**Lab Sheet 1**

**Sabaragamuwa University of Sri Lanka**

**Faculty of Computing**

**Software Engineering**

**SE6103 - Parallel and Distributed Systems**

Name : R.D.M. Perera

Reg No : 19APSE4287

Academic Period : 3rd Year 2nd Semester

Due Date : 18/11/2024

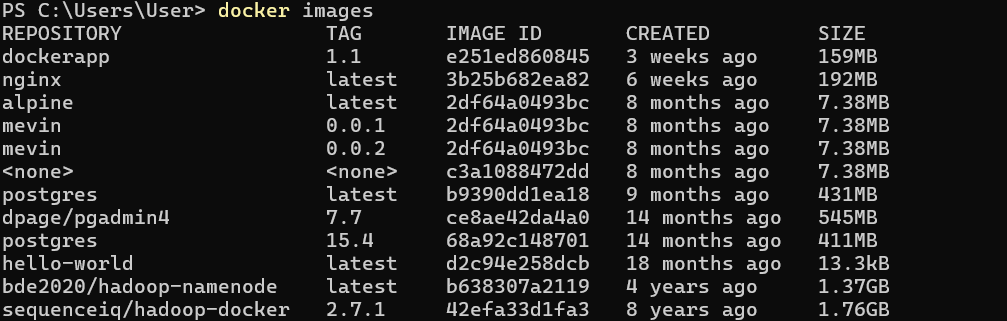
**Check Docker Version**

docker –version



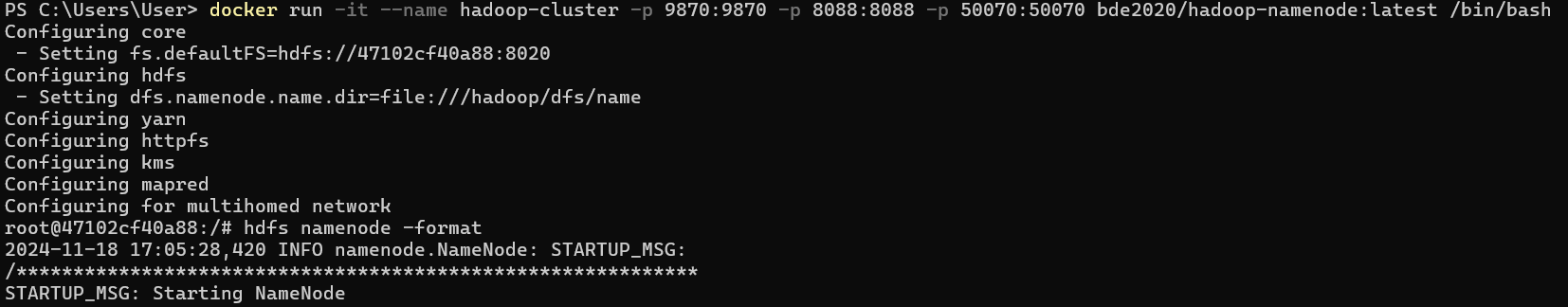
**Checking the Images which are already downloaded**

docker images



**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

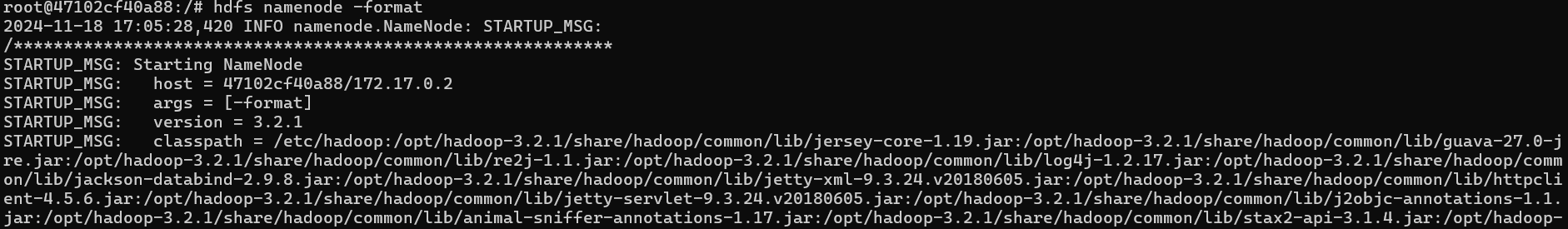


**Configure the Hadoop file system**

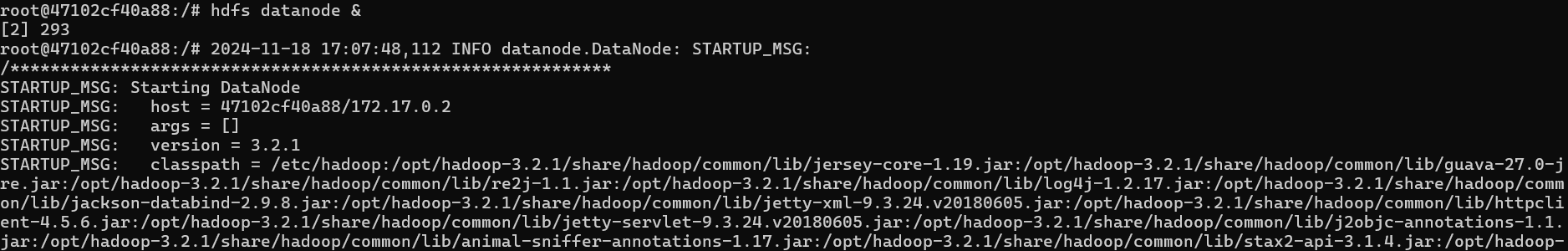
hdfs namenode -format

hdfs namenode &

hdfs datanode &

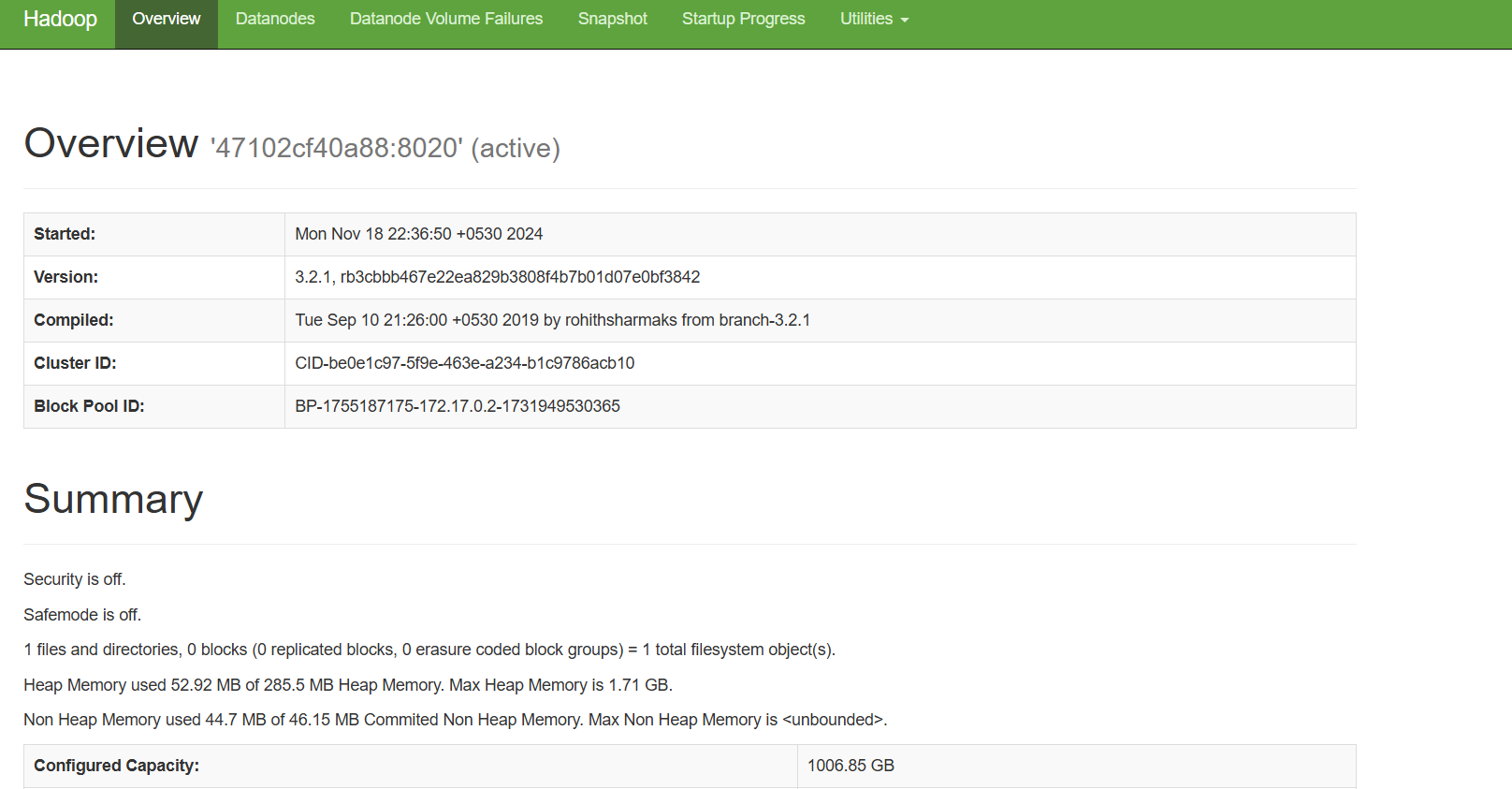






**Check the Hadoop Health using local port**

<http://localhost:9870/dfshealth.html#tab-overview>



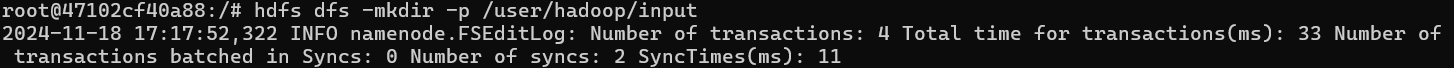
**Start the node manager and resource manager**

yarn nodemanager &

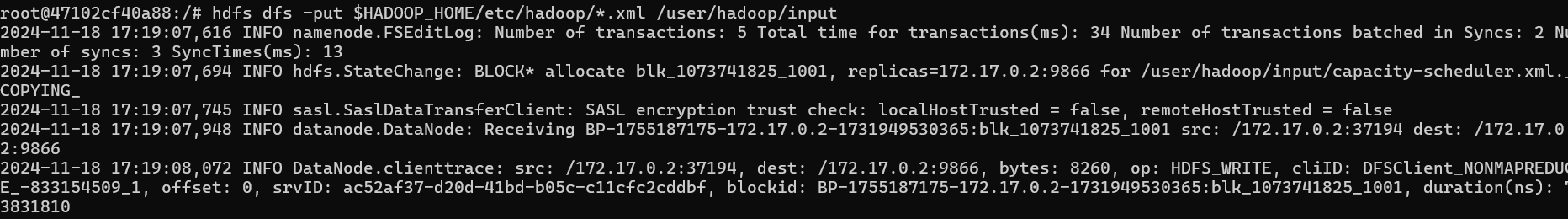
yarn resourcemanager &

**Upload sample data to hdfs**

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

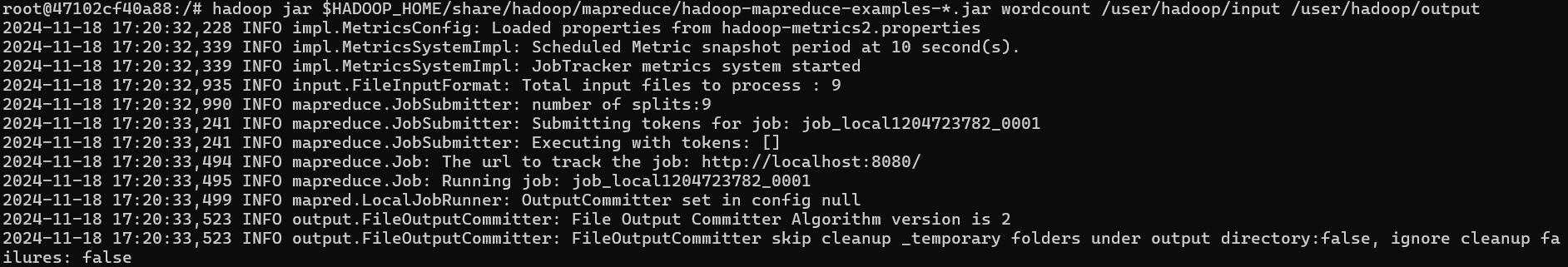


hdfs dfs -put $HADOOP\_HOME/etc/hadoop/\*.xml /user/hadoop/input



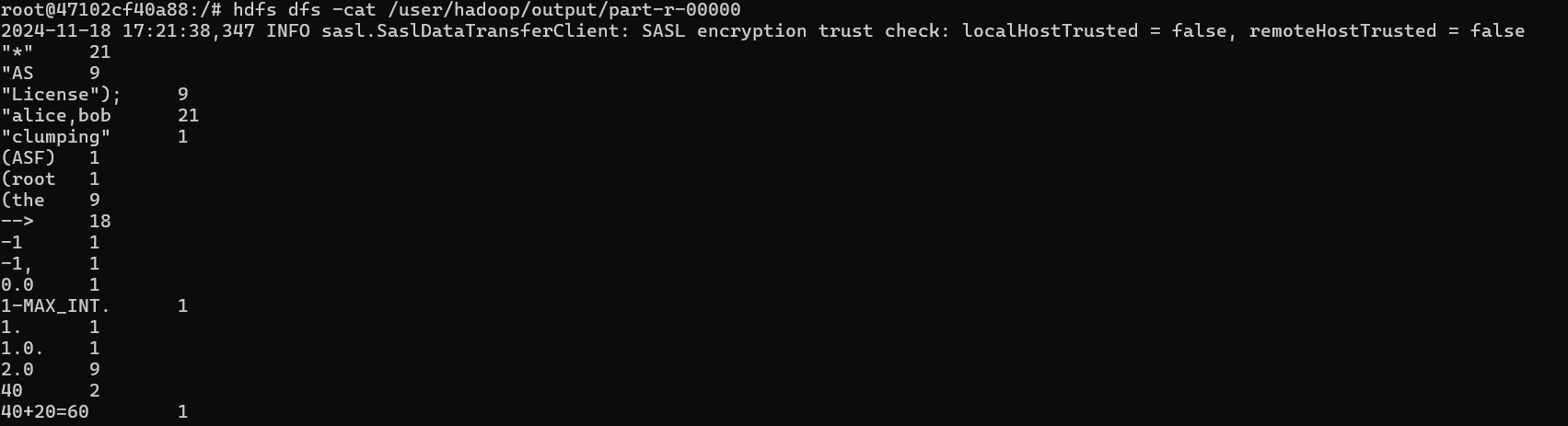
**Run the word counter job**

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



**Check the output**

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



**Stop the Cluster**

exit

docker stop hadoop-cluster

