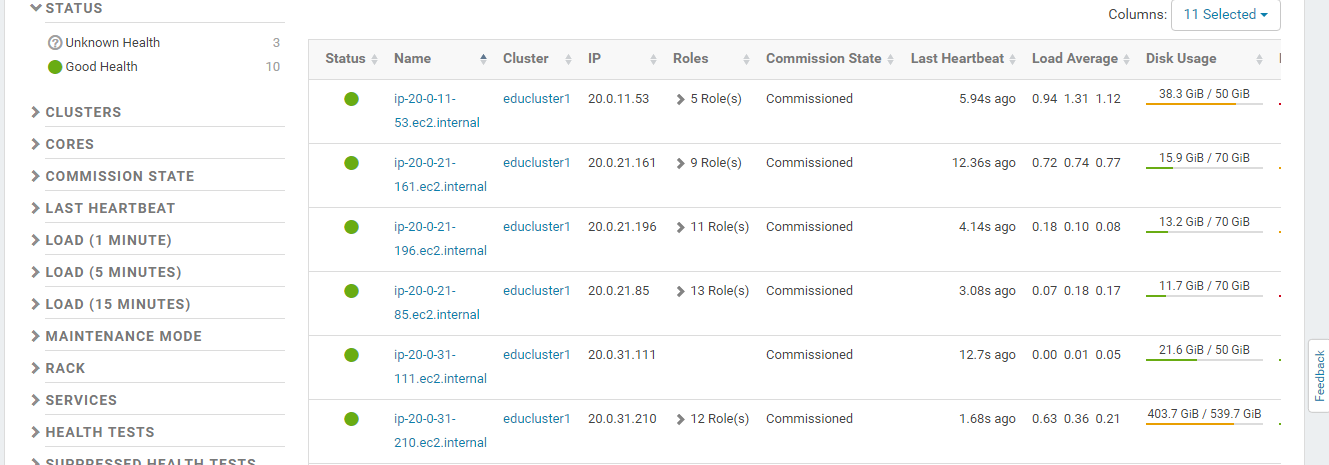
1. Do you use Cloudera Manager to monitor and manage the jobs, cluster, etc.?

Cloudera manager gives the status of the health of the cluster

and the commands executed for cluster management

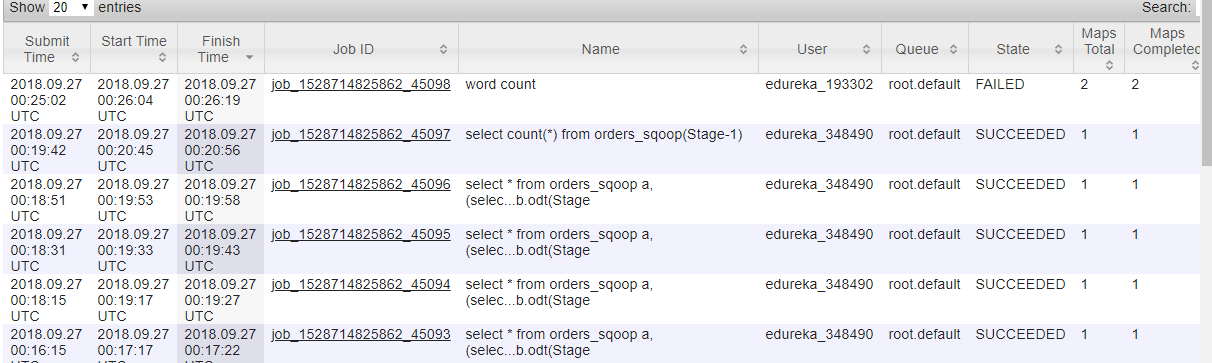
metastore log directories of all services

all hosts-

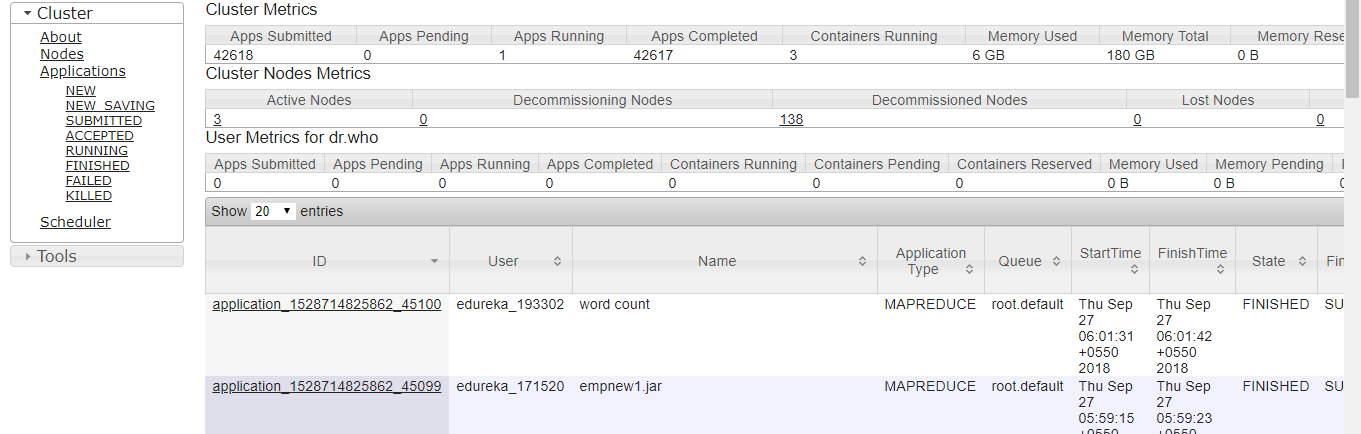


Roll of hosts-

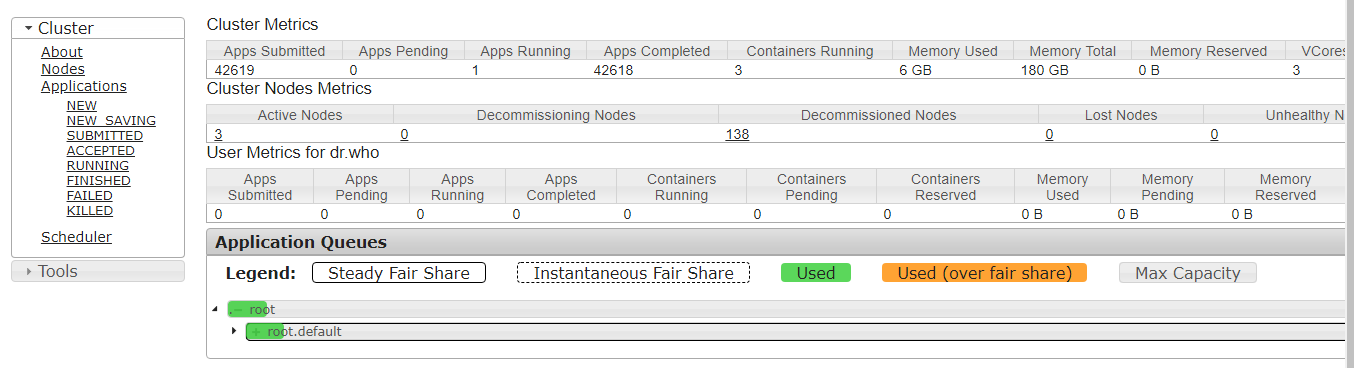
Map Reduce -job history



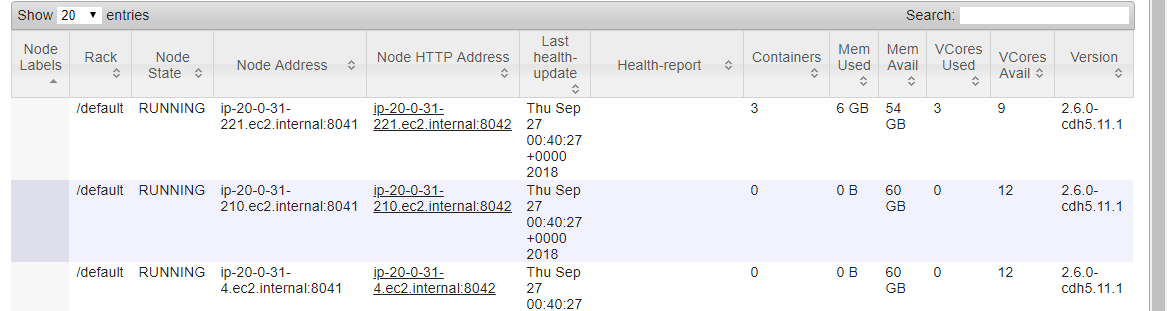
Resource Manager



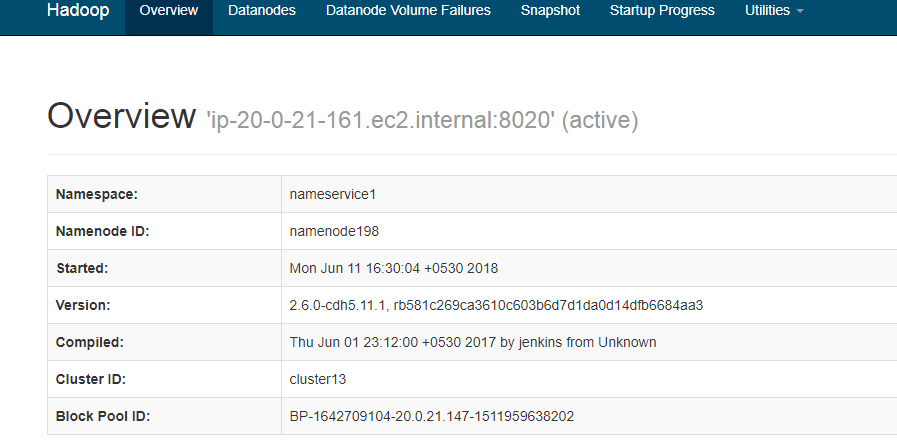
Schedular-

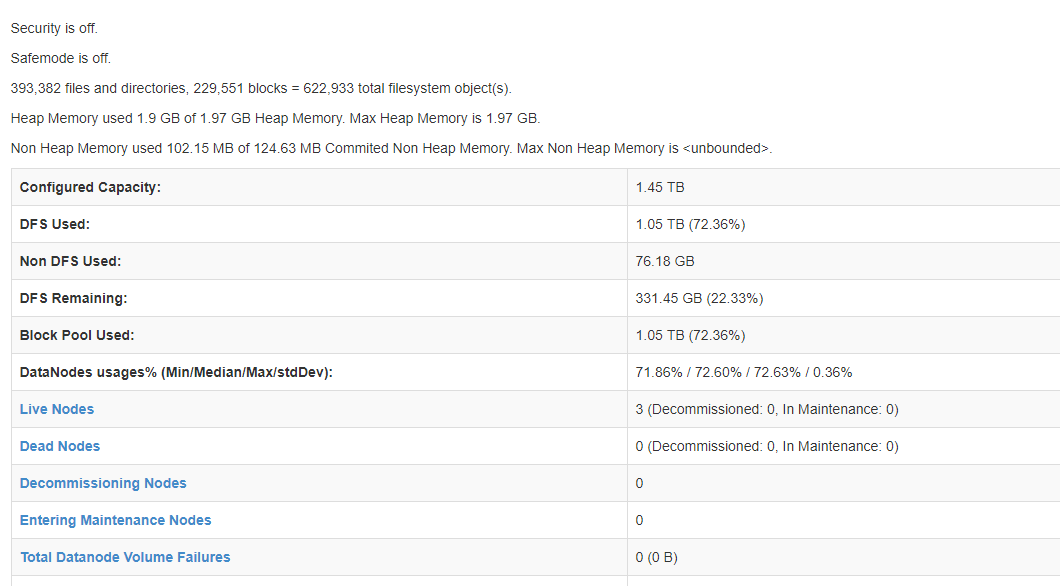


Node status-



Name node web ui





How many Map slots & reducer slots configured in each DataNode? (he didn’t ask this)

What is your cluster size?

What is the DataNode configurations? (RAM, CPU core, Disk size)

What is the NameNode configurations? (RAM, CPU core, Disk size)

Did you write any Hive UDFs?

What are the actions in Action Node?

What is Combiner? Where does it fit and give an example? Preferably from your project.

What is Partitioner? Why do you need it and give an example? Preferably from your project.

How about your experience in Cloudera Manager?

What are the actions in Action Node?