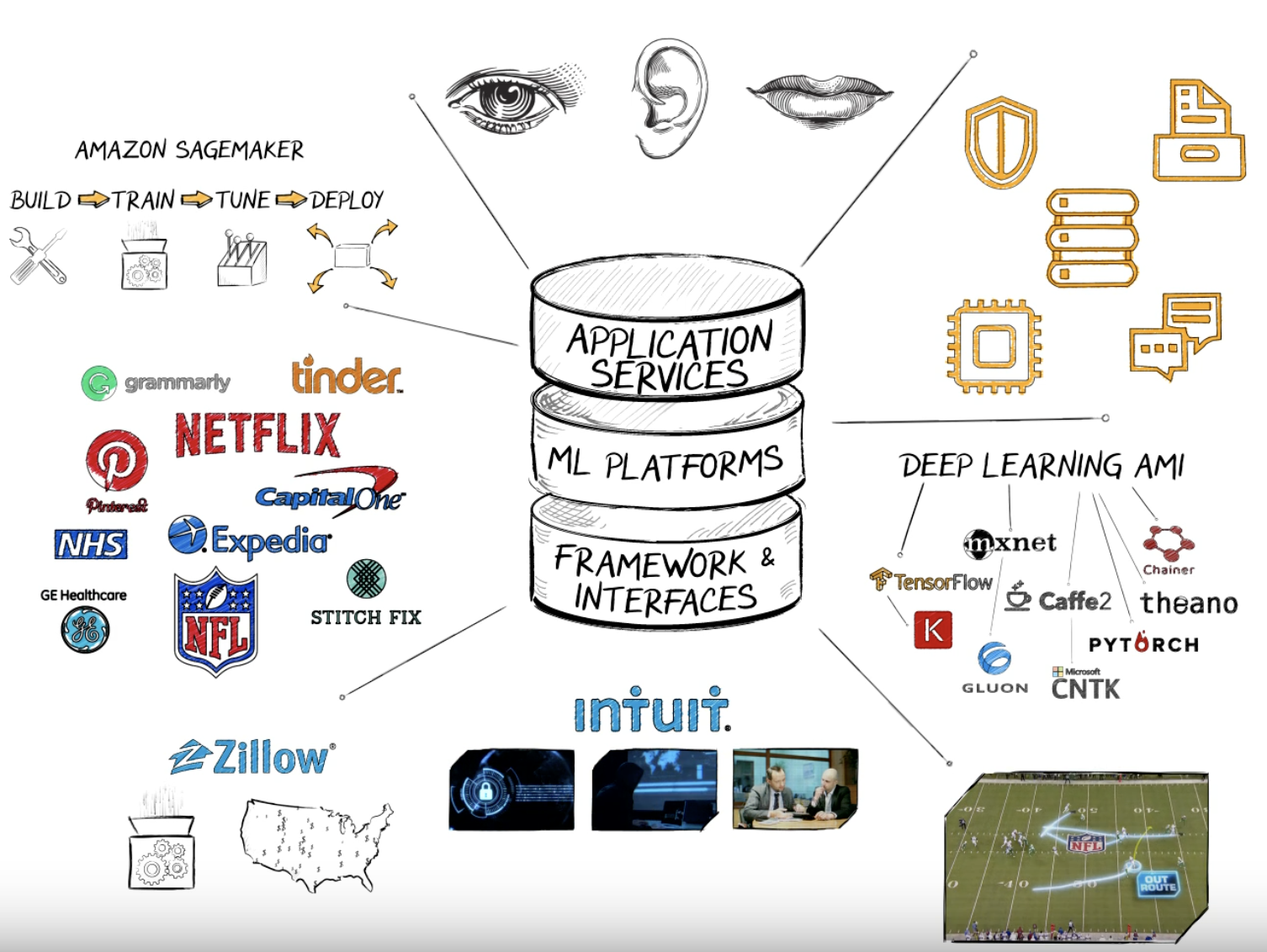
**AWS Tutorial: Analyze Big Data with Hadoop**

Time spent: 70min

**What I learned:**

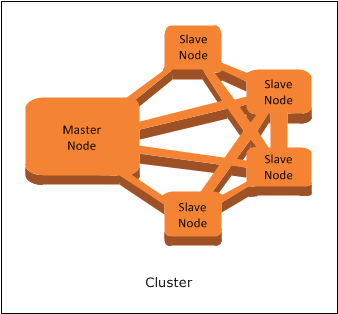
In the overview video, I learned that there was three layers on the AWS ML.

**AWS Machine Learning Overview**



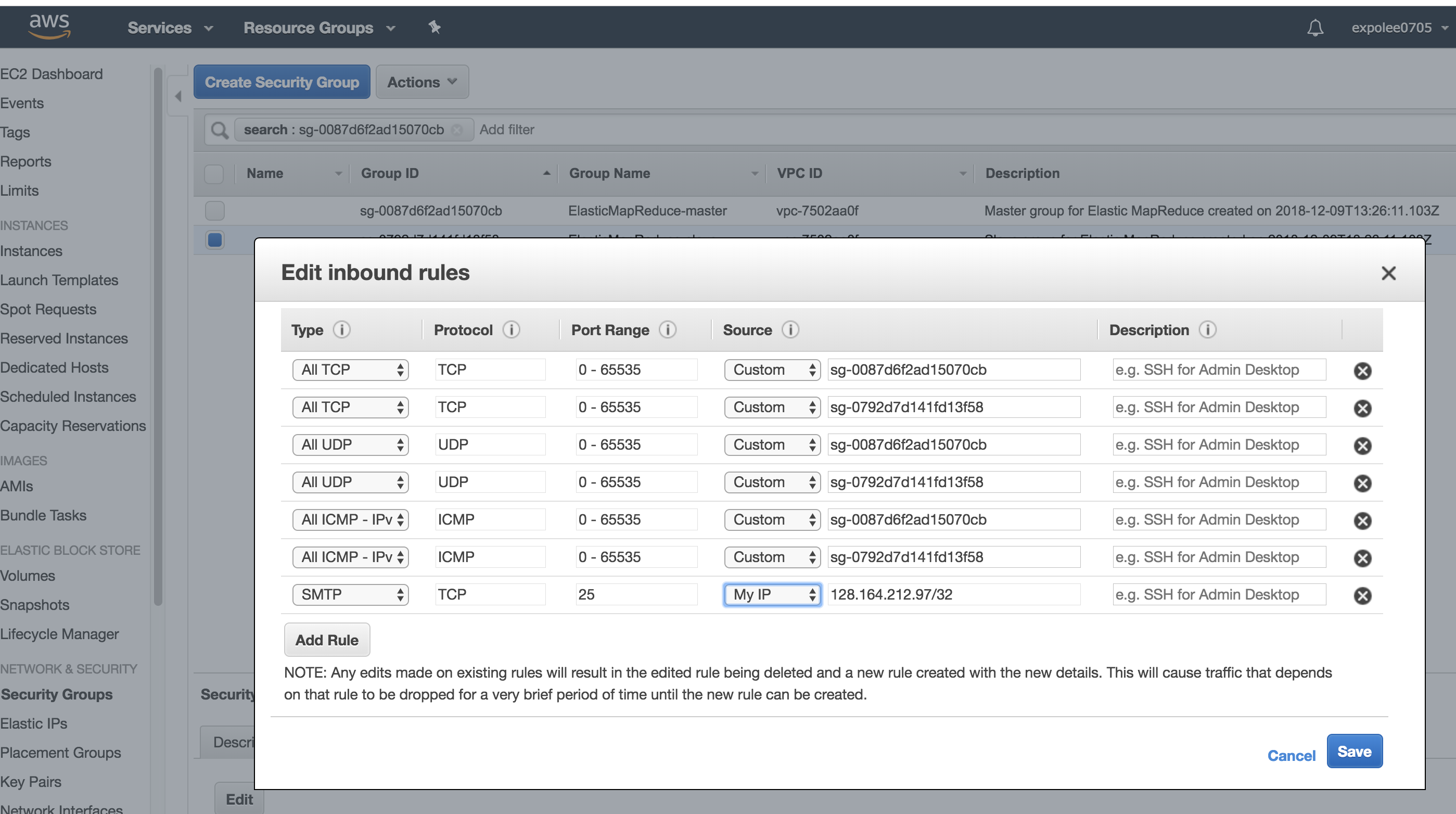
**AWS EMR Overview**

The central component of Amazon EMR is the *cluster*. A cluster is a collection of Amazon Elastic Compute Cloud (Amazon EC2) instances. Each instance in the cluster is called a *node*. Each node has a role within the cluster, referred to as the *node type*. Amazon EMR also installs different software components on each node type, giving each node a role in a distributed application like Apache Hadoop.

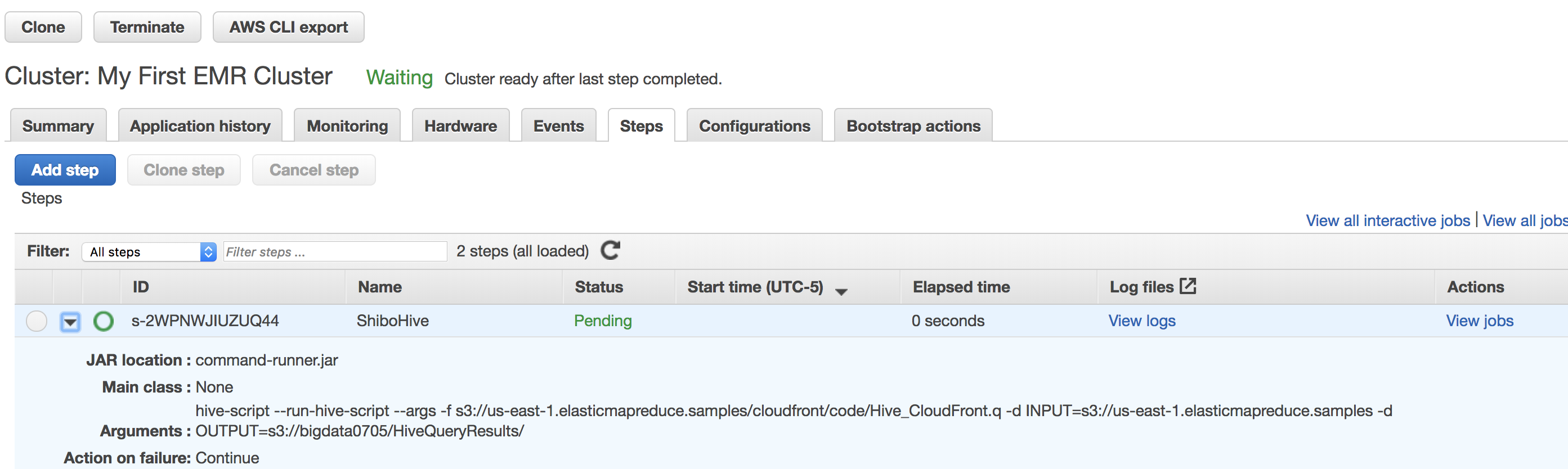


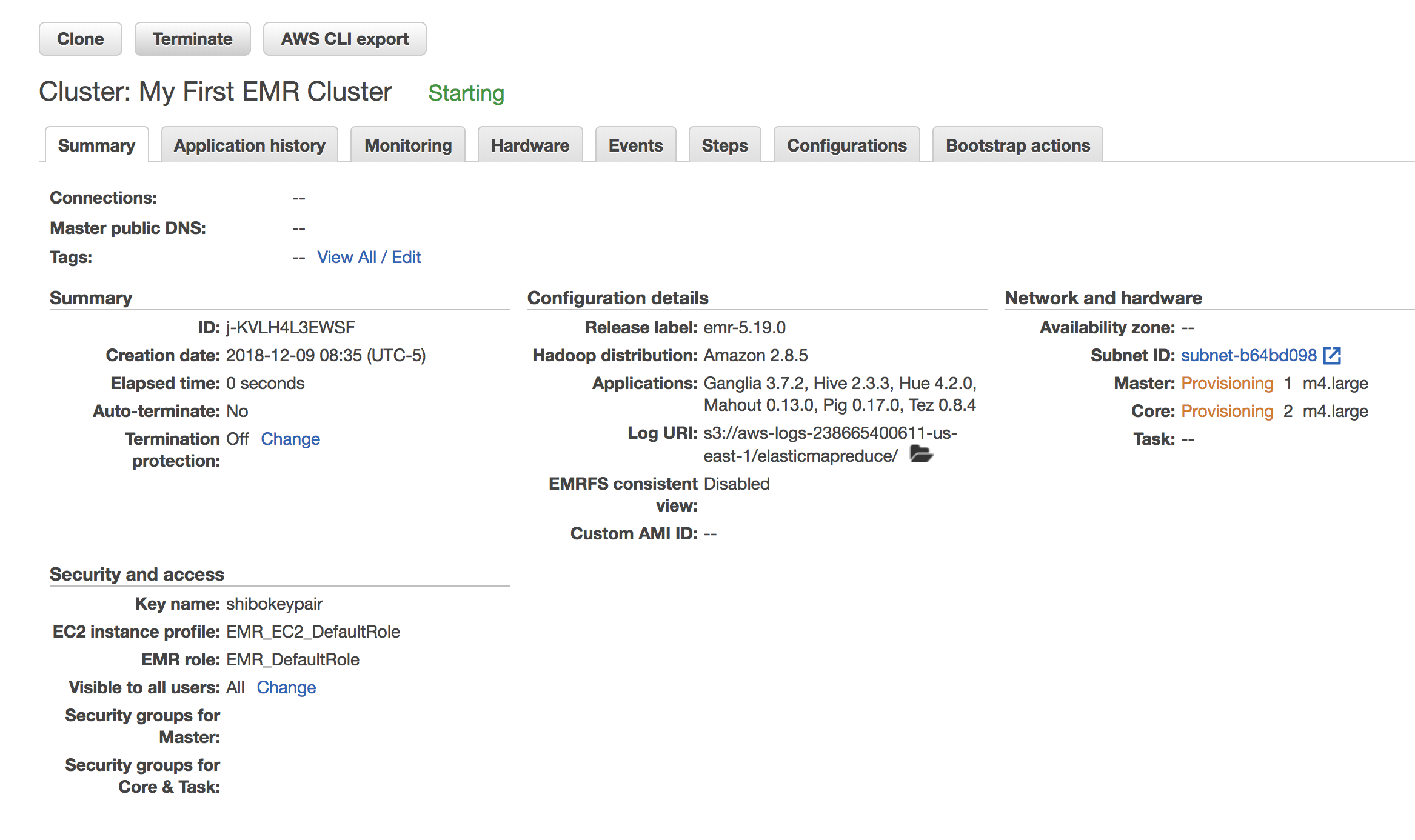
**How to create my cluster**

**Add SSH Access**

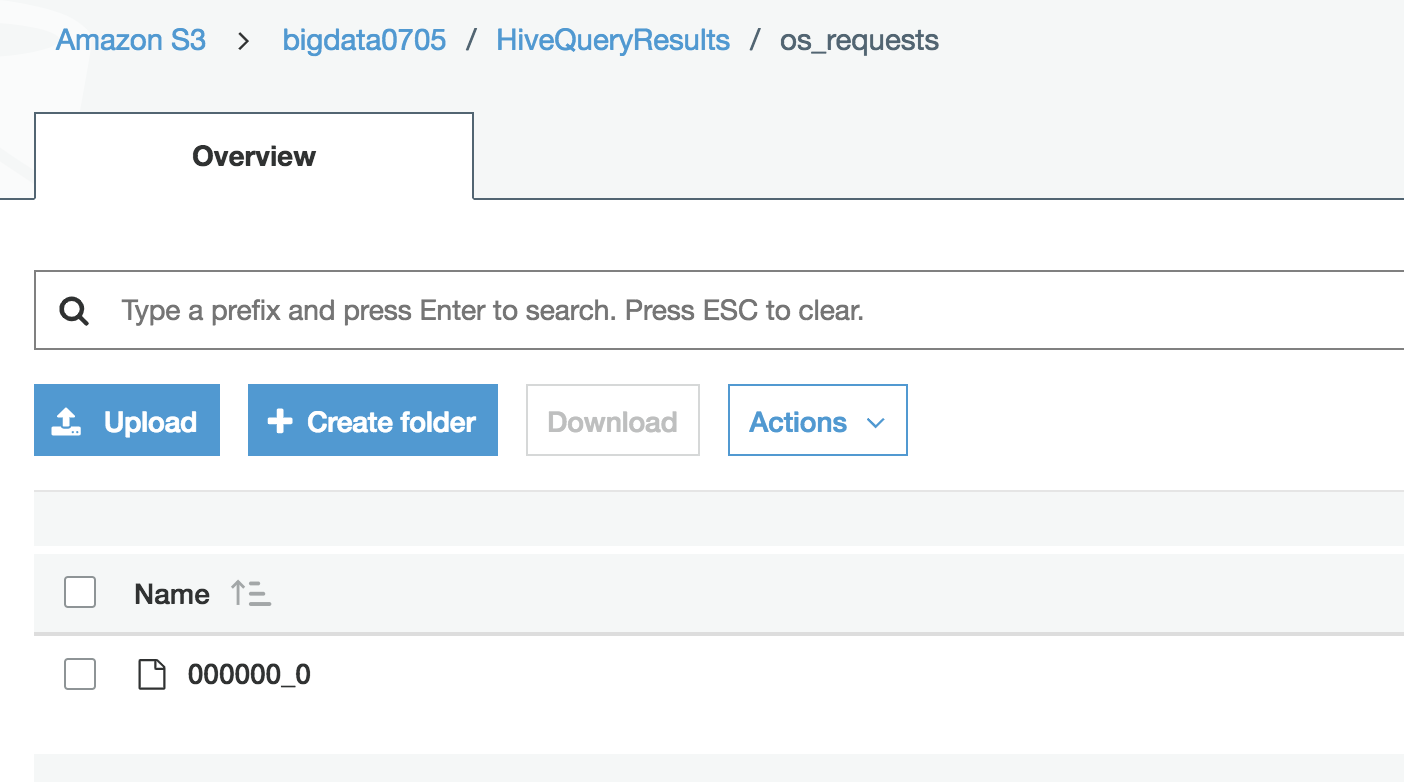
****

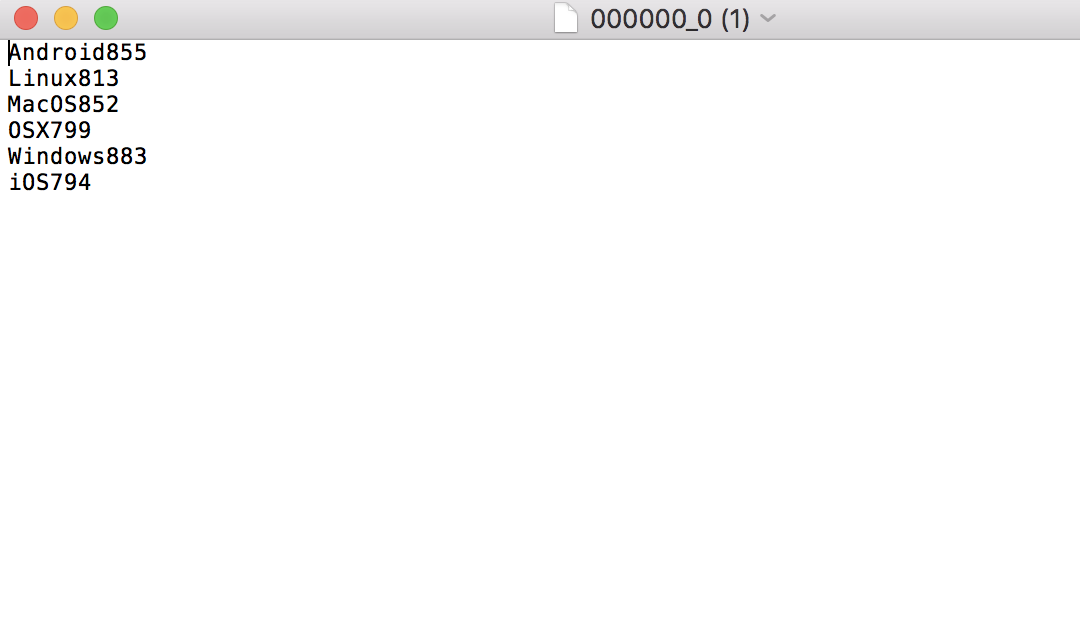
**Start the script by adding steps**

****



**Get the result in S3**





**Summary**

In this tutorial, I learned the mechanism and structure of AWS EMR. With the help of it, I finished the data process by the Hadoop.