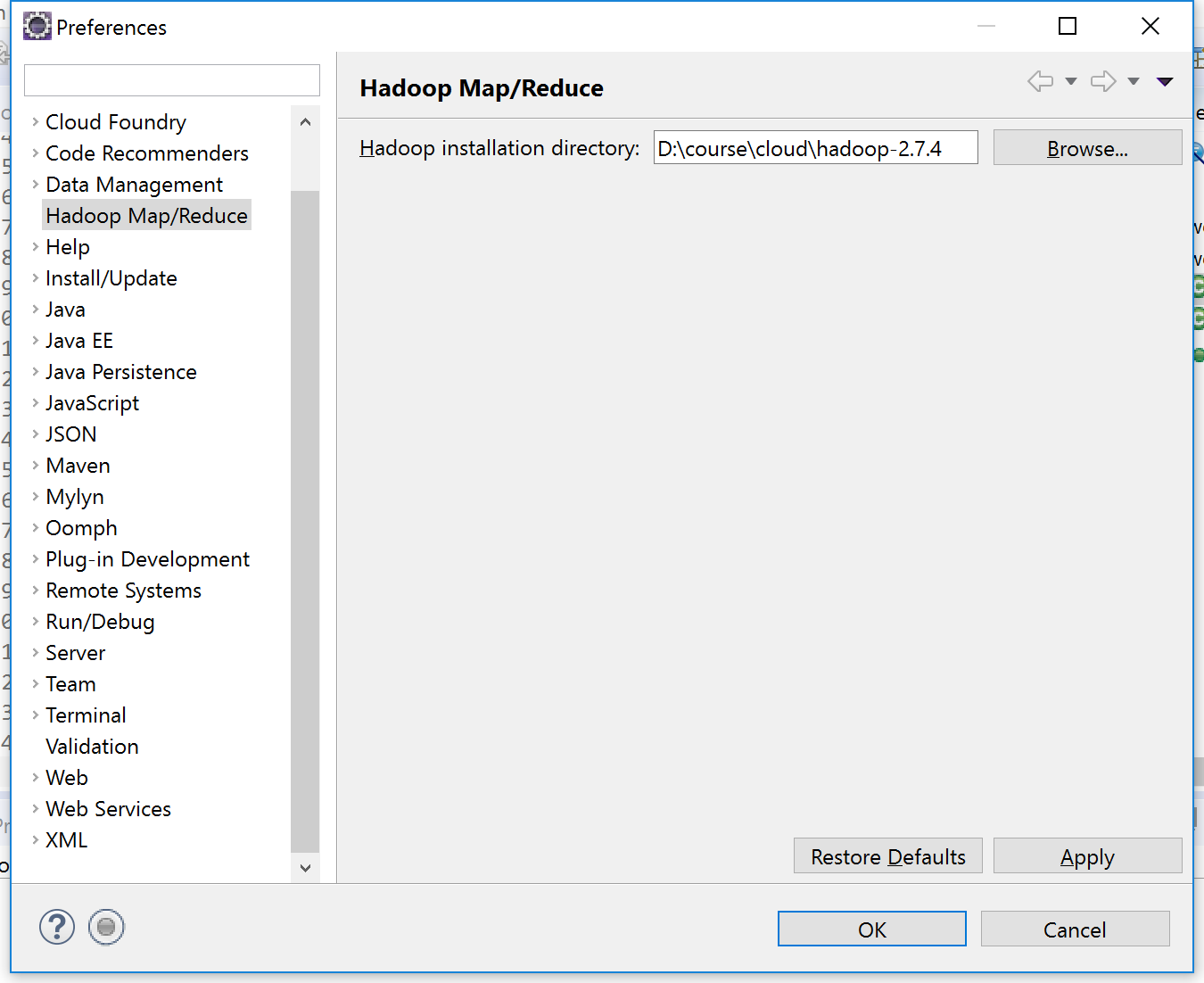
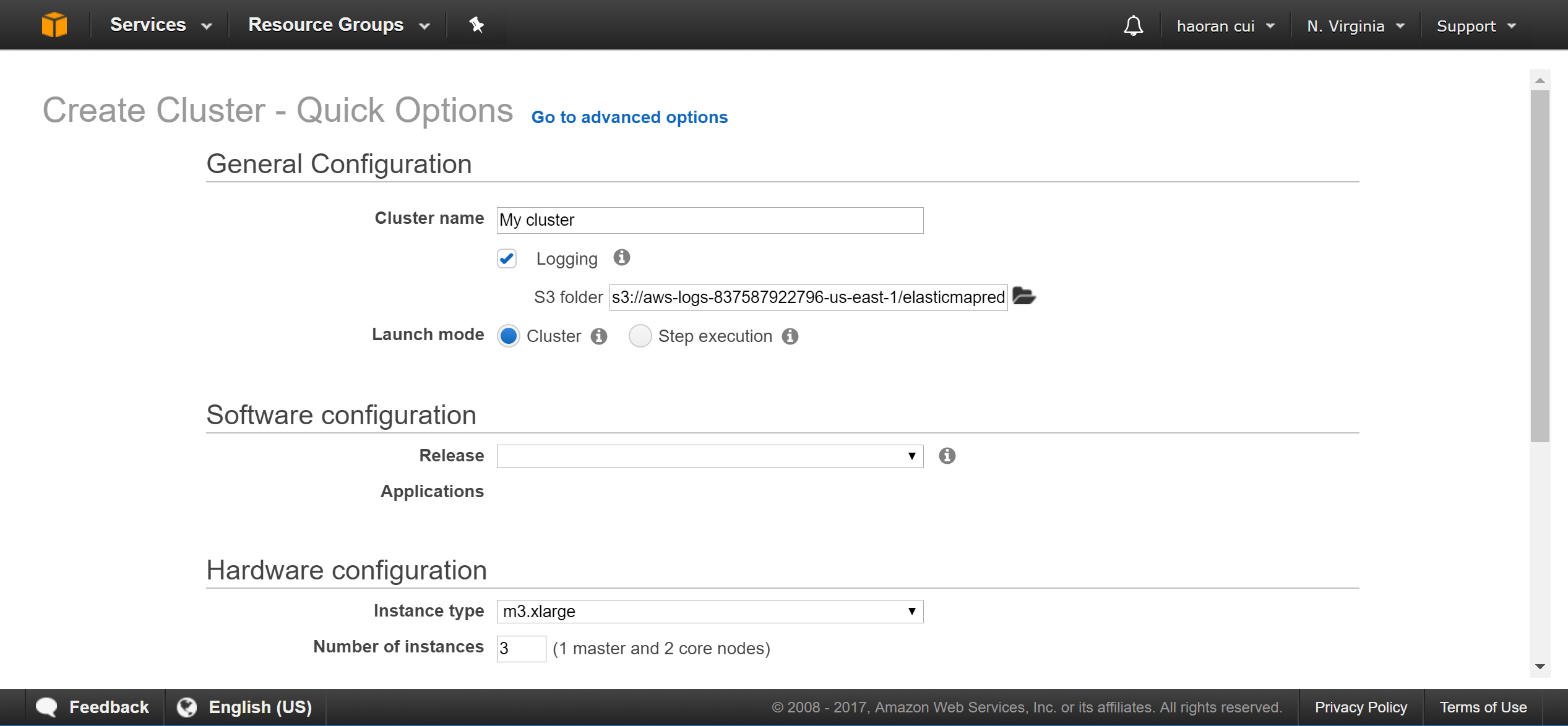
I build up a Hadoop environment on my laptop, down load Hadoop, eclipse, and a plugin for Hadoop on eclipse.



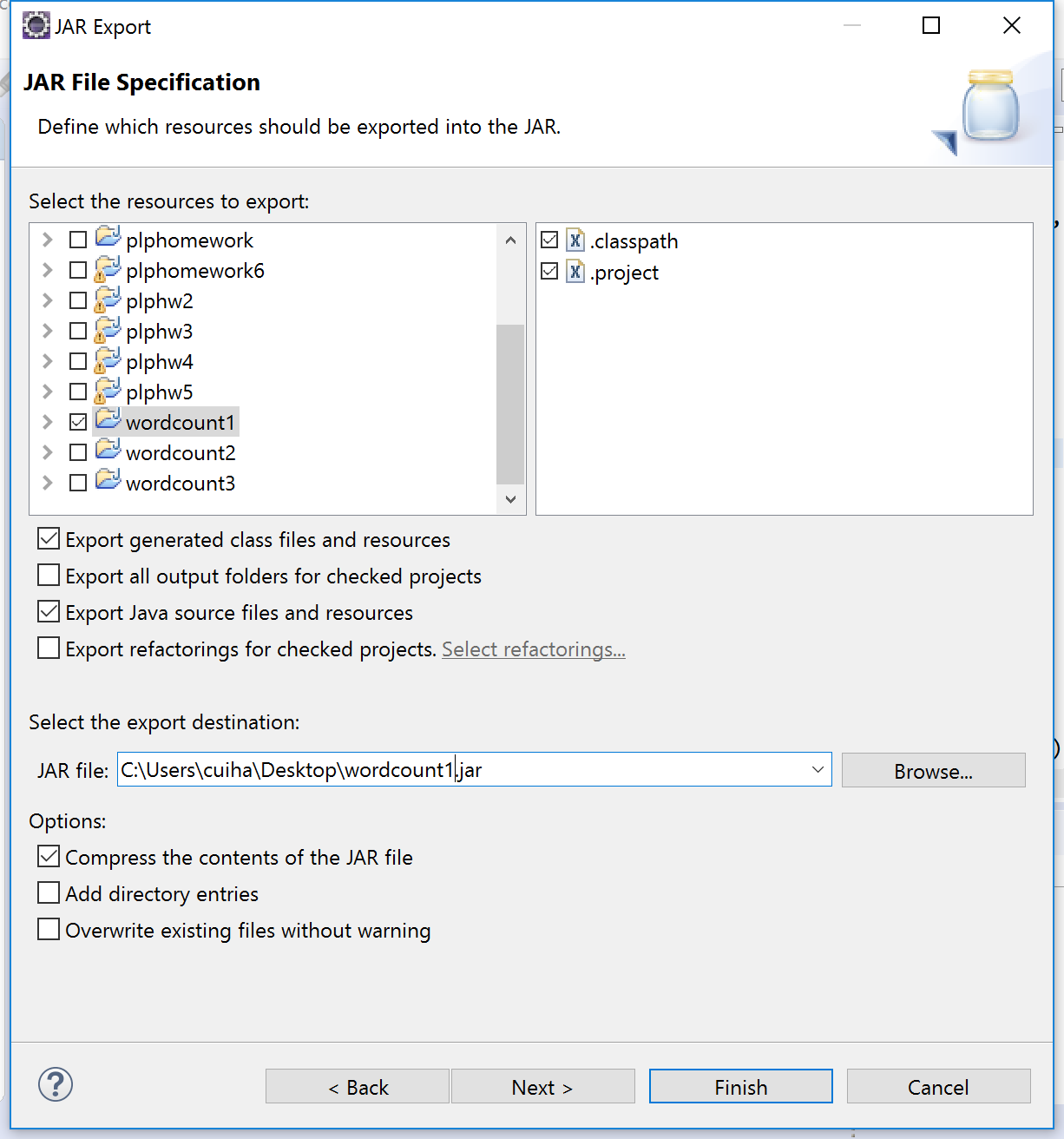
I still got some problem on my local Hadoop, I can’t run on my local, but I can write it with jars from Hadoop.

And I find out that I can run it on AWS with EMR it was easy.

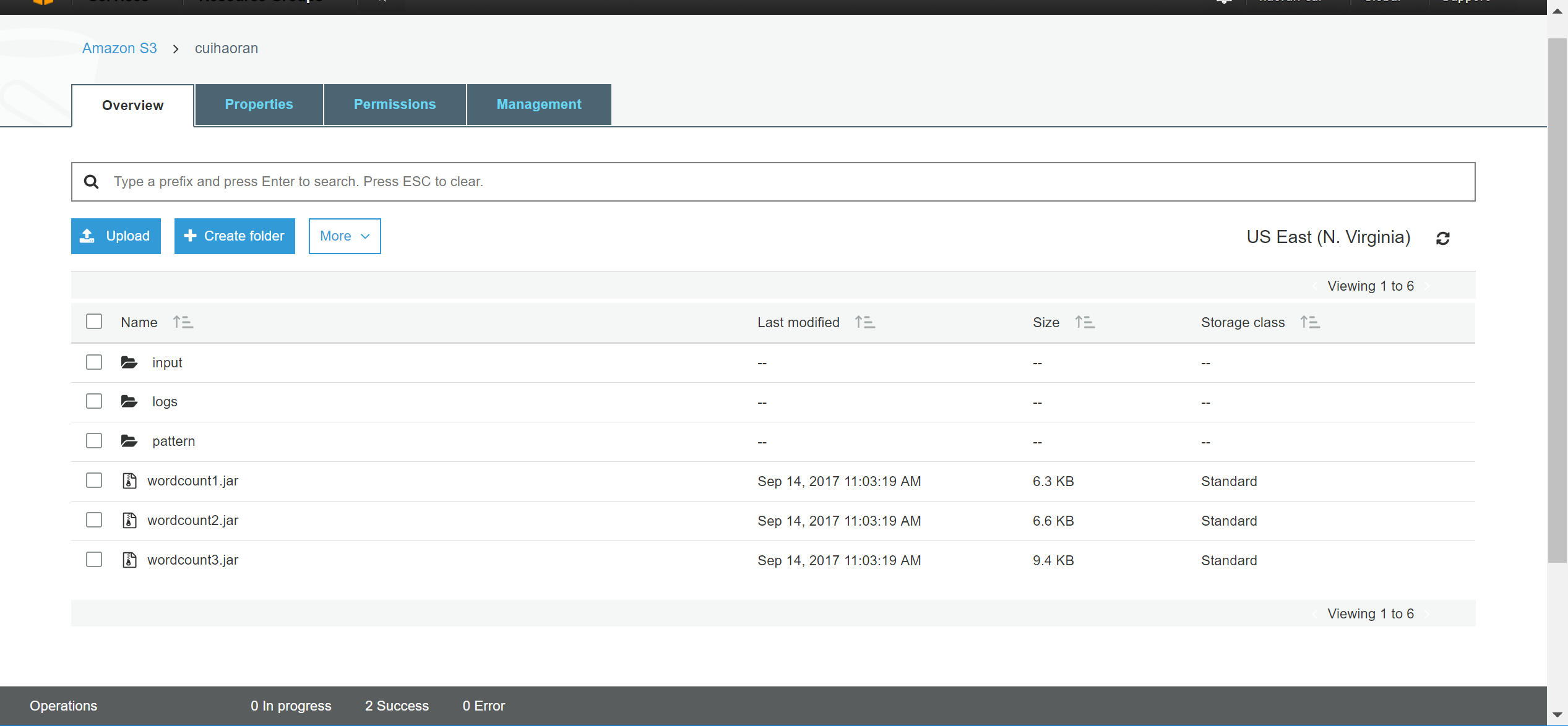
First set up clusters.



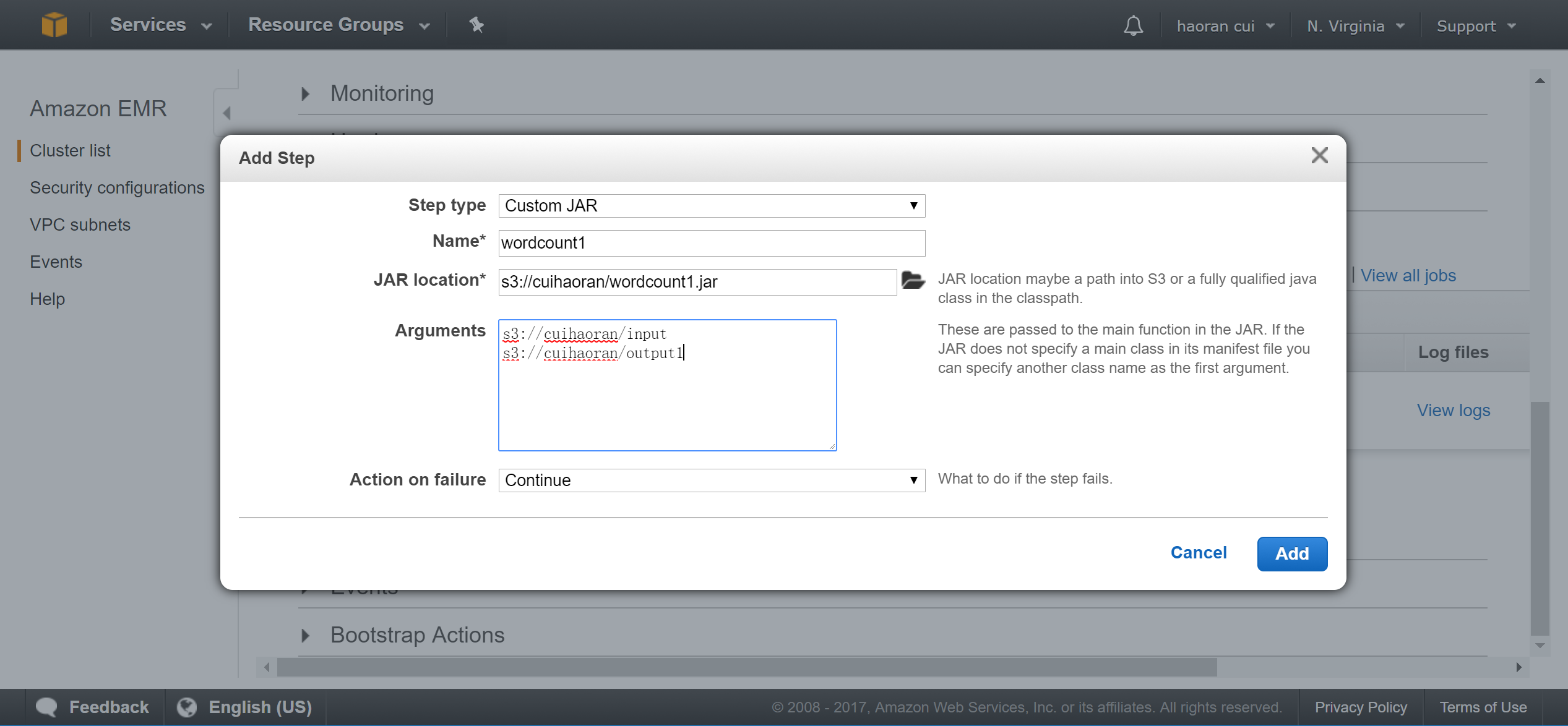
Then we should export our code in to a jar.

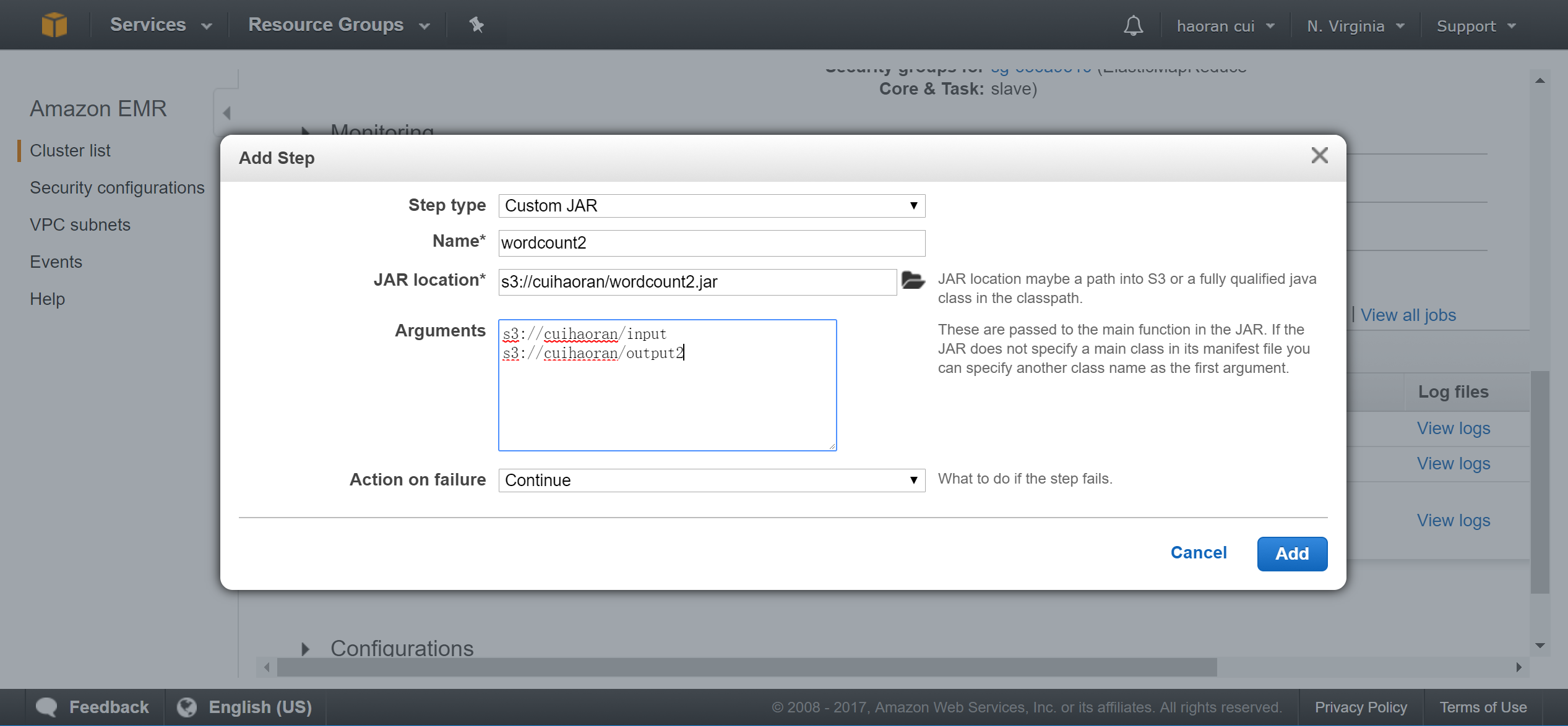


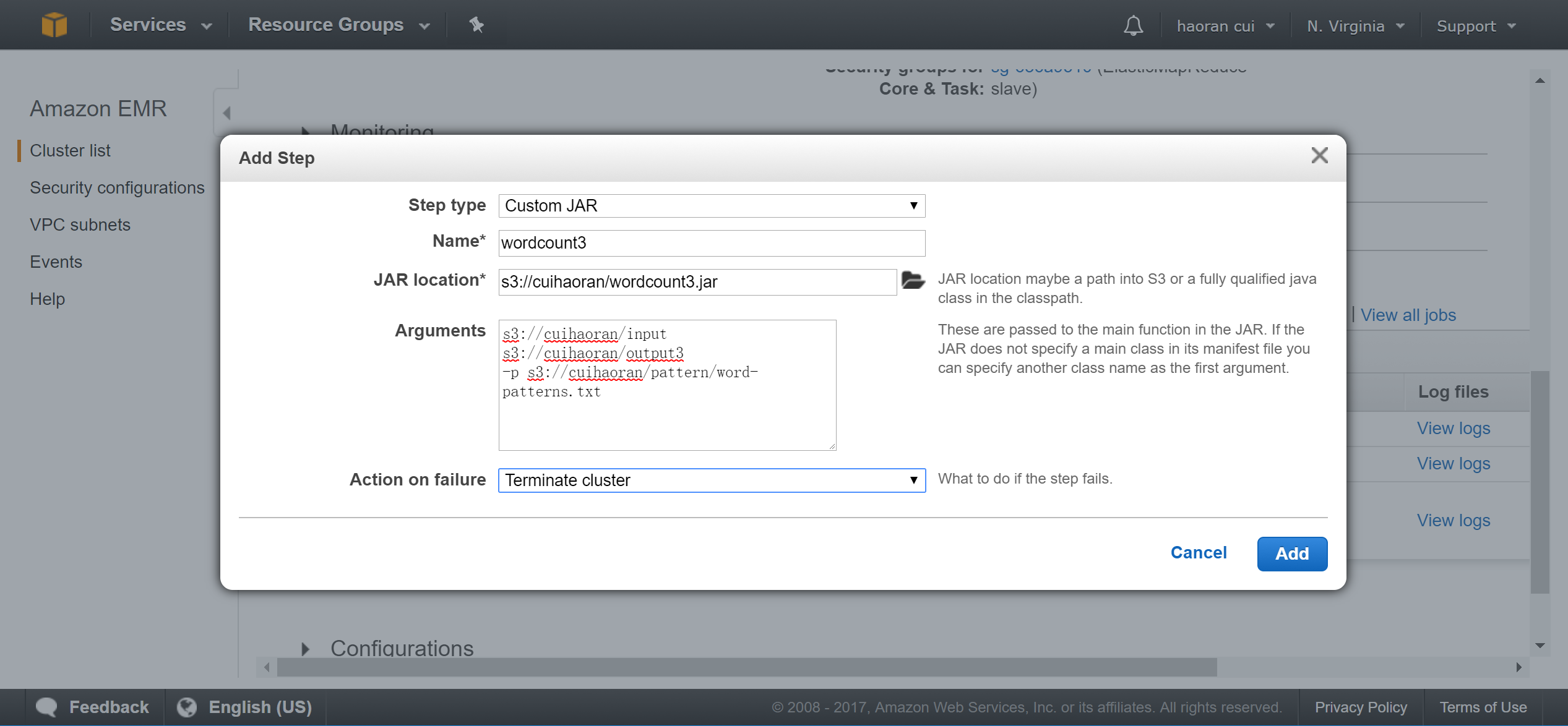
We have to up load our jars and data source ( I have increased the data source 10 times )to s3.

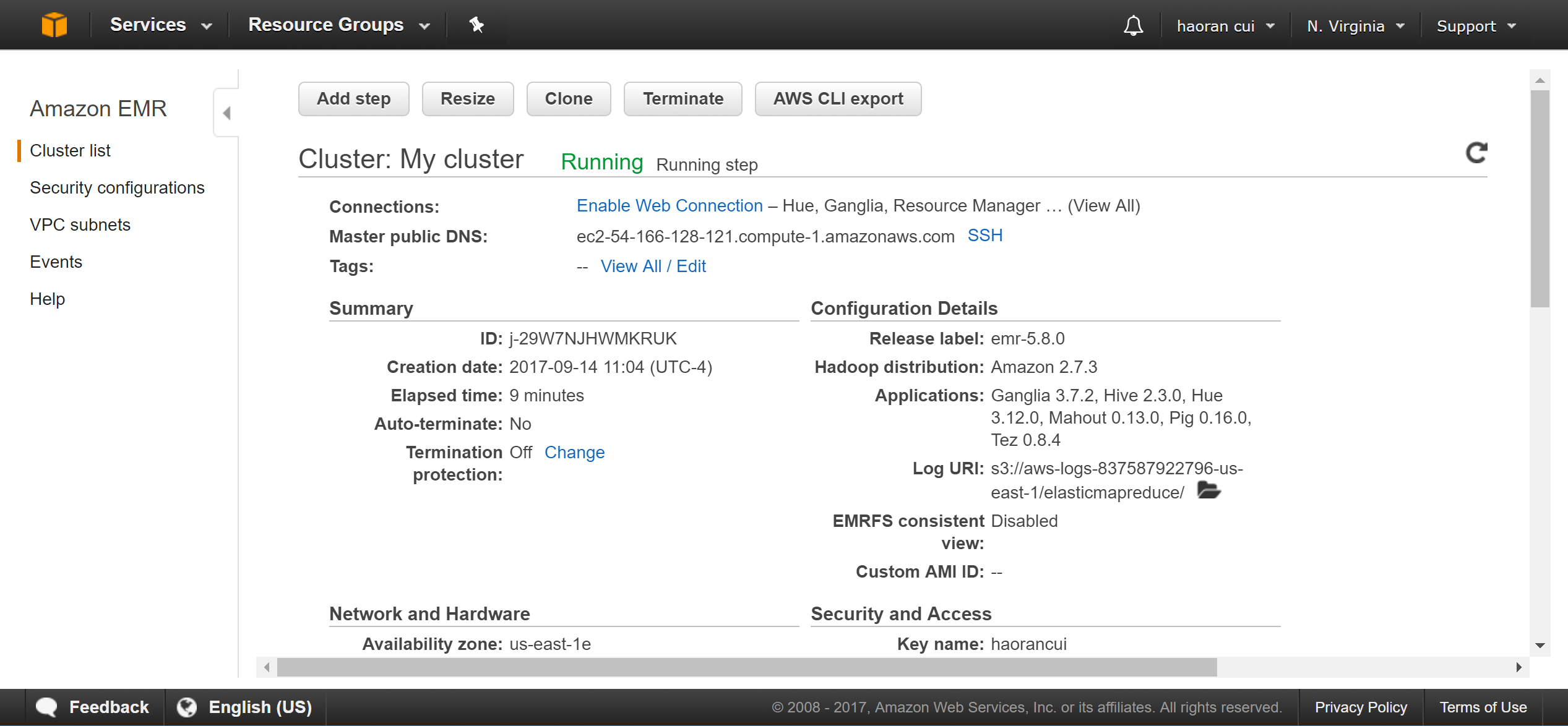


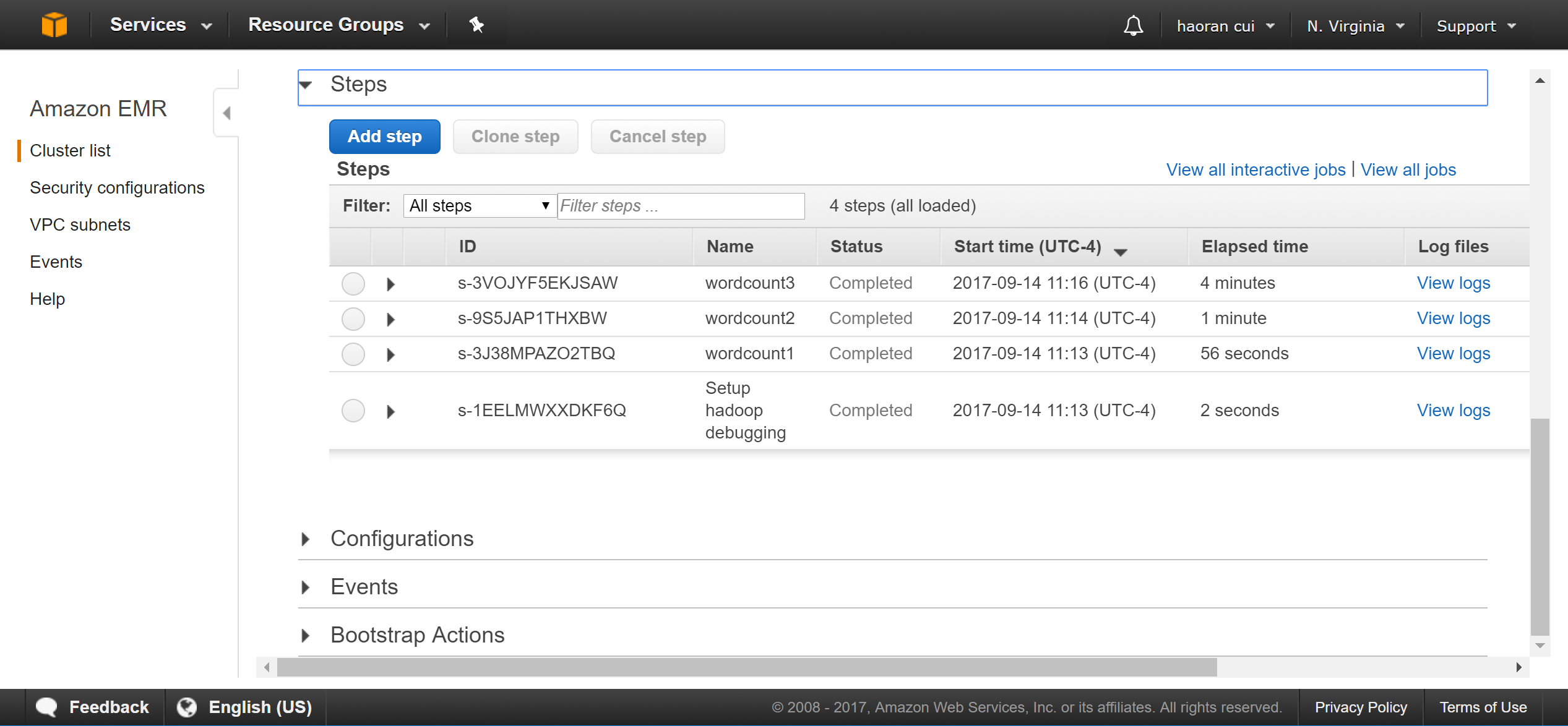
Then we add our jar to EMR by add steps, and write the arguments.







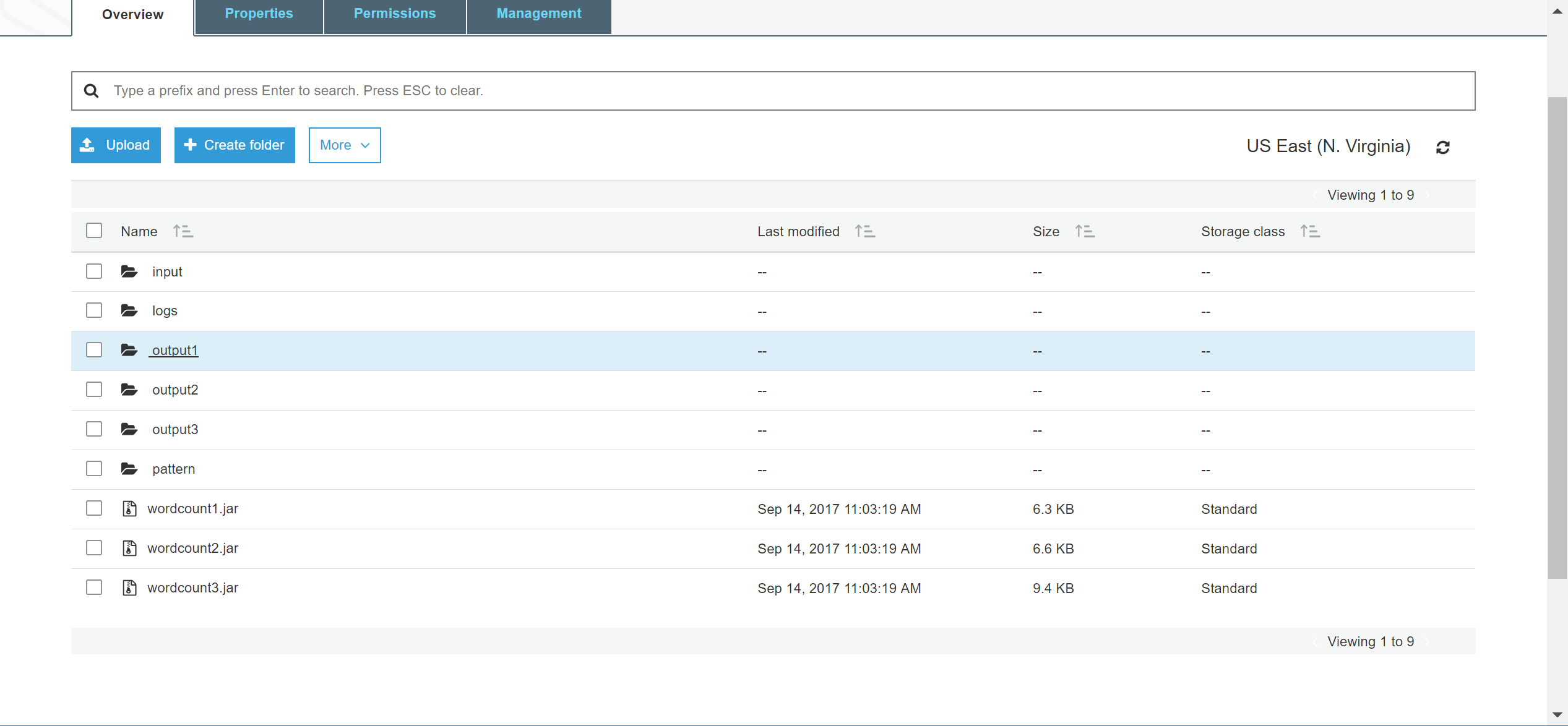


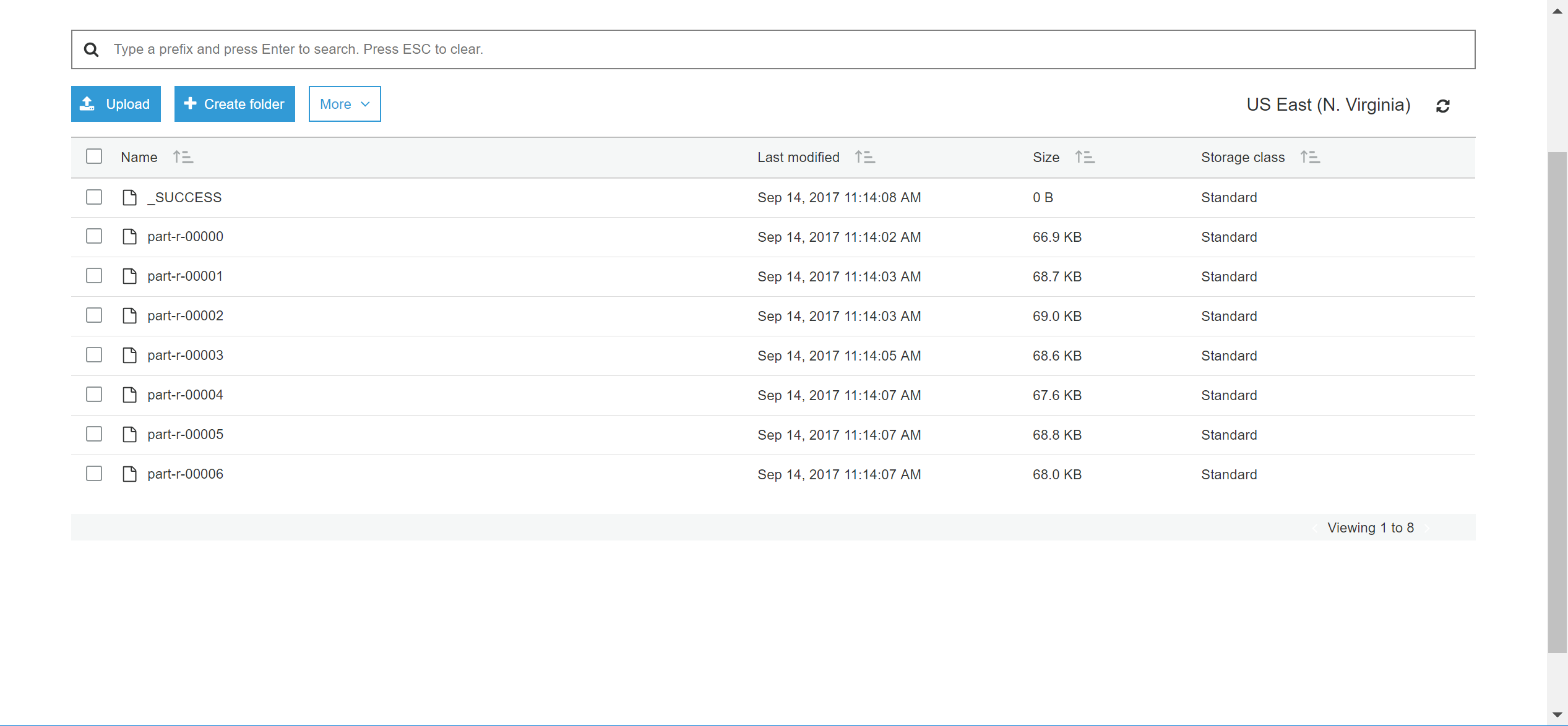


Result will be in the output folder once it complete.

Here are the results for this homework.

Task 1:





Result:

<https://s3.amazonaws.com/cuihaoran/output1/_SUCCESS>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00000>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00001>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00002>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00003>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00004>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00005>

<https://s3.amazonaws.com/cuihaoran/output1/part-r-00006>

Task 2:

<https://s3.amazonaws.com/cuihaoran/output2/_SUCCESS>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00000>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00001>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00002>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00003>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00004>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00005>

<https://s3.amazonaws.com/cuihaoran/output2/part-r-00006>

Task 3:

<https://s3.amazonaws.com/cuihaoran/output3/_SUCCESS>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00000>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00001>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00002>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00003>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00004>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00005>

<https://s3.amazonaws.com/cuihaoran/output3/part-r-00006>