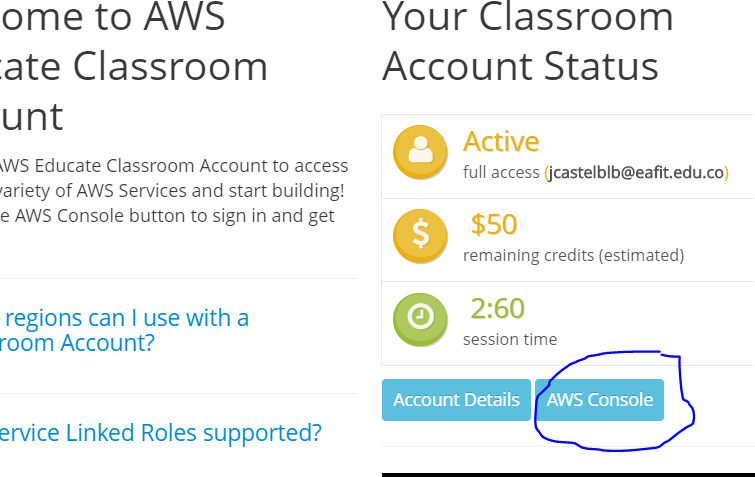
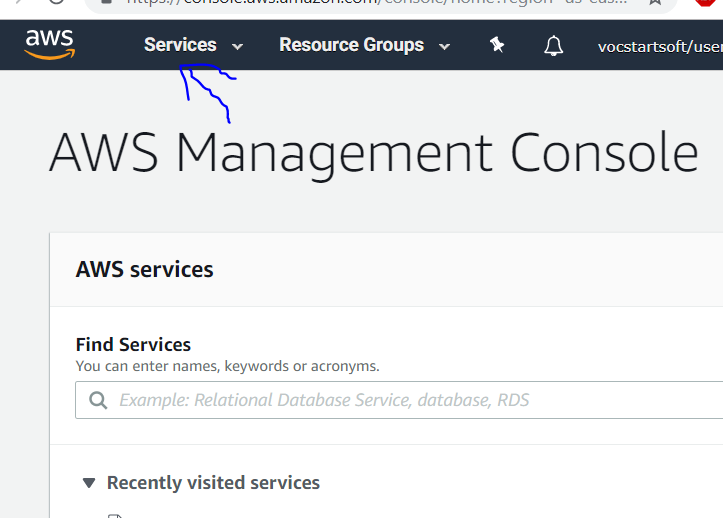
Paso a Paso:

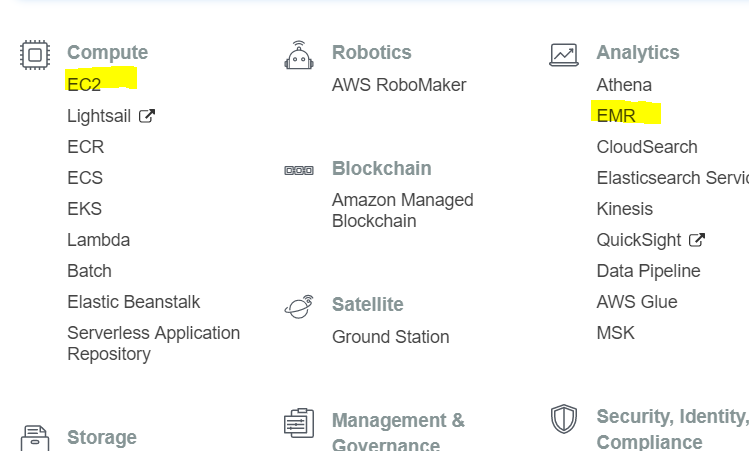
1. Ingreso a <https://www.awseducate.com/signin/SiteLogin>
2. Luego le damos click a My Classrooms o a AWS Account
3. Click Go to Classroom y Continuar
4. Seleccionamos AWS Console



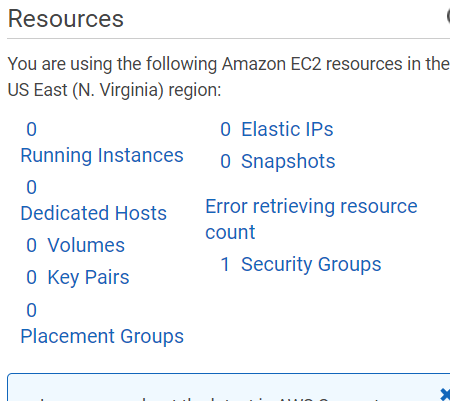
1. Seleccionamos Services



1. Tendremos en cuenta EC2 y EMR

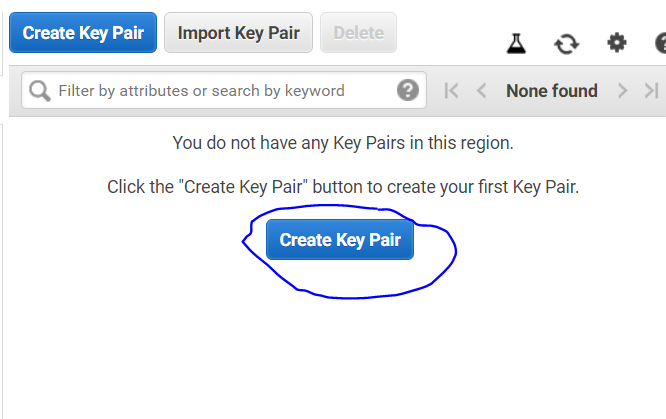


1. Vamos a entrar primero a EC2



Podemos observar que tenemos Instancias corriendo y 0 keys (No se tienen archivos de autenticación en la máquina que vamos a crear.

1. Seleccionamos “key Pairs” para poder darle seguridad de acceso a nuestra máquina



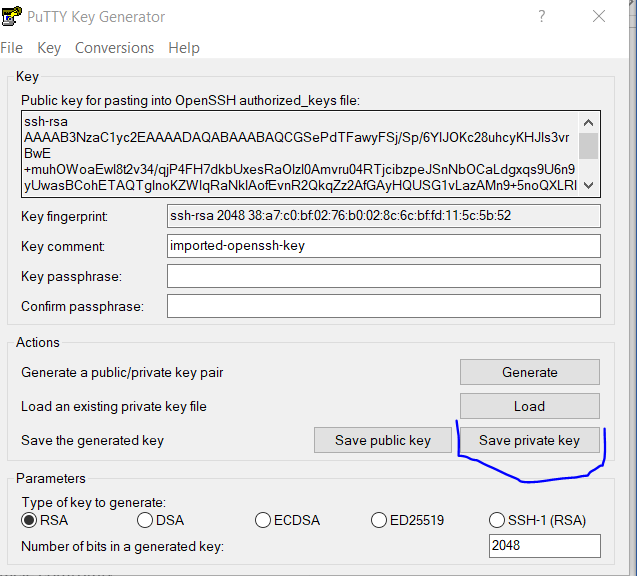
Le damos un nombre cualquiera que haga referencia a la llave; ejemplo “FinWork” y se descargará un archivo “FinWork.pem”

Nota: Necesitaremos instalar Putty: <https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>



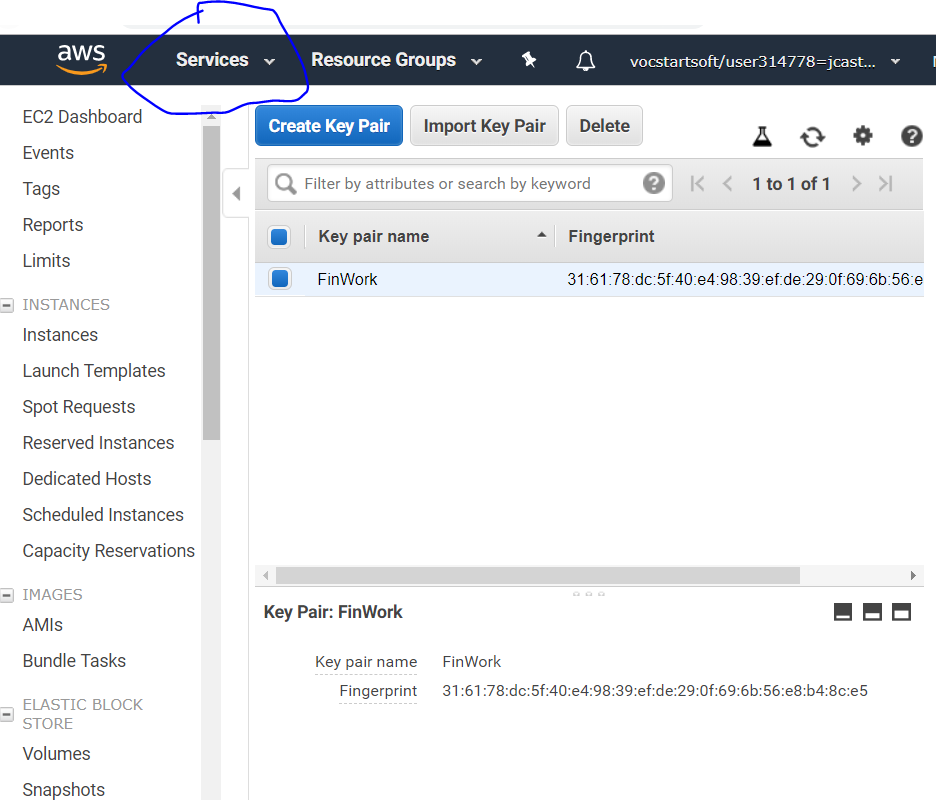
1. Abrimos Puttygen y cargamos el a rchivo “.pem”

Y nos debe de quedar algo como lo siguiente:



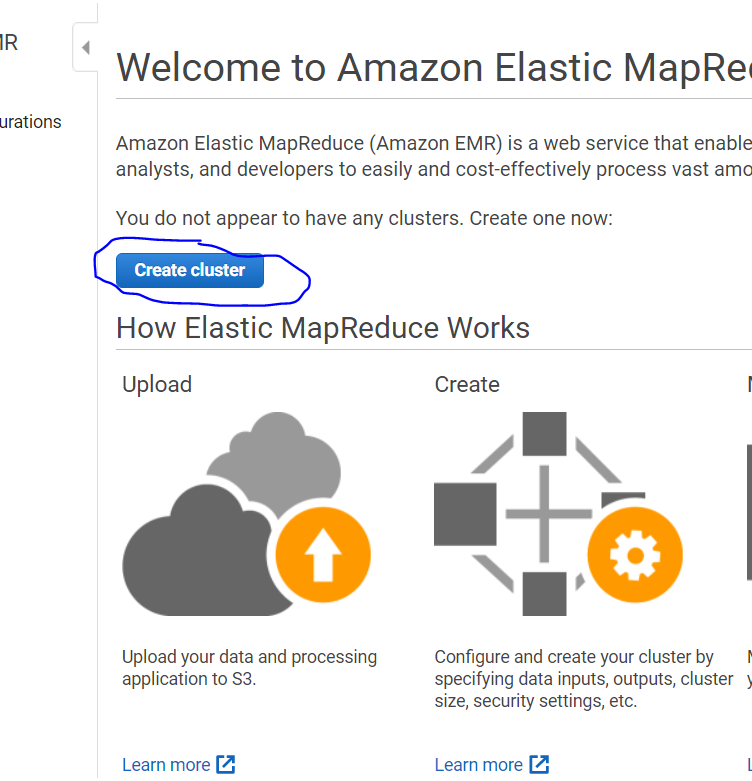
Le damos click a “Save private key” y lo guardamos con el mismo nombre y cerramos.

1. Luego volvemos a AWS

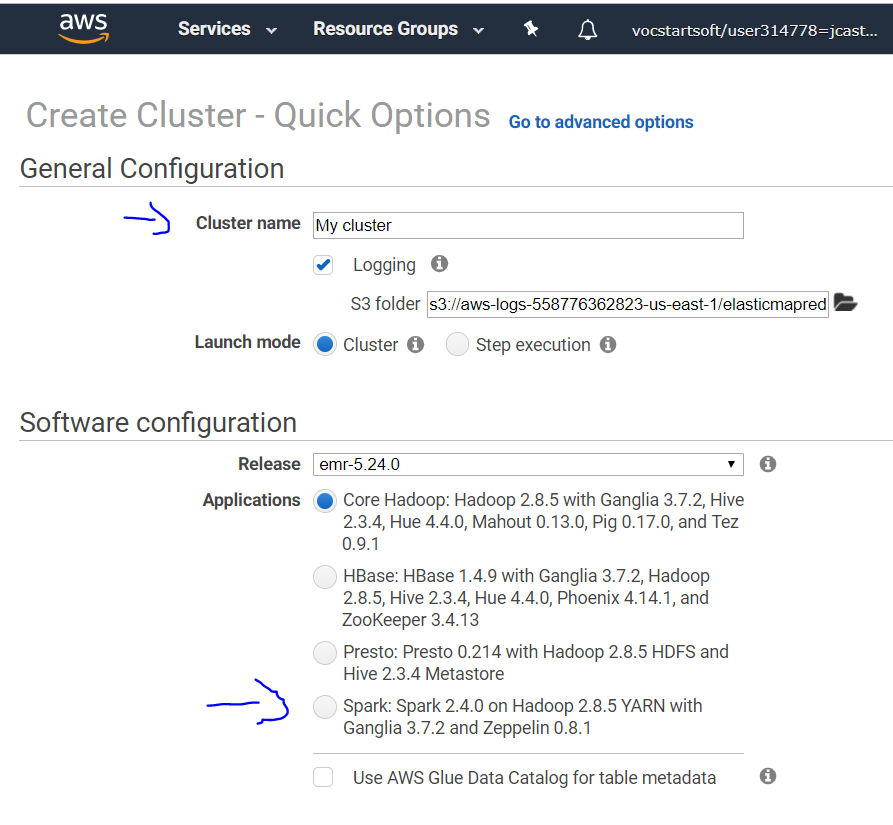


Y seleccionamos “services”

1. Vamos a entrar ahora a EMR y seleccionamos “Create cluster”

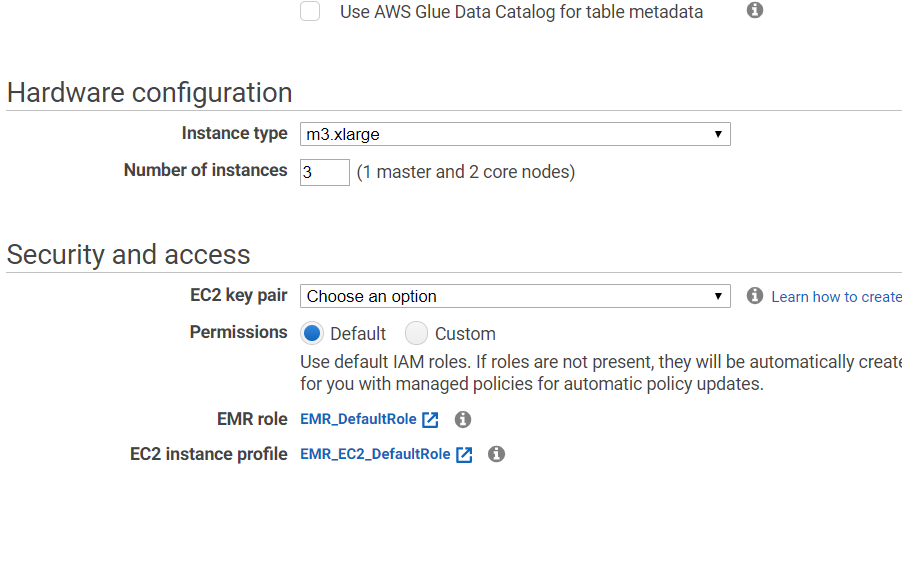


1. Luego:



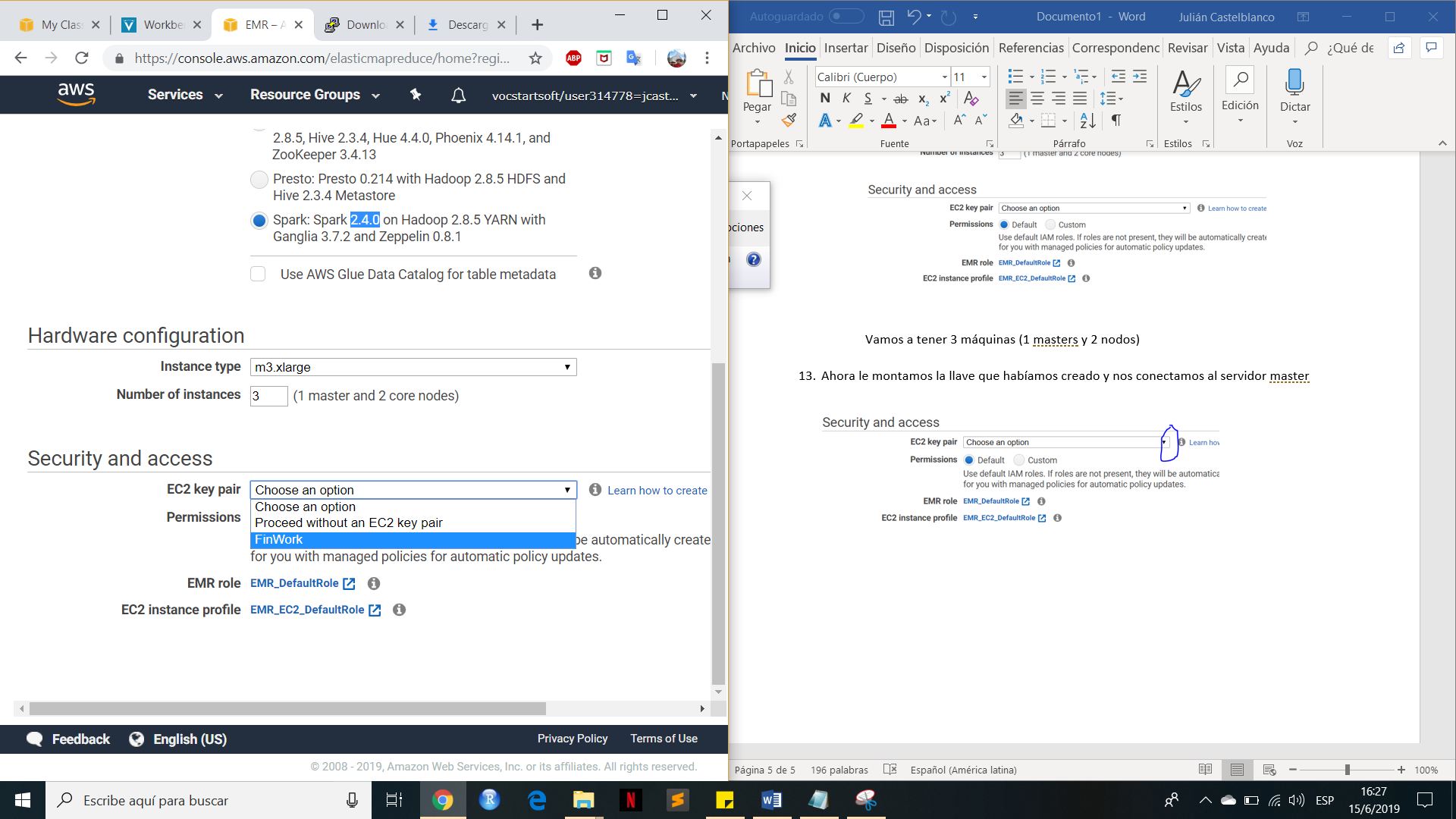
Le otorgamos un nombre y seleccionamos la versión de Spark

Nota:



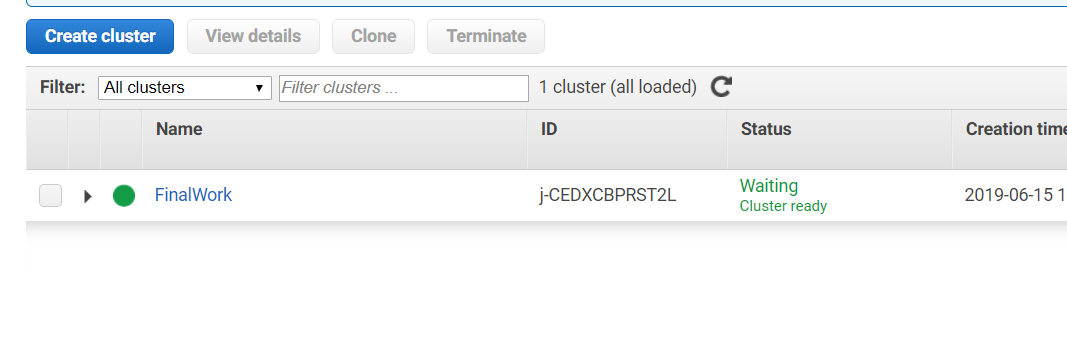
Vamos a tener 3 máquinas (1 masters y 2 nodos)

1. Ahora le montamos la llave que habíamos creado y nos conectamos al servidor master

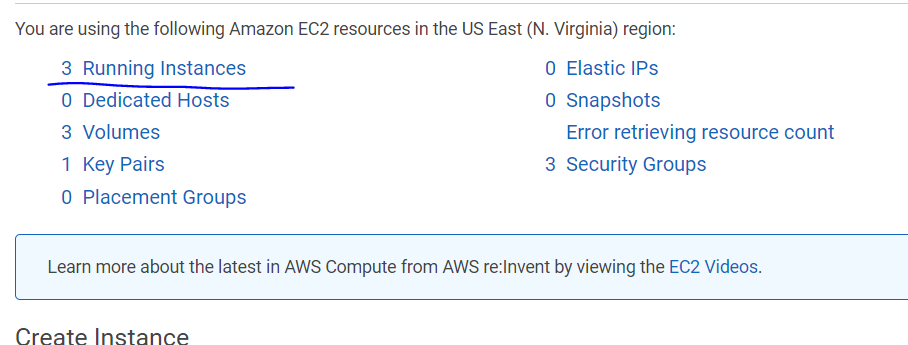


1. Luego le damos Create cluster. Con esto ya tendremos creadas nuestras máquinas virtuales en “Cluster”. Entendamos como Cluster como un grupo de máquinas que trabajan al mismo tiempo y si alguno llega a caerse los otros siguen trabajando.

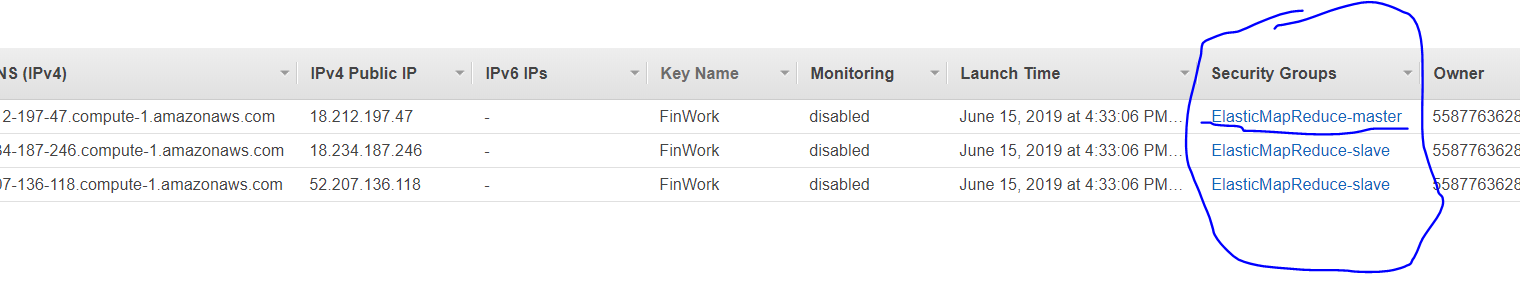
Si realizamos todo bien, saldría:



1. Luego nos dirigimos a Services y seleccionamos EC2 y nos deben salir tres instancias corriendo y le damos click

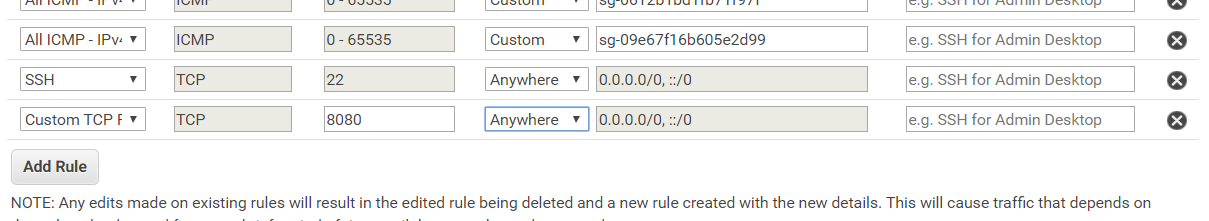


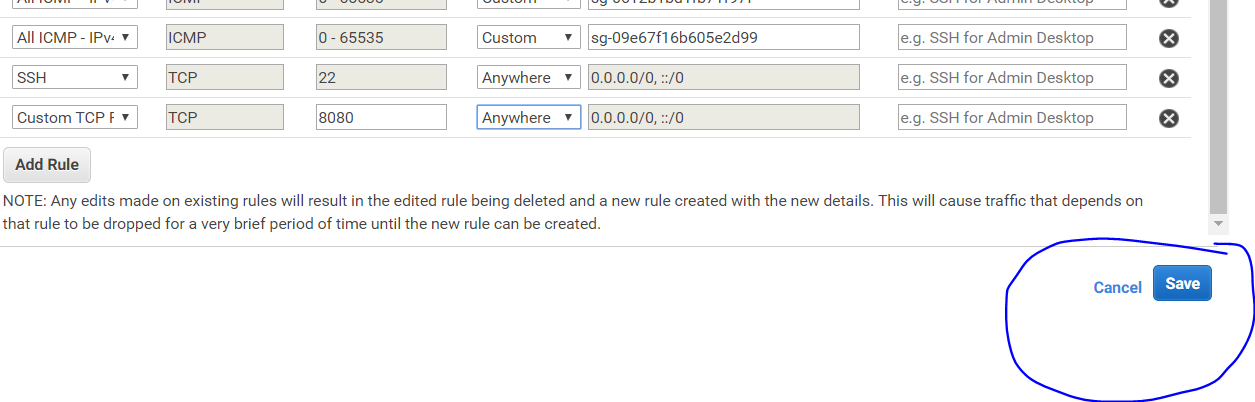
1. Al final encontraremos el Cluster Master y le damos click



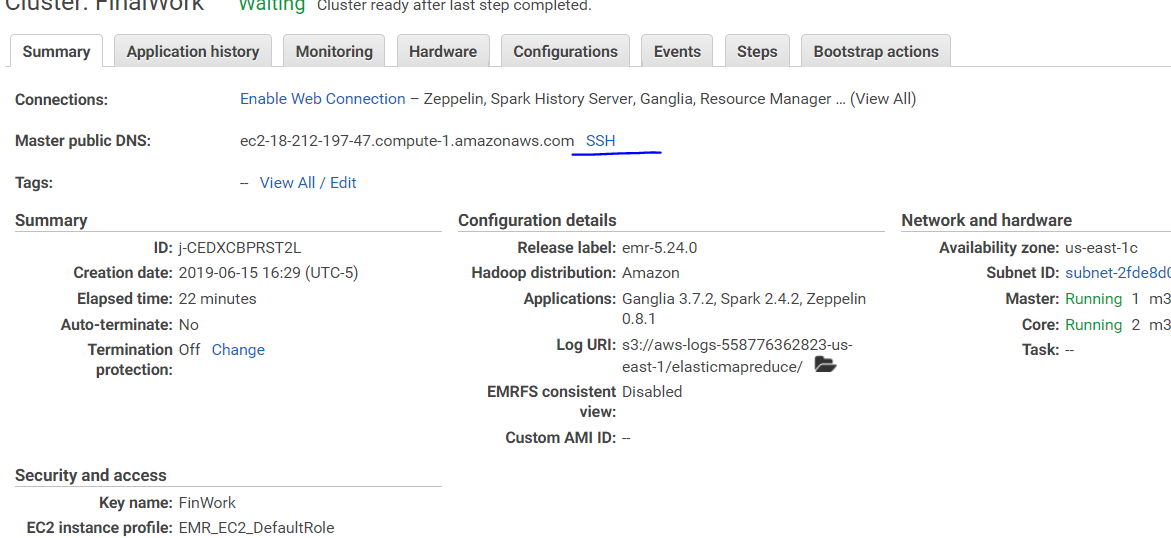
1. Luego le damos editar y agregamos 2 Rules

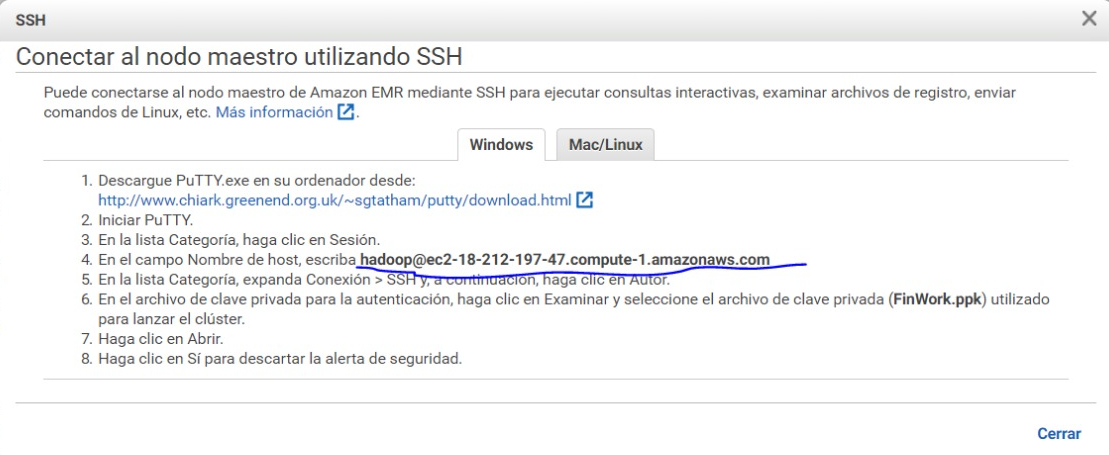
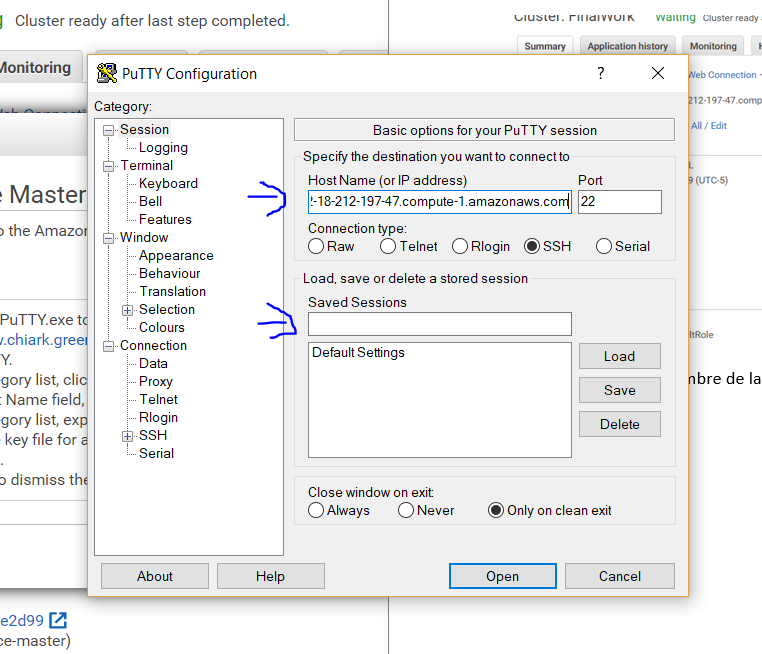
((SSH -> puerto 22) y (Custom TCP Rule -> puerto 8080))

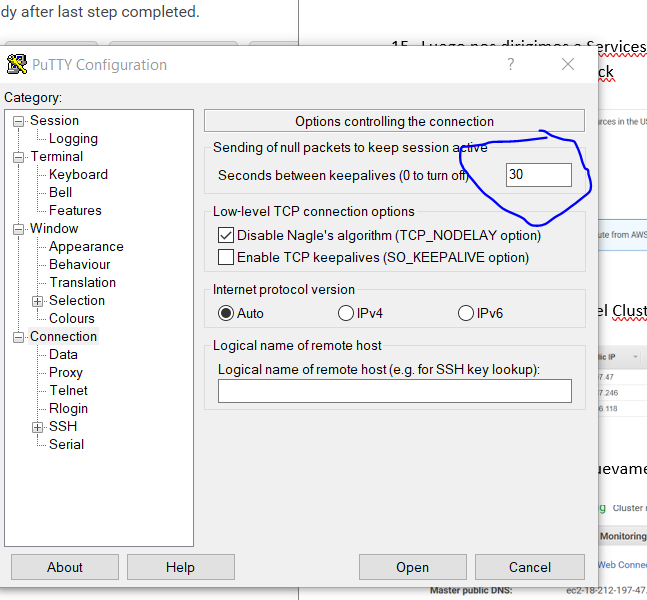




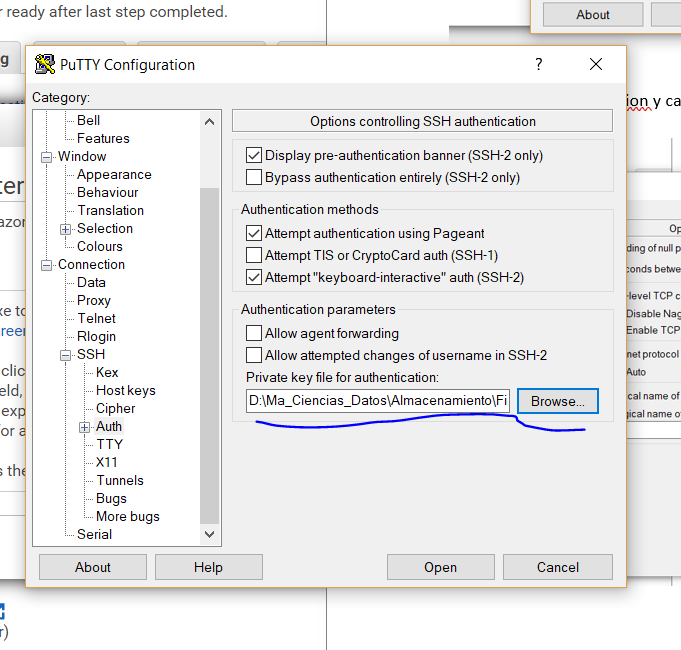
1. Luego seleccionamos nuevamente el master en EMR



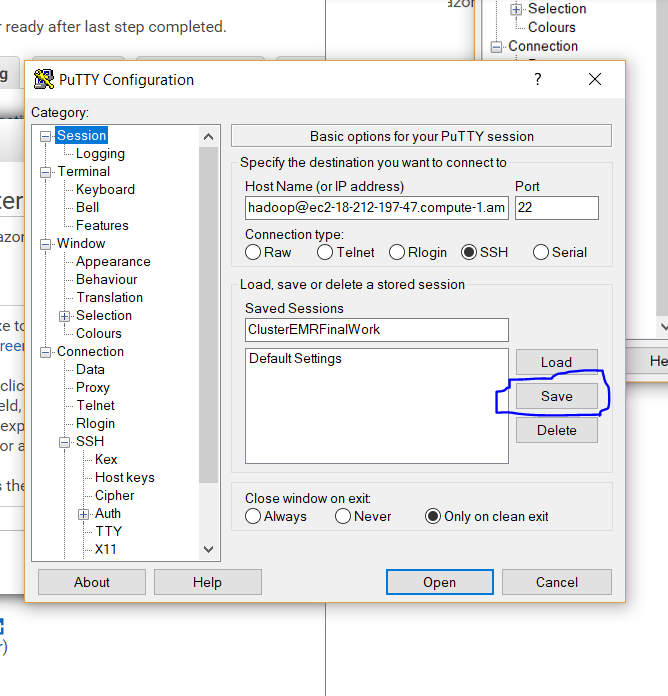
1. 
2. Y luego copiamos el nombre de la máquina, con este ingresaremos desde putty
3. Y llenamos
4. Luego vamos a Connection y cambiamos por el tiempo que deseamos que mantenga viva la sesión



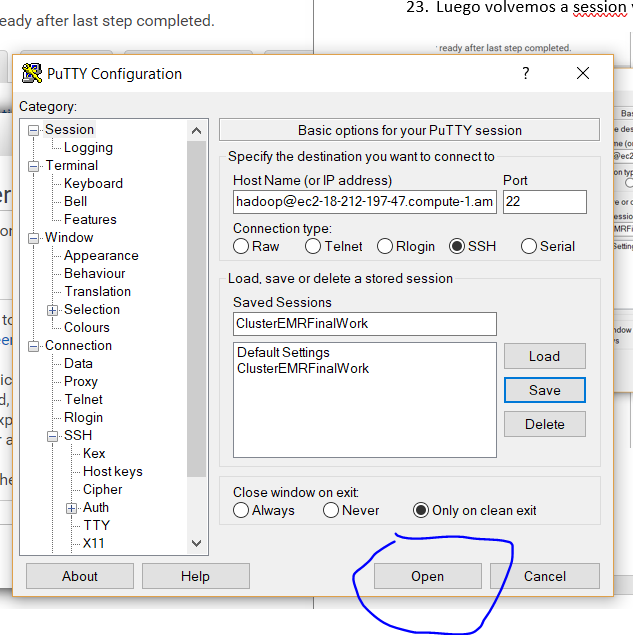
1. Luego en SSH – Auth subimos la .ppk que habíamos generado anteriormente



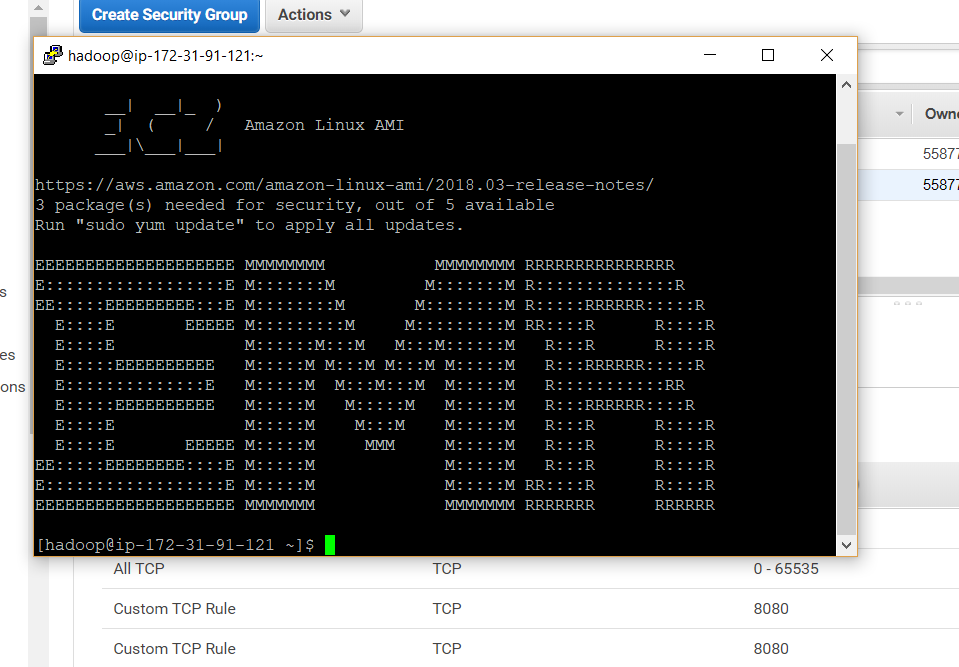
1. Luego volvemos a session y guardamos



1. Luego Abrimos



1. Si realizamos todo bien, nos debe salir algo como esto



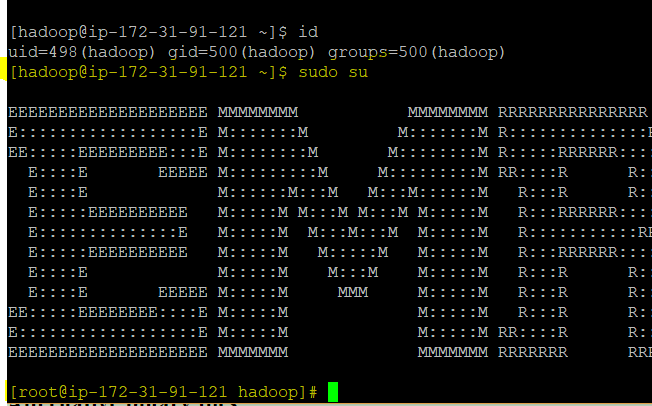
1. Se sigue las instrucciones del siguiente link <https://github.com/st1800eafit/class20191/blob/master/aws/jupyter-install.md>

**contectarse al master del EMR-AWS:**

sudo su (Esto se hace para conectar el usuario root)

passwd hadoop

// coloquele el password que desee a hadoop

# su – hadoop

Se coloca password **hadoop**

## instalar node:

Vamos a instalar los siguientes pasos:

yum install -y gcc-c++ make

curl -sL https://rpm.nodesource.com/setup\_6.x | sudo -E bash -

yum install -y nodejs

npm install -g configurable-http-proxy

## instalar virtualenv:

su - hadoop

virtualenv env --python=python3

cd env

source bin/activate

**deactivate (Esta instrucción desactiva el ambiente virtual)**

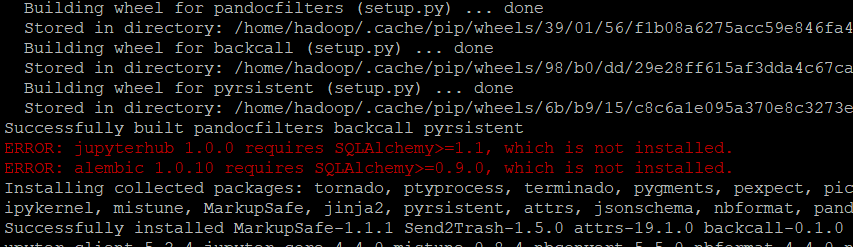
## instalar jupyterhub:

python3 -m pip install jupyterhub

npm install configurable-http-proxy

python3 -m pip install notebook

En esta instalacion se observa un error, el cual no es de alarmar



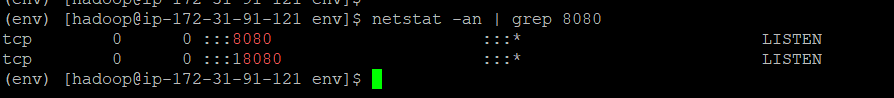
**instalar dependencias:**

easy\_install jinja2 sqlalchemy tornado notebook

**ejecutar jupyterhub:**

jupyterhub --port 8080 &

Luego de esta instalación se la instrucción **netstat -an | grep 8080, con el fin de verificar que el proceso este corriendo en el puerto 8080**



Se abre Google crome y se coloc la url de la maquina creada (es el mismo nombre de maquina que se utilizo en putty) en el puerto 8080 como se muestra a continuación.