

This computer acts as the master of the cluster, both for HADOOP and for SPARK.

Spark tasks must be launched through it.

teaa-master-cluster

teaa-worker01-cluster

teaa-worker02-cluster

teaa-worker03-cluster

teaa-worker04-cluster

teaa-worker20-cluster

2 CPU cores
6 GB of RAM
40 GB of Disk

2 CPU cores
6 GB of RAM
40 GB of Disk

2 CPU cores
6 GB of RAM
40 GB of Disk

2 CPU cores
6 GB of RAM
40 GB of Disk

2 CPU cores
6 GB of RAM
40 GB of Disk

HDFS: $40\text{GB} \times 20 = 800\text{GB}$ — Spark: $2 \text{ CPU cores} \times 20 \text{ worker nodes} = 40 \text{ CPU cores}$ with a total of 120 GB of RAM