Practical 3

Big Data Analytics

2CS702

Mistry Unnat

20BCE515

Department of Computer Science and Engineering

Institute of Technology

Nirma University

Ahmedabad

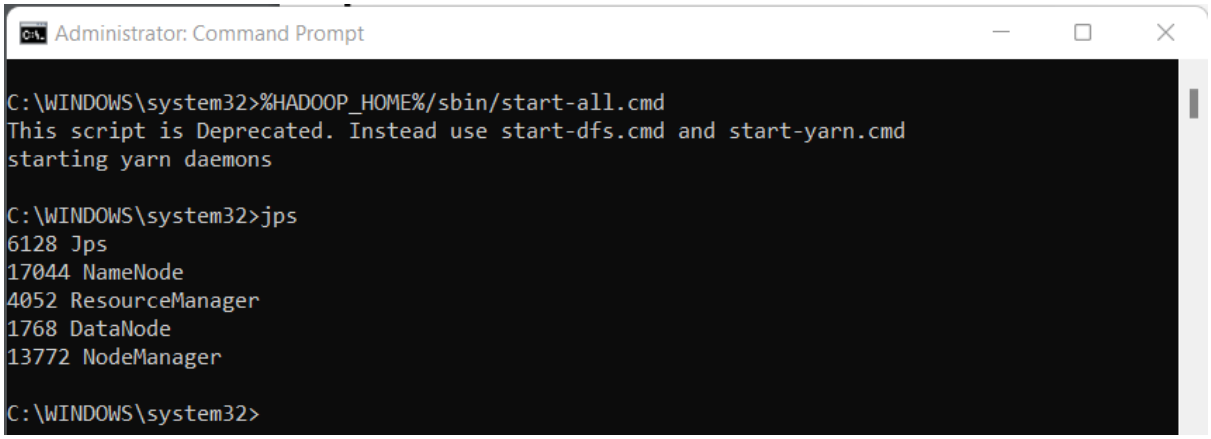
AIM : Setup Single Node Hadoop Cluster and Apply HDFS Commands on single node Hadoop Cluster

Commands :

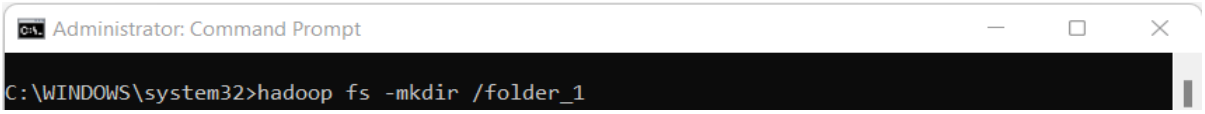
Starting the daemons... and checking if they are up and working

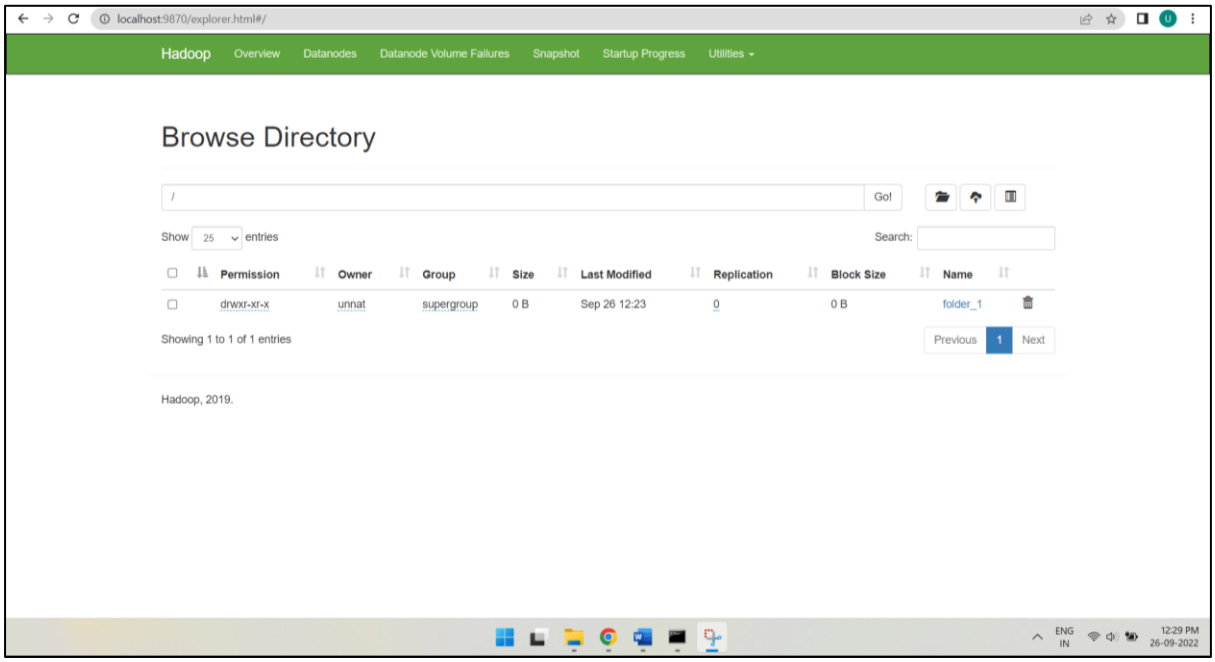
Commands used :

1. %HADOOP_HOME%/sbin/start-all.cmd
2. jps

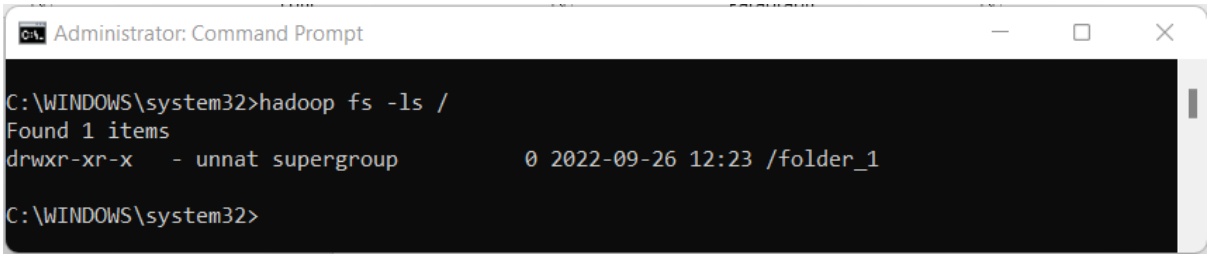


1. **mkdir :** creates directory





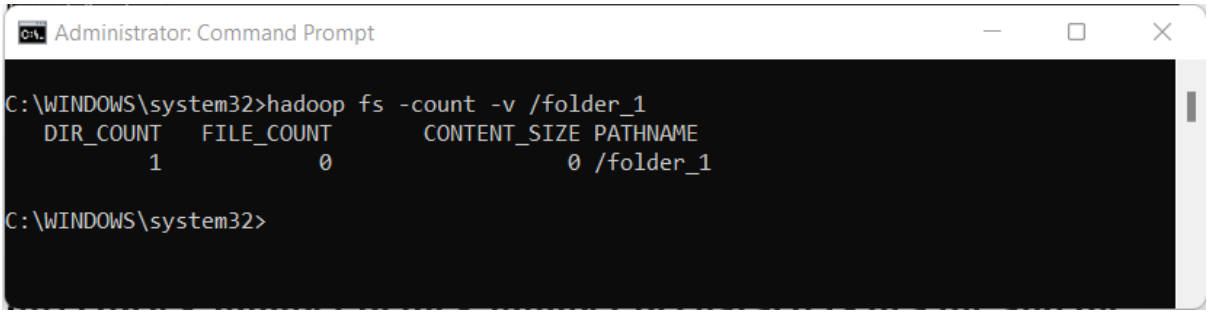
2. **ls :** lists files and folders



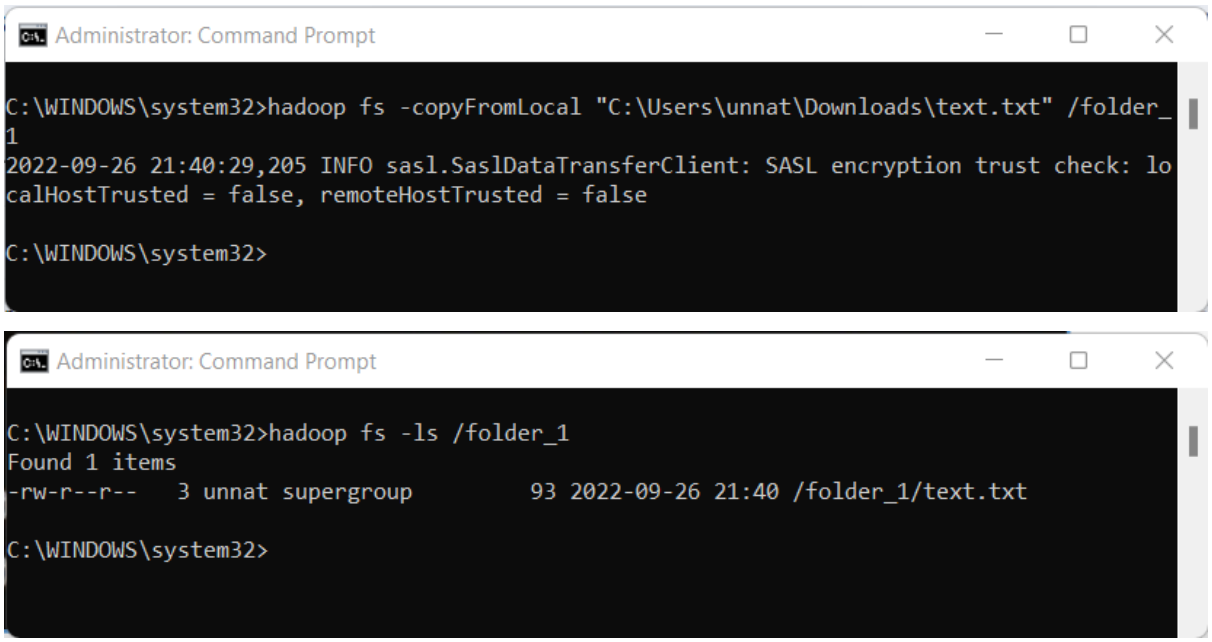
3. **chmod :** To change file and directory permissions, use the command chmod (change mode).

Permission	Files	Directories
r	can **read** the file	can **ls** the directory
w	can **write** the file	can **modify** the directory's contents
x	can **execute** the file	can **cd** to the directory

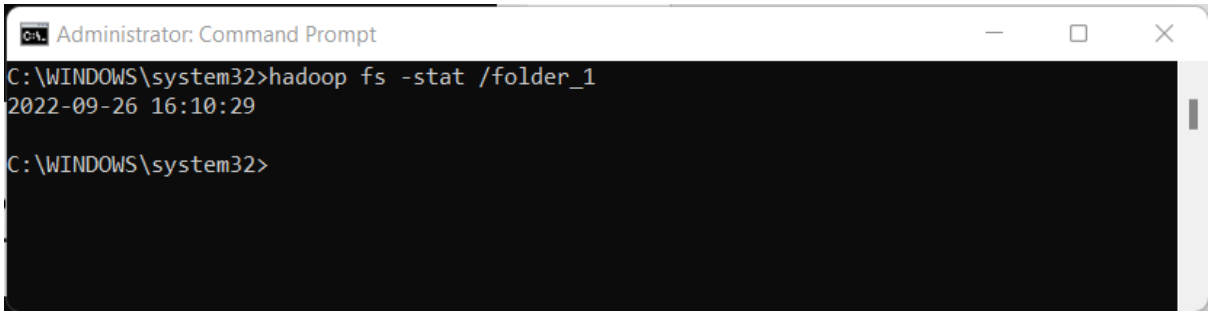
4. **count :** Hadoop HDFS count option is used to count a number of directories, number of files, number of characters in a file and file size.



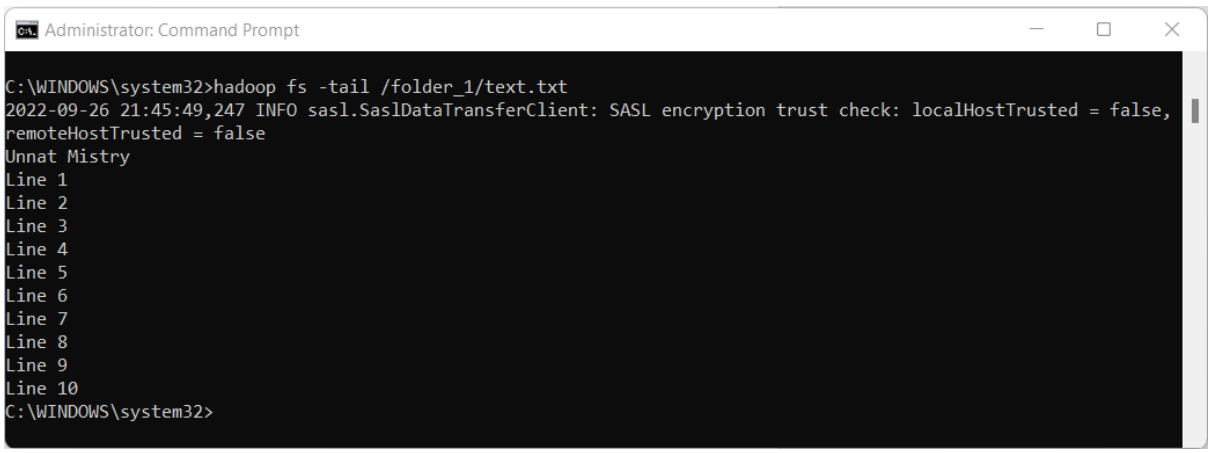
5. **copyFromLocal :** Hadoop copyFromLocal command is used to copy the file from your local file system to the HDFS(Hadoop Distributed File System).



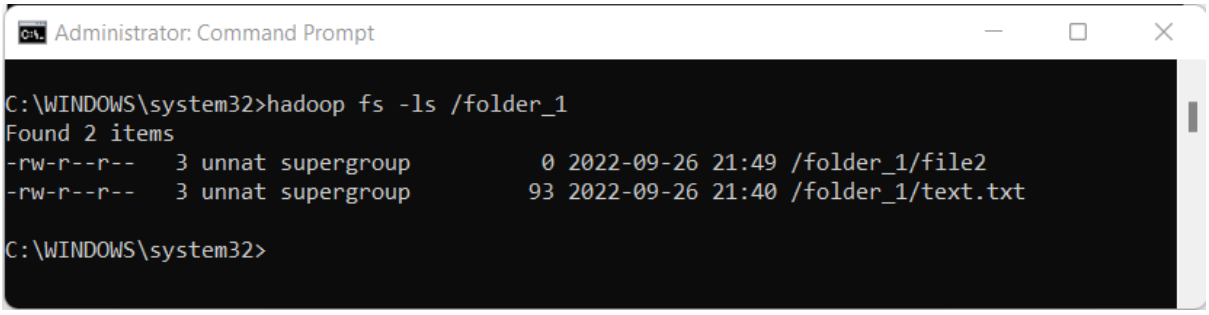
6. **stat :** Print statistics about the file/directory at <path> in the specified format.



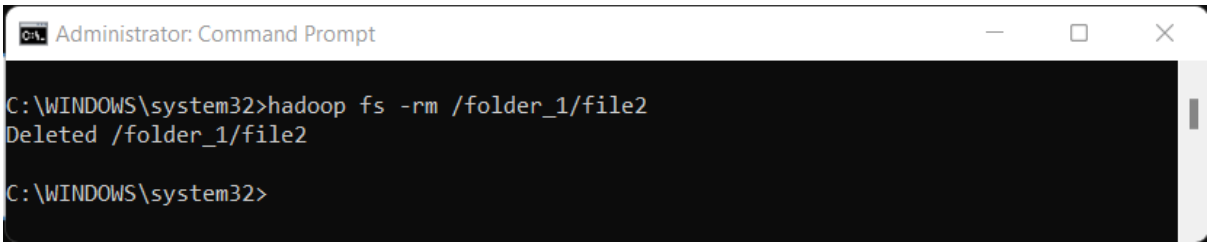
7. **tail :** The Hadoop fs shell tail command shows the last 1KB of a file on console or stdout.



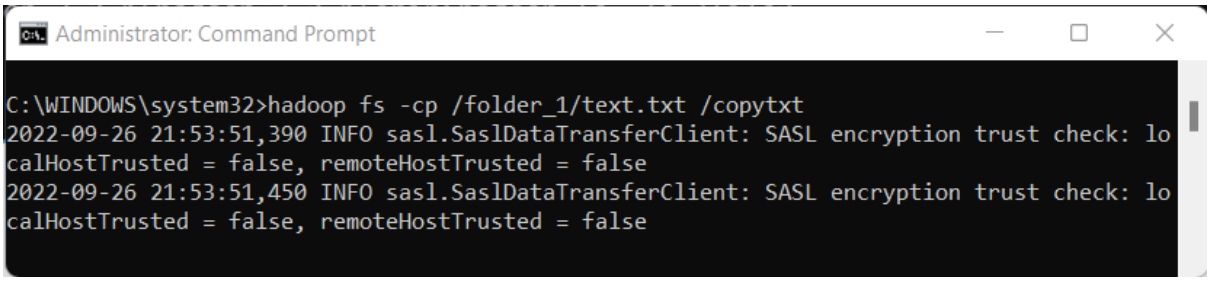
8. **touchz :** touchz command creates a file in HDFS with file size equals to 0 byte. The directory is the name of the directory where we will create the file, and filename is the name of the new file we are going to create.

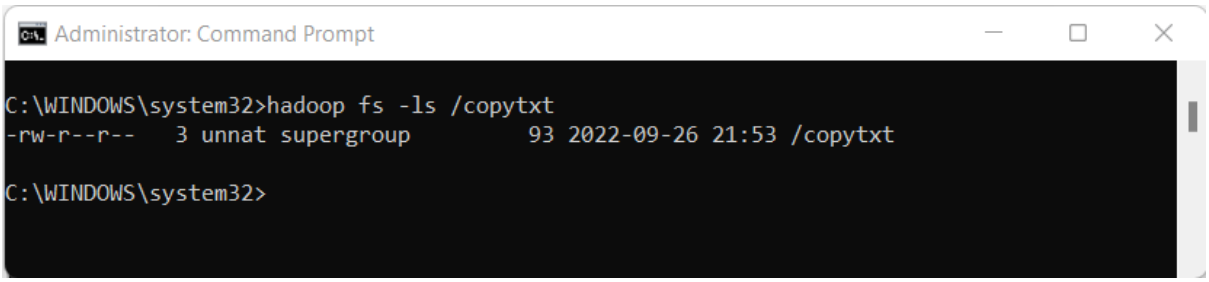


9. **rm :** This command is similar to the Linux rm command, and it is used for removing a file from the HDFS file system.

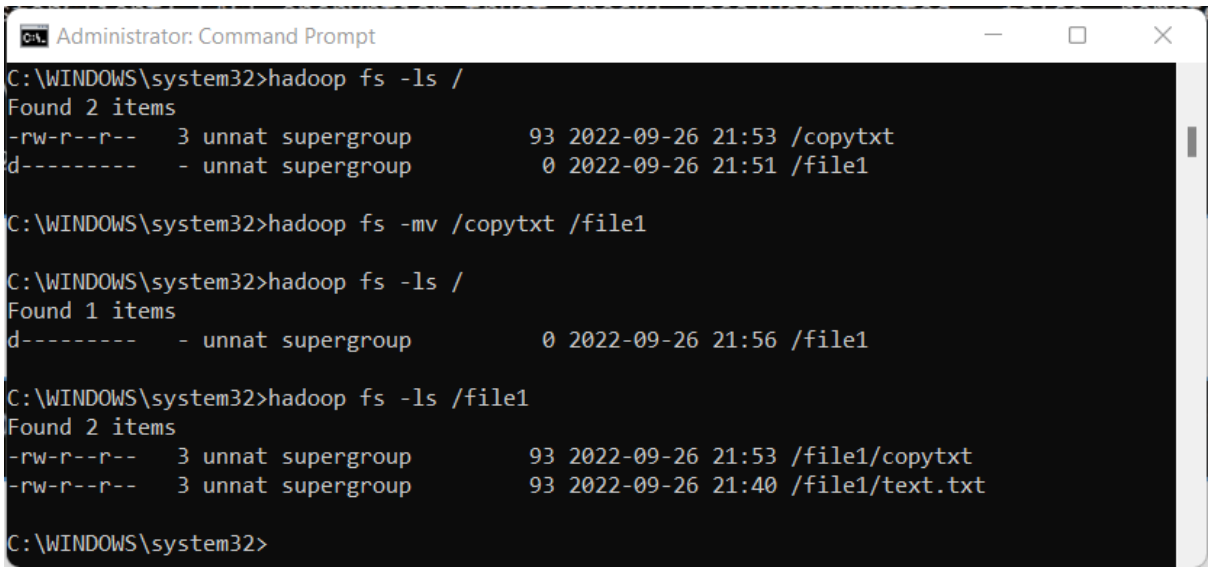


10. **cp :** The cp command copies a file from one directory to another directory within the HDFS.





11. mv : moves files from source to destination.



Conclusion :

In this practical, we setup a single node Hadoop cluster and applied HDFS commands on single node Hadoop cluster.