

**Aim :** File Manipulation in HDFS.

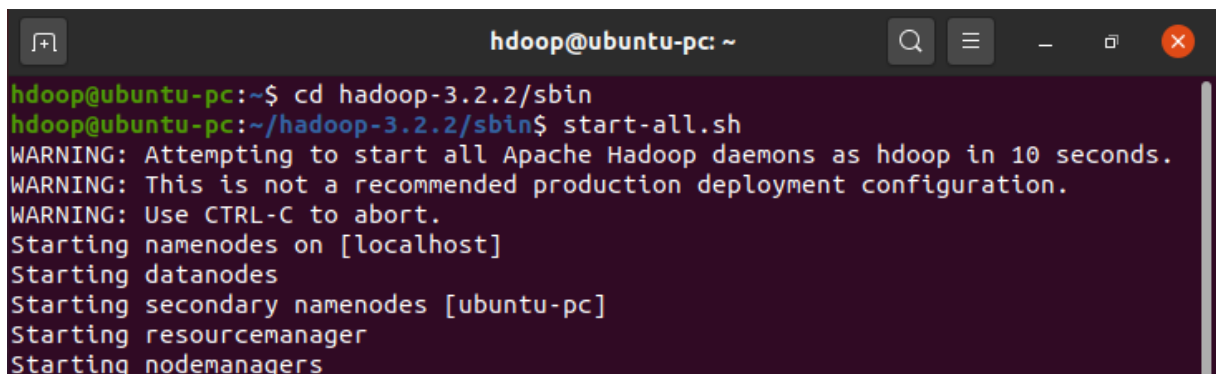
**Objective:** To carry out file manipulation in hdfs.

**Tools Used: Hadoop**

**Actual Work Done:**

To use HDFS commands, start the Hadoop services using the following command:

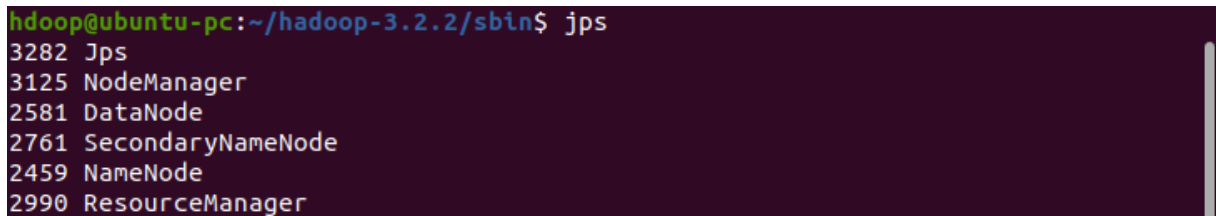
sbin/start-all.sh

A terminal window titled 'hadoop@ubuntu-pc: ~' showing the execution of 'start-all.sh' in the directory '~/.hadoop-3.2.2/sbin'. The output includes several warnings about starting all daemons and a non-recommended configuration, followed by the successful starting of namenodes, datanodes, secondary namenodes, resource manager, and node managers.

```
hadoop@ubuntu-pc:~$ cd hadoop-3.2.2/sbin
hadoop@ubuntu-pc:~/hadoop-3.2.2/sbin$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [ubuntu-pc]
Starting resourcemanager
Starting nodemanagers
```

To check if Hadoop is up and running:

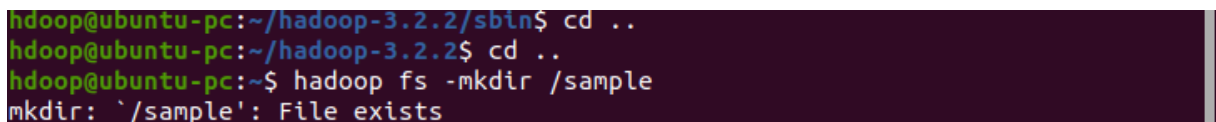
Jps

A terminal window showing the output of the 'jps' command. It lists the Java processes running: Jps, NodeManager, DataNode, SecondaryNameNode, NameNode, and ResourceManager with their respective PIDs.

```
hadoop@ubuntu-pc:~/hadoop-3.2.2/sbin$ jps
3282 Jps
3125 NodeManager
2581 DataNode
2761 SecondaryNameNode
2459 NameNode
2990 ResourceManager
```

mkdir:

To create a directory, similar to Unix ls command.

A terminal window showing the execution of 'mkdir /sample' using the 'hadoop fs' command. The output indicates that the directory already exists.

```
hadoop@ubuntu-pc:~/hadoop-3.2.2/sbin$ cd ..
hadoop@ubuntu-pc:~/hadoop-3.2.2$ cd ..
hadoop@ubuntu-pc:~$ hadoop fs -mkdir /sample
mkdir: `/sample': File exists
```

cp:

Copy files from one directory to another within HDFS, similar to Unix cp command.

```
hadoop@ubuntu-pc:~$ hadoop fs -cp /sample/text.txt /sample1
hadoop@ubuntu-pc:~$ hadoop fs -rm -r /sample1
Deleted /sample1
```

**rm:**

Remove a file from HDFS, similar to Unix rm command. This command does not delete directories. For recursive delete, use command -rm -r.

```
hadoop@ubuntu-pc:~$ hadoop fs -cp /sample/text.txt /sample1
hadoop@ubuntu-pc:~$ hadoop fs -rm -r /sample1
Deleted /sample1
```

**Put:**

Transfer and store a data file from local systems to the Hadoop file system using the put command.

```
hadoop@ubuntu-pc:~$ hadoop fs -put /home/sunbeam/text.txt /sample
hadoop@ubuntu-pc:~$ hadoop fs -cat text.txt
```

**Cat:**

Initially, view the data from HDFS using **cat** command.

```
hadoop@ubuntu-pc:~$ hadoop fs -cat /sample/text.txt
Hadoop Framework.
HDFS is storage for Hadoop.
```

**Get:**

Gets the file from HDFS to the local file system using **get** command.

```
hadoop@ubuntu-pc:~$ hadoop fs -get /sample/text.txt /home/sunbeam
get: `/home/sunbeam/text.txt': File exists
```

You can shut down the HDFS by using the following command.

```
hadoop@ubuntu-pc:~/hadoop-3.2.2/sbin$ stop-all.sh
WARNING: Stopping all Apache Hadoop daemons as hdoop in 10 seconds.
WARNING: Use CTRL-C to abort.
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [ubuntu-pc]
Stopping nodemanagers
Stopping resourcemanager
```

**Conclusion/Outcome:** Thus, I have carried out file Manipulation successfully.