**Assignment**

**Sabaragamuwa University of Sri Lanka**

**Faculty of Computing**

**Department of Software Engineering**

**SE6103 - Parallel and Distributed Systems**

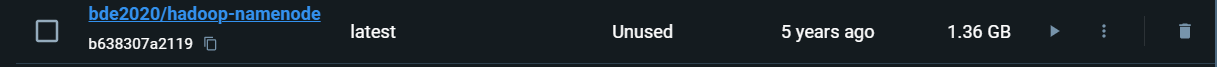
Name: G.V. Wimalagunasekara

Reg. No: 19APSE4295

Academic Period: 3rd Year 2nd Semester

**Question 02**

1. docker pull bde2020/hadoop-namenode:latest – Already Installed



1. docker run -it --name hadoop-cluster -p 9870:9870 -p 8088:8088 -p 50070:50070 bde2020/hadoop-namenode:latest /bin/bash



1. Run namenode

hdfs namenode -format

hdfs namenode &

A screen shot of a computer screen

Description automatically generated

Run datanode

hdfs datanode &

A black screen with white text

Description automatically generated

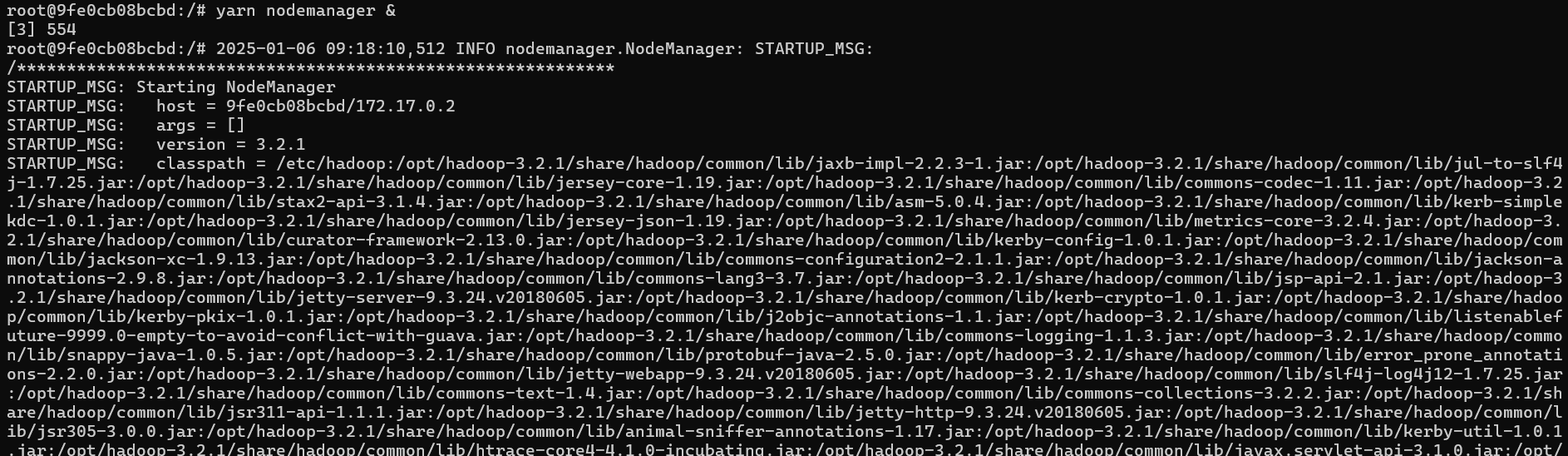
1. access to Hadoop namenode and datanode from http//:localhost:9870

A screenshot of a computer

Description automatically generated

1. Start node manager

yarn nodemanager &



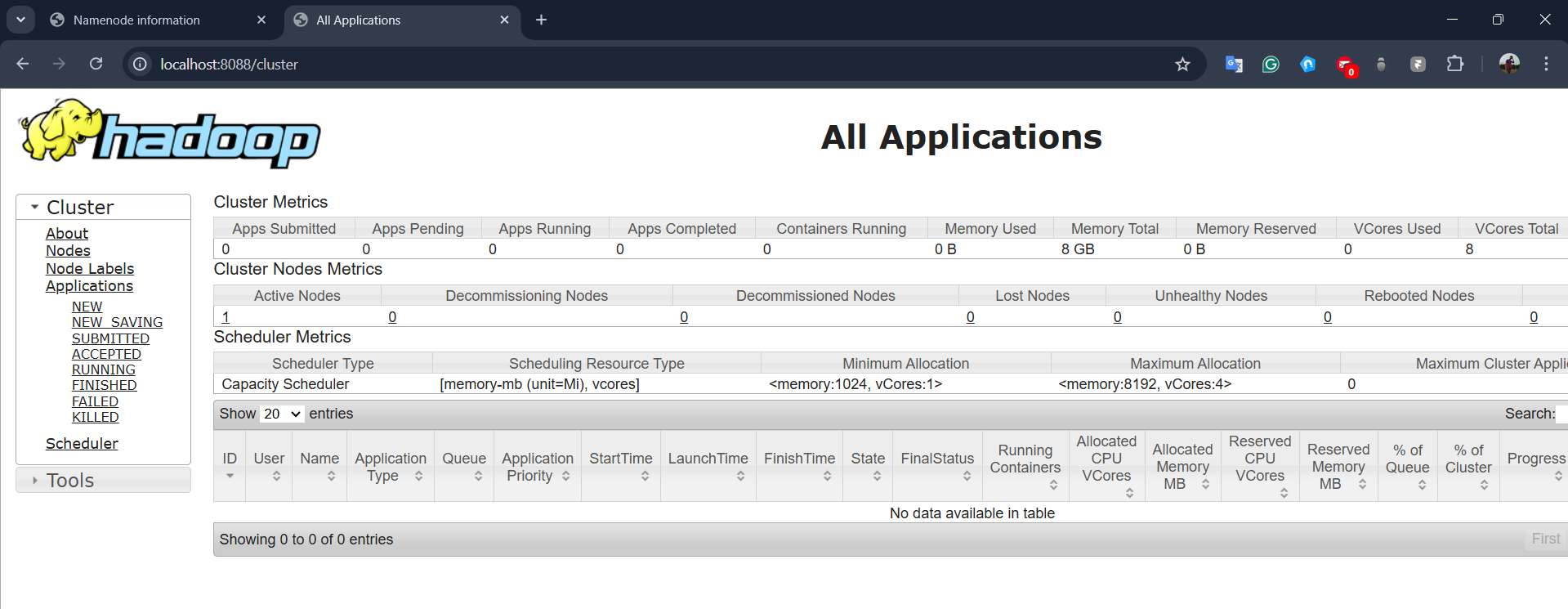
Start resource manager

yarn resourcemanager &

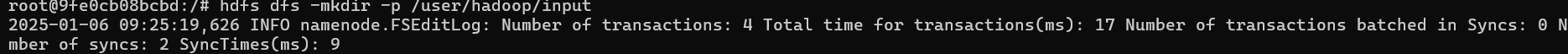
A black and white screen with white text

Description automatically generated

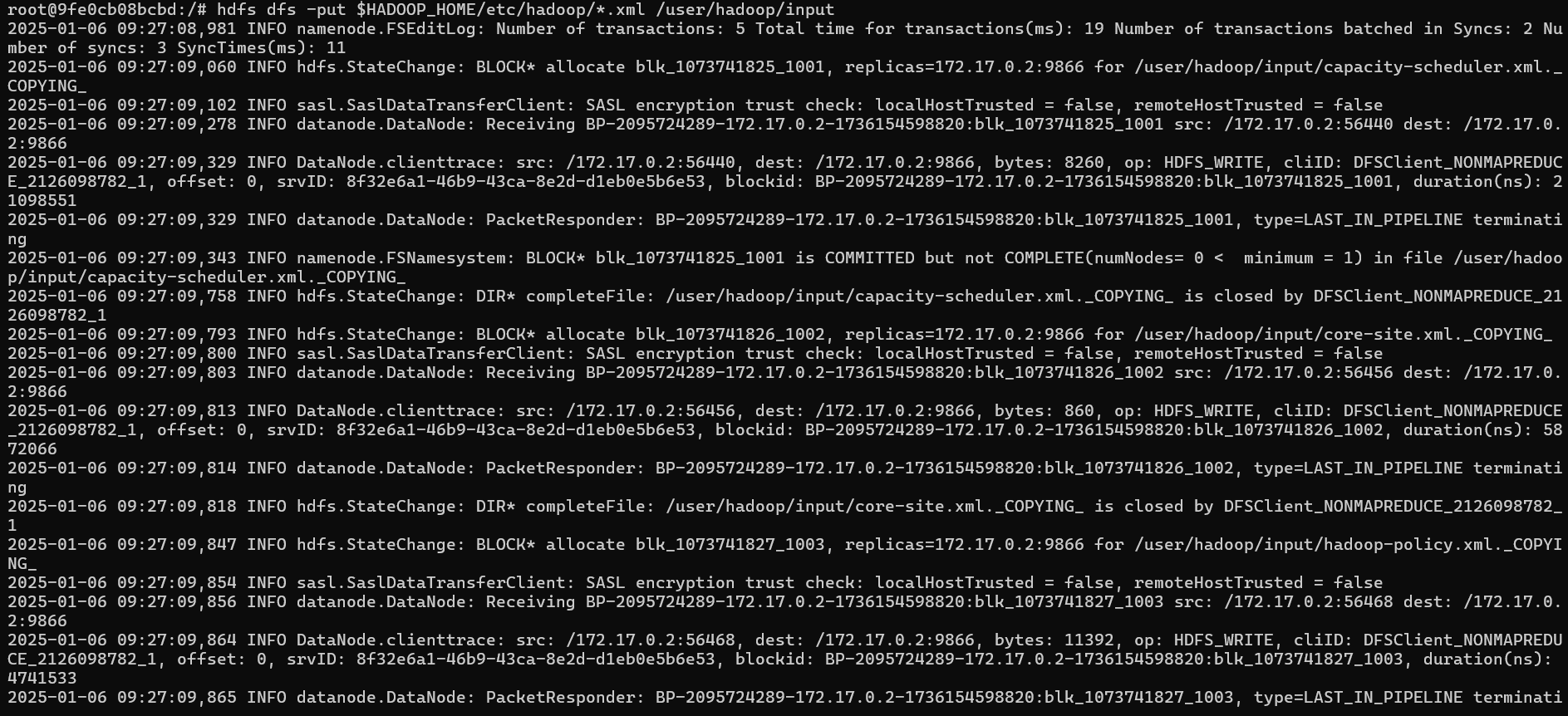
Then you can access to the Hadoop resource manager



1. hdfs dfs -mkdir -p /user/hadoop/input



hdfs dfs -put $HADOOP\_HOME/etc/hadoop/\*.xml /user/hadoop/input



1. hadoop jar $HADOOP\_HOME/share/hadoop/mapreduce/hadoop-mapreduce-examples-\*.jar wordcount /user/hadoop/input /user/hadoop/output

A screenshot of a computer

Description automatically generated

hdfs dfs -cat /user/hadoop/output/part-r-00000