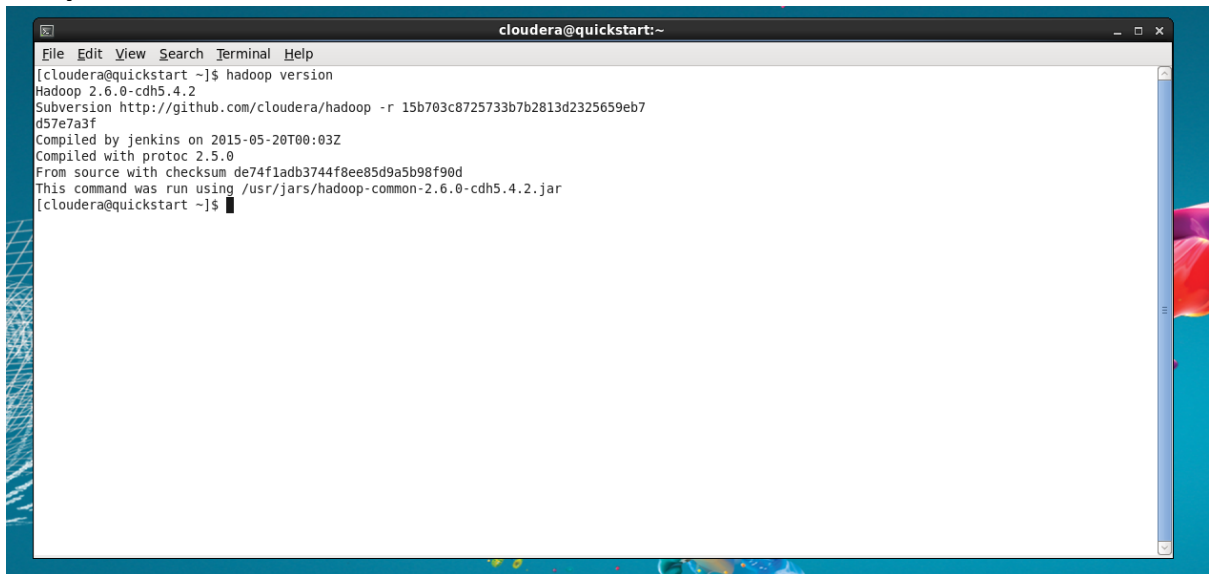# *Experiment – 02*

**Aim: -** To implement Various Hadoop HDFS Commands.

- **Command –**
  hadoop version

  **Description –**
  Shows the current Hadoop version running.

  **Output –**



- **Command –**
  hdfs dfs -ls /

  **Description –**
  This command is used to list all the files.

  **Output –**

```
                                    cloudera@quickstart:~
File  Edit  View  Search  Terminal  Help
[cloudera@quickstart ~]$ hadoop version
Hadoop 2.6.0-cdh5.4.2
Subversion http://github.com/cloudera/hadoop -r 15b703c8725733b7b2813d2325659eb7
d57e7a3f
Compiled by jenkins on 2015-05-20T00:03Z
Compiled with protoc 2.5.0
From source with checksum de74f1adb3744f8ee85d9a5b98f90d
This command was run using /usr/jars/hadoop-common-2.6.0-cdh5.4.2.jar
[cloudera@quickstart ~]$ hdfs dfs -ls /
Found 5 items
drwxr-xr-x   - hbase supergroup          0 2021-06-05 23:47 /hbase
drwxr-xr-x   - solr  solr                0 2015-06-09 03:38 /solr
drwxrwxrwx   - hdfs  supergroup          0 2021-06-05 23:58 /tmp
drwxr-xr-x   - hdfs  supergroup          0 2015-06-09 03:38 /user
drwxr-xr-x   - hdfs  supergroup          0 2015-06-09 03:36 /var
[cloudera@quickstart ~]$
```

- **Command –**
  hadoop fs -ls /

## Description –

This command is used to list all the files in generic file system.

## Output –

```
[cloudera@quickstart ~]$ hadoop fs -ls /
Found 5 items
drwxr-xr-x   - hbase supergroup          0 2021-06-05 23:47 /hbase
drwxr-xr-x   - solr  solr                0 2015-06-09 03:38 /solr
drwxrwxrwx   - hdfs  supergroup          0 2021-06-05 23:58 /tmp
drwxr-xr-x   - hdfs  supergroup          0 2015-06-09 03:38 /user
drwxr-xr-x   - hdfs  supergroup          0 2015-06-09 03:36 /var
[cloudera@quickstart ~]$
```
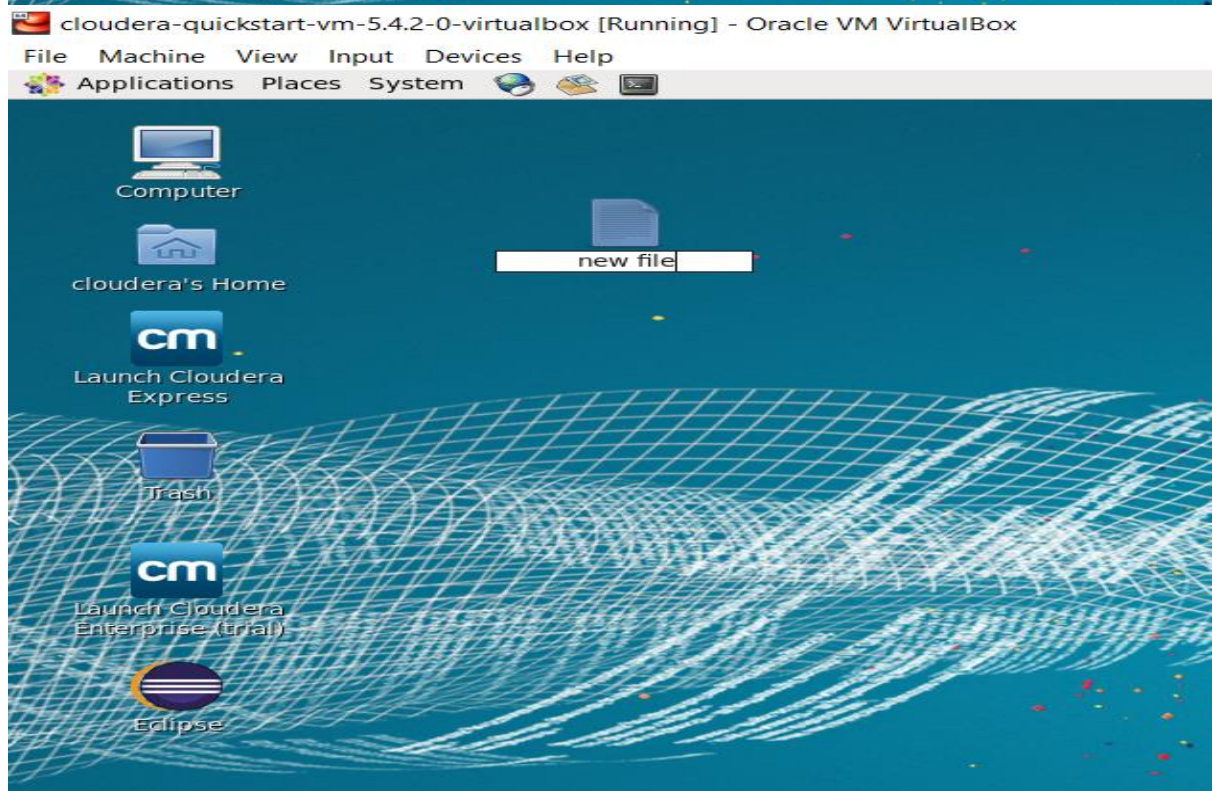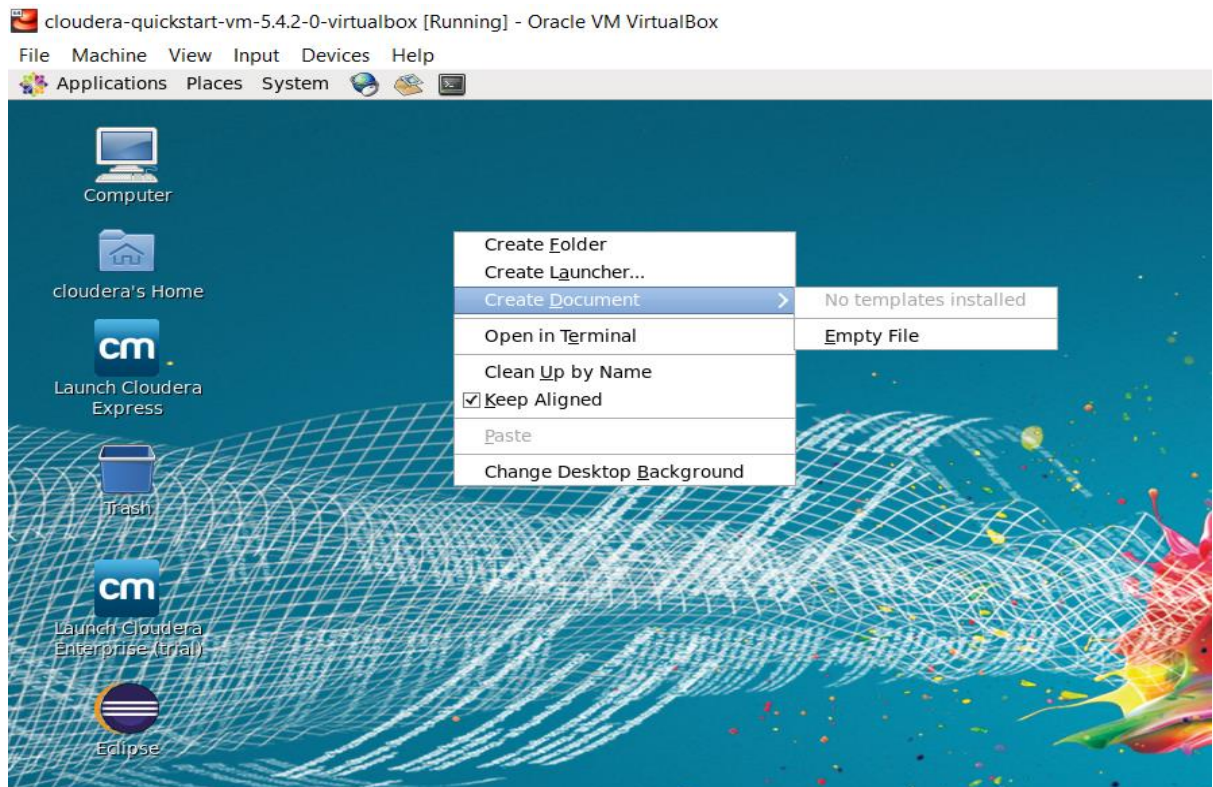
- **Create a new file named as prc 1 and add the following text paragraph in it.**
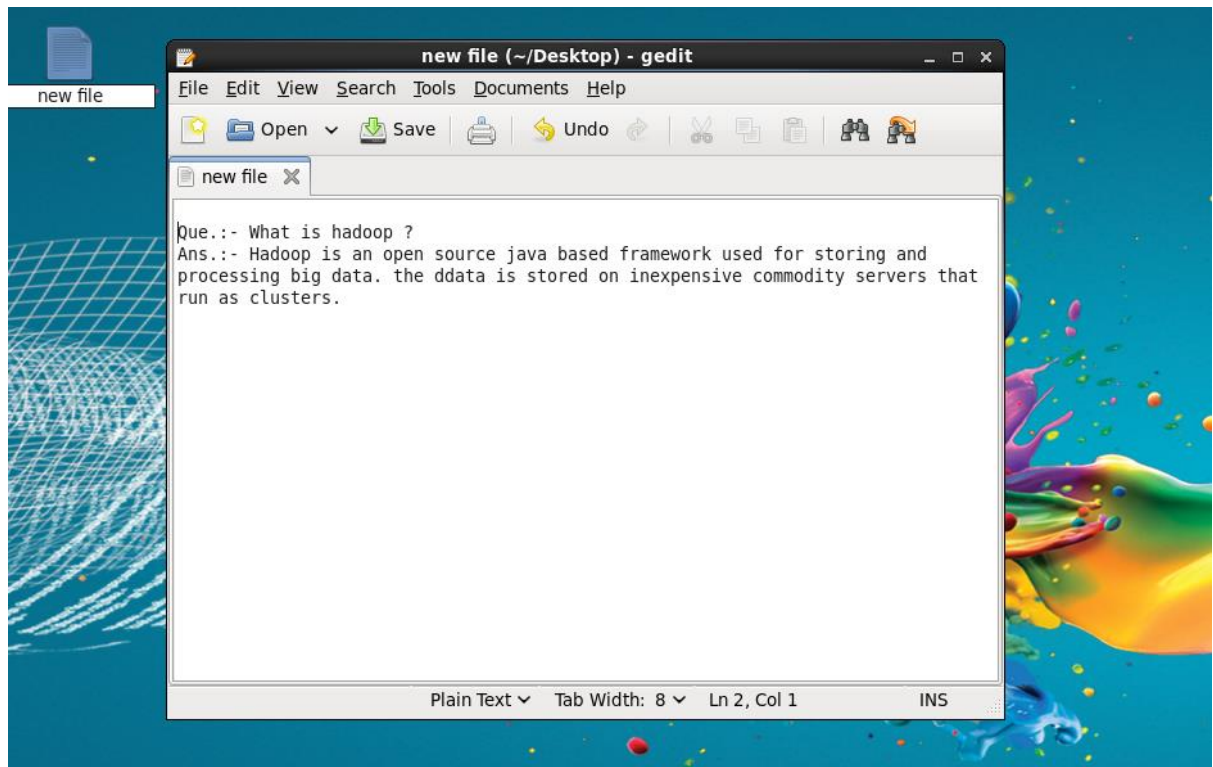
- **Command –**
  hdfs dfs -mkdir /rjc
  hdfs dfs -copyFromLocal /home/cloudera/Desktop/newfile /rjc
  hdfs dfs -ls/rjc

  **Description –**
  **mkdir**:  create a directory. In Hadoop *dfs* there is no home directory by default. So let's first create it.
  **copyFromLocal**: copy files/folders from local file system to hdfs store specified path.

  **Output –**

- **Command –**

  hdfs dfs -put /home/cloudera/Desktop/newfile  /
  hdfs dfs -ls /

  **Description –**

  **put**: shift files/folders from local file system to hdfs store specified path.

  **Output –**

  ```
  [cloudera@quickstart ~]$ hdfs dfs -put /home/cloudera/Desktop/newfile /
  [cloudera@quickstart ~]$ hdfs dfs -ls /
  Found 7 items
  drwxr-xr-x   - hbase    supergroup          0 2021-06-05 23:47 /hbase
  -rw-r--r--   1 cloudera supergroup        196 2021-06-06 00:45 /newfile
  drwxr-xr-x   - cloudera supergroup          0 2021-06-06 00:38 /rjc
  drwxr-xr-x   - solr     solr                0 2015-06-09 03:38 /solr
  drwxrwxrwx   - hdfs     supergroup          0 2021-06-05 23:58 /tmp
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:38 /user
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:36 /var
  [cloudera@quickstart ~]$
  ```

- **Command –**

  ls
  ls Desktop /

  **Description –**

  **ls Desktop /**: display list of files using with type

  **Output –**

  ```
  [cloudera@quickstart ~]$ hdfs dfs -ls /
  Found 7 items
  drwxr-xr-x   - hbase    supergroup          0 2021-06-05 23:47 /hbase
  -rw-r--r--   1 cloudera supergroup        196 2021-06-06 00:45 /newfile
  drwxr-xr-x   - cloudera supergroup          0 2021-06-06 00:38 /rjc
  drwxr-xr-x   - solr     solr                0 2015-06-09 03:38 /solr
  drwxrwxrwx   - hdfs     supergroup          0 2021-06-05 23:58 /tmp
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:38 /user
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:36 /var
  [cloudera@quickstart ~]$ ls
  cloudera-manager  Desktop    Downloads  enterprise-deployment.json  kerberos  Music     Pub
  cm_api.py         Documents  eclipse    express-deployment.json     lib       Pictures  Tem
  [cloudera@quickstart ~]$ ls Desktop/
  Eclipse.desktop  Enterprise.desktop  Express.desktop  newfile  new file~
  [cloudera@quickstart ~]$
  ```

- **Command –**
  hdfs dfs -copyToLocal /rjc/newfile /home/cloudera/Desktop
  ls Desktop/

  ### Description –
  **copyToLocal** : Copy files/folders from hdfs store path specified to local file system.

  ### Output –
  ```
  [cloudera@quickstart ~]$ hdfs dfs -copyToLocal /rjc/newfile /home/cloudera/Desktop
  copyToLocal: `/home/cloudera/Desktop/newfile': File exists
  [cloudera@quickstart ~]$
  ```

- **Command –**
  hdfs dfs -get /rjc/newfile /home/cloudera/Desktop/new
  hdfs dfs -cat /rjc/newfile

  ### Description –
  **get:** Shift files/folders from hdfs store path specified to local file system.

  **cat:** Print file contents**.**

  ### Output –
  ```
  [cloudera@quickstart ~]$ hdfs dfs -copyToLocal /rjc/newfile /home/cloudera/Desktop
  copyToLocal: `/home/cloudera/Desktop/newfile': File exists
  [cloudera@quickstart ~]$ hdfs dfs -get /rjc/newfile /home/cloudera/Desktop/new
  [cloudera@quickstart ~]$ hdfs dfs -cat /rjc/newfile

  Que.:- What is hadoop ?
  Ans.:- Hadoop is an open source java based framework used for storing and processing big data. the ddata is stored on inexpensive commodity servers
   that run as clusters.
  [cloudera@quickstart ~]$
  ```
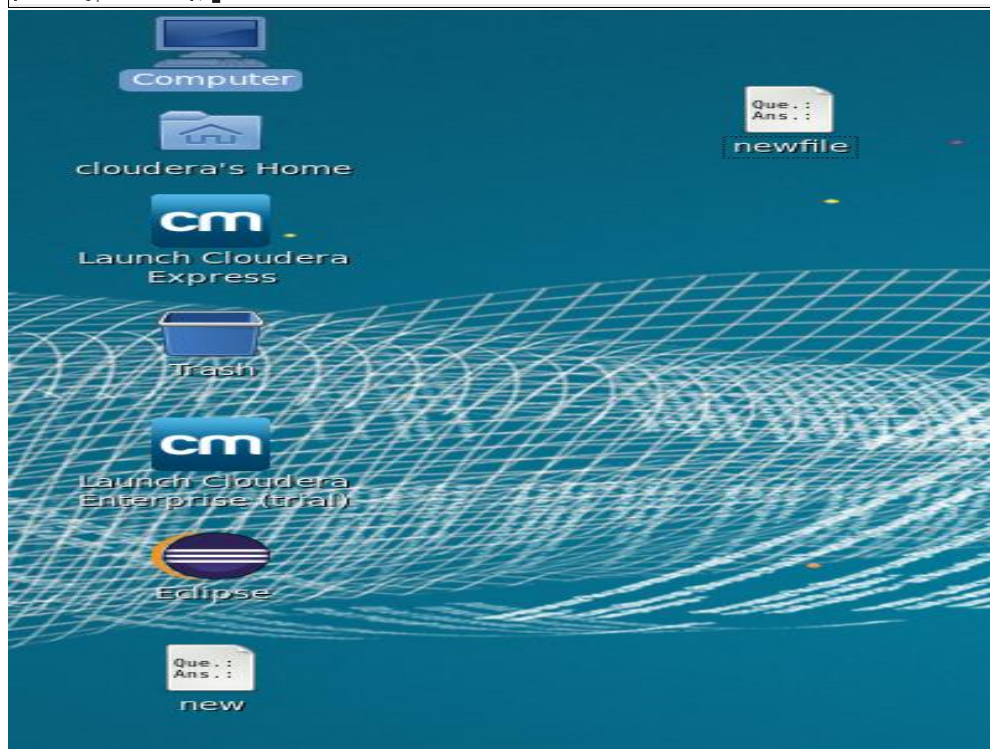
  

- **Command –**
  hdfs dfs -mkdir /newdir
  hdfs dfs -ls /newdir
  hdfs dfs -cp /rjc/newfile /newdir
  hdfs dfs -ls /newdir

  **Description –**
  **cp:** This command is used to copy files within hdfs.

  **Output –**
  ```
  [cloudera@quickstart ~]$ hdfs dfs -mkdir /newdir
  [cloudera@quickstart ~]$ hdfs dfs -ls /newdir
  [cloudera@quickstart ~]$ hdfs dfs -cp /rjc/newfile /newdir
  [cloudera@quickstart ~]$ hdfs dfs -ls /newdir
  Found 1 items
  -rw-r--r--   1 cloudera supergroup        196 2021-06-06 01:08 /newdir/newfile
  [cloudera@quickstart ~]$
  ```

- **Command –**
  hdfs dfs -mkdir /output_new
  hdfs dfs -ls /output_new
  hdfs dfs -mv /rjc/newfile /output_new
  hdfs dfs -ls /output_new

  **Description –**
  **mv:** This command is used to move files within hdfs.

  **Output –**
  ```
  [cloudera@quickstart ~]$ hdfs dfs -ls /output_new
  ls: `/output_new': No such file or directory
  [cloudera@quickstart ~]$ hdfs dfs -mkdir /output_new
  [cloudera@quickstart ~]$ hdfs dfs -ls /output_new
  [cloudera@quickstart ~]$ hdfs dfs -mv /rjc/newfile /output_new
  [cloudera@quickstart ~]$ hdfs dfs -ls /rjc
  [cloudera@quickstart ~]$ hdfs dfs -ls /output_new
  Found 1 items
  -rw-r--r--   1 cloudera supergroup        196 2021-06-06 00:38 /output_new/newfile
  [cloudera@quickstart ~]$
  ```

- **Command –**
  hdfs dfs -rm /output_new/newfile
  hdfs dfs -rm -r /output_new

  **Description –**
  **rm**: Remove file from directory
  **rm -r**: Remove directory

  **Output –**

```
[cloudera@quickstart ~]$ hdfs dfs -rm /output_new/newfile
21/06/06 01:16:14 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /output_new/newfile
[cloudera@quickstart ~]$ hdfs dfs -ls /output_new
[cloudera@quickstart ~]$
```

- **Create a new file named as manoj and add the following text paragraph in it.**
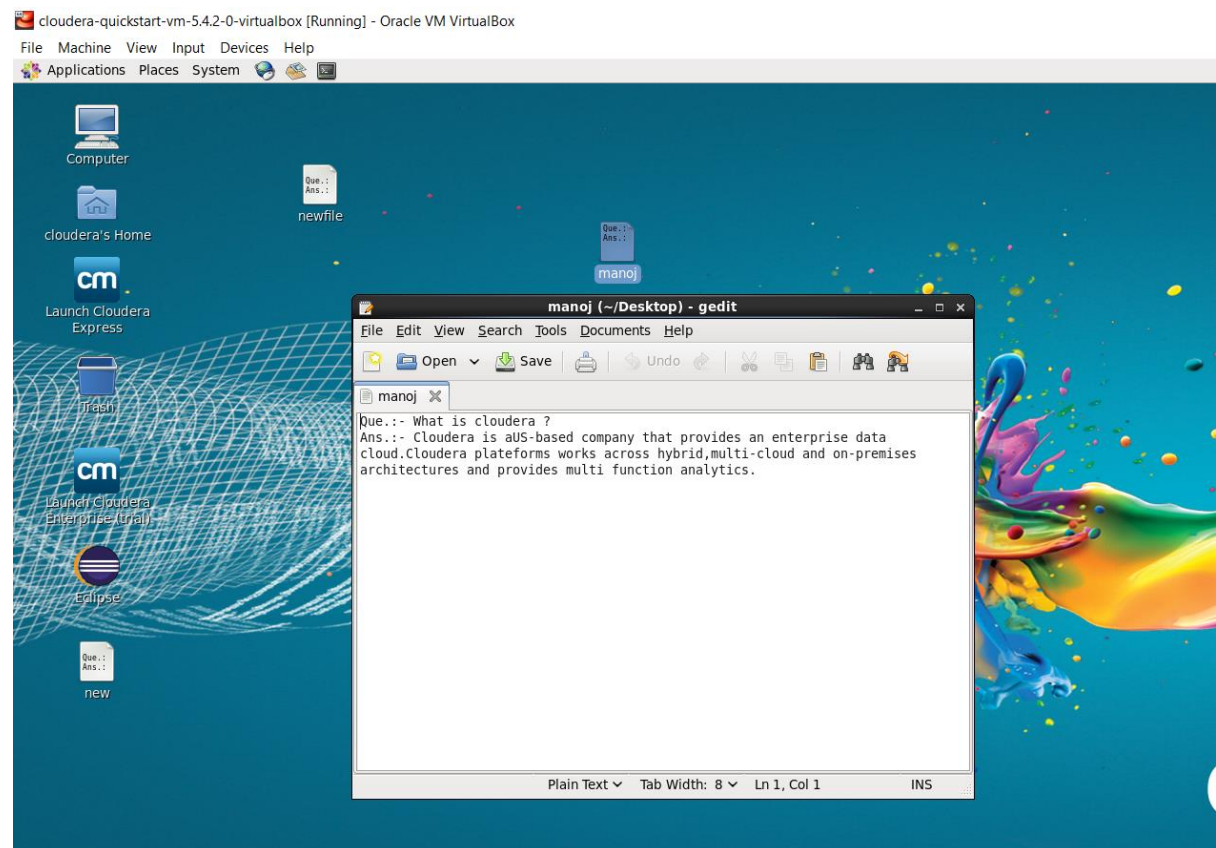
- **Command –**

  hdfs dfs - moveFromLocal / home/cloudera/Desktop/manoj /manoj

  **Description –**

  **moveFromLocal:** This command will move file from local to hdfs.

  **Output –**

  ```
  [cloudera@quickstart ~]$ hdfs dfs -moveFromLocal /home/cloudera/Desktop/manoj /manoj
  [cloudera@quickstart ~]$ hdfs dfs -ls /
  Found 10 items
  drwxr-xr-x   - hbase    supergroup          0 2021-06-05 23:47 /hbase
  -rw-r--r--   1 cloudera supergroup        224 2021-06-06 01:30 /manoj
  drwxr-xr-x   - cloudera supergroup          0 2021-06-06 01:08 /newdir
  -rw-r--r--   1 cloudera supergroup        196 2021-06-06 00:45 /newfile
  drwxr-xr-x   - cloudera supergroup          0 2021-06-06 01:16 /output_new
  drwxr-xr-x   - cloudera supergroup          0 2021-06-06 01:13 /rjc
  drwxr-xr-x   - solr     solr               0 2015-06-09 03:38 /solr
  drwxrwxrwx   - hdfs     supergroup          0 2021-06-05 23:58 /tmp
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:38 /user
  drwxr-xr-x   - hdfs     supergroup          0 2015-06-09 03:36 /var
  [cloudera@quickstart ~]$
  ```

- **Command –**

  hdfs dfs - tail /manoj

  **Description –**

  **tail /:** It returns the last kilobyte of the data.

  **Output –**

  ```
  [cloudera@quickstart ~]$ hdfs dfs -tail /manoj
  Que.:- What is cloudera ?
  Ans.:- Cloudera is aUS-based company that provides an enterprise data cloud.Cloudera plateforms works across hybrid,multi-cloud and on-premises arc
  hitectures and provides multi function analytics.
  [cloudera@quickstart ~]$
  ```

- **Command –**

  hdfs dfs – expunge
  hdfs dfs - setrep 4 /manoj

  **Description –**

  **expunge:** This command is used to empty the trash available in an HDFS system.
  **setrep:** This command is used to change the replication factor of a file/directory in HDFS.
  By default it is 3 for anything which is stored in HDFS.

  **Output –**

  ```
  [cloudera@quickstart ~]$ hdfs dfs -expunge
  21/06/06 01:33:12 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
  [cloudera@quickstart ~]$ hdfs dfs -setrep 4 /manoj
  Replication 4 set: /manoj
  [cloudera@quickstart ~]$
  ```

- **Command –**
  hdfs dfs -du /manoj

  **Description –**
  **du:** It will give the size of each file in directory.

  **Output –**
  ```
  [cloudera@quickstart ~]$ hdfs dfs -du /manoj
  224  896  /manoj
  [cloudera@quickstart ~]$ 
  ```

- **Command –**
  hdfs dfs -df
  hdfs dfs -df -h

  **Description –**
  **df -h:** Display details of that files like memory usage size on disk in human readable format.

  **Output –**
  ```
  [cloudera@quickstart ~]$ hdfs dfs -df
  Filesystem                         Size         Used      Available  Use%
  hdfs://quickstart.cloudera:8020  58665738240  394653696  48141094912   1%
  [cloudera@quickstart ~]$ hdfs dfs -df -h
  Filesystem                        Size      Used  Available  Use%
  hdfs://quickstart.cloudera:8020  54.6 G  376.4 M    44.8 G    1%
  [cloudera@quickstart ~]$ 
  ```

- **Command –**
  hdfs fsck /manoj -files

  **Description –**
  **Fsck:** It is used to check the health of the file system, to find missing files, over replicated, under replicated and corrupted blocks.

  **Output –**

```
[cloudera@quickstart ~]$ hdfs fsck /manoj -files
Connecting to namenode via http://quickstart.cloudera:50070
FSCK started by cloudera (auth:SIMPLE) from /10.0.2.15 for path /manoj at Sun Jun 06 02:05:06 PDT 2021
/manoj 224 bytes, 1 block(s):  Under replicated BP-989008105-127.0.0.1-1433846136903:blk_1073742209_1385. Target Replicas is 4 but found 1 replica(
s).
Status: HEALTHY
 Total size:    224 B
 Total dirs:    0
 Total files:   1
 Total symlinks:            0
 Total blocks (validated):      1 (avg. block size 224 B)
 Minimally replicated blocks:   1 (100.0 %)
 Over-replicated blocks:        0 (0.0 %)
 Under-replicated blocks:       1 (100.0 %)
 Mis-replicated blocks:         0 (0.0 %)
 Default replication factor:    1
 Average block replication:     1.0
 Corrupt blocks:                0
 Missing replicas:              3 (75.0 %)
 Number of data-nodes:          1
 Number of racks:               1
FSCK ended at Sun Jun 06 02:05:06 PDT 2021 in 34 milliseconds


The filesystem under path '/manoj' is HEALTHY
[cloudera@quickstart ~]$
```

- **Command –**
  hdfs dfs -mkdir / inputnew
  hdfs dfs - ls / inputnew
  hdfs dfs -touchz / inputnew /file1
  hdfs dfs - ls / inputnew

## Description –

**touchz:** It creates an empty file.

## Output –

```
[cloudera@quickstart ~]$ hdfs dfs -mkdir /inputnew
[cloudera@quickstart ~]$ hdfs dfs -ls /inputnew
[cloudera@quickstart ~]$ hdfs dfs -touchz /inputnew /file1
touchz: `/inputnew': Is a directory
[cloudera@quickstart ~]$ hdfs dfs -touchz /inputnew/file1
[cloudera@quickstart ~]$ hdfs dfs -ls /inputnew
Found 1 items
-rw-r--r--   1 cloudera supergroup          0 2021-06-06 02:07 /inputnew/file1
[cloudera@quickstart ~]$
```