**4조 튜토리얼**

Configuration

### Step 1: Install Git Bash, AWS CLI

1) Install Git Bash

- Git Bash Install URL(<http://www.techoism.com/how-to-install-git-bash-on-windows/>)

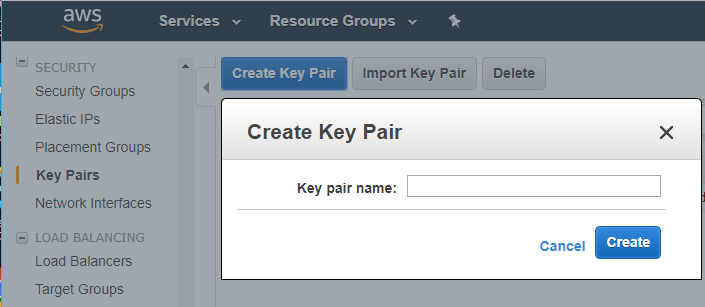
2) AWS CLI Install

 - Window용 AWS CLI Download(<https://s3.amazonaws.com/aws-cli/AWSCLI64.msi>)

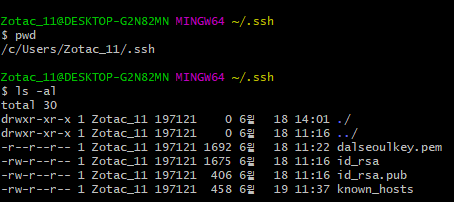
- 개인 PC 에서 aws의 EMR master 서버에 Terminal로 접속하여 작업 수행을 위하여 Command Line Interface 설치

3) Create Key Pair for AWS Connection Confirm

 - key Pair 생성(.pem) 및 Local PC에 Download



.pem 파일을 ~/.ssh 폴더에 copy



.ssh 폴더가 없을 시 ssh-keygen 실행하여 생성

### Step 2: AWS Configure to connect to AWS

1) AWS Connection을 위하여 아래와 같이 설정

$ aws configure

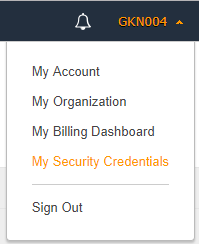
AWS Access Key ID [None]: \*\*\*\*\*\*\*\*\*\*\*\*

AWS Secret Access Key [None]: \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

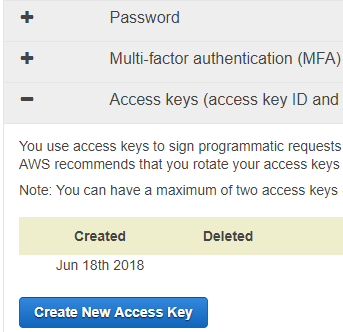
Default region name [None]: ap-northeast-2

Default output format [None]: json

- Access Key ID 및 Secret Access Key는 아래와 같이 획득 한다.



My Security Credentials Click



Create New Access Key Click 하여 생성 한다.

### Step 3: yelp data Download and upload to s3

1)  Download data and Extract

- Data set URL and information about the data set:<https://www.yelp.com/dataset>

- Download **yelp\_dataset.tar.gz** file

- untar yelp\_dataset.tar.gz file(tar -zxvf yelp\_dataset.tar)

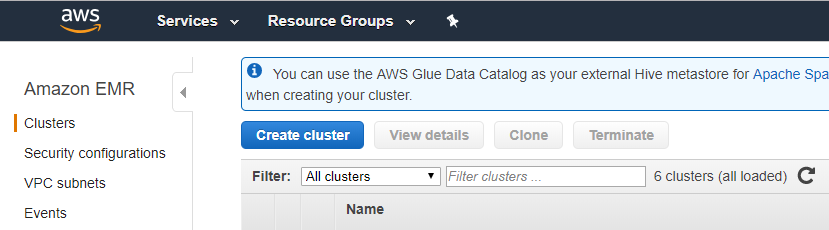
2) Upload data to s3

- $ aws s3 cp photos.json s3://yelp/photos

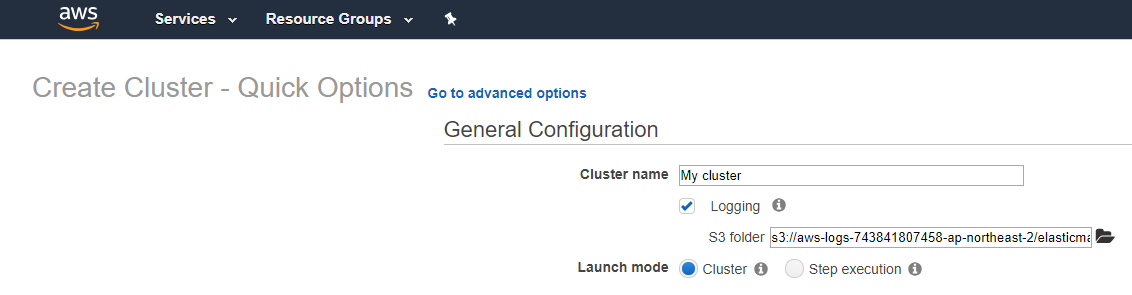
**AWS EMR Cluster Create and Launch**

### Step 1: Create EMR Cluster

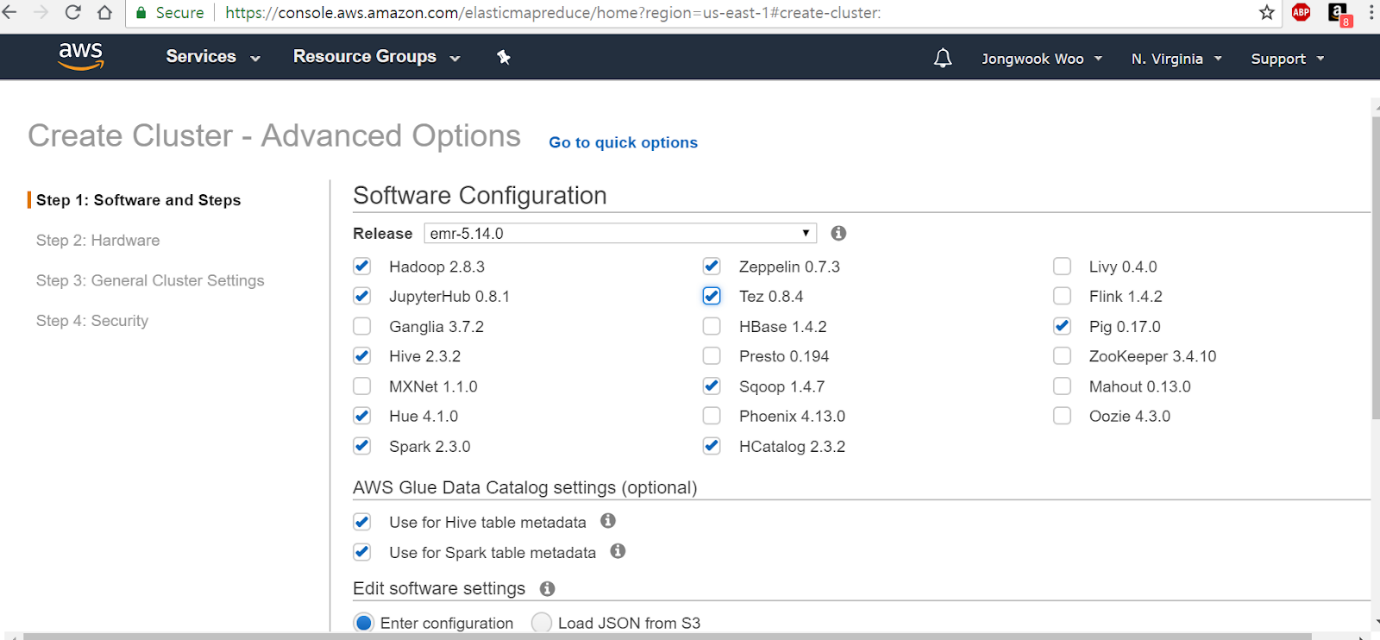
Create Cluster를 Click 한다.

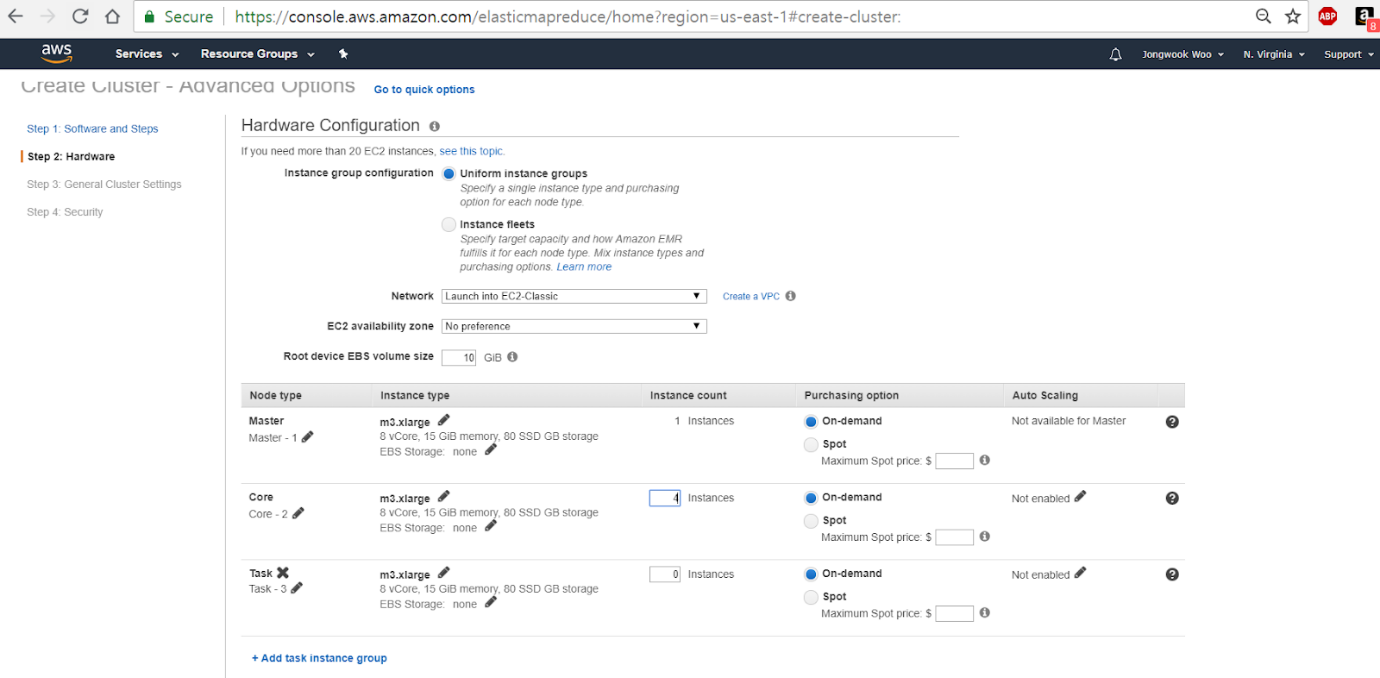


Go to Advanced options 를 Click 한다.



아래와 같이 설정을 한다.





### Step 2: Launch EMR Cluster

Creation 생성 이후 일정시간 지난후 Launch가 된다.

**Setup Web Connection**

### Step 1: Open an SSH Tunnel to the Amazon EMR Master Node

1) SSH Tunnel 생성을 위하여 아래 명령 실행

ssh -i ~/.ssh/dalseoulkey.pem -ND 8157 [hadoop@ec2-13-209-18-59.ap-northeast-2.compute.amazonaws.com](mailto:hadoop@ec2-13-209-18-59.ap-northeast-2.compute.amazonaws.com) &

### Step 2: Configure a proxy management tool

1) FoxyProxy Download and Install (<http://foxyproxy.mozdev.org/downloads.html>)

2) Restart Chrome after installing FoxyProxy

3) Using a text editor create a file named foxyproxy-settings.xml containing the following:

<?xml version="1.0" encoding="UTF-8"?>  
<foxyproxy>  
    <proxies>  
        <proxy name="emr-socks-proxy" id="2322596116" notes="" fromSubscription="false" enabled="true" mode="manual" selectedTabIndex="2" lastresort="false" animatedIcons="true" includeInCycle="true" color="#0055E5" proxyDNS="true" noInternalIPs="false" autoconfMode="pac" clearCacheBeforeUse="false" disableCache="false" clearCookiesBeforeUse="false" rejectCookies="false">  
            <matches>  
                <match enabled="true" name="\*ec2\*.amazonaws.com\*" pattern="\*ec2\*.amazonaws.com\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="\*ec2\*.compute\*" pattern="\*ec2\*.compute\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="10.\*" pattern="http://10.\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="\*10\*.amazonaws.com\*" pattern="\*10\*.amazonaws.com\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="\*10\*.compute\*" pattern="\*10\*.compute\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="\*.compute.internal\*" pattern="\*.compute.internal\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
                <match enabled="true" name="\*.ec2.internal\*" pattern="\*.ec2.internal\*" isRegEx="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" />  
            </matches>  
            <manualconf host="localhost" port="8157" socksversion="5" isSocks="true" username="" password="" domain="" />  
        </proxy>  
    </proxies>  
</foxyproxy>

4) Click on the FoxyProxy icon in the toolbar and select Options.

5) Click Import/Export.

6) Click Choose File, select foxyproxy-settings.xml, and click Open.

7) In the Import FoxyProxy Settings dialog, click Add.

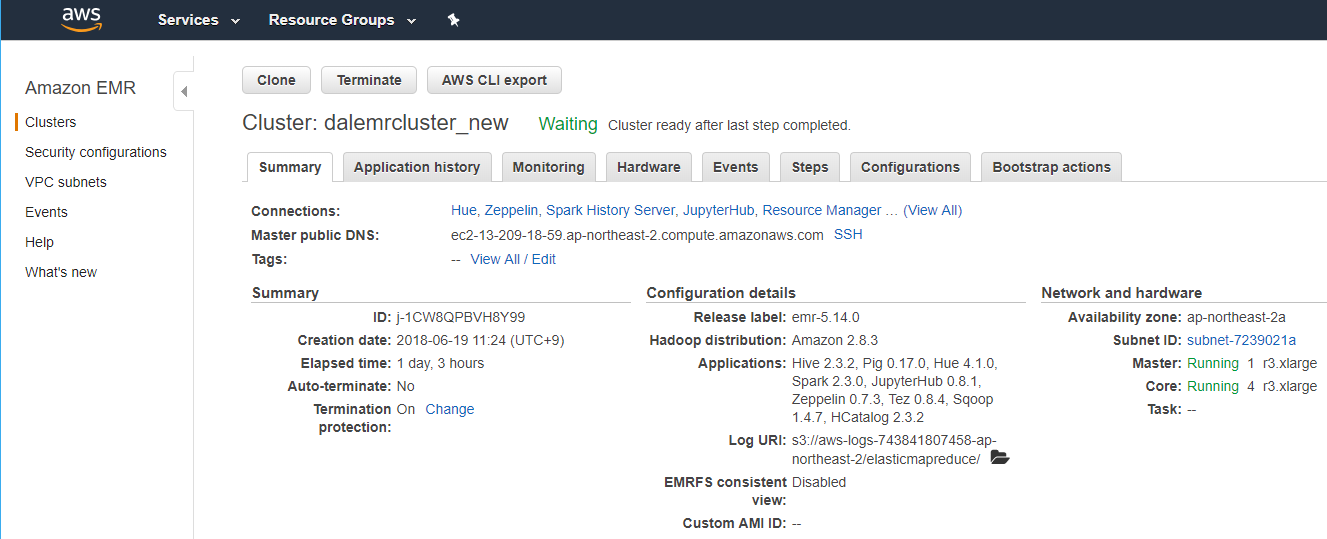
8) At the top of the page, for Proxy mode, choose Use proxies based on their pre-defined patterns and priorities

9) To open the web interfaces, in your browser's address bar, type *master-public-dns* followed by the port number or URL.

**Execute Hue and Zeppelin**

### Step 1: Amazon EMR Cluster 화면을 통하여 실행

아래 화면에서 Hue, Zeppelin Hyper Link를 click 하여 실행



### Step 2: Hue and Zeppelin 실행 확인

