Big Data Analytics

**Lab Practical and date** – Practical 3, Monday 10th August 2020

**Name and Roll Number**- Het Shah, 17BIT103

**Practical Objective**- Setup single node Hadoop cluster and apply HDFS commands on single node Hadoop Cluster.

**Steps Involved-**

We installed Hadoop 3.2.1 single node and learned the basic HDFC commands

**Background**

**Hadoop**

Apache Hadoop is a collection of open-source software utilities that facilitate using a network of many computers to solve problems involving massive amounts of data and computation. It provides a software framework for distributed storage and processing of big data using the MapReduce programming model

**HDFS**

The Hadoop Distributed File System (HDFS) is a distributed file system designed to run on commodity hardware. It has many similarities with existing distributed file systems. However, the differences from other distributed file systems are significant. HDFS is highly fault-tolerant and is designed to be deployed on low-cost hardware. HDFS provides high throughput access to application data and is suitable for applications that have large data sets. HDFS relaxes a few POSIX requirements to enable streaming access to file system data. HDFS was originally built as infrastructure for the Apache Nutch web search engine project. HDFS is now an Apache Hadoop subproject.

**Hadoop Installation**

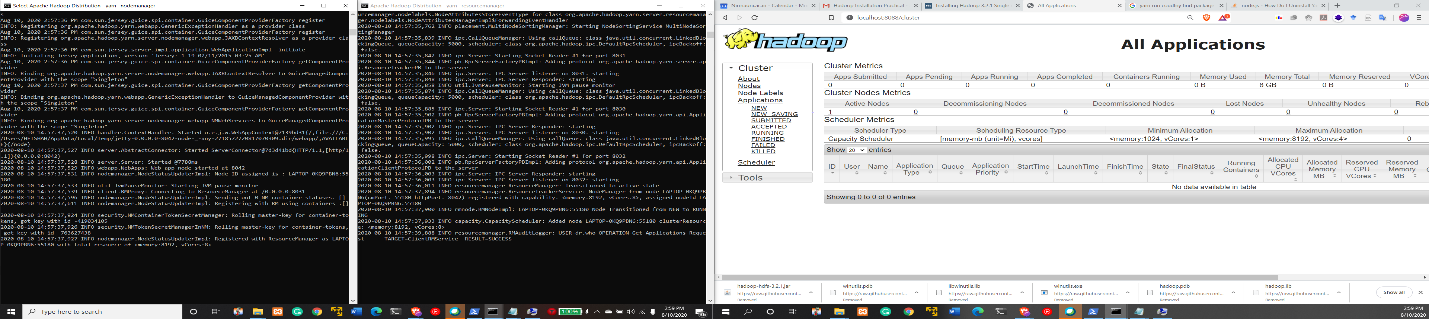
1. **Hadoop -version**, shows the Hadoop version installed in the system



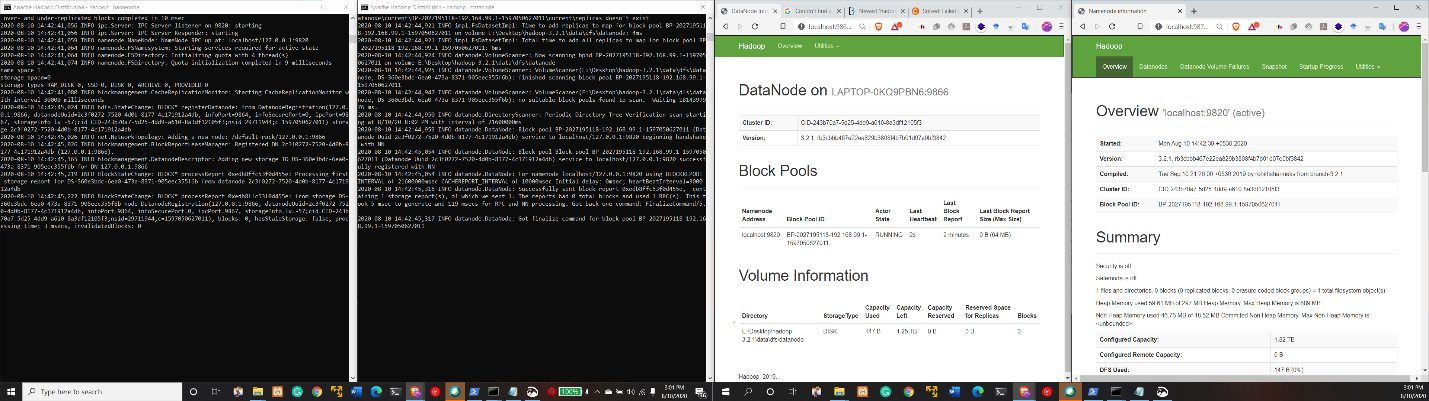
1. **Jps-** lists the instrumented Java HotSpot VMs on the target system



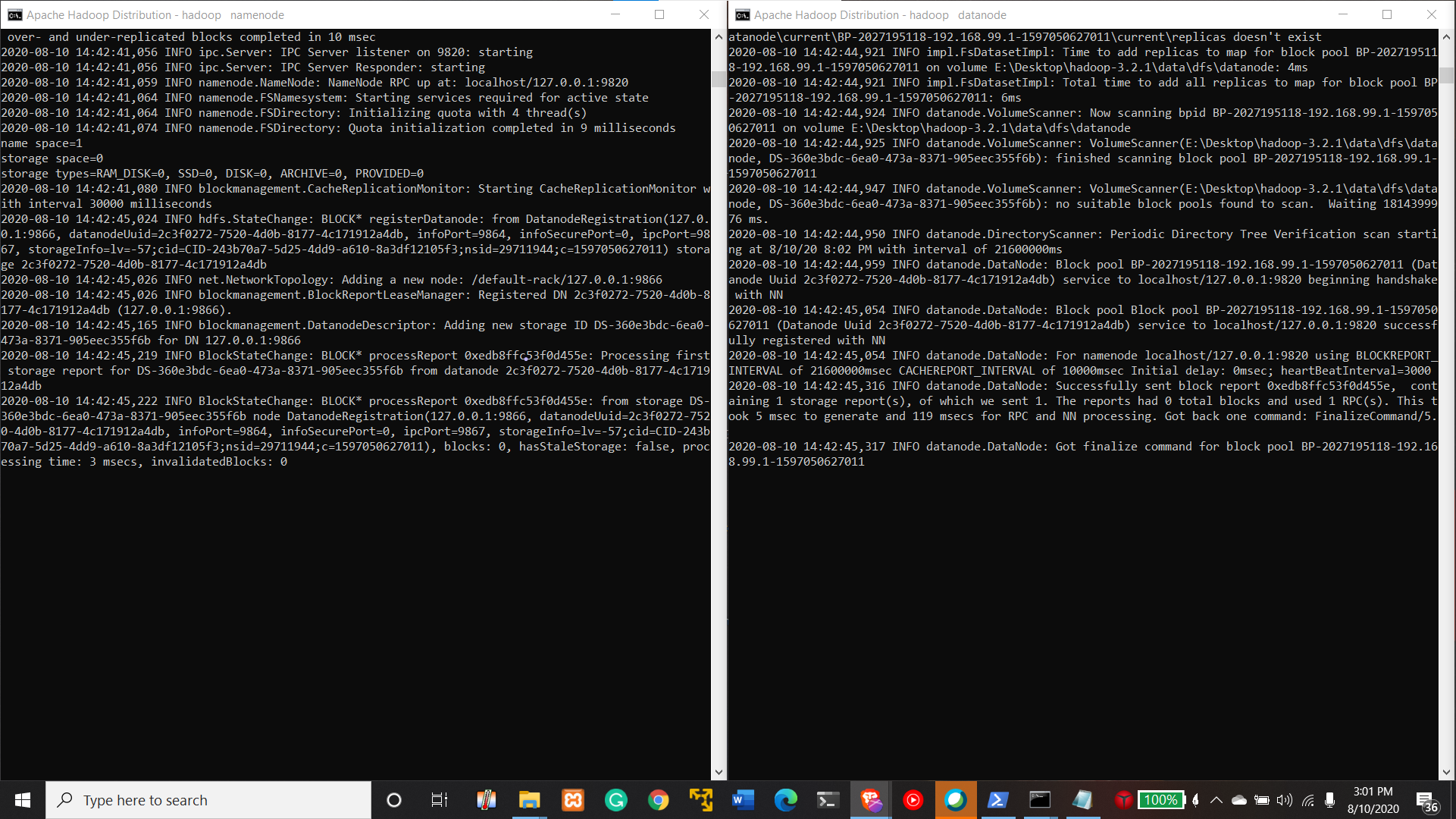
1. Localhost:8080/cluster



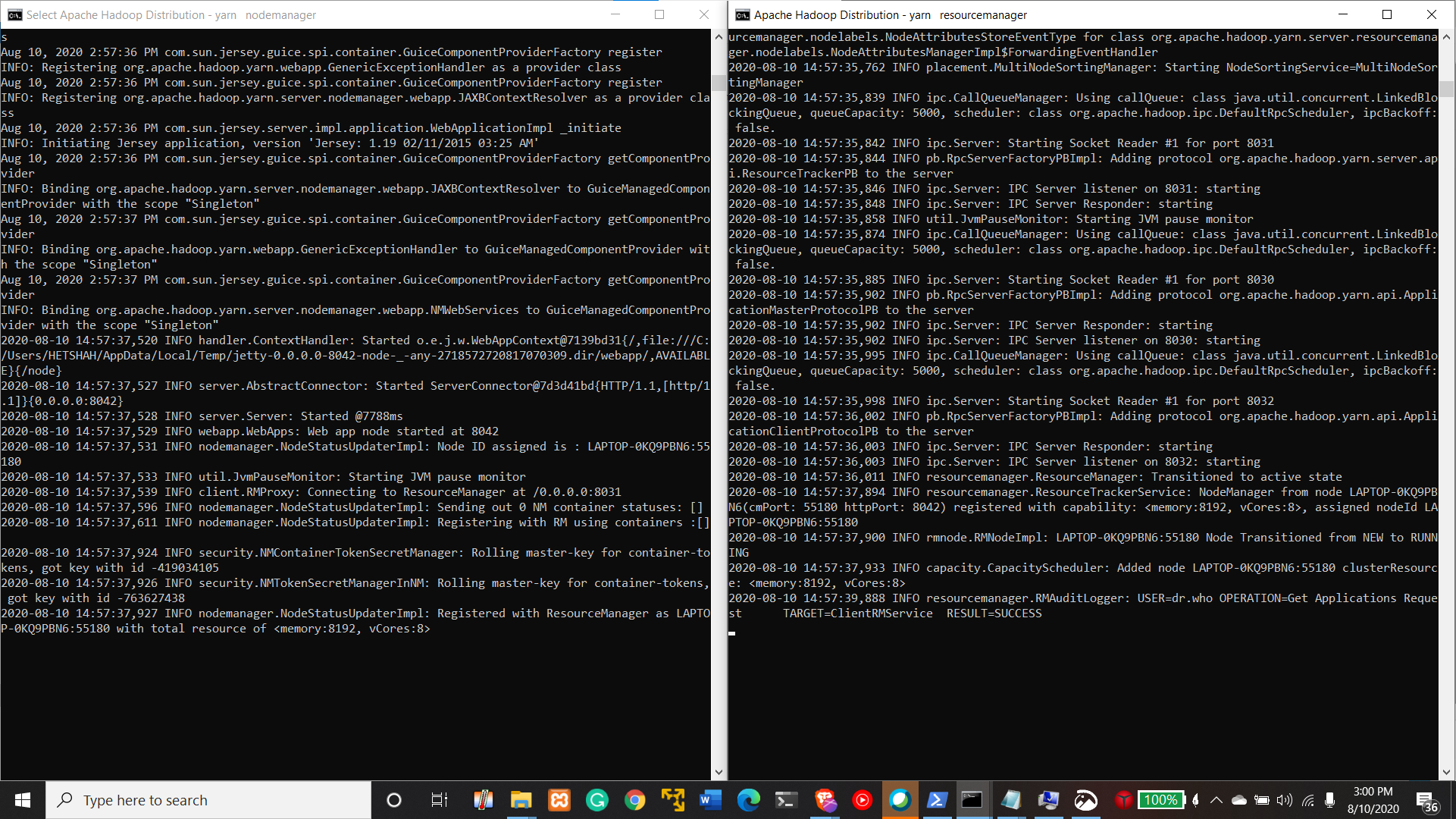
1. Localhost 9820 and Localhost 9860 showing the data done and the system overview



1. Hadoop **NameNode** and the **DataNode** running side by side

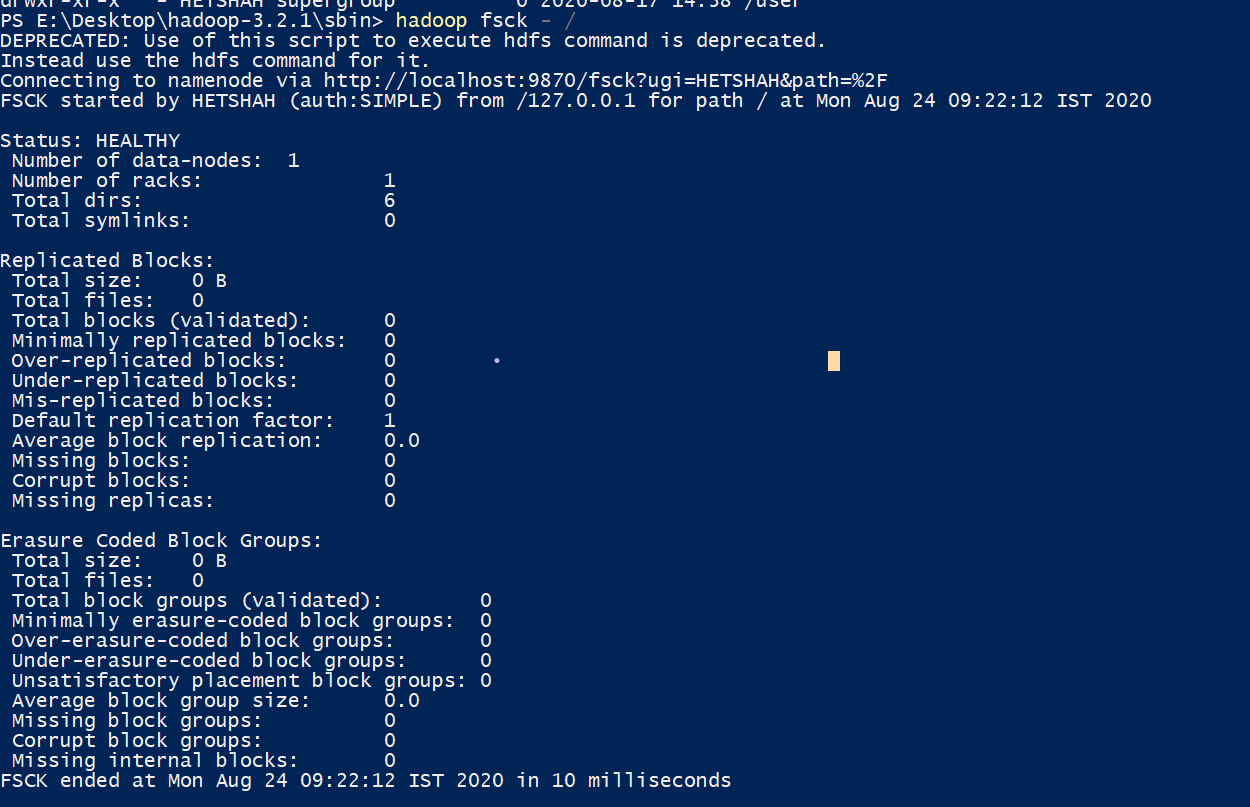


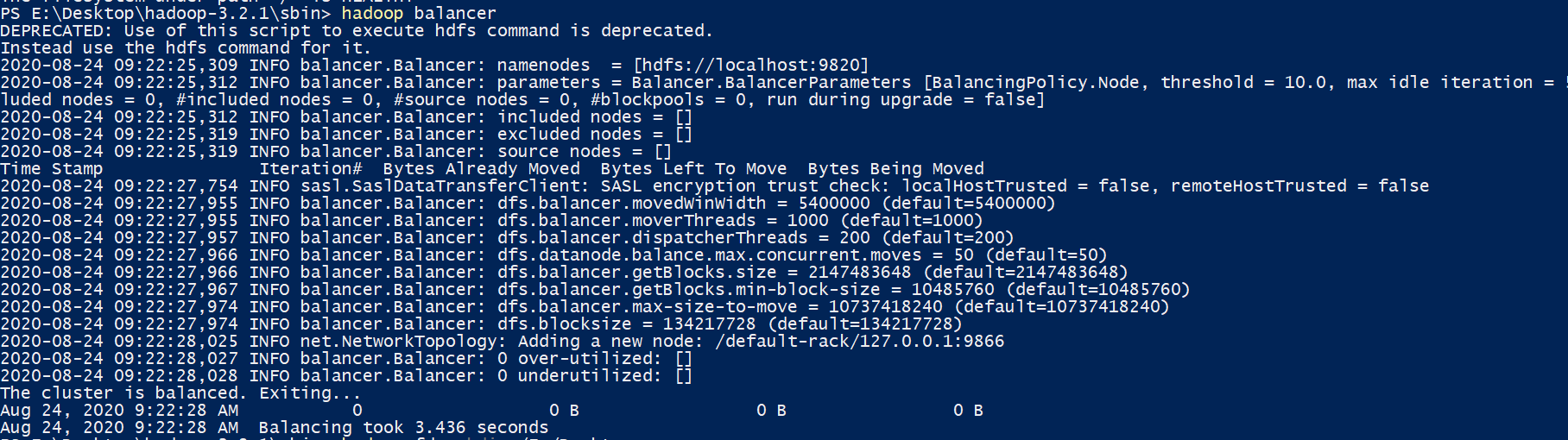
1. Hadoop **Yarn Manager** is open



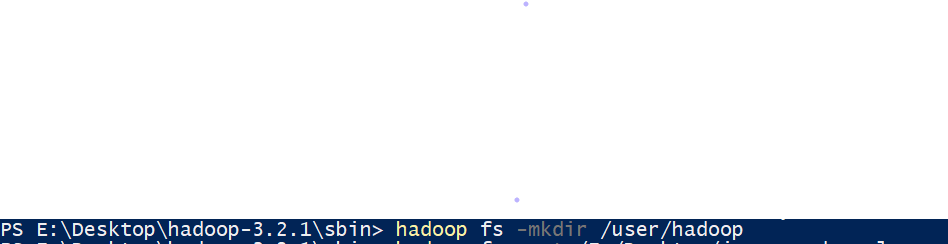
**HDFS commands**

**Hadoop fsck -/**

****

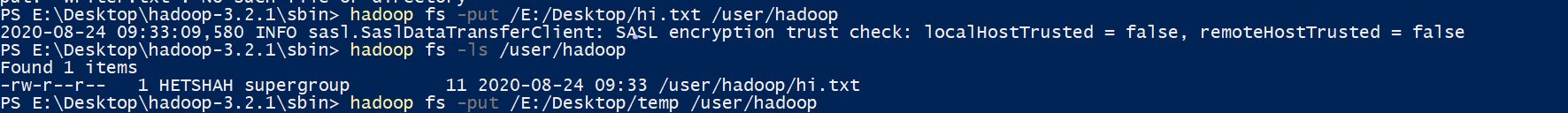
**Hadoop Balancer**

**Hadoop fs -mkdir /user/hadoop**

****

**Hadoop fs -put /E:/Desktop/hi.txt /user/Hadoop**

**Hadoop fs -ls /user/Hadoop**

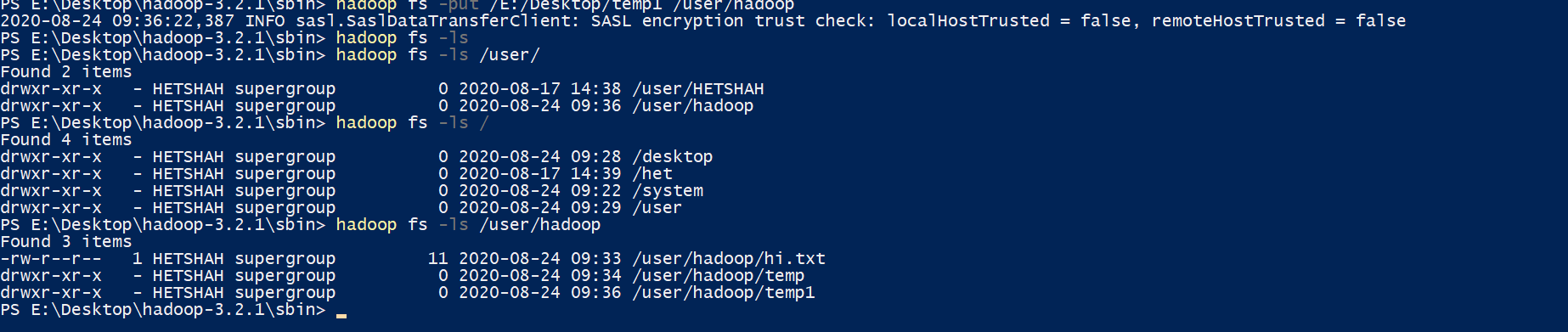
****

**Hadoop fs -put /E:/Desktop/temp /user/Hadoop**

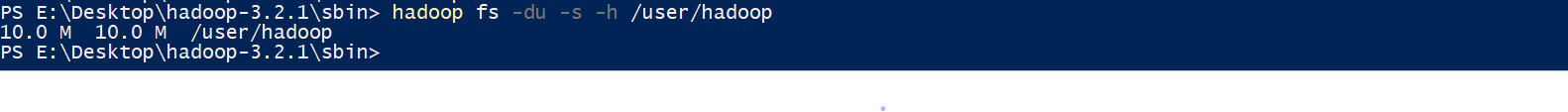
**Hadoop fs -ls**

**Hadoop fs -ls /user/**

**Hadoop fs -ls /user/hadoop**

****

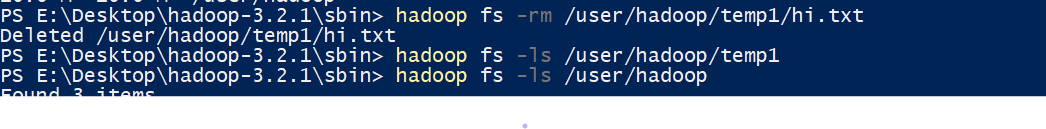
**Hadoop fs -du -s h /user/hadoop**

****

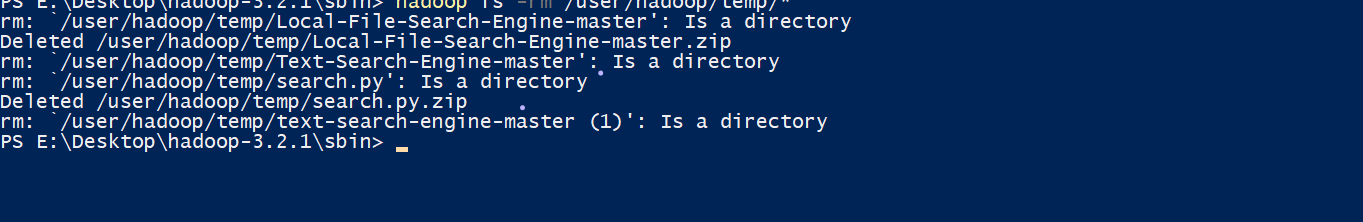
**Hadoop fs -rm /user/Hadoop/temp1/hi.txt**

**Hadoop fs -ls /user/Hadoop/temp1**

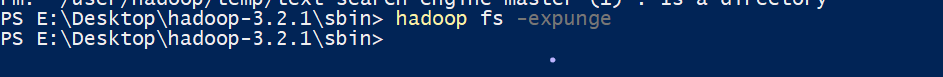
**Hadoop -ls user/Hadoop**

****

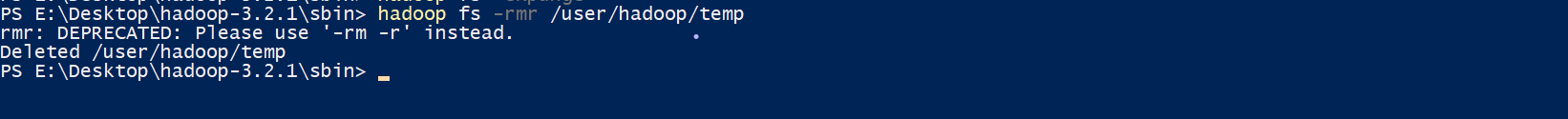
**Hadoop fs -rm /Hadoop/temp**

****

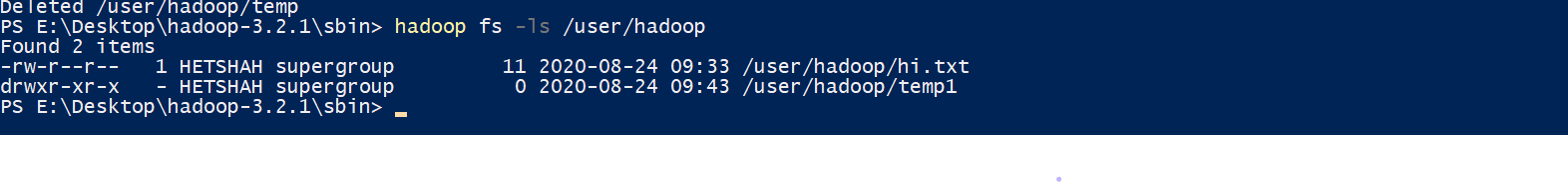
**Hadoop fs -expunge**

****

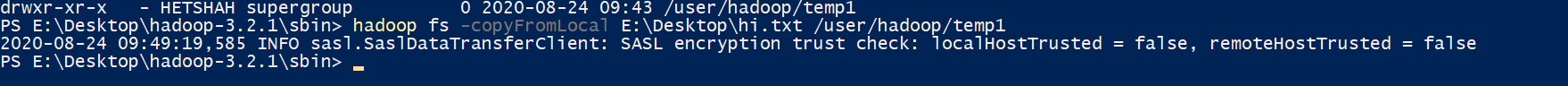
**Hadoop fs -rmr /user/Hadoop/temp**

****

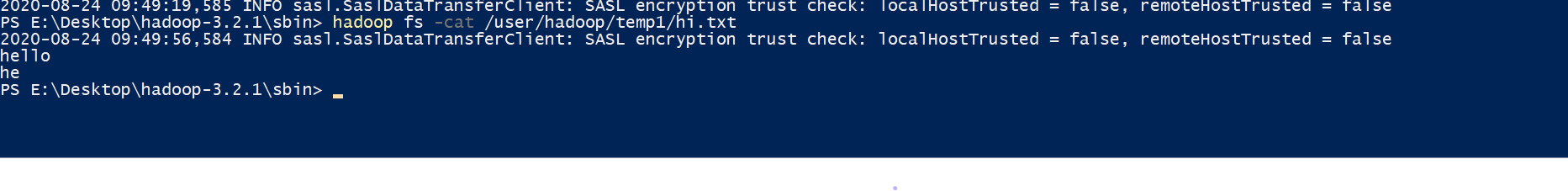
**Hadoop fs -ls /user/hadoop**

****

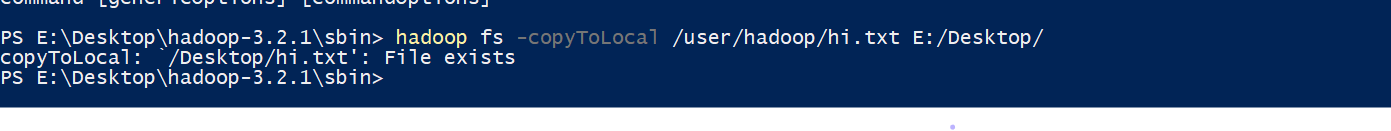
**Hadoop fs -copyfromLocal E:\Desktop\hi.txt /user/Hadoop/temp1**

****

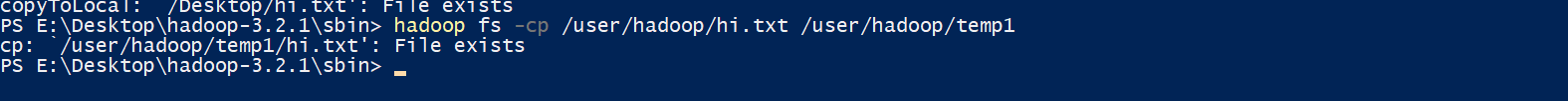
**Hadoop fs cat /user/Hadoop/temp1/hi.txt**

****

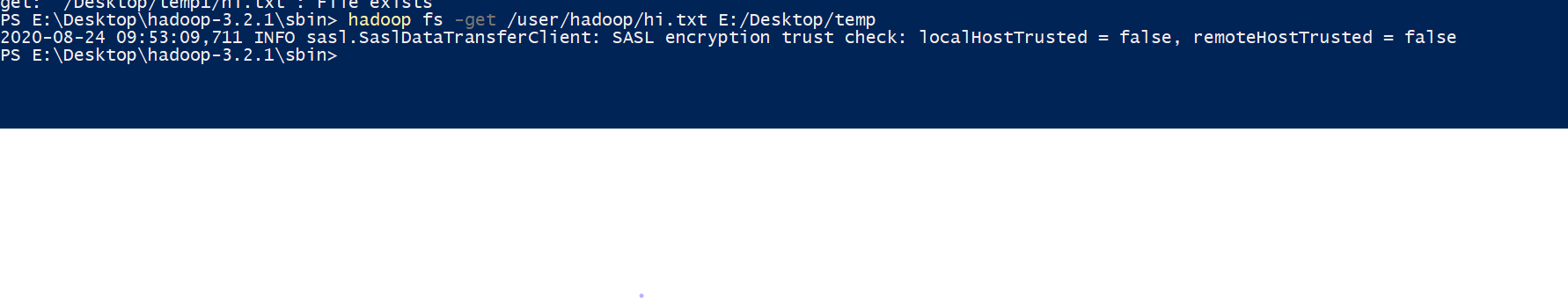
**Hadoop fs -copytoLocal /user/Hadoop/hi.txt E:/Desktop/hi.txt**

****

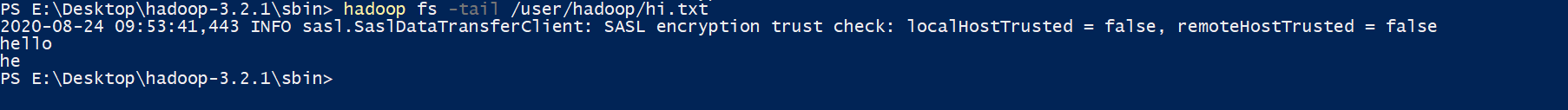
**Hadoop fs -cp /user/Hadoop/hi.txt /user/Hadoop/temp1**

****

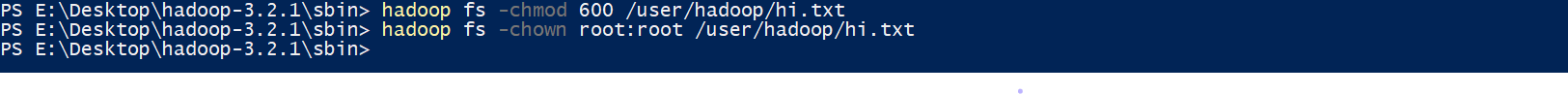
**Hadoop fs -get /user/Hadoop/hi.txt /user/Hadoop/temp**

****

**Hadoop fs -tail /user/Hadoop/hi.txt**

****

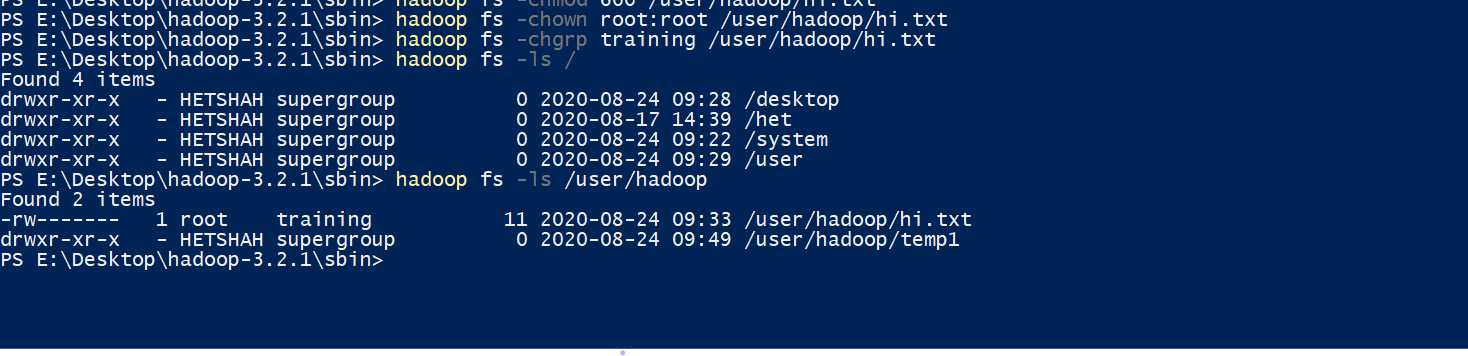
**Hadoop fs -chmod 600 /user/Hadoop/hi.txt**

**Hadoop fs -chown root:root /user/Hadoop/hi.txt**

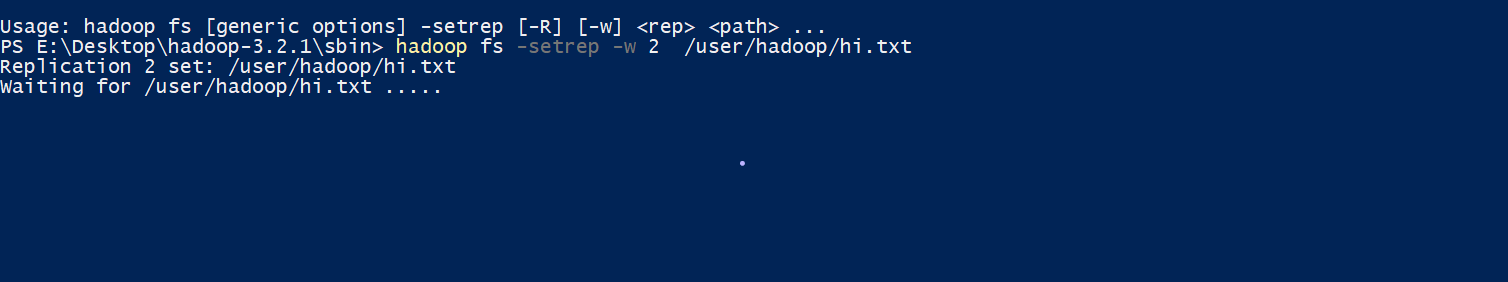
**Hadoop fs -chgrp training /user/Hadoop/hi.txt**

**Hadoop fs -ls/**

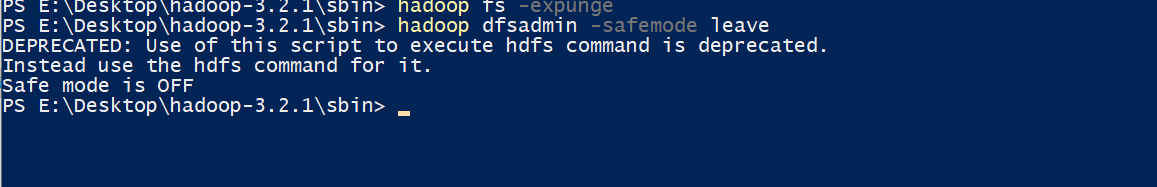
**Hadoop fs -ls /user/hadoop**

****

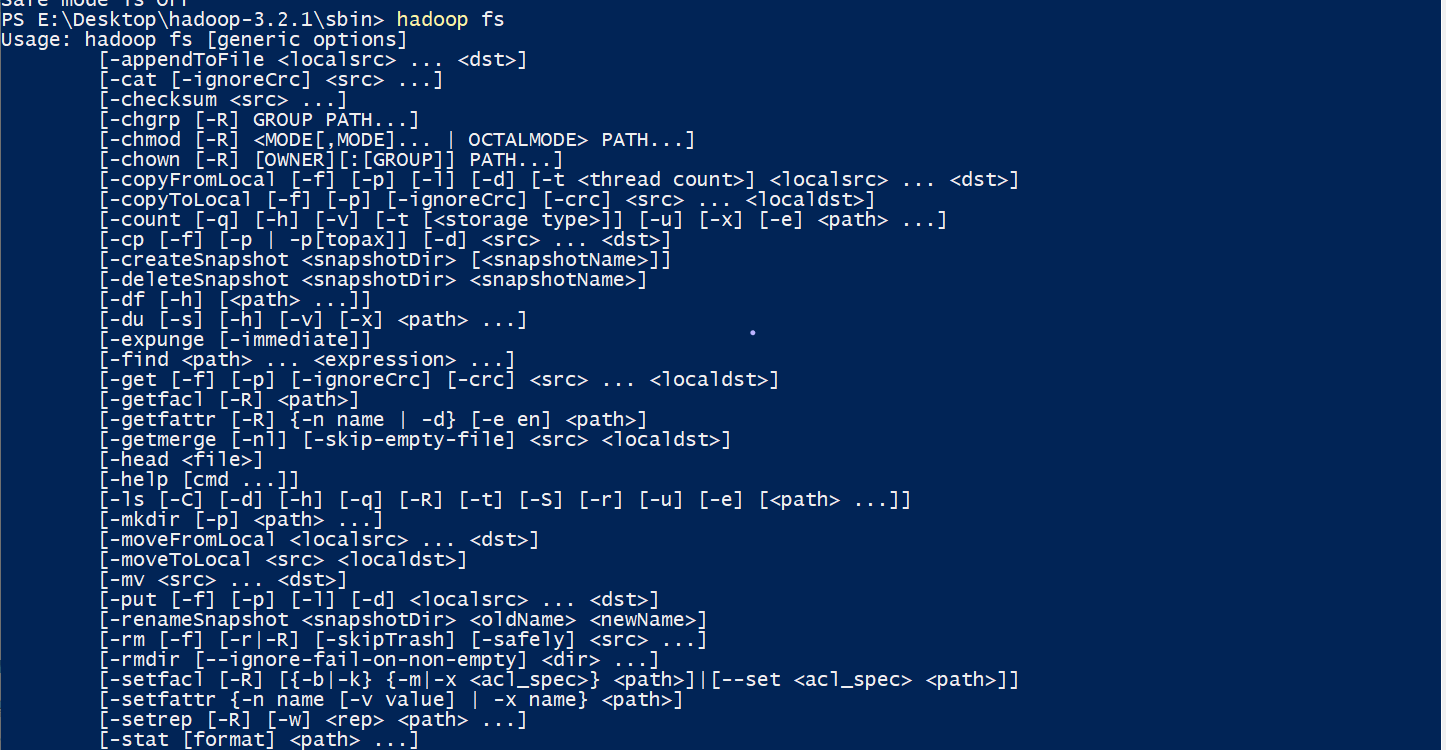
**Hadoop fs -setrep -w 2 /user/Hadoop/hi.txt**

****

**Hadoop dfsadmin -safemode leave**

****

**Hadoop fs**

****

**Conclusion**

In this practical we learned how to install Hadoop in our systems and we also used the HDFS commands for file manipulation to and fro from the system to the HDFS and vice versa.