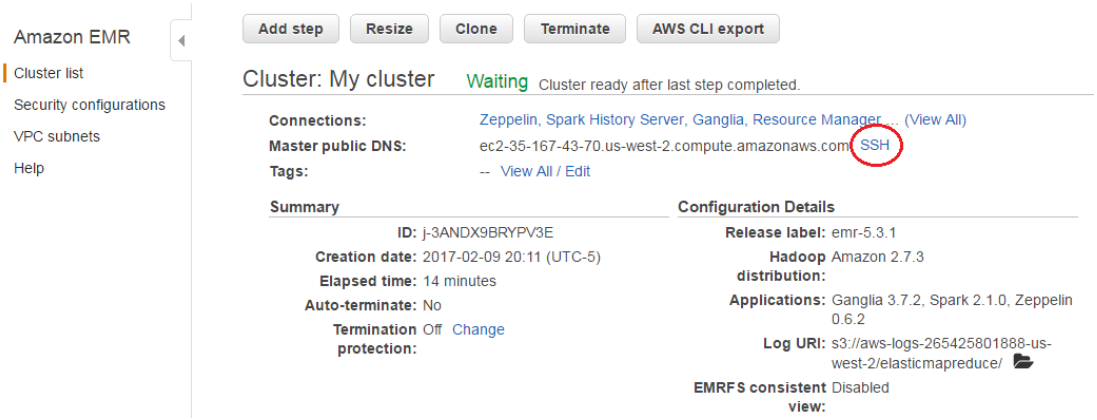
1. Go to the AWS EMR console
2. Click **Create cluster**
3. Under Software configuration, Applications select **Spark: Spark 2.1.0 on Hadoop 2.7.3 YARN with Ganglia 3.7.2 and Zeppelin 0.6.2**
4. Under Hardware configuration, change **Number of Instances** to **1.**
5. Under Security and access, for EC2 Key pair, select the key pair you created in Lab 1 (you will need to have either the ppk (Windows) or pem (Mac) private key file stored on your local computer).
6. Click **Create cluster**
7. Wait for your cluster to start up. After a few minutes, the Master public DNS field for your cluster should be filled in (you can refresh by clicking the cycle icon in the upper right corner) as shown in the picture below.
8. Click the **SSH** link (see picture below).



9. Follow the instructions in the popup window to ssh into the master node of your EMR cluster (you can select instructions for either Windows or Mac by clicking the tabs in the popup).
10. Note: you will have to wait until the cluster is finished booting up before you can run spark programs using pyspark or spark-submit. You can check if they have been loaded yet by typing “which pyspark” or “which spark-submit”.