**Practical 1**

**Aim : Basic Commands of Linux and Hadoop**

**Hadoop HDFS Commands**

1. **Hadoop Version->hadoop version**

The Hadoop fs shell command version prints the Hadoop version.



**2) LS Command**

**->hdfs dfs -ls /**

HDFS Command to display the list of Files and Directories in HDFS. It Lists the contents of the directory specified by path,showing the names,permissions, owner, size and modification datefor each entry.

**hdfs dfs is the command that is specific to HDFS.**

|  |
| --- |
|  |
|  |  |

**->hadoop fs -ls /**

hadoop fs is more “generic” command that allows you to interact with multiple file systems

including Hadoop. we are using the ls command to enlist the files and directories present in

HDFS. The Hadoop fs shell command ls displays a list of the contents of a directory specified in

the path provided by the user. It shows the name, permissions, owner, size, and modification date

for each file or directories in the specified directory.

Using the ls command, we can check for the directories in HDFS.

**2) MKDIR Command**

HDFS Command to create the directory in HDFS.

Usage: hdfs dfs –mkdir /directory\_name

Here I am trying to create a directory named “rjc” in HDFS.

|  |
| --- |
|  |
|  |  |

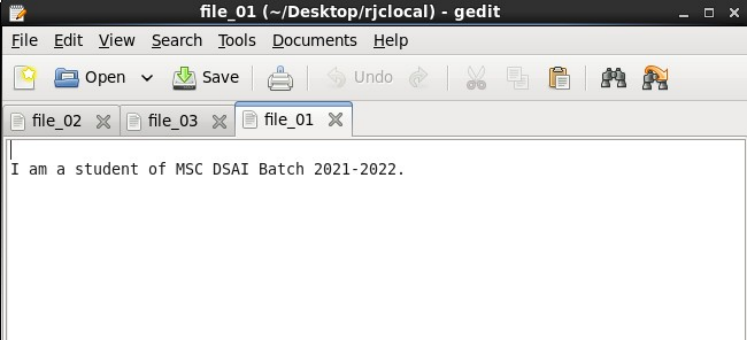
After creating ,Using the ls command, we can check for the directories in HDFS or Using ls command we listed the directory ‘rjc” created usingmkdir.

**3) copyFromLocal Command**

First we will Create a document.

Steps: Right click anywhere on desktop->empty file->file\_01

Put some information in the file file\_03.



Now we are trying to copy the ‘file\_01’ file present in the local file system to the ‘rjc’

directory of Hadoop. Below command copies the file from the local file system to HDFS.

|  |
| --- |
|  |
|  |  |

**4) If getting any error due to permissions**

Use-> export HADOOP\_USER\_NAME=hdfs

**5) Put Command**

-> hdfs dfs -put /home/cloudera/Desktop/ file\_02 /rjc

Here in this example, we are trying to copy “file\_02” of the local file system to the Hadoop file system.

The Hadoop fs shell command put is similar to the copyFromLocal, which copies files or directory from the local filesystem to the destination in the Hadoop filesystem.

Now using the ls command, we can check for the directories in HDFS.

|  |
| --- |
|  |
|  |  |

**6) copyToLocal Command**

**->hdfs dfs -copyToLocal /rjc/Sample /home/cloudera/Desktop**

**copyToLocal** command copies the file from HDFS to the local file system

Here in this example, we are trying to copy the ‘file\_01 file present in the rjc directory of HDFS

to the local file system.

|  |
| --- |
|  |
|  |  |

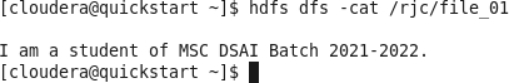
Deleted Sample file from desktop. If it is already exist. And then again run the command.

**7) CAT Command**

**->hdfs dfs -cat /rjc/file\_01**

we are using the cat command to display the content of the ‘Sample\_01’ file present in rjc directory of HDFS

The cat command reads the file in HDFS and displays the content of the file on console or stdout**.**



**8) Cp Command**

First we will create ‘rjcnew’ inside hdfs and then we copy ‘file\_01 file which is present in ‘rjc’ folder inside this ‘rjcnew directory in hdfs.

->hdfs dfs -cp /rjc/Sample /newdir

We are copying the ‘sample’ present in rjc directory in HDFS to the rjcnew of HDFS.

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

|  |
| --- |
|  |
|  |  |

|  |
| --- |
|  |
|  |  |

**9) MV Command**

->hdfs dfs -mv /rjc/file\_02 /rjcnew

we have a directory ‘rjc’ in HDFS. We are using mv command to move the rjc directory to the rjcnew directory in HDFS.

The HDFS mv command moves the files or directories from the source to a destination within HDFS.

|  |
| --- |
|  |
|  |  |

**10) RM Command**

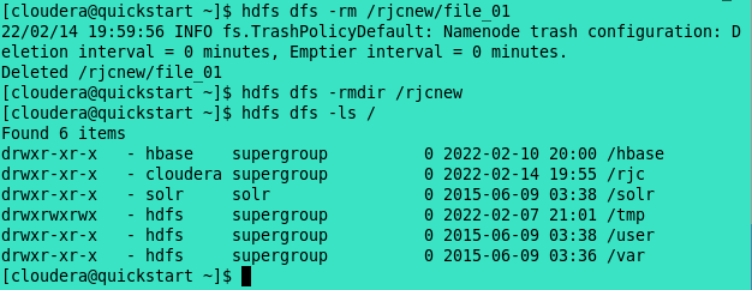
**->hdfs dfs –rm /rjcnew /file name**

The hadoop dfs -rm command deletes objects and directories full of objects.

|  |
| --- |
|  |
|  |  |

**11)->hdfs dfs -rm -r /filename or hdfs dfs –rmdir /filename**

In case, we want to delete a directory which contains files, -rm will not be able to delete the directory. In that case we can use recursive option for removing all the files from the directory following by removing the directory when it is empty**.**



**12) MoveFromLocal Command**

**->hdfs dfs -moveFromLocal /home/cloudera/Desktop/new /outputdir**

The Hadoop fs shell command moveFromLocal moves the file or directory from the local filesystem to the destination in Hadoop HDFS.

|  |
| --- |
|  |
|  |  |

**13) Tail Command**

**->hdfs dfs -tail /rjc/file\_01**

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

|  |
| --- |
|  |
|  |  |

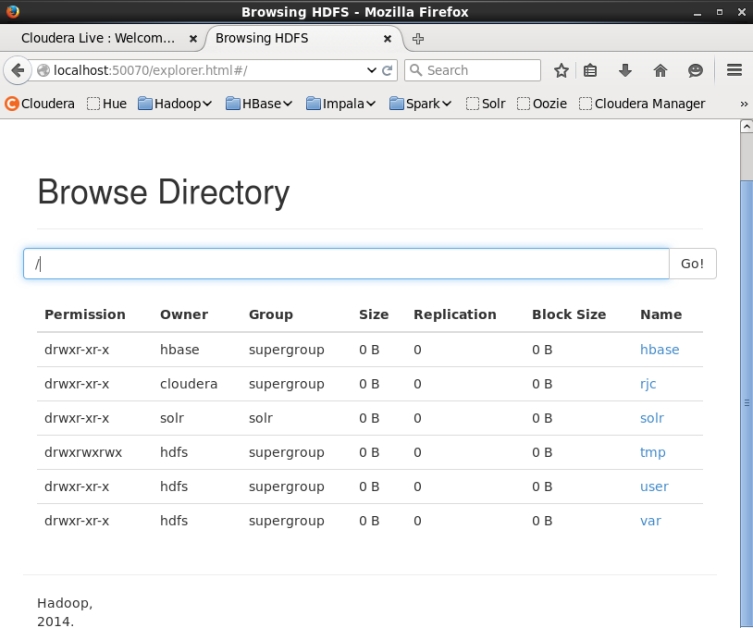
**14) Expunge Command**

**->hdfs dfs -expunge**

This command is used to empty the trash available in an HDFS system.

|  |
| --- |
|  |
|  |  |

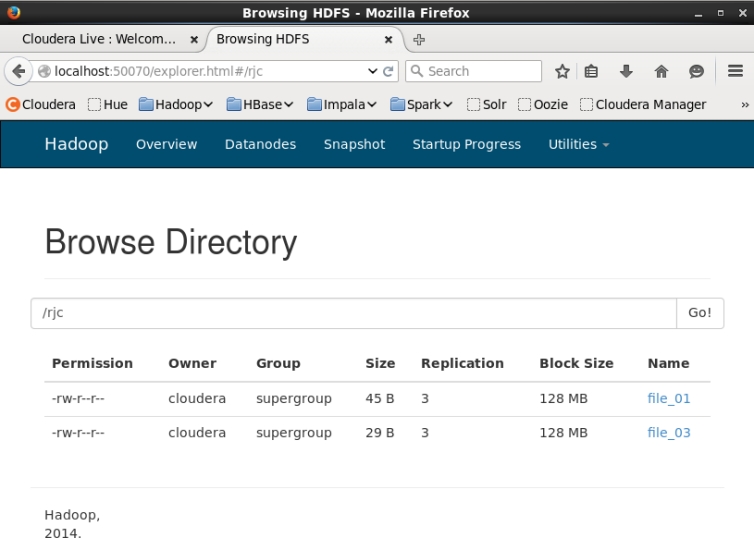
**15) Replication Command**

Earlier Replication Number of file\_01 is 1**.**

**->hdfs dfs -setrep 3 /output\_abc**

This command is used to change the replication factor of a file to a specific count instead of the default replication factor for the remaining in the HDFS file system.

|  |
| --- |
|  |
|  |  |

Now Replication factor will become 3. 

**16) DU Command**

**->hdfs dfs -du /rjc**

Use the hdfs du command to get the size of a directory in HDFS du stands for disk usage.

Since the replication factor of file\_01 was set as 3 earlier we could see that 45\*3=135 and for file\_02 will become 29\*3 =87.

|  |
| --- |
|  |
|  |  |

**17) appendToFile Command**

This command is used to append the text in the file on the Hadoop. We can write the text with the help of echo command.

|  |
| --- |
|  |
|  |  |

**18) Df command**

**->hdfs dfs -df**

To get all the space related details of the Hadoop File System we can use df command. It provides the information regarding the amount of space used and amount of space available on the currently mounted filesystem.

|  |
| --- |
|  |
|  |  |

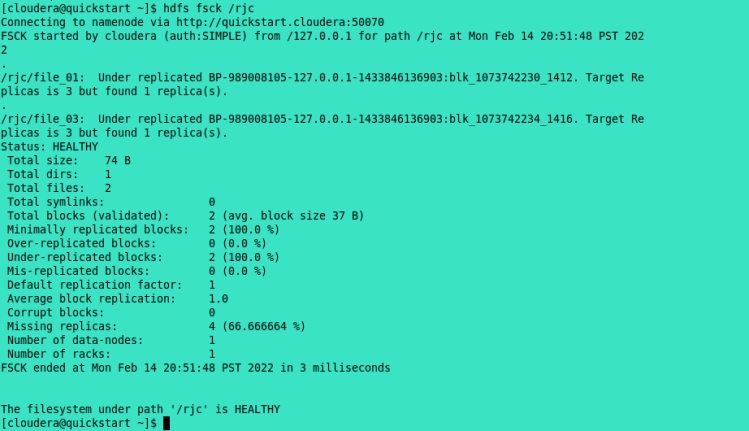
**->hdfs dfs -df -h**

|  |
| --- |
|  |
|  |  |

With h parameter the information is human readable.

**19) Fsck command**

The fsck Hadoop command is used to check the health of the HDFS. It moves a corrupted file to the lost+found directory. It deletes the corrupted files present in HDFS. It prints the files being checked**.**

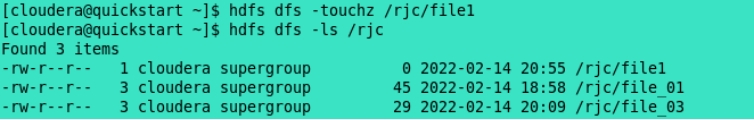


**->hdfs fsck /rjc -files**

|  |
| --- |
|  |
|  |  |

**20) Touchz Command**

It creates an empty file.



**21) Stat Command**

The Hadoop fs shell command stat prints the statistics about the file or directory in the specified Format.

It shows the recent date of modification.

->hdfs dfs -stat /rjc

**22) ->hdfs dfs -stat %b /rjc/file\_01**

|  |
| --- |
|  |
|  |  |

%b shows byte size of file

**23) Checksum Command**

Checksum property, which defaults to 512 bytes. The chunk size is stored as metadata in the crc file, so the file can be read back correctly even if the setting for the chunk size has change.

|  |
| --- |
|  |
|  |  |

## 24) Help command

**->hdfs dfs -help mkdir**

|  |
| --- |
|  |
|  |  |

Shows the syntax of whereas commands