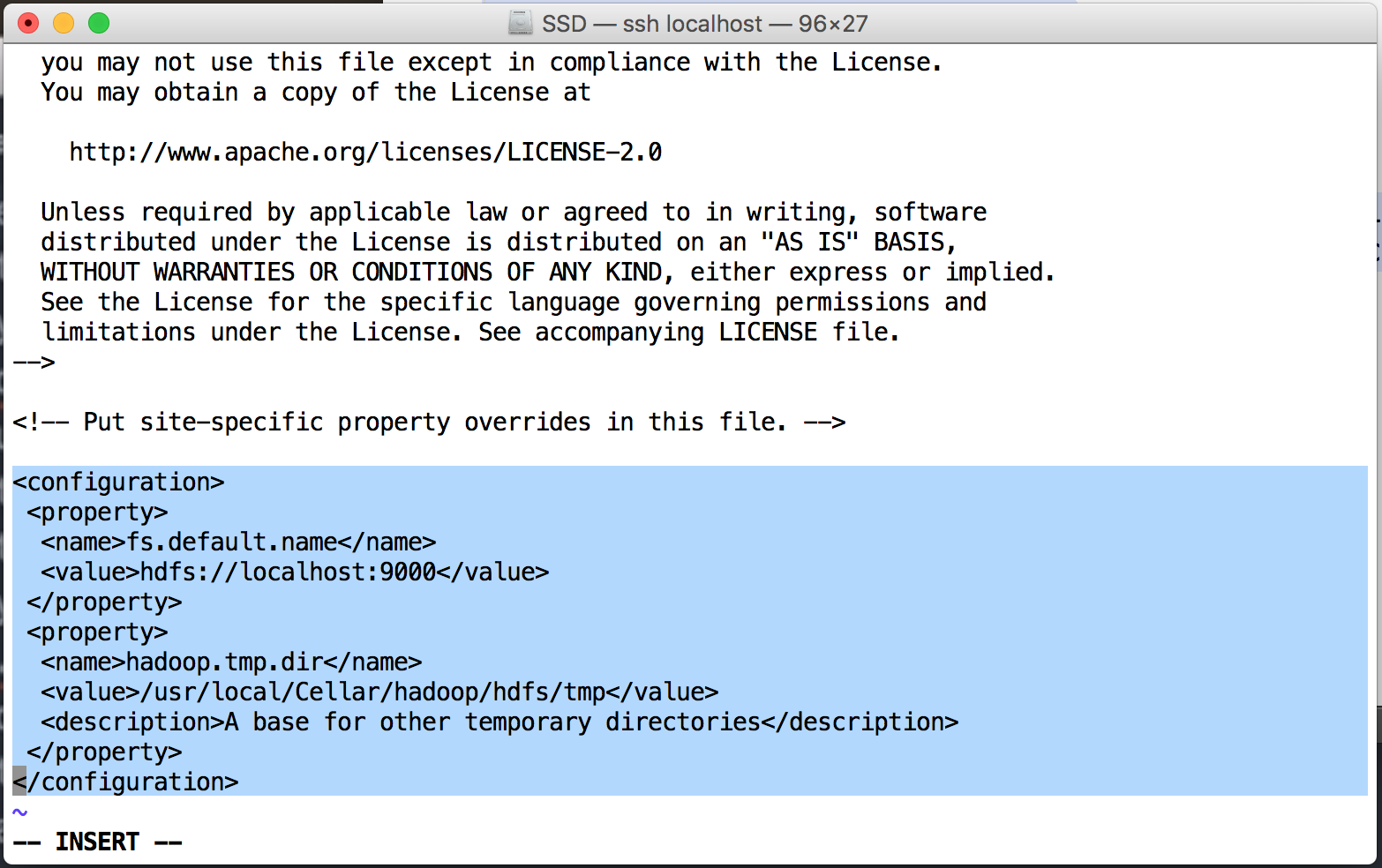
<property>

<name>fs.default.name</name>

<value>hdfs://localhost:9000</value>

</property>

</configuration>



**mapred-site.xml**

<configuration>

<property>

<name>mapred.job.tracker</name>

<value>localhost:9010</value>

</property>

<property>

<name>mapred.tasktracker.map.tasks.maximum</name>

<value>1</value>

</property>

<property>

<name>mapred.tasktracker.reduce.tasks.maximum</name>

<value>1</value>

</property>

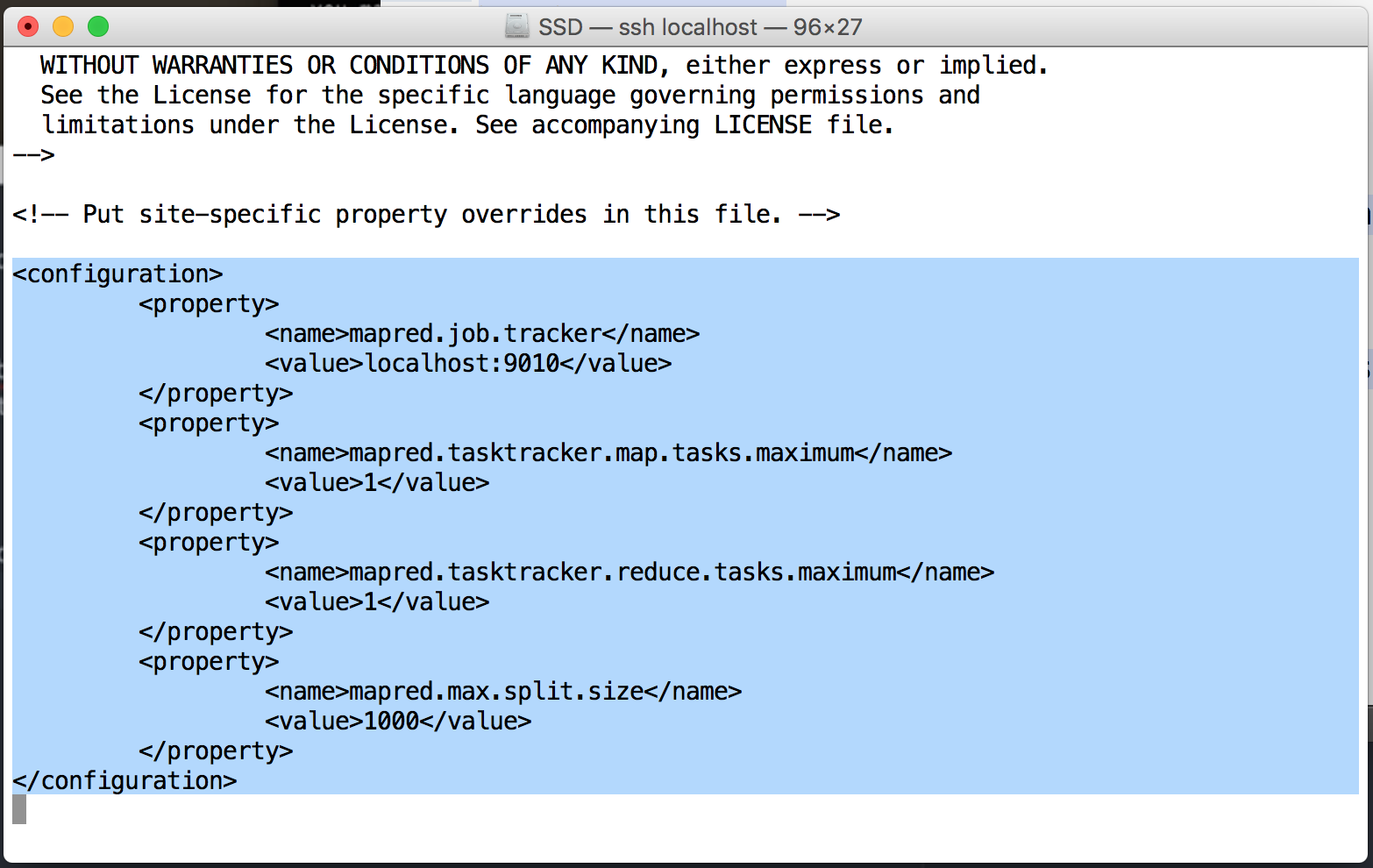
<property>

<name>mapred.max.split.size</name>

<value>1000</value>

</property>

</configuration>



**hdfs-site.xml**

<configuration>

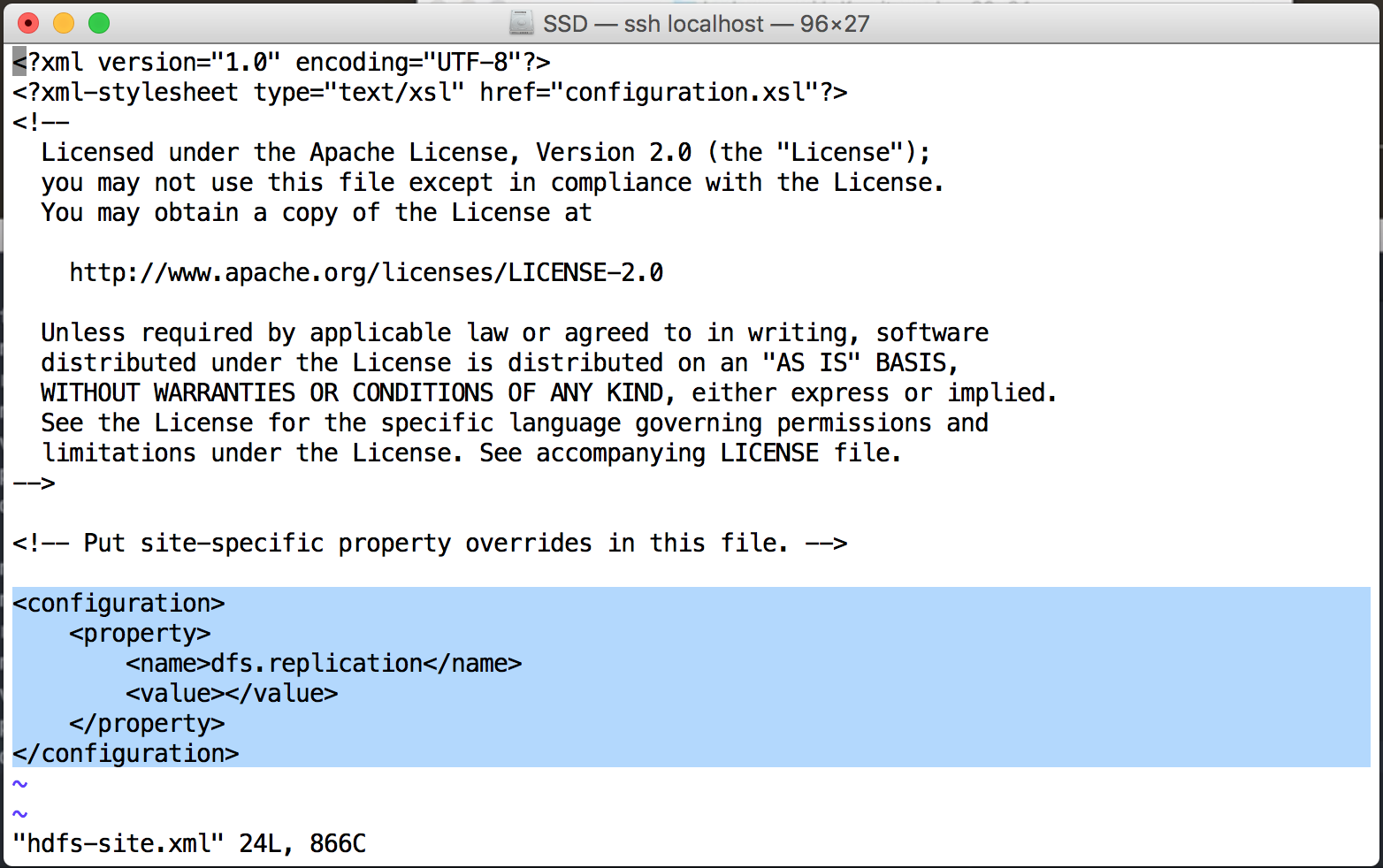
<property>

<name>dfs.replication</name>

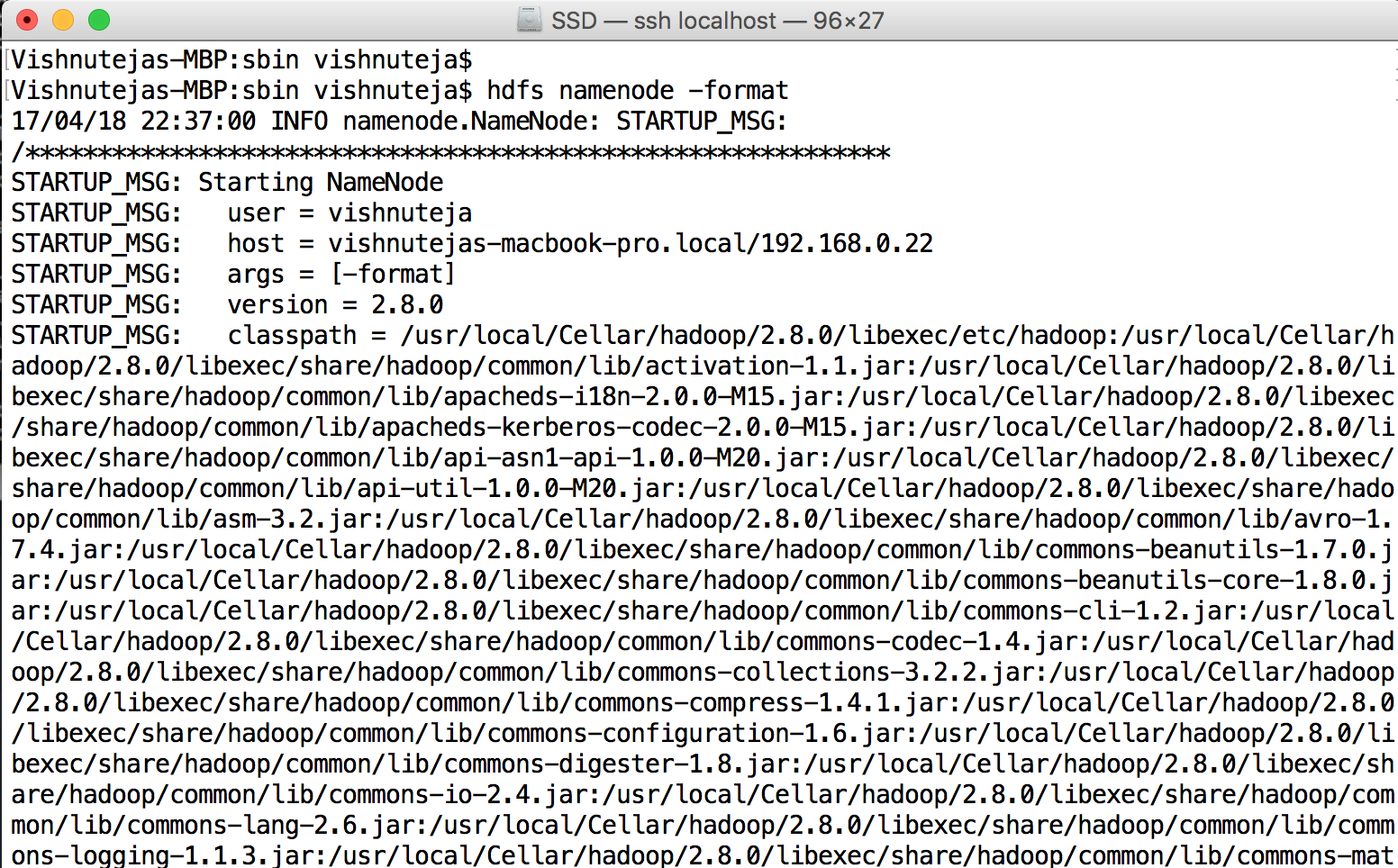
<value>1</value>

</property>

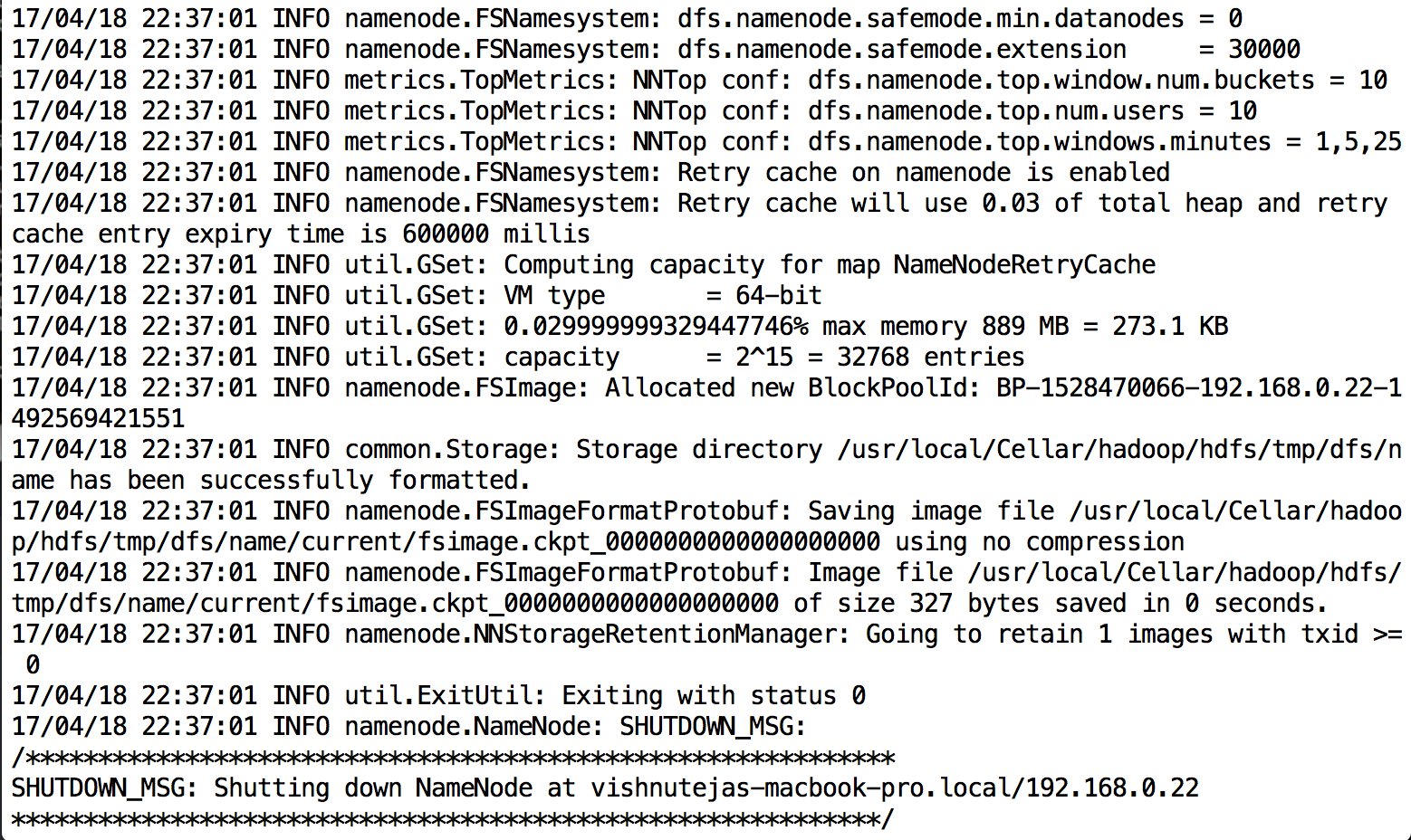
</configuration>



Format Hadoop filesystem

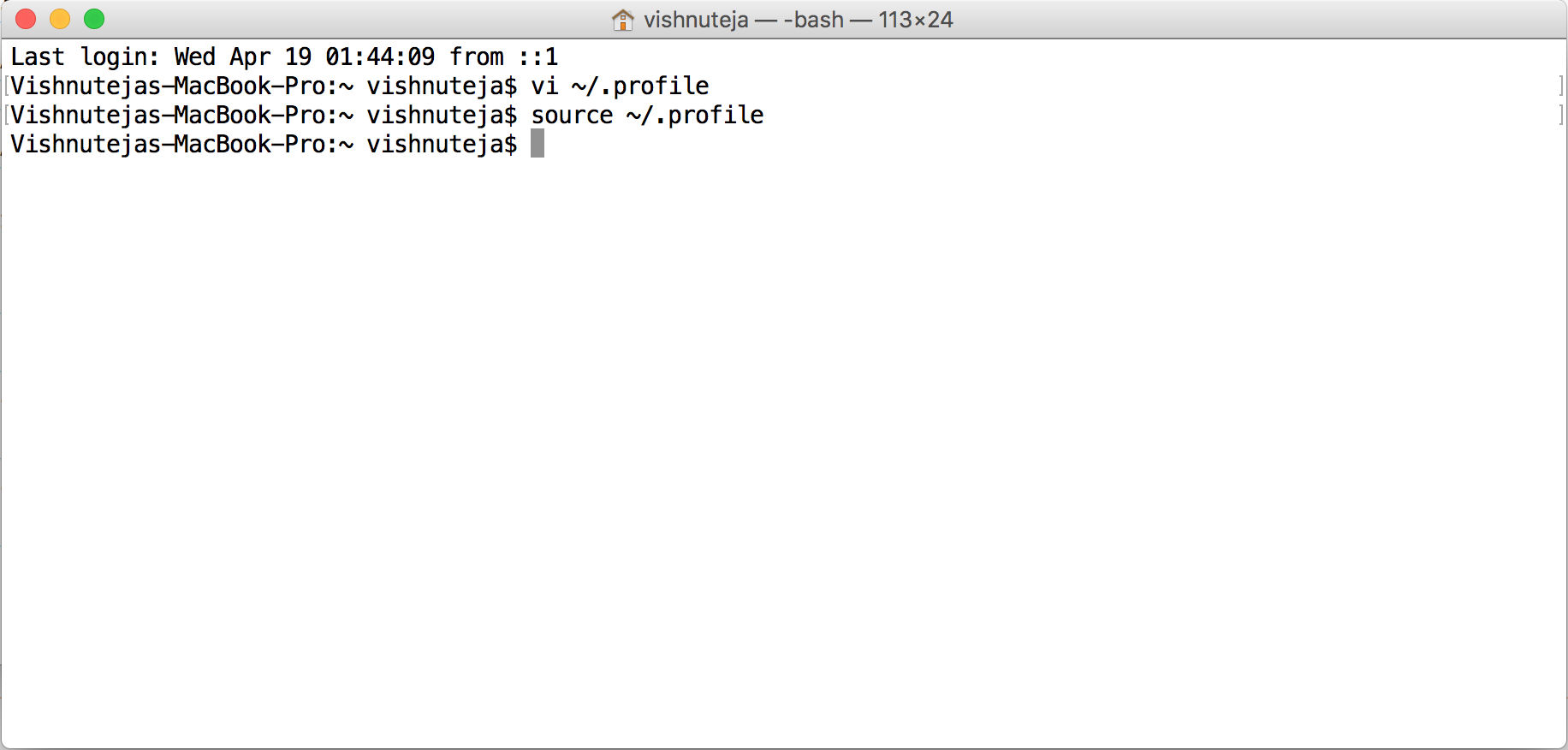


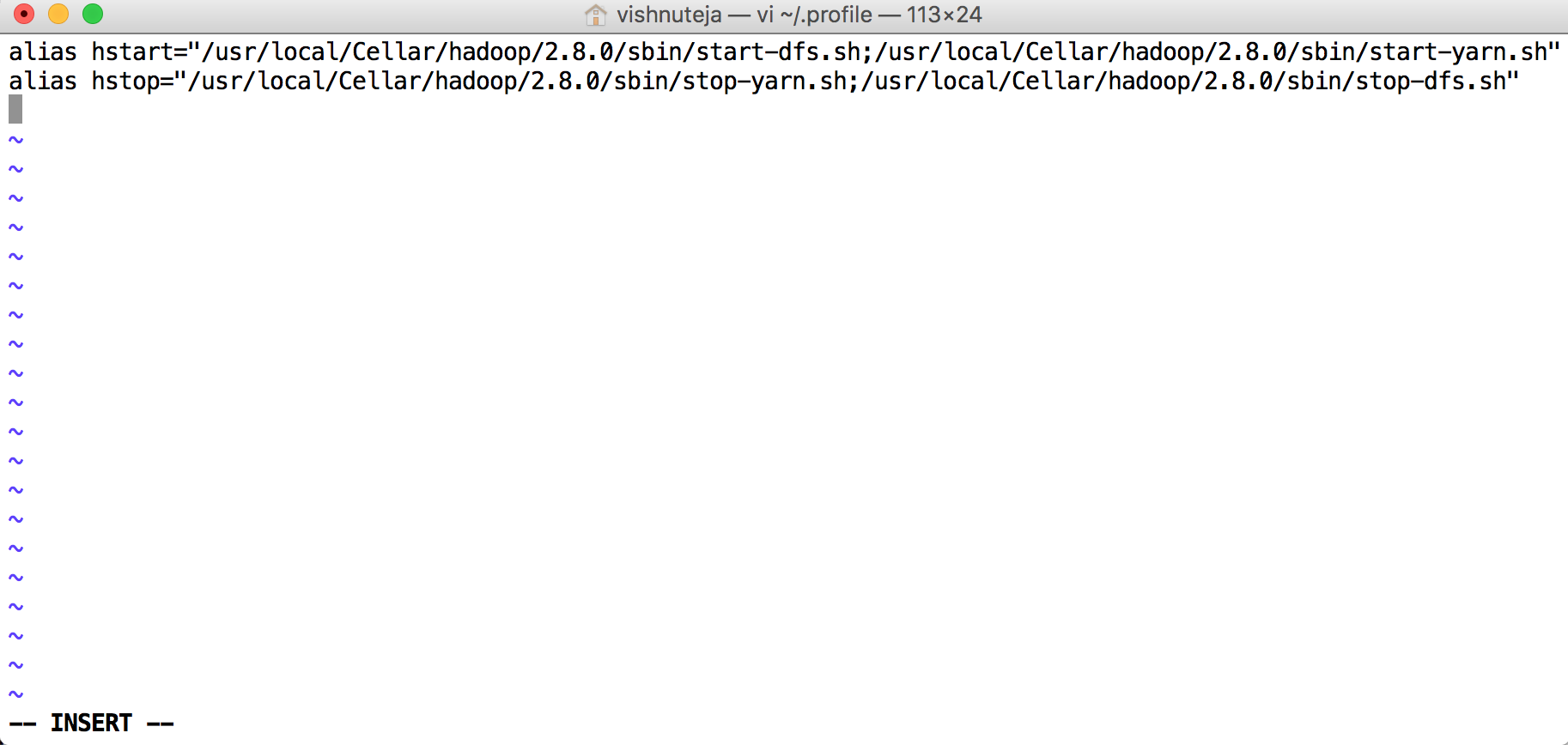
…

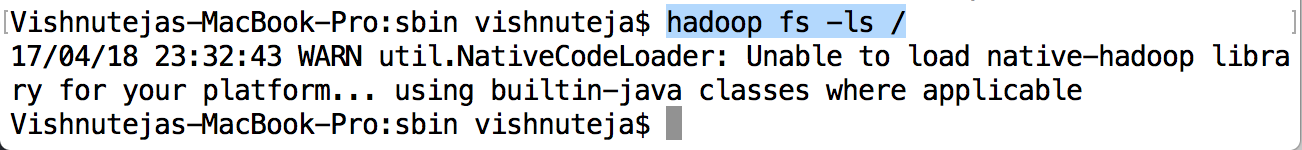


alias hstart="/usr/local/Cellar/hadoop/2.8.0/sbin/start-dfs.sh;/usr/local/Cellar/hadoop/2.8.0/sbin/start-yarn.sh"

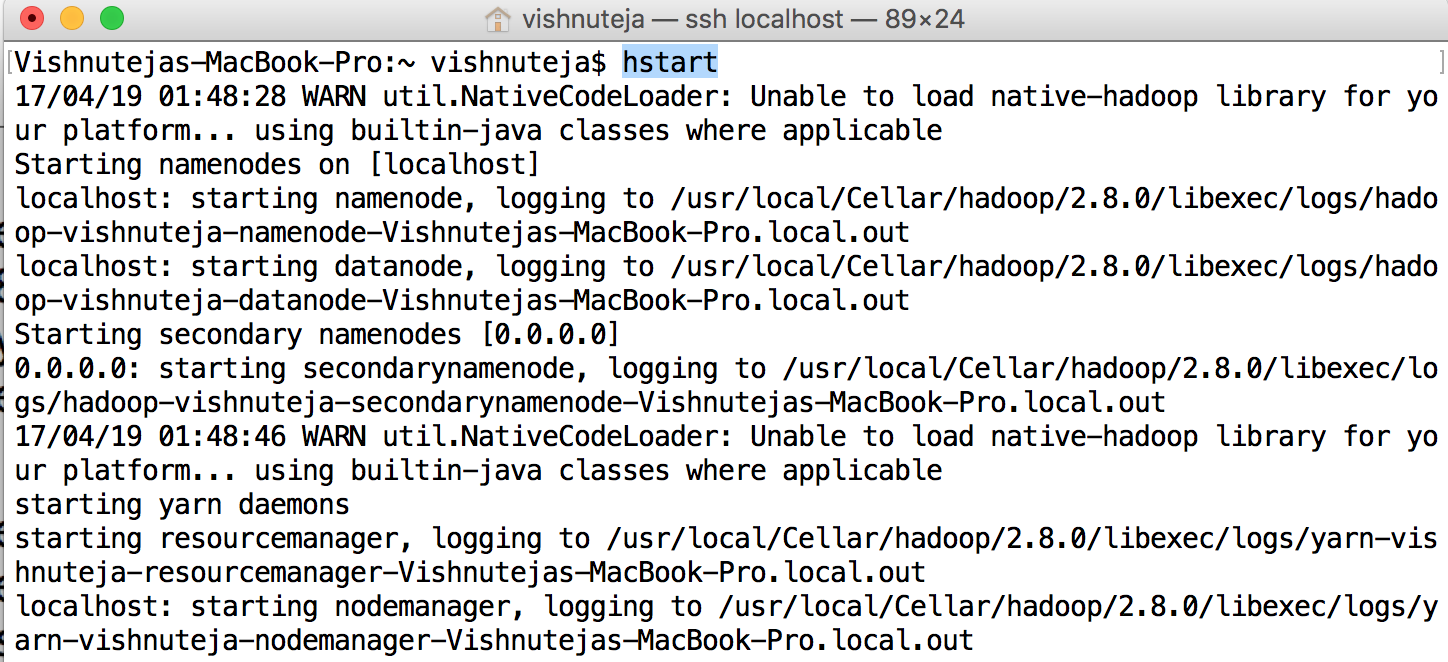
alias hstop="/usr/local/Cellar/hadoop/2.8.0/sbin/stop-yarn.sh;/usr/local/Cellar/hadoop/2.8.0/sbin/stop-dfs.sh"



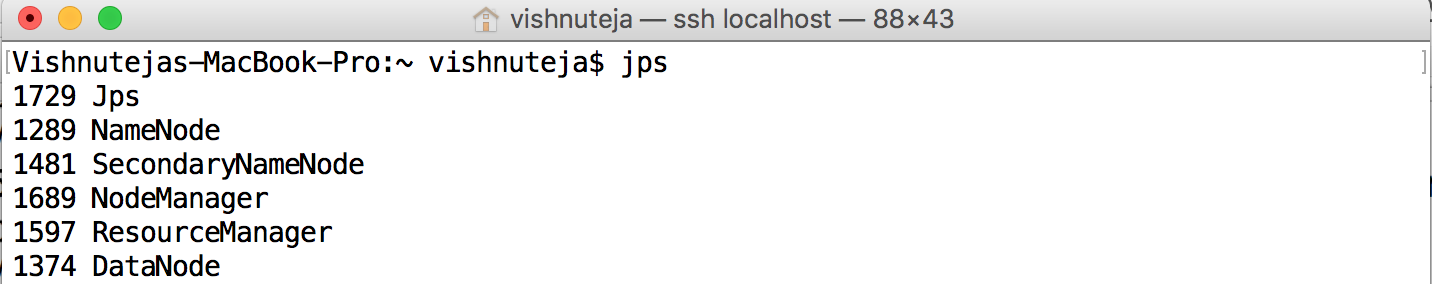


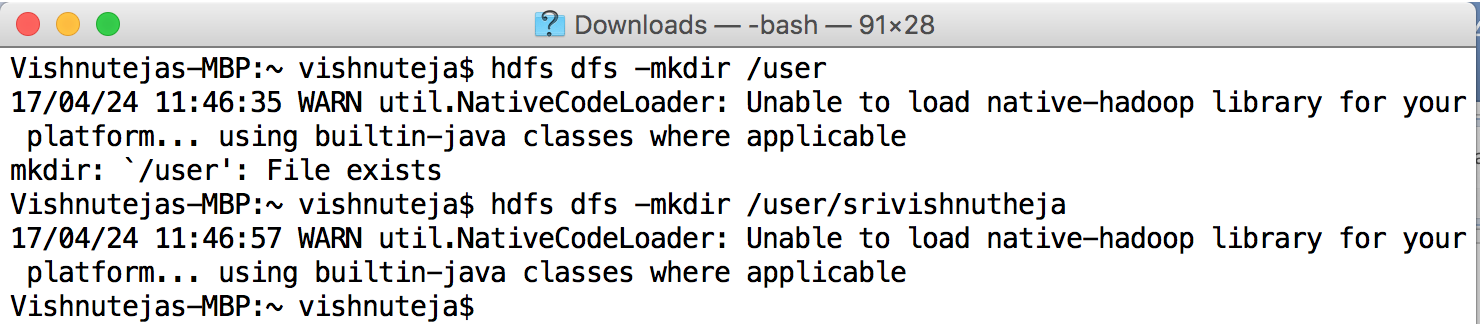


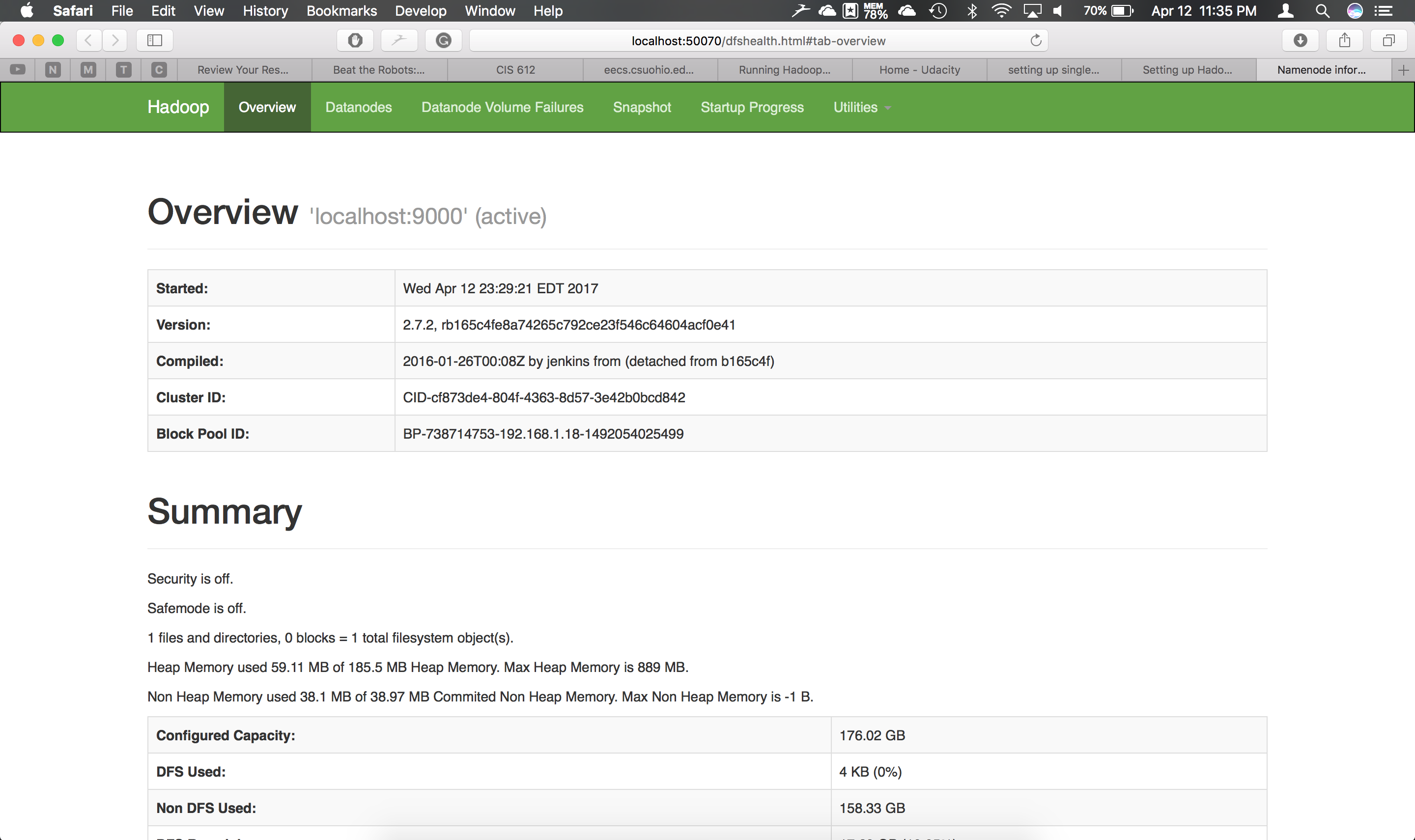
Start Hadoop

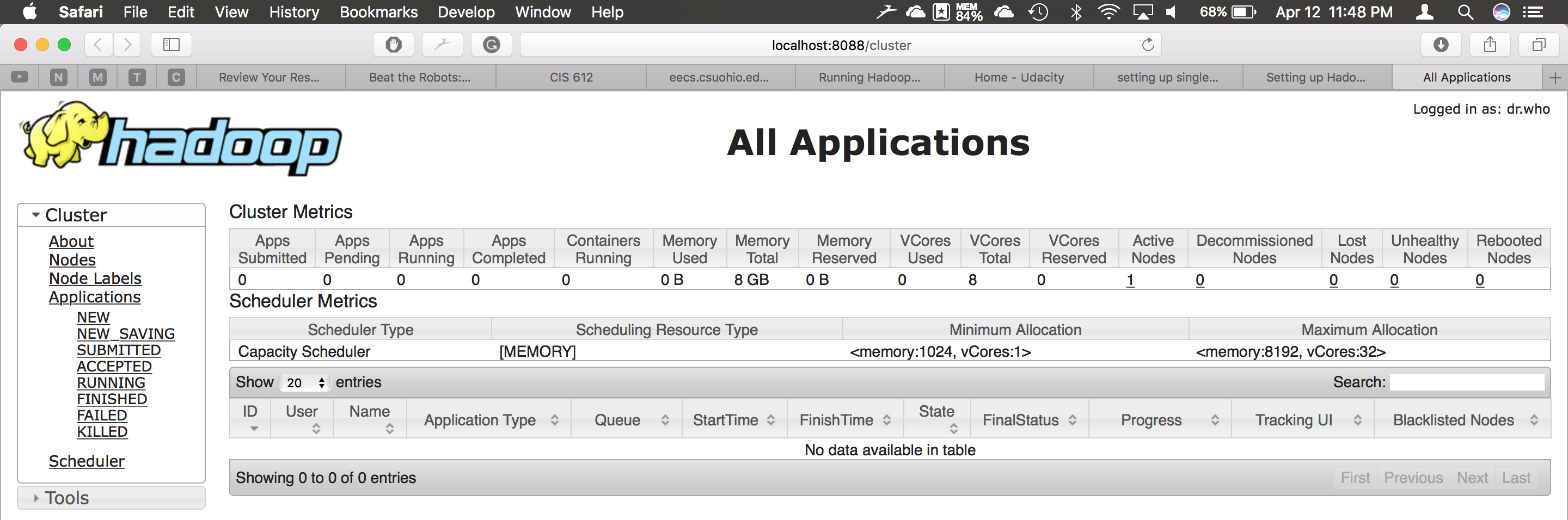


Making sure that all the Hadoop process are running

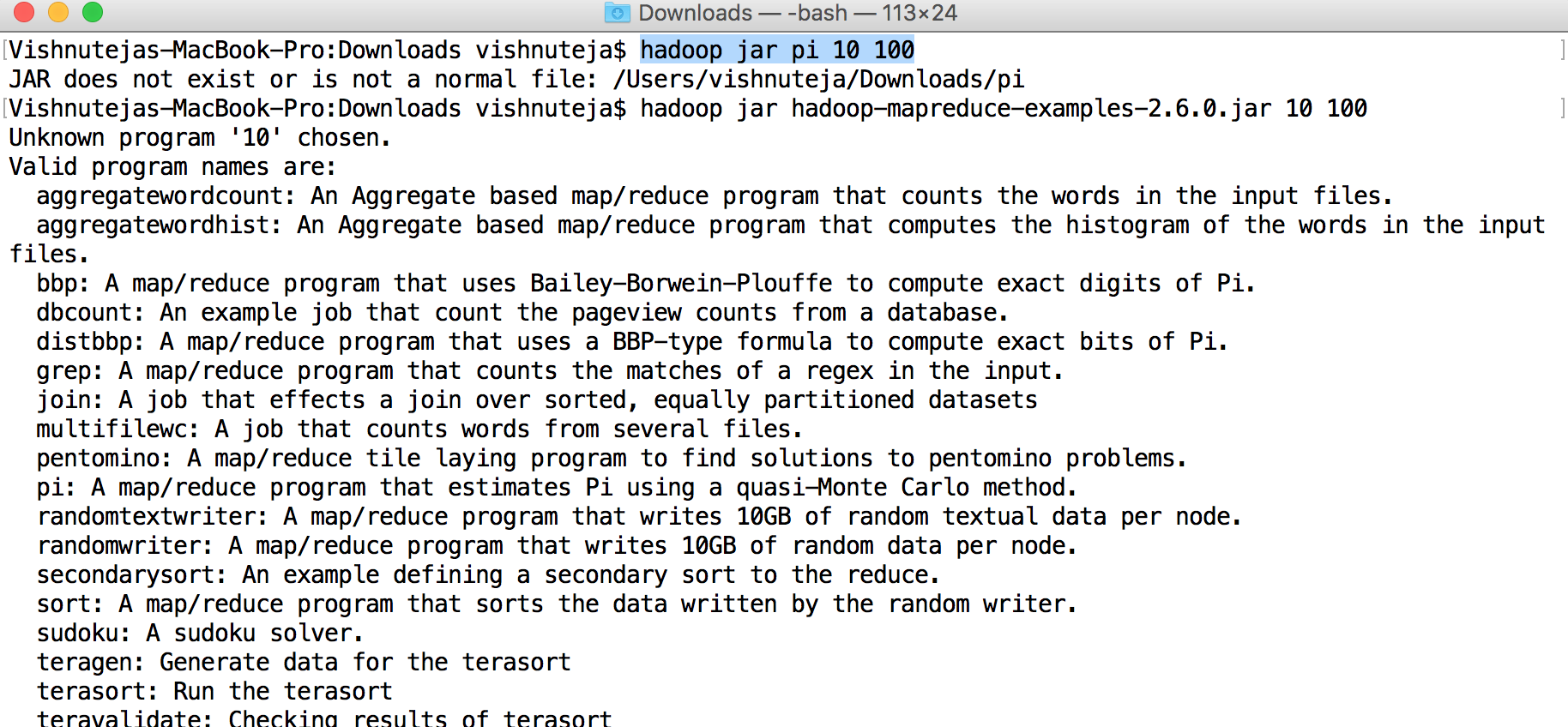




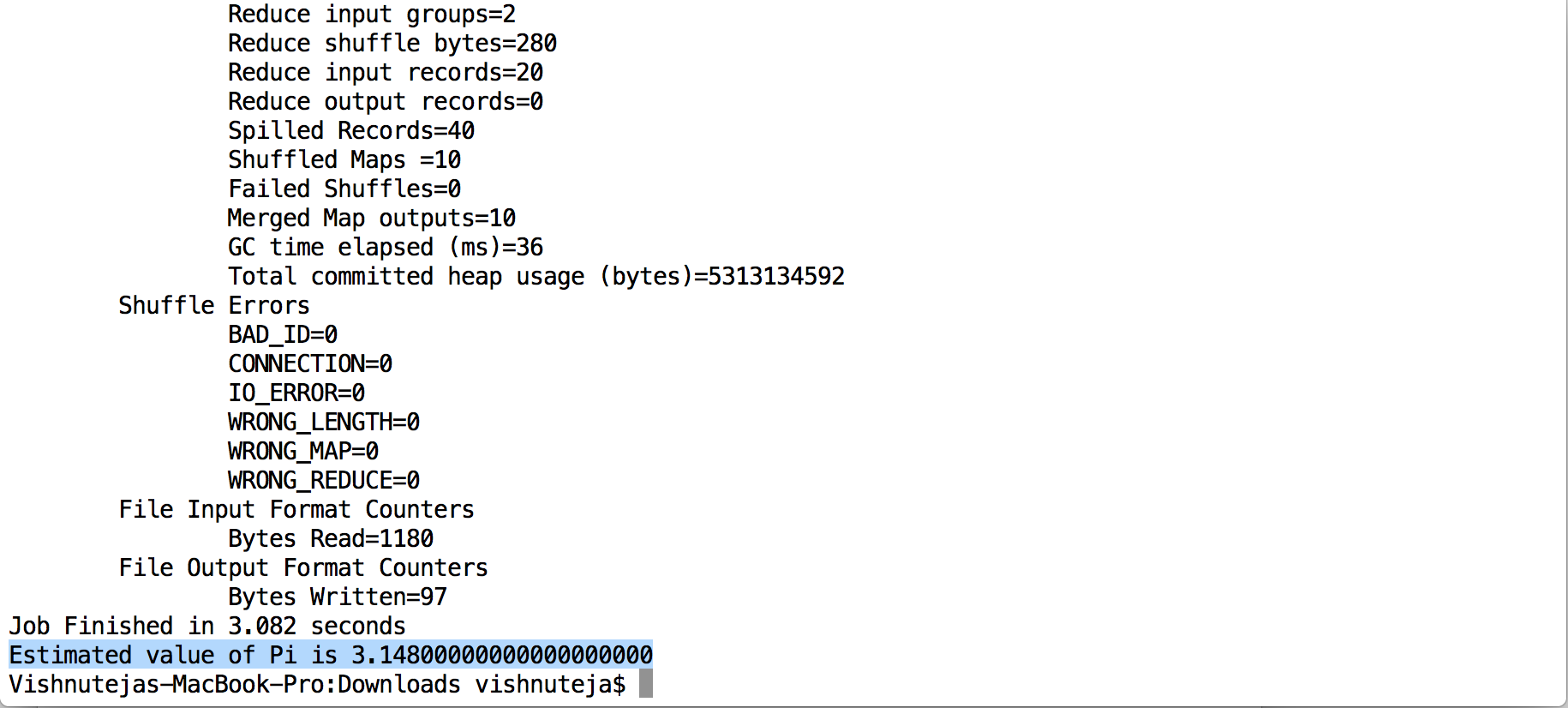




Running a Hadoop Example

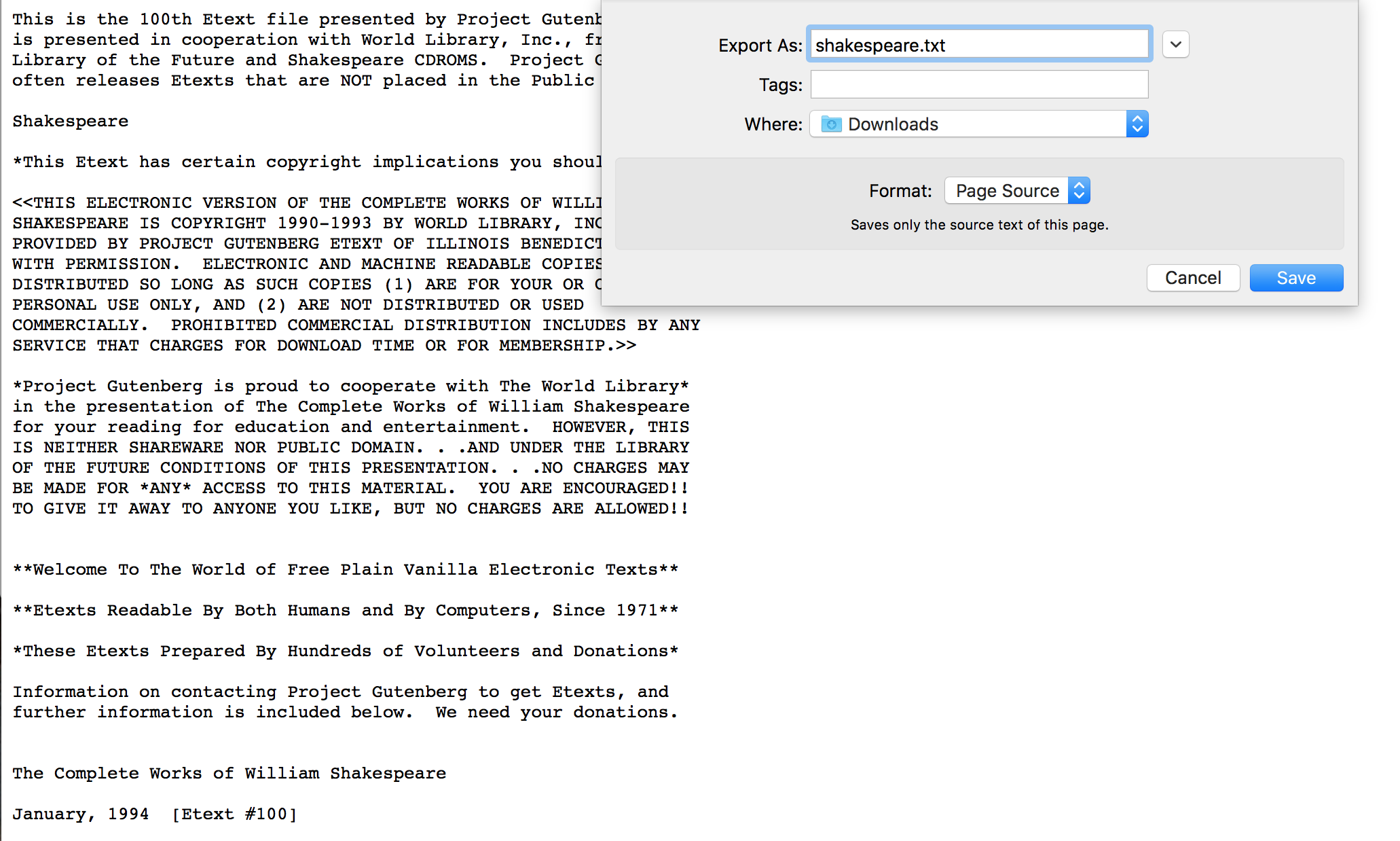


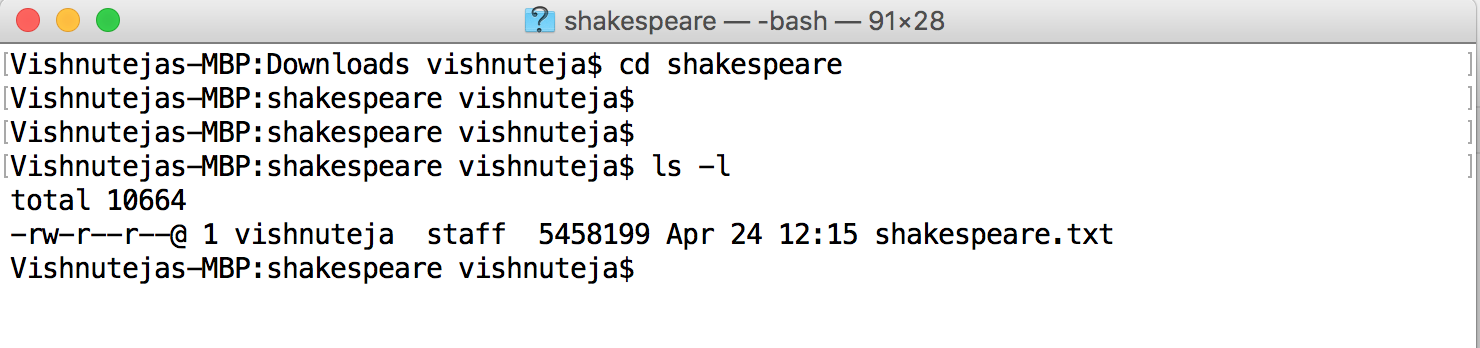
…….



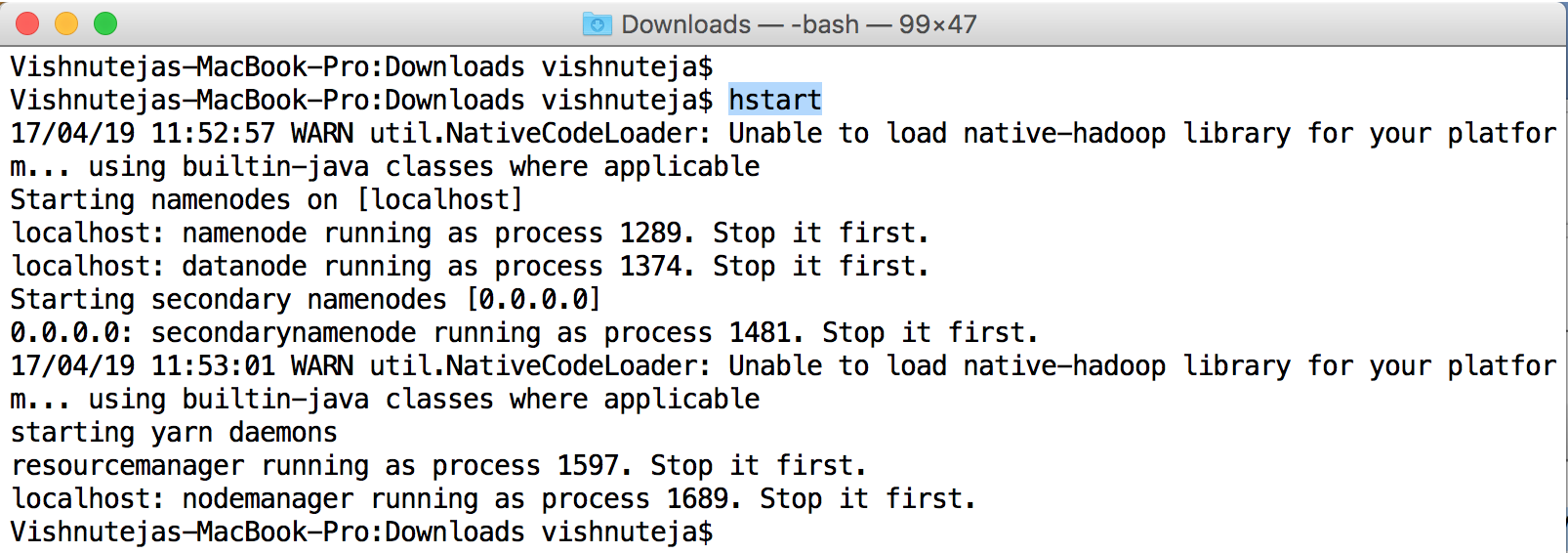
Downloaded Shakespeare work <https://ocw.mit.edu/ans7870/6/6.006/s08/lecturenotes/files/t8.shakespeare.txt>

into /Downloads/Shakespeare

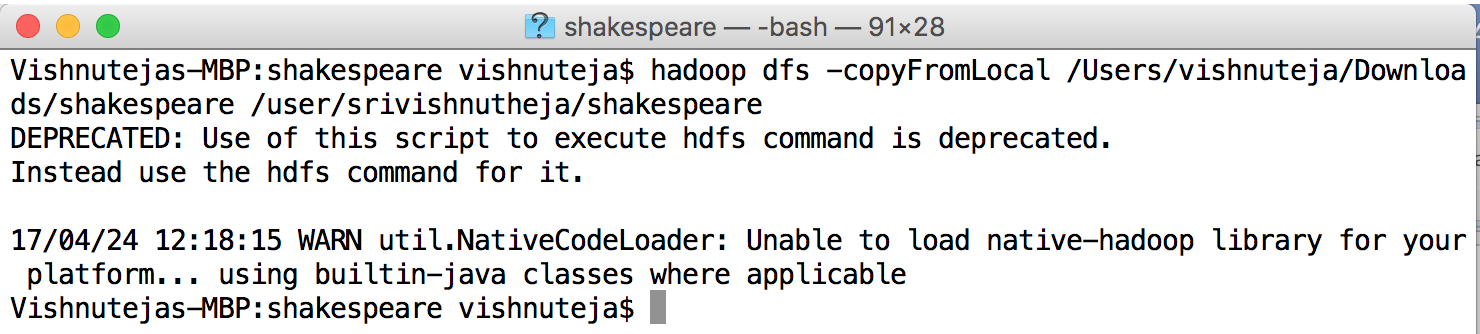


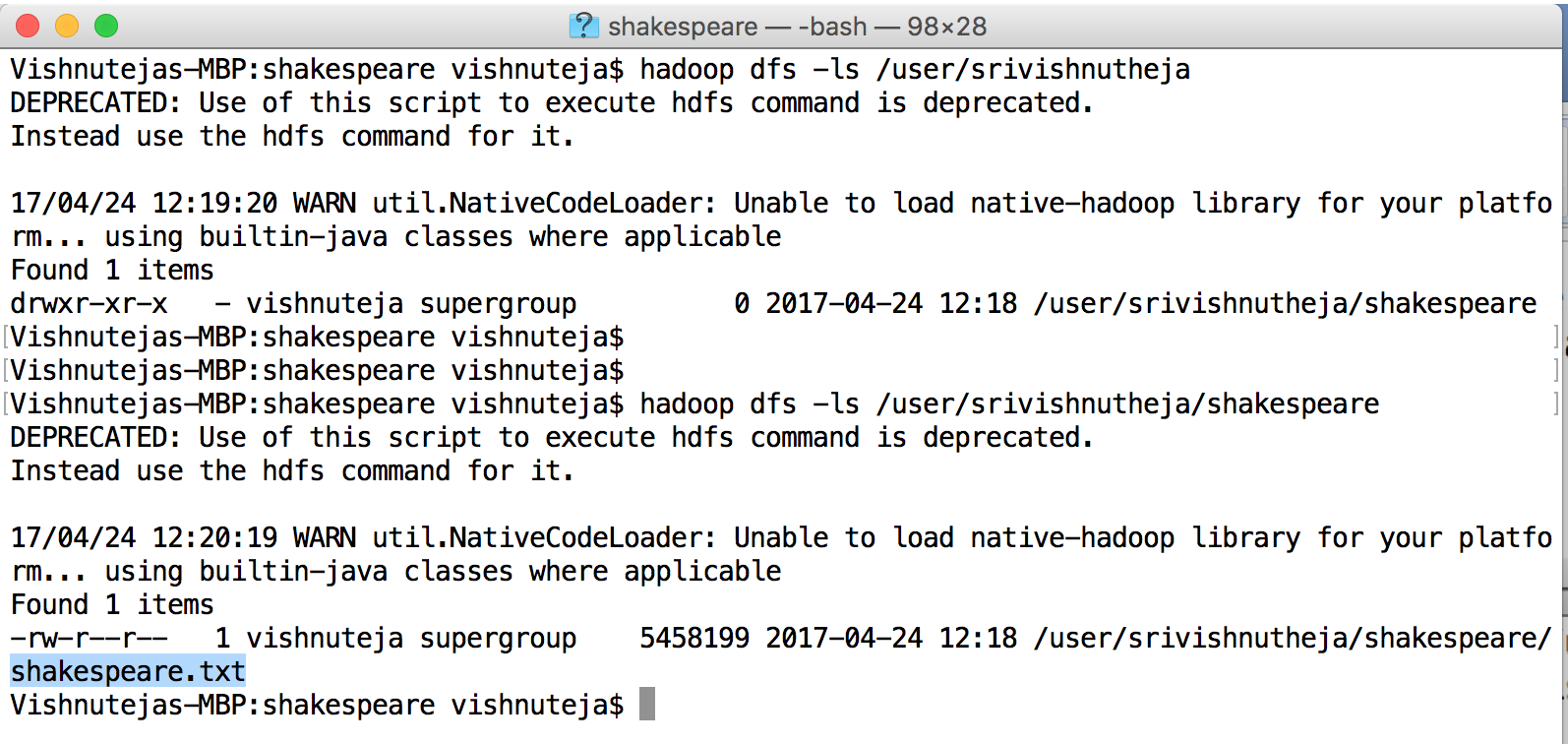


Restart the HDFS



hadoop dfs -copyFromLocal /Users/vishnuteja/Downloads/shakespeare /user/srivishnutheja/shakespeare





Vishnutejas-MBP:Downloads vishnuteja$ hadoop jar Hadoop-mapreduce-examples-2.6.0.jar wordcount /user/srivishnutheja/shakespeare /user/srivishnutheja/shakespeare-output

17/04/24 12:23:59 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

17/04/24 12:24:00 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session-id

17/04/24 12:24:00 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=

17/04/24 12:24:00 INFO input.FileInputFormat: Total input files to process : 1

17/04/24 12:24:00 INFO mapreduce.JobSubmitter: number of splits:1

17/04/24 12:24:01 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_local554011606\_0001

17/04/24 12:24:01 INFO mapreduce.Job: The url to track the job: http://localhost:8080/

17/04/24 12:24:01 INFO mapreduce.Job: Running job: job\_local554011606\_0001

17/04/24 12:24:01 INFO mapred.LocalJobRunner: OutputCommitter set in config null

17/04/24 12:24:01 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1

17/04/24 12:24:01 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

17/04/24 12:24:01 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCommitter

17/04/24 12:24:01 INFO mapred.LocalJobRunner: Waiting for map tasks

17/04/24 12:24:01 INFO mapred.LocalJobRunner: Starting task: attempt\_local554011606\_0001\_m\_000000\_0

17/04/24 12:24:01 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1

17/04/24 12:24:01 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup \_temporary folders under output directory:false, ignore cleanup failures: false

17/04/24 12:24:01 INFO util.ProcfsBasedProcessTree: ProcfsBasedProcessTree currently is supported only on Linux.

17/04/24 12:24:01 INFO mapred.Task:  Using ResourceCalculatorProcessTree : null

17/04/24 12:24:01 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/user/srivishnutheja/shakespeare/shakespeare.txt:0+5458199

17/04/24 12:24:01 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)

17/04/24 12:24:01 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100

17/04/24 12:24:01 INFO mapred.MapTask: soft limit at 83886080

17/04/24 12:24:01 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600

17/04/24 12:24:01 INFO mapred.MapTask: kvstart = 26214396; length = 6553600

17/04/24 12:24:01 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer

17/04/24 12:24:02 INFO mapreduce.Job: Job job\_local554011606\_0001 running in uber mode : false

17/04/24 12:24:02 INFO mapreduce.Job:  map 0% reduce 0%

17/04/24 12:24:02 INFO mapred.LocalJobRunner:

17/04/24 12:24:02 INFO mapred.MapTask: Starting flush of map output

17/04/24 12:24:02 INFO mapred.MapTask: Spilling map output

17/04/24 12:24:02 INFO mapred.MapTask: bufstart = 0; bufend = 8546434; bufvoid = 104857600

17/04/24 12:24:02 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 22609100(90436400); length = 3605297/6553600

17/04/24 12:24:03 INFO mapred.MapTask: Finished spill 0

17/04/24 12:24:03 INFO mapred.Task: Task:attempt\_local554011606\_0001\_m\_000000\_0 is done. And is in the process of committing

17/04/24 12:24:03 INFO mapred.LocalJobRunner: map

17/04/24 12:24:03 INFO mapred.Task: Task 'attempt\_local554011606\_0001\_m\_000000\_0' done.

17/04/24 12:24:03 INFO mapred.LocalJobRunner: Finishing task: attempt\_local554011606\_0001\_m\_000000\_0

17/04/24 12:24:03 INFO mapred.LocalJobRunner: map task executor complete.

17/04/24 12:24:03 INFO mapred.LocalJobRunner: Waiting for reduce tasks

17/04/24 12:24:03 INFO mapred.LocalJobRunner: Starting task:

…

…

…

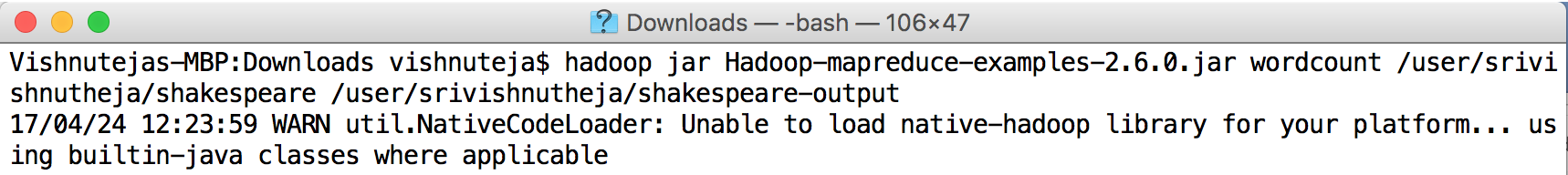
17/04/24 12:24:04 INFO mapred.LocalJobRunner: Finishing task: attempt\_local554011606\_0001\_r\_000000\_0

17/04/24 12:24:04 INFO mapred.LocalJobRunner: reduce task executor complete.

17/04/24 12:24:04 INFO mapreduce.Job:  map 100% reduce 100%

17/04/24 12:24:04 INFO mapreduce.Job: Job job\_local554011606\_0001 completed successfully

17/04/24 12:24:04 INFO mapreduce.Job: Counters: 35



..

