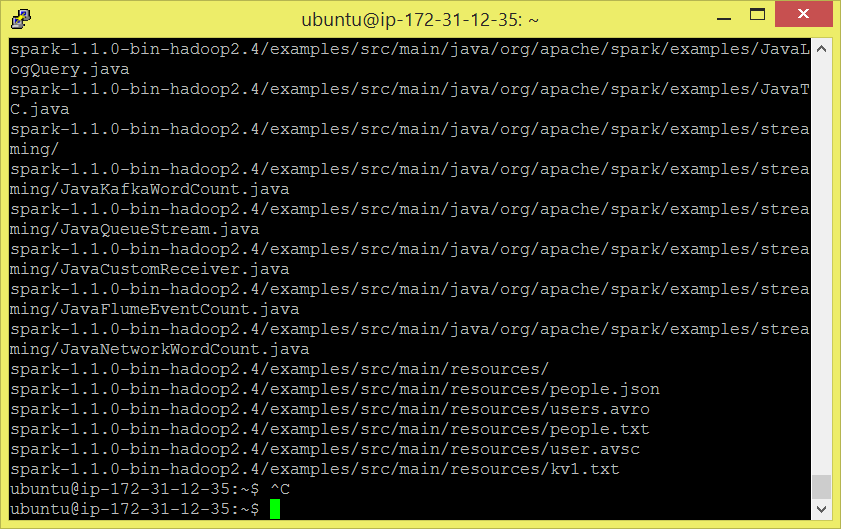
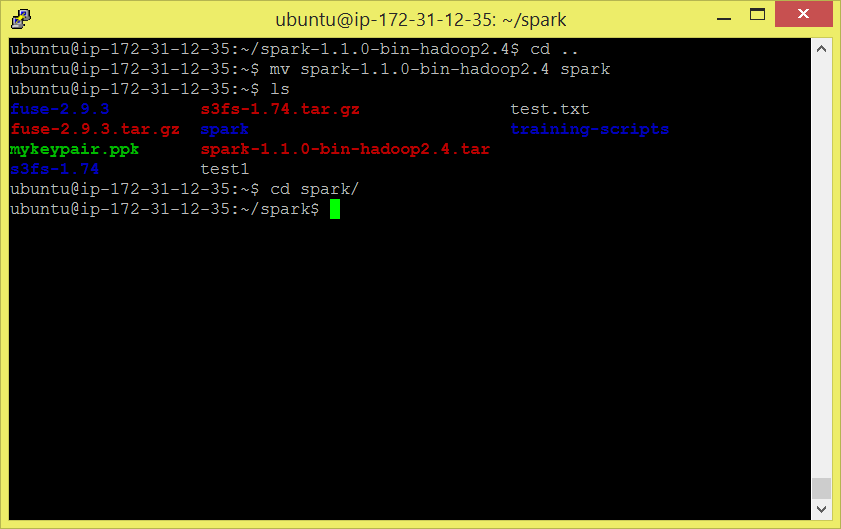


Wget <http://d3kbcqa49mib13.cloudfront.net/spark-1.1.0-bin-hadoop2.4.tgz>

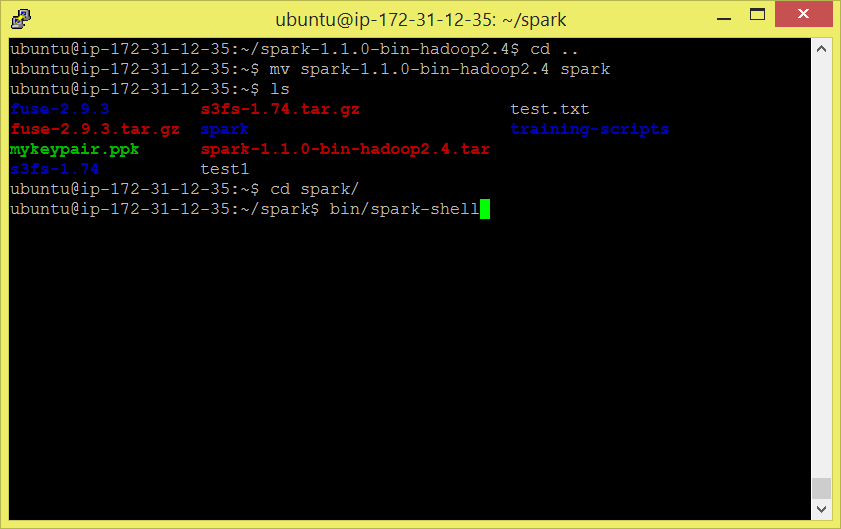
ubuntu@ip-172-31-12-35:~$ gzip -d spark-1.1.0-bin-hadoop2.4.tgz

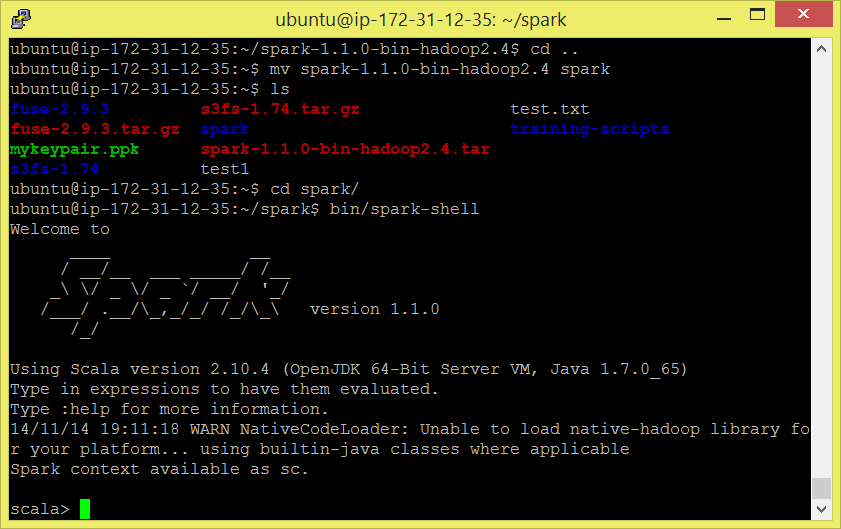
ubuntu@ip-172-31-12-35:~$ tar xvf spark-1.1.0-bin-hadoop2.4.tar



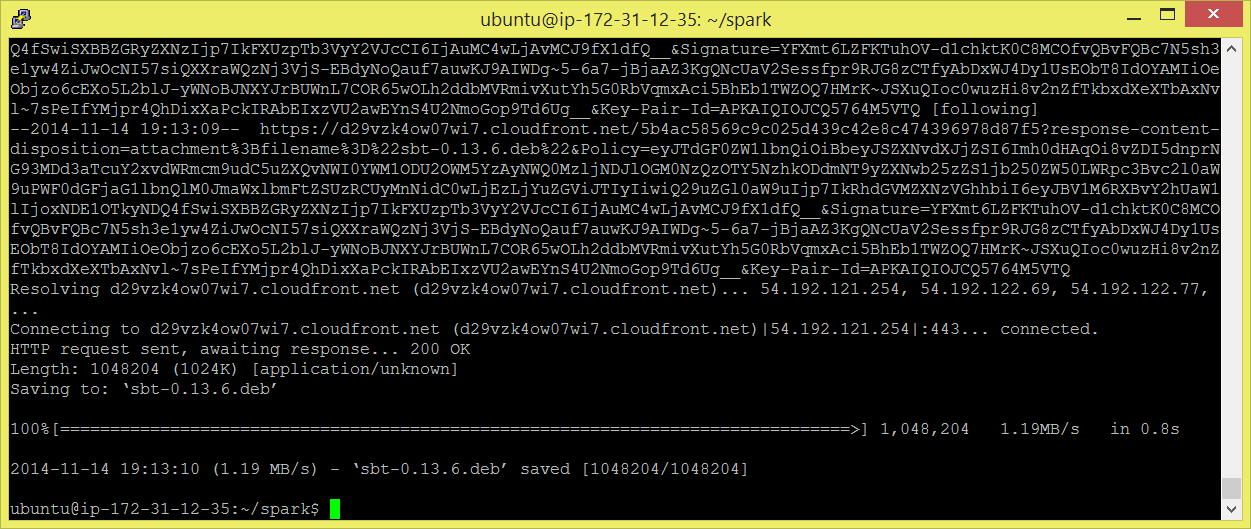


bin/spark-shell





wget <https://dl.bintray.com/sbt/debian/sbt-0.13.6.deb>



sudo dpkg --install sbt-0.13.6.deb



[http://spark.apache.org/docs/latest/quick-start.html#standalone-applications](http://spark.apache.org/docs/latest/quick-start.html" \l "standalone-applications)

vi simple.sbt

name := "Simple Project"

version := "1.0"

scalaVersion := "2.10.4"

libraryDependencies += "org.apache.spark" %% "spark-core" % "1.1.0"



*# Your directory layout should look like this*

$ find .

.

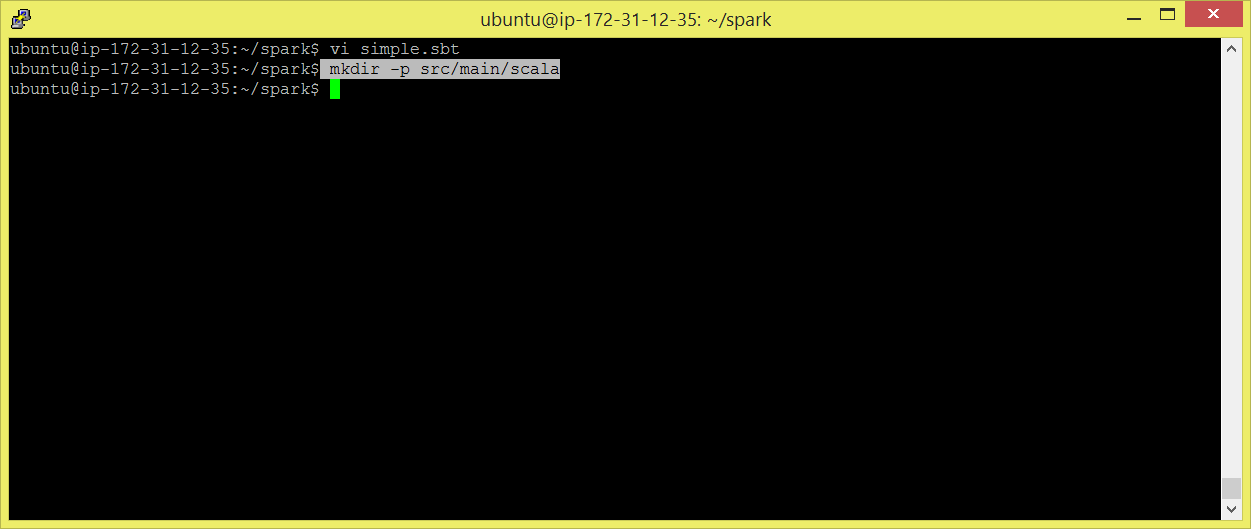
./simple.sbt

./src

./src/main

./src/main/scala

./src/main/scala/SimpleApp.scala

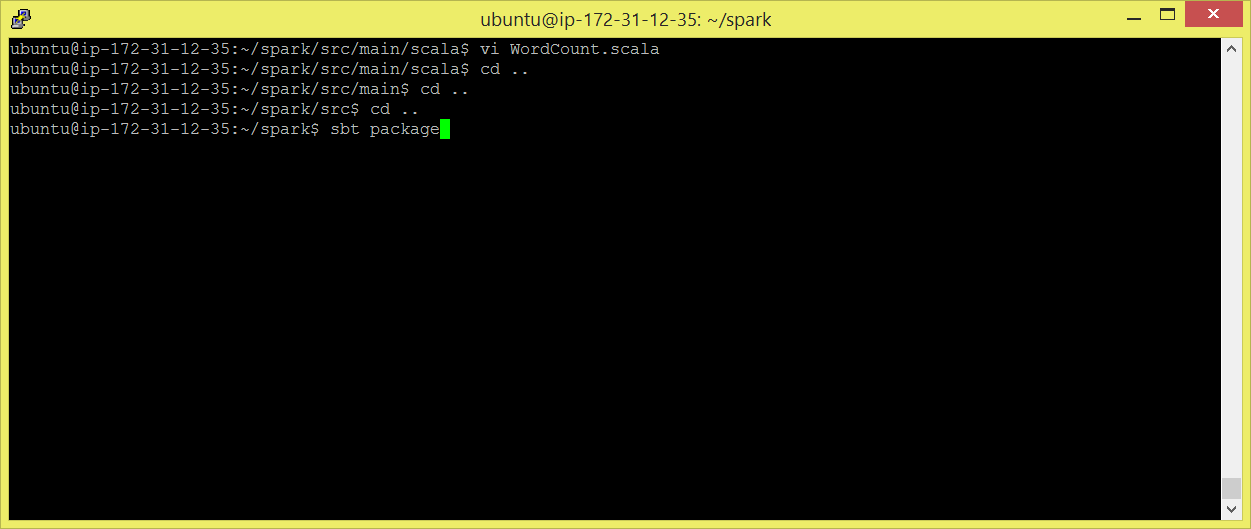


Vi Wordcount.scala

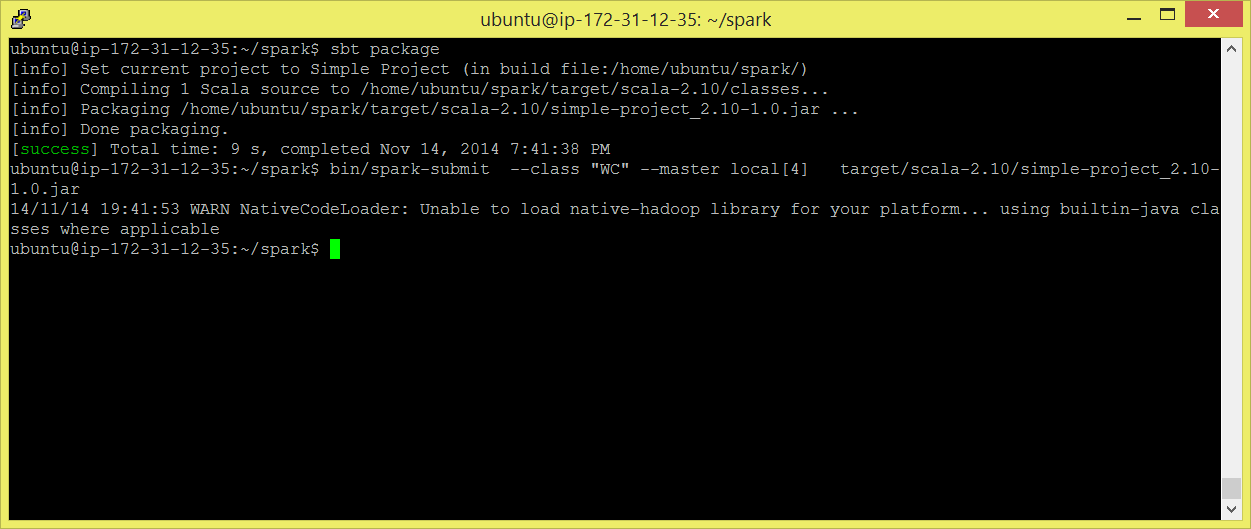


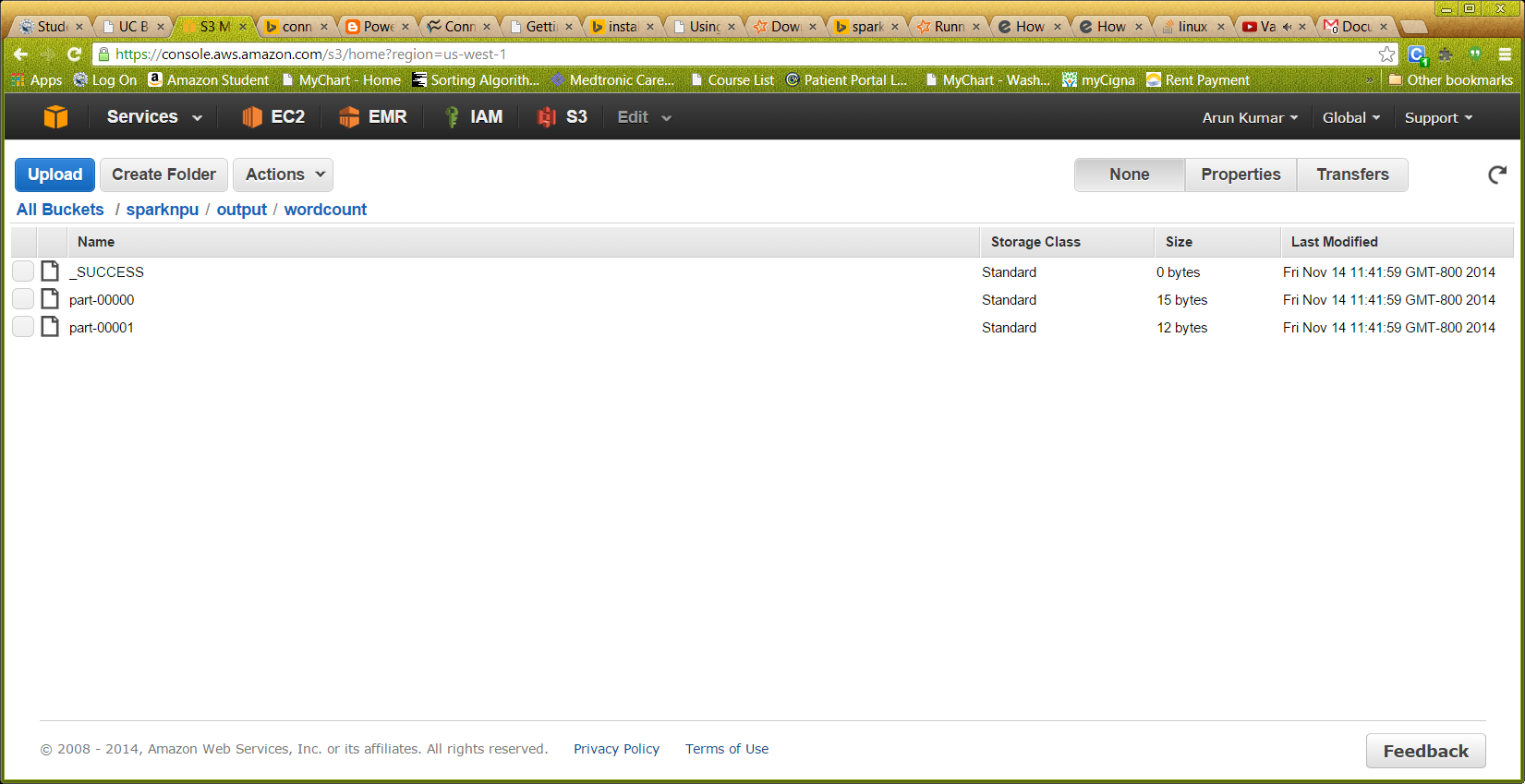
*# Package a jar containing your application*

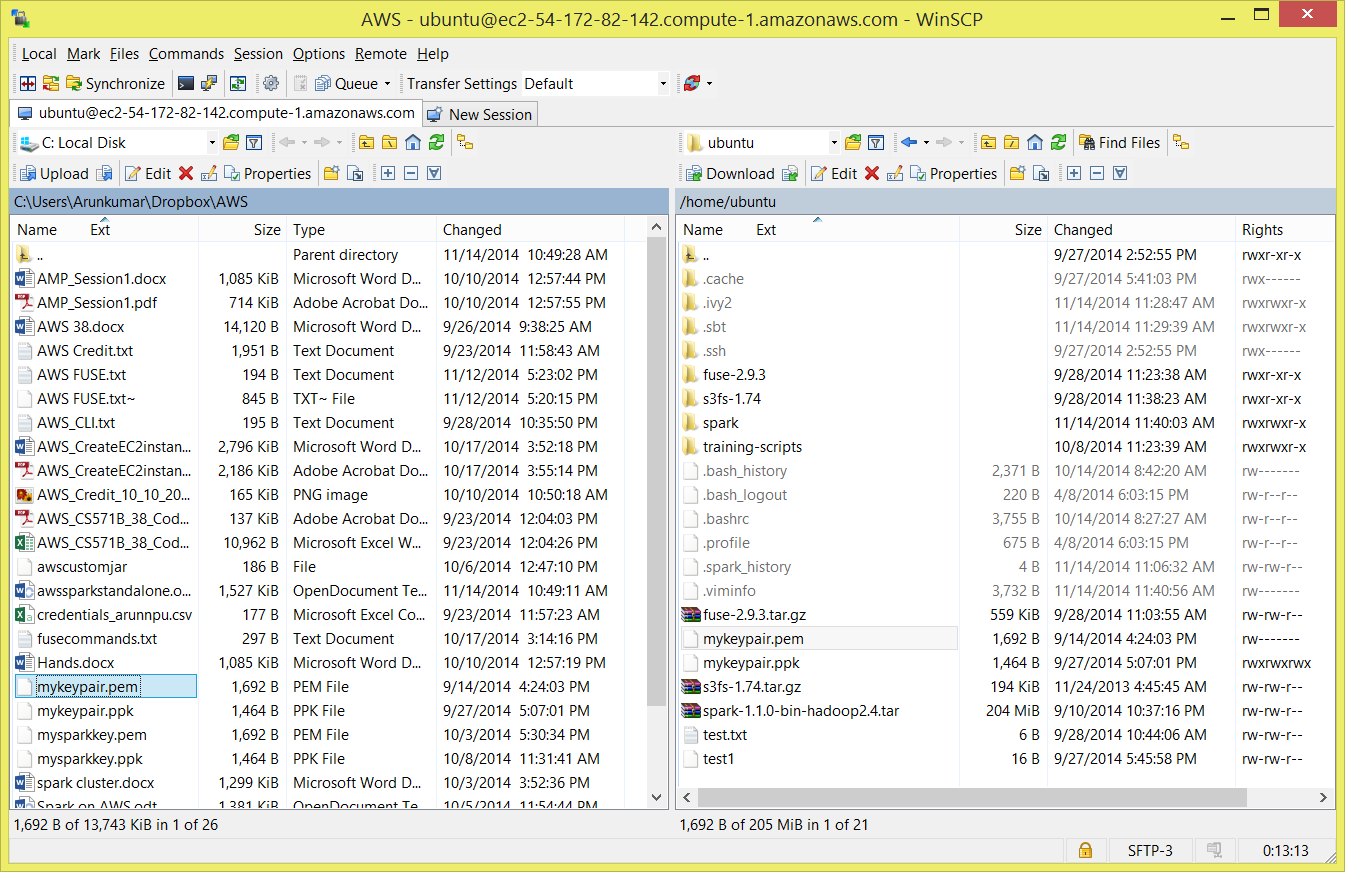
sbt package



bin/spark-submit --class "WC" --master local[4] target/scala-2.10/simple-project\_2.10-1.0.jar



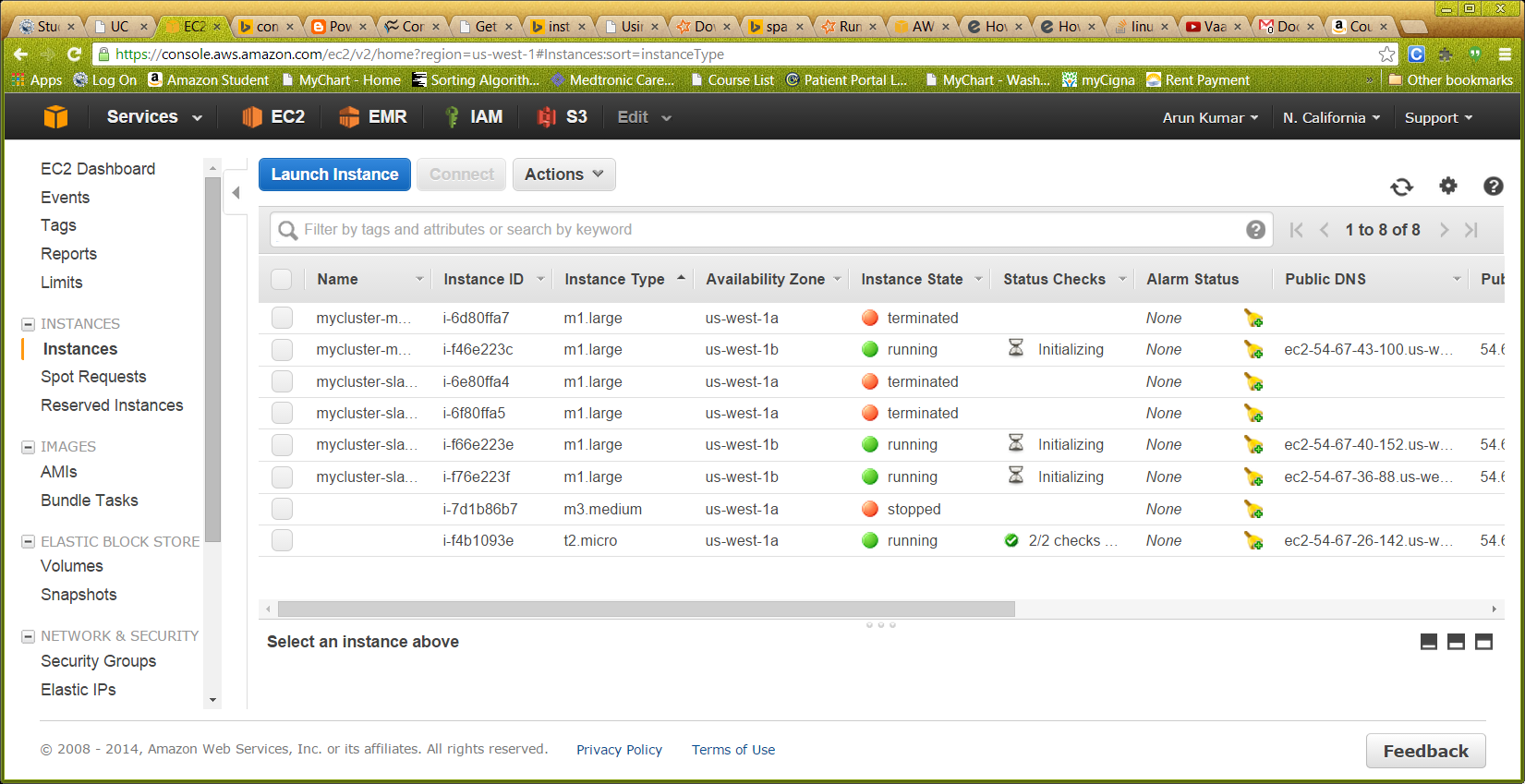


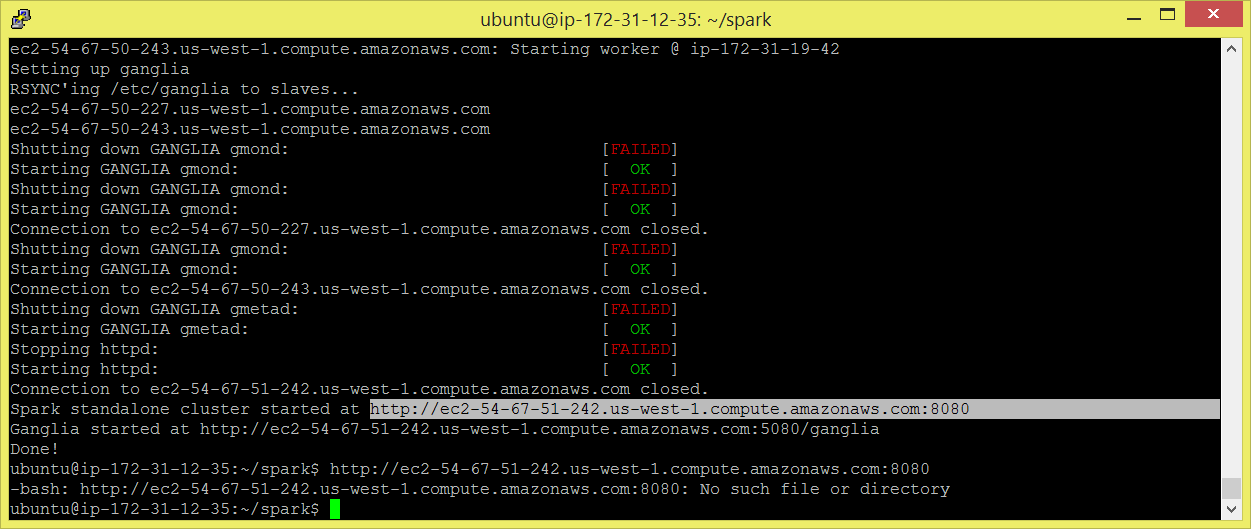


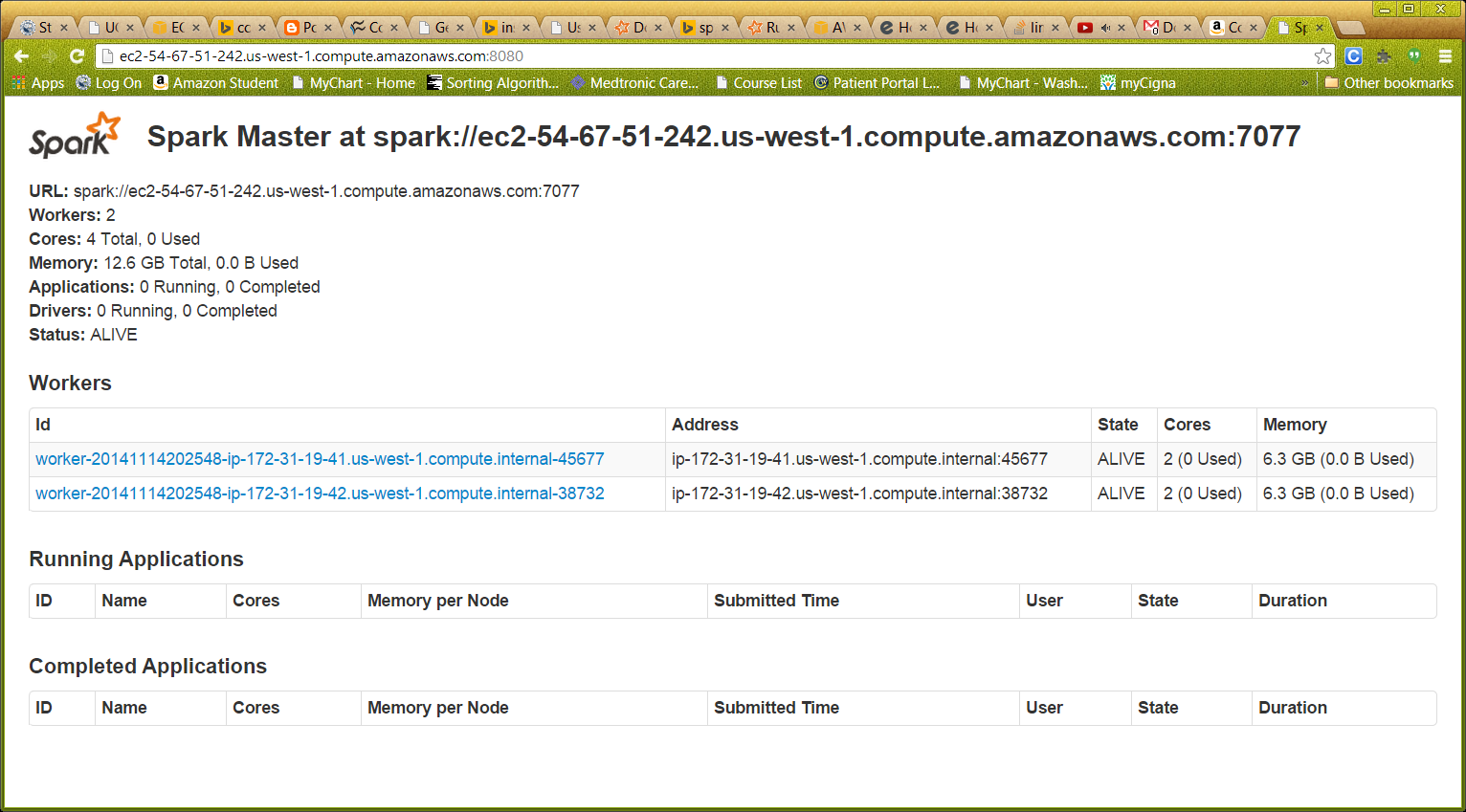
**Launching a Cluster**

ec2/spark-ec2 -k mykeypair -i ~/mykeypair.pem -s 2 --region us-west-1 -- wait 240 launch mycluster

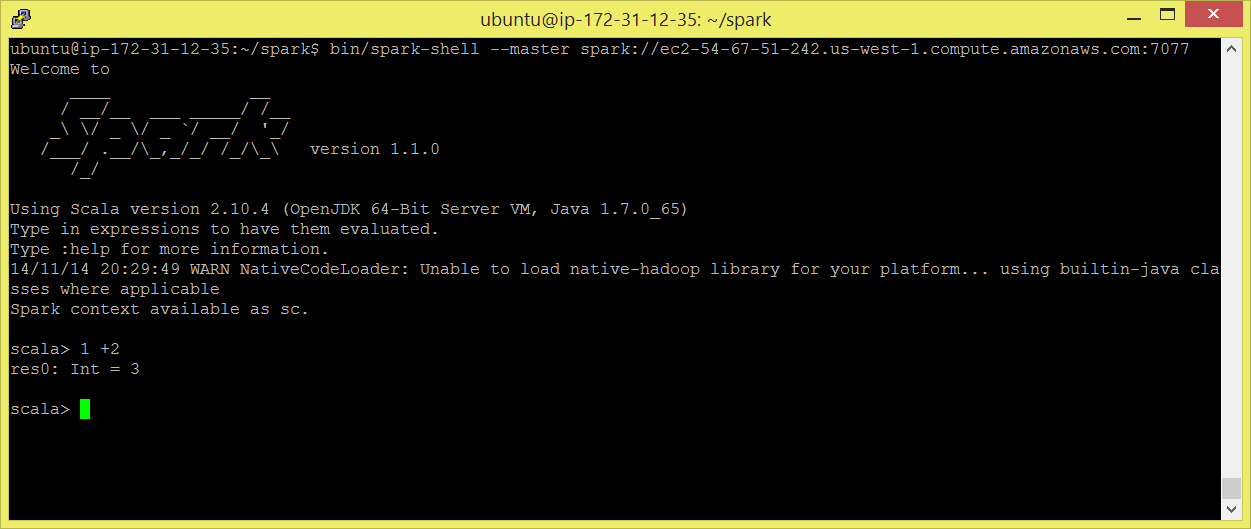


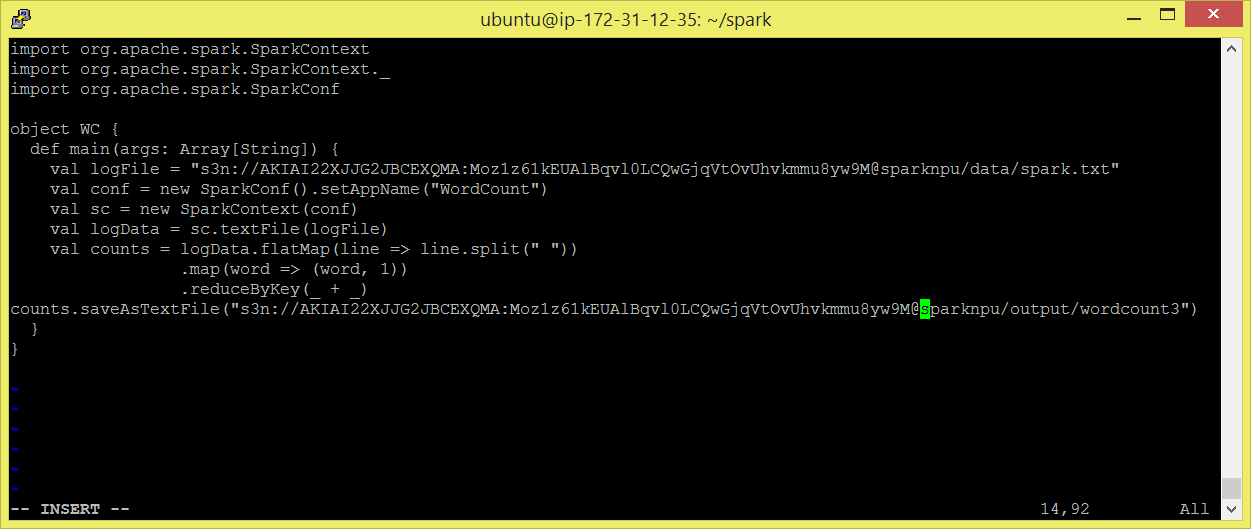


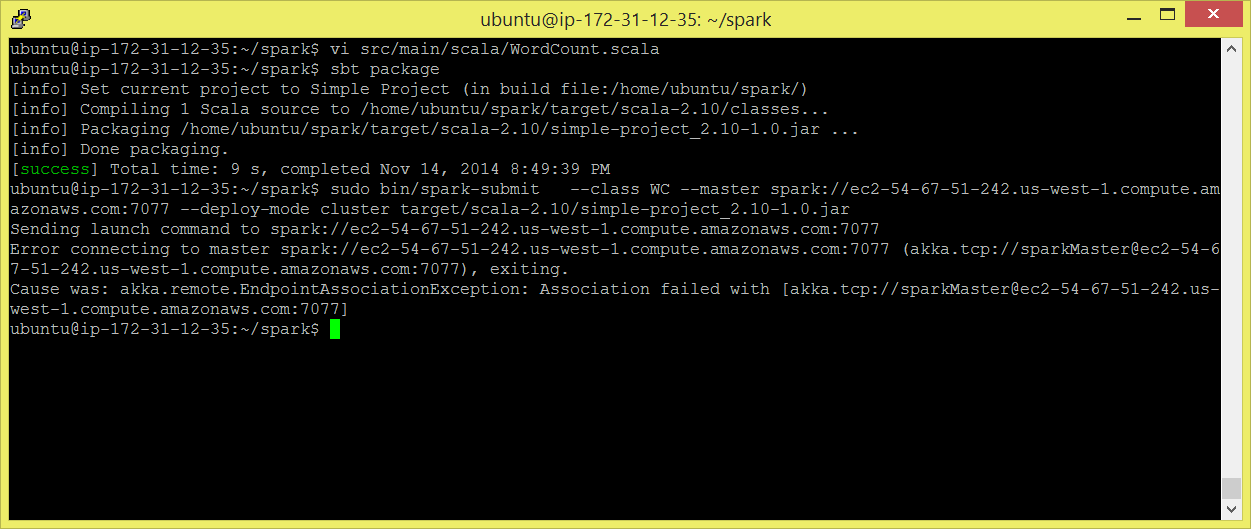


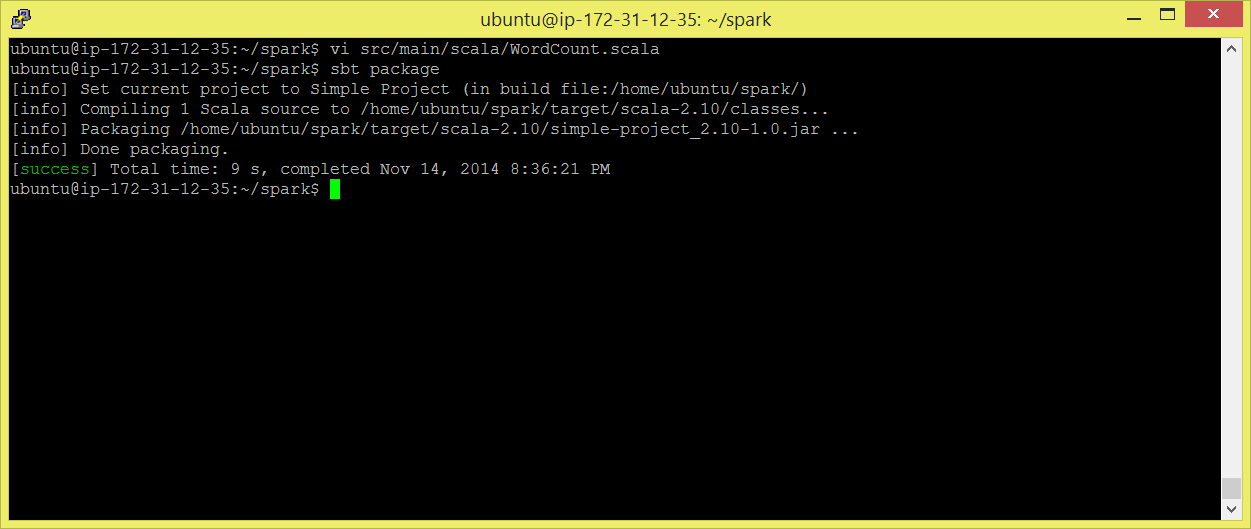


bin/spark-shell --master **spark://ec2-54-67-51-242.us-west-1.compute.amazonaws.com:7077**



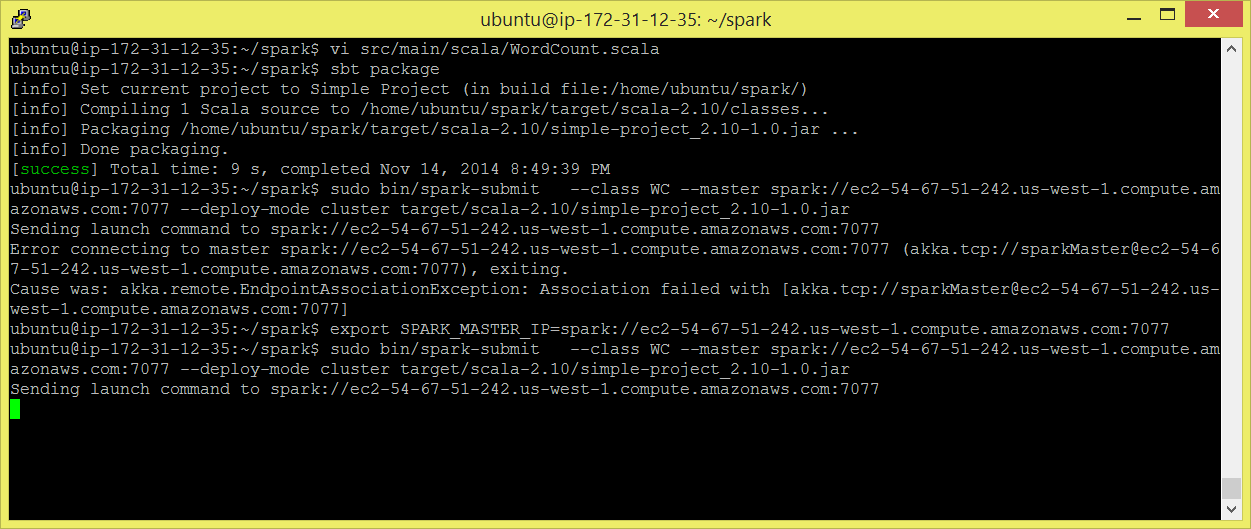






bin/spark-submit --class WC --master spark://ec2-54-67-51-242.us-west-1.compute.amazonaws.com:7077 --deploy-mode cluster target/scala-2.10/simple-project\_2.10-1.0.jar

AKIAI22XJJG2JBCEXQMA: Moz1z61kEUAlBqvl0LCQwGjqVtOvUhvkmmu8yw9M



./spark-ec2 -k mykeypair -i mykeypair.per -- region us-west-1 login mycluster

bin/spark-submit --class WC --master spark://ec2-54-67-51-242.us-west-1.compute.amazonaws.com:7077 --deploy-mode cluster s3n:// AKIAI22XJJG2JBCEXQMA: Moz1z61kEUAlBqvl0LCQwGjqVtOvUhvkmmu8yw9M@sparknpu/jar/ simple-project\_2.10-1.0

bin/spark-submit --class WCArg --master spark:[**//ec2-54-67-46-155.us-west-1.compute.amazonaws.com:**](http://ec2-54-67-46-155.us-west-1.compute.amazonaws.com:8080/)**7077** --deploy-mode cluster s3n://AKIAI22XJJG2JBCEXQMA:Moz1z61kEUAlBqvl0LCQwGjqVtOvUhvkmmu8yw9M@sparknpu/Jar/spark.jar

s3n://AKIAI22XJJG2JBCEXQMA:Moz1z61kEUAlBqvl0LCQwGjqVtOvUhvkmmu8yw9M@sparknpu/data/spark.txt

s3n://AKIAI22XJJG2JBCEXQMA:Moz1z61kEUAlBqvl0LCQwGjqVtOvUhvkmmu8yw9M@sparknpu/output/sparkwordcount

sudo bin/spark-submit --master spark://ec2-54-172-4-3.compute-1.amazonaws.com:7077 --

deploy-mode cluster --class edu.spark.WordLengthCount

s3n://Access\_key\_ID:Secret\_access\_key @10657bucket/jars/SparkWordLegthCount.jar

s3n://Access\_key\_ID:Secret\_access\_key @10657bucket/input1/File1.txt

s3n://Access\_key\_ID:Secret\_access\_key @10657bucket/OUTPUT/