Name : W.H.M.L.S.Bandara

Reg. No : 19APSE4298

Academic Period : 3rd Year 2nd Semester

Degree Program : Software Engineering

Due Date : 18/11/2024

**Lab Sheet 01**

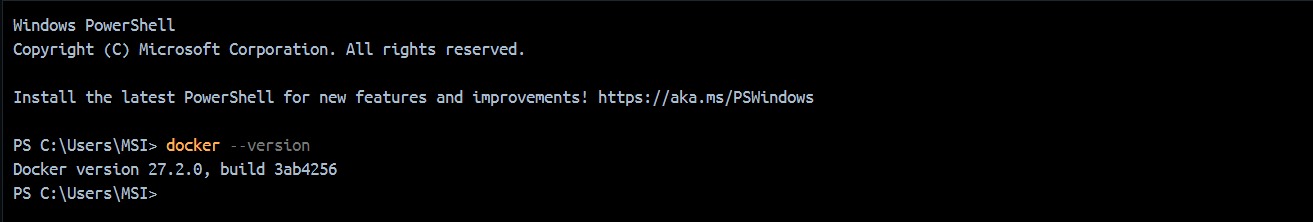
Sabaragamuwa University of Sri Lanka

Faculty of Computing

Computing & Information Systems

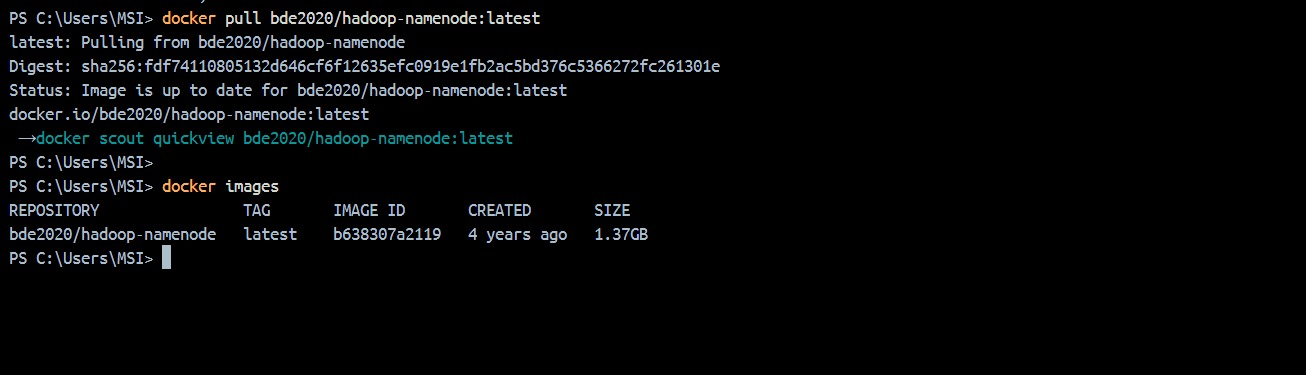
SE 6103- Parallel and Distributed Systems

1. **Check the docker version**



1. **Pull the Hadoop image**

* *docker pull bde2020/hadoop-namenode:latest*



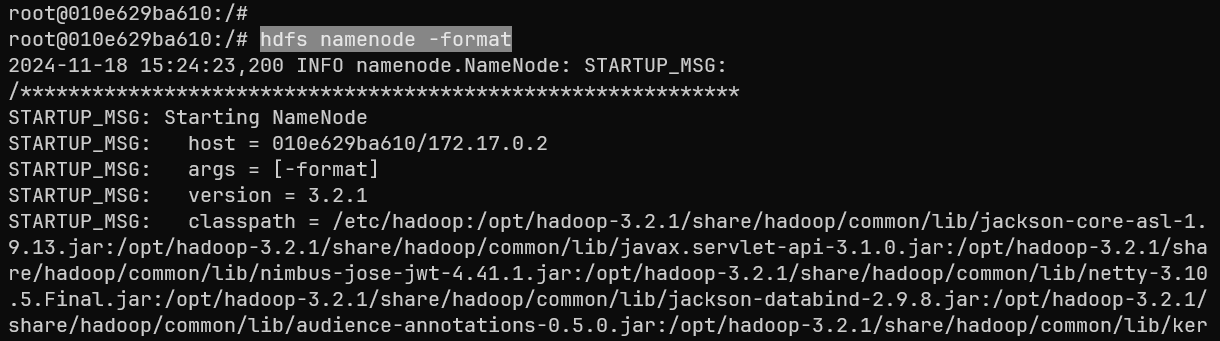
1. **Run the docker container**

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



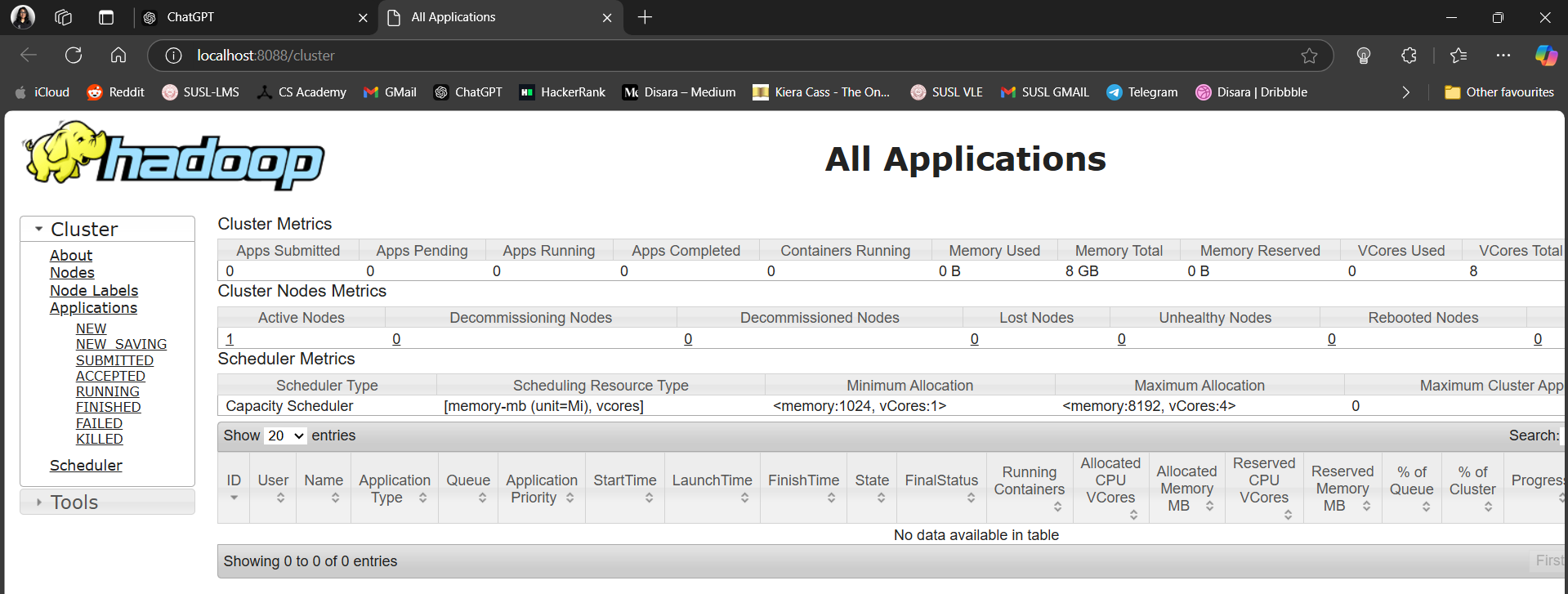
1. **Configure the Hadoop file system**

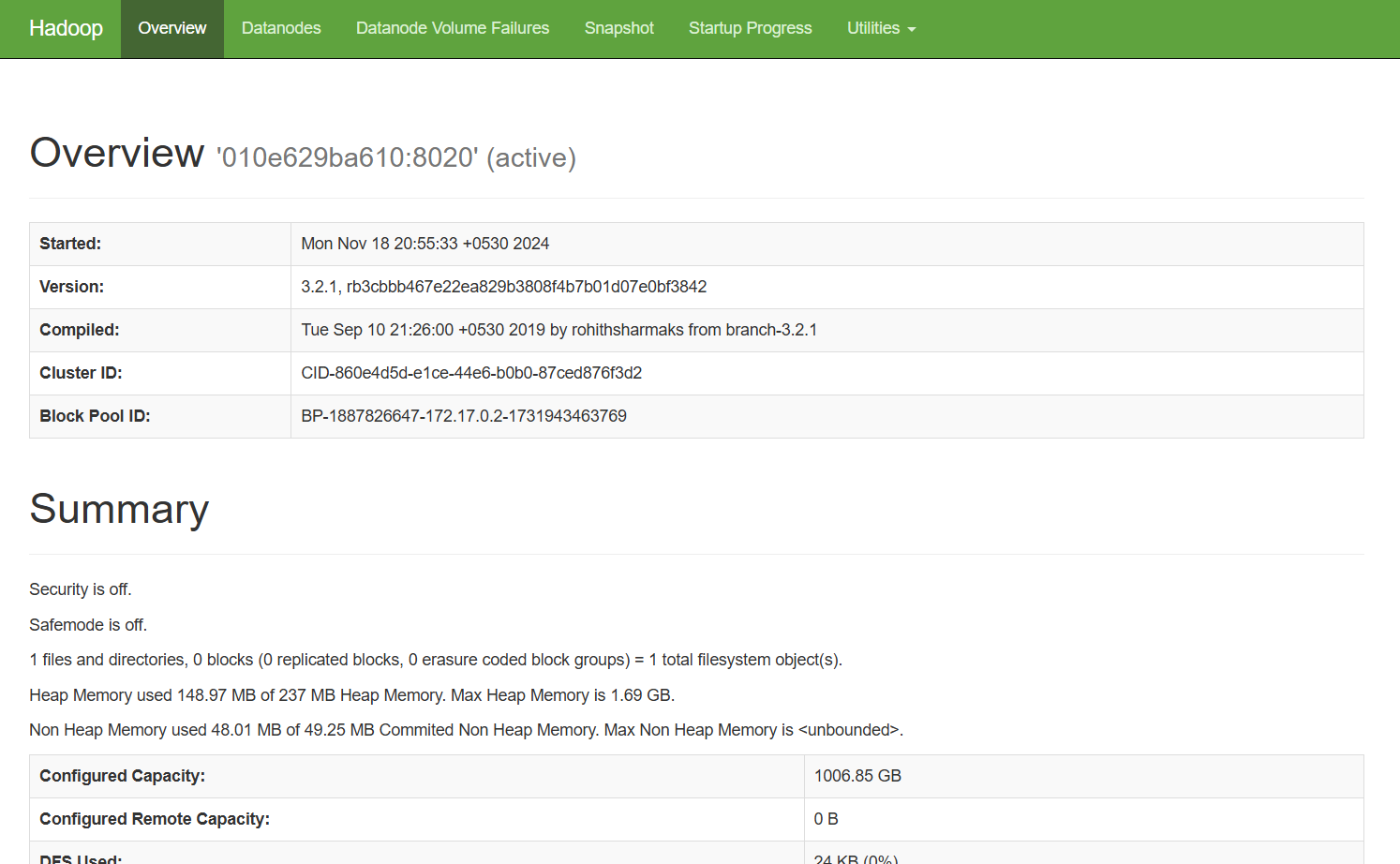
* *hdfs namenode -format*
* *hdfs namenode &*
* *hdfs datanode &*



1. **Access the Hadoop web interface and check the Hadoop health using the local**

**port**



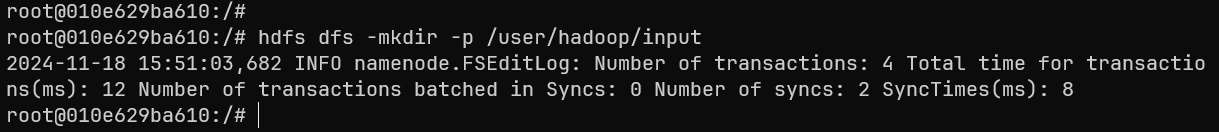


1. **Start the node manager and resource manager**

* *yarn nodemanager &*
* *yarn resourcemanager &*

1. **Upload sample data to HDFS**

* *hdfs dfs -mkdir -p /user/hadoop/input*



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

A screen shot of a computer

Description automatically generated

1. **Run the word counter job**

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

A screenshot of a computer program

Description automatically generated

1. **Check the output**

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

A computer screen with white text

Description automatically generated

1. **Stop the cluster**

* *docker stop hadoop-cluster*