

**Bigdata Technologies WEEK 1**

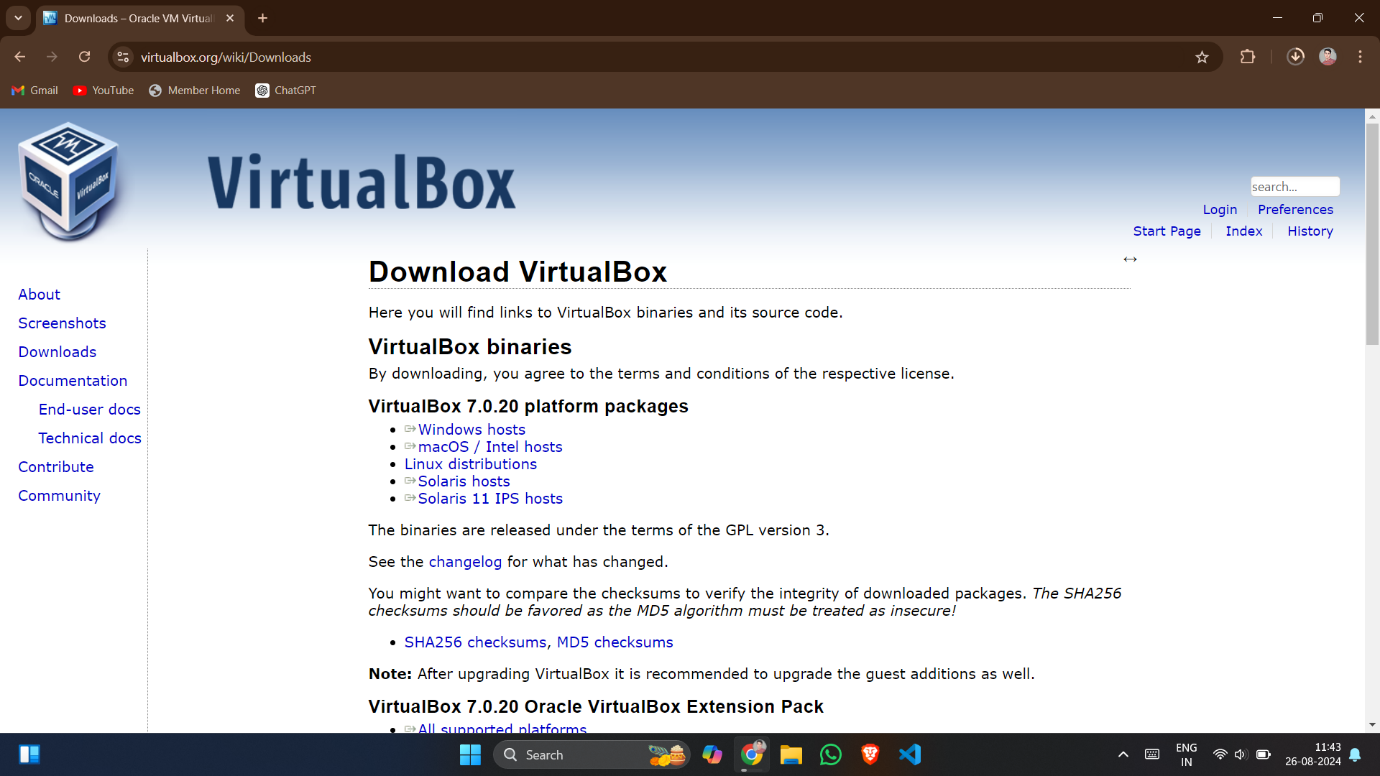
**PROBLEM STATEMENT :**

Oracle virtual VM and Cloudera installation.

**Website:**

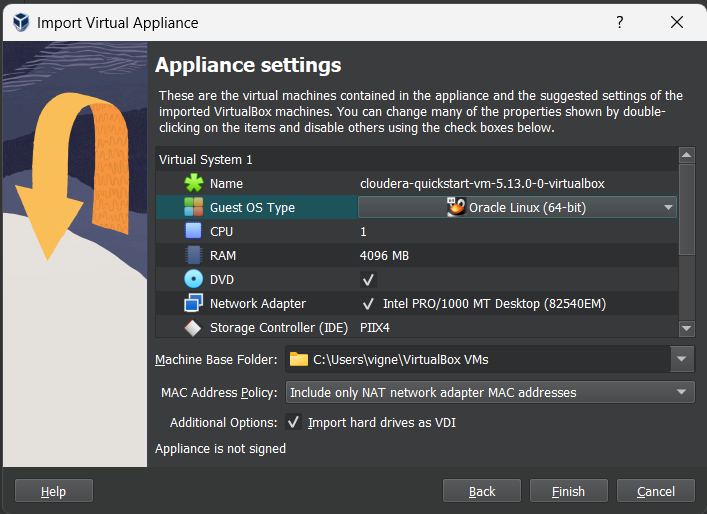
Oracle VM VirtualBox **Steps:**

1. Install Oracle VM VirtualBox and Cloudera from the official websites.
2. Now install the VM VirtualBox in the system and also extract all cloudera files into the respective local drive.

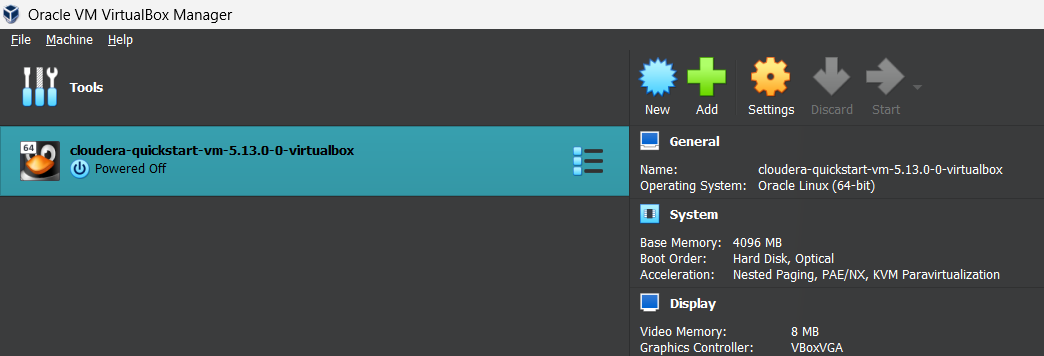


1. Now open the VM VirtualBox and import the cloudera file(15kb) and make changes in the OS while importing.



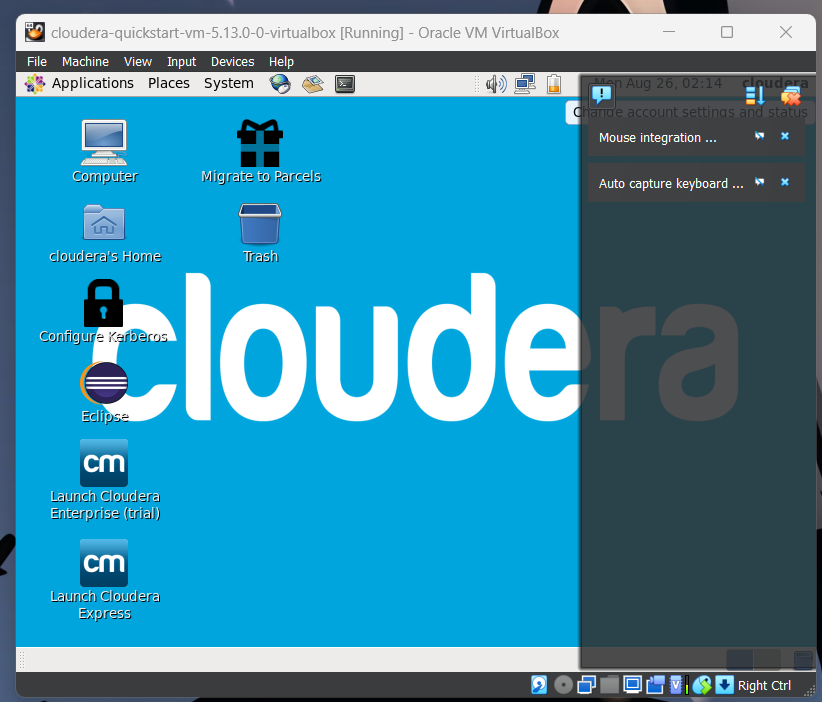


4.Now after the importing of the cloudera is completed, open the cloudera in the VM VirtualBox.



1. As the cloudera is running on the console, you can perform all the Hadoop operations in the cloudera terminal.

**Final Output Console:**



**2.Problem Statement:**

**Create a script that demonstrate the use of basic linux commands such as ls,cd,mkdir etc to manage files and directories and various operations**

**1. What command is used to list the files and directories in the current directory?**

Ls -command lists all files and directories in the current working directory.

**2. How do you create a new directory in Linux?**

mkdir directory\_name - is used to create a new directory. Replace directory\_name with the desired name of your directory**.**

**3. Which command is used to change the current directory?**

cd directory\_name - changes the current directory to the specified one. Use cd .. to move up one directory level.

**4. How can you display the contents of a text file in the terminal?**

cat filename - is used to display the contents of a text file. Replace filename with the name of the file.

**5. How do you copy a file from one location to another?**

cp source\_file destination\_directory - copies a file from the source location to the destination. Replace source\_file with the file you want to copy and destination\_directory with the target directory.

**6. Which command is used to move or rename a file?**

mv old\_name new\_name - is used to move or rename a file or directory. Replace old\_name with the current name and new\_name with the new name or new location.

**7. How do you delete a file in Linux?**

rm filename - removes a file. Replace filename with the name of the file you want to delete.

**8. What command is used to find out your current working directory?**

Pwd - prints the current working directory path.

**9. How do you display the first few lines of a file?**

head filename - shows the first 10 lines of a file by default. You can specify the number of lines using -n (e.g., head -n 5 filename).

**10. Question: How can you check the disk usage of files and directories?**

du

The `du` command estimates the disk space used by files and directories. Use `du -h` for a human-readable format**.**

**11. Which command shows the current user’s login name?**

Who am i

**12.How do you create an empty file in Linux?**

touch filename

**13.How can you delete a directory and all its contents?**

rm -r directory\_name

**14.. Question: What command is used to count the number of lines, words, and characters in a file?**

wc filename

**15.. Question: How do you display the current date and time?**

date

**16.Which command shows the last few lines of a file?**

tail filename

**17.How do you search for a specific pattern in a file?**

grep "pattern" filename

**18.Which command will show the amount of free and used memory in the system?**

**free**

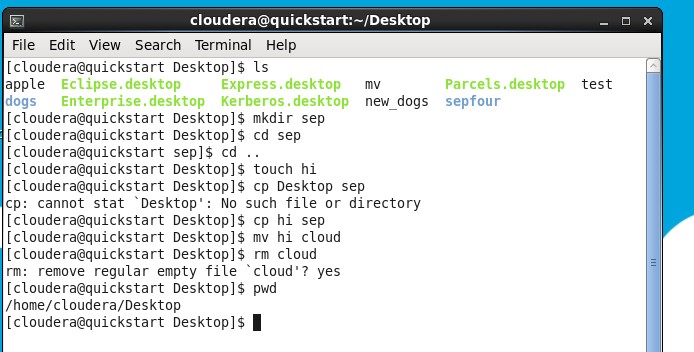
**19. What command would you use to compress a file using gzip?**

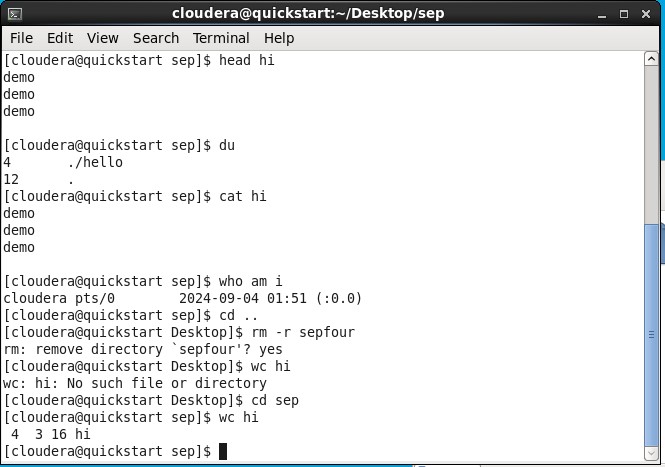
gzip filename

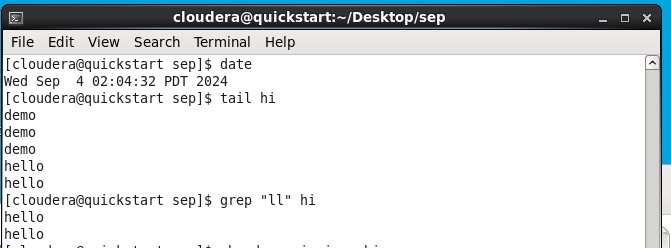
**20. Which command is used to display your IP address?**

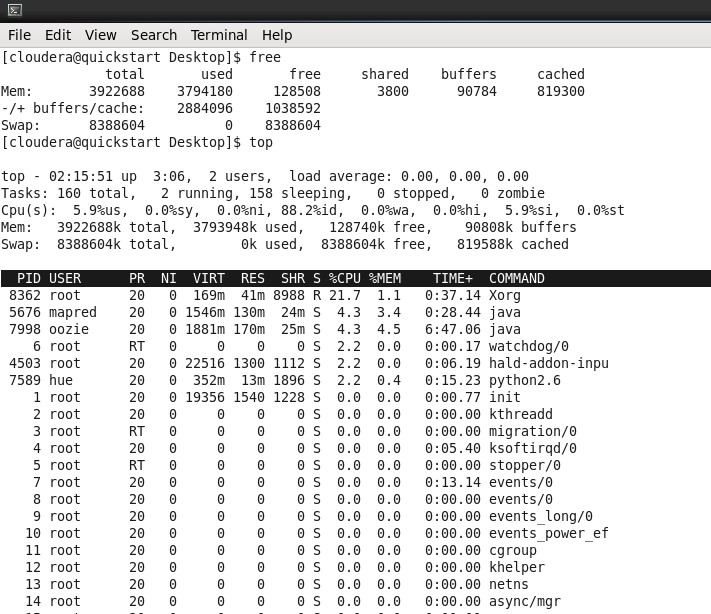
ifconfig

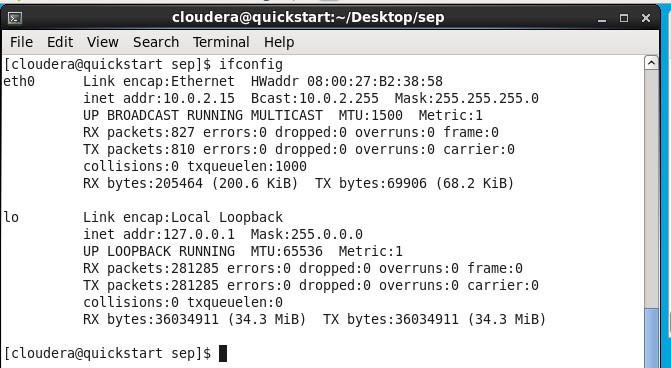
**ls,mkdir,cd,touch,cp,mv,rm,pwd:**

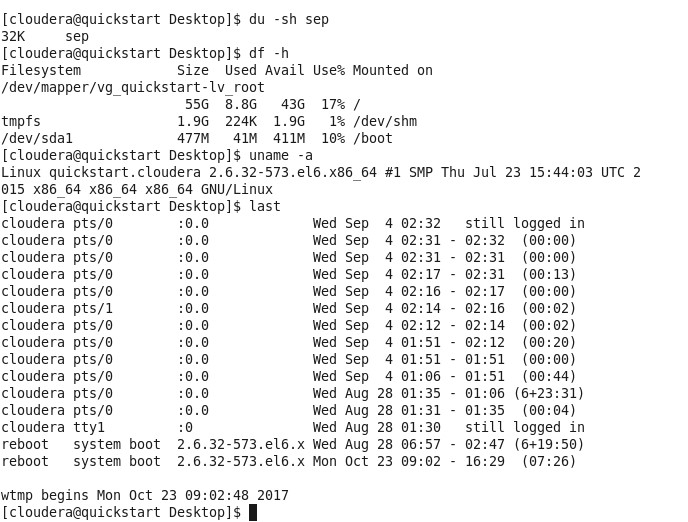
****

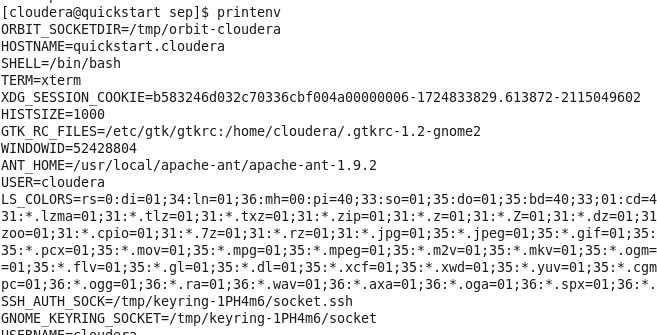
**Head,du,cat,who am I,wc:**

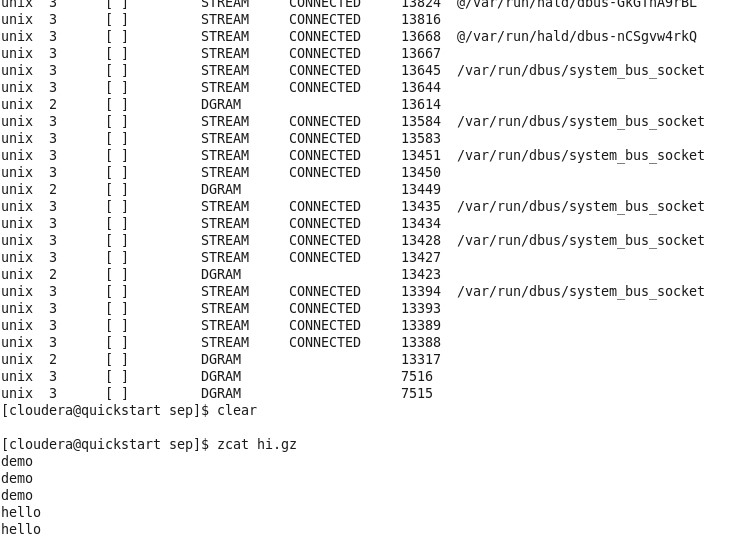
**date,tail,grep,free,top:**

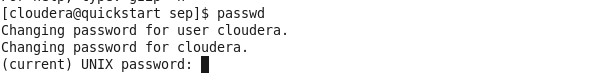


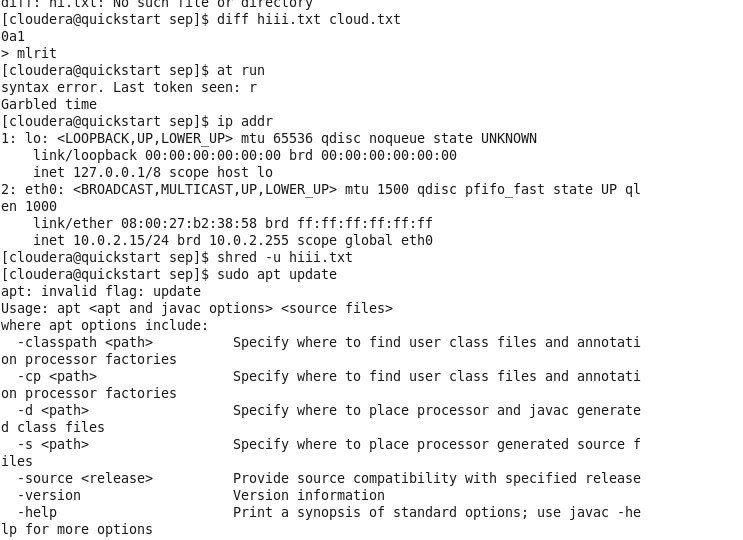
**Ifconfig,df -h,uname -a,printenv:**

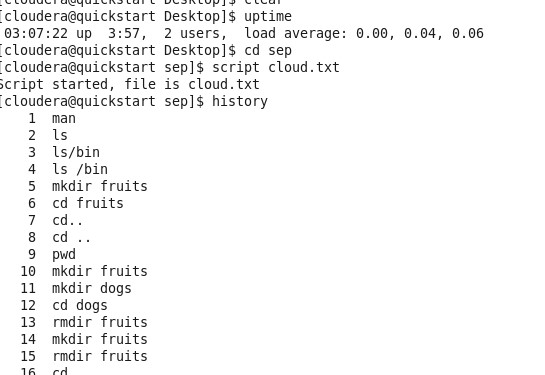




**Clear,zcat,passwd,diff,history:**







**3.Problem Statement:** **Write a script that demonstrate the basic HDFS shell commands**

**1. How do you list all files and directories in a specific HDFS directory?**

hdfs dfs -ls /path/to/directory

**2. Which command is used to display the contents of a file in HDFS?**

hdfs dfs -cat /path/to/file

**3. How can you create a new directory in HDFS?**

hdfs dfs -mkdir /path/to/new\_directory

**4. What command is used to copy files from the local file system to HDFS?**

hdfs dfs -put local\_file /path/to/hdfs\_directory

**5. How do you retrieve a file from HDFS to the local file system?**

hdfs dfs -get /path/to/hdfs\_file /local/destination\_directory

**6. Which command can be used to remove a file or directory in HDFS?**

hdfs dfs -rm /path/to/hdfs\_file\_or\_directory

**7. How do you delete a non-empty directory in HDFS?**

hdfs dfs -rm -r /path/to/non\_empty\_directory

**8. What command is used to check the disk usage of a directory in HDFS?**

hdfs dfs -du -h /path/to/directory

**9. How do you move or rename a file in HDFS?**

hdfs dfs -mv /path/to/source /path/to/destination

**10. Which command will display the last part of a file in HDFS?**

hdfs dfs -tail /path/to/file

**11. How do you merge multiple files into a single file in HDFS?**

hdfs dfs -getmerge /path/to/source\_directory /path/to/destination\_file

**12. What command is used to change the owner of a file or directory in HDFS?**

hdfs dfs -chown new\_owner:new\_group /path/to/file\_or\_directory

**13. Question: How do you change the permissions of a file in HDFS?**

hdfs dfs -chmod 755 /path/to/file\_or\_directory

**14. Which command will change the group ownership of a file or directory in HDFS?**

hdfs dfs -chgrp new\_group /path/to/file\_or\_directory

**15. How can you set the replication factor for a file in HDFS?**

hdfs dfs -setrep -w replication\_factor /path/to/file

**16. What command is used to display the status of a file or directory in HDFS?**

hdfs dfs -stat /path/to/file\_or\_directory

**17. How do you copy files from one HDFS location to another?**

hdfs dfs -cp /source/path /destination/path

**18. Which command is used to count the number of directories, files, and bytes in HDFS?**

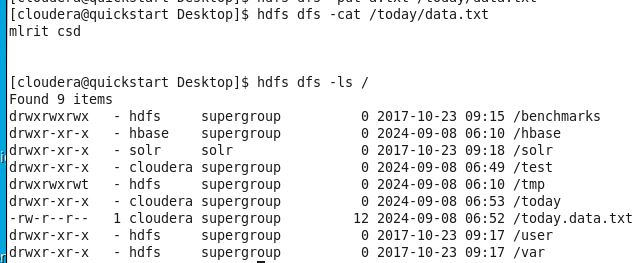
hdfs dfs -count /path/to/directory

**19. Question: How do you check the checksum of a file in HDFS?**

hdfs dfs -checksum /path/to/file

**20. Question: Which command is used to display the HDFS usage summary?**

hdfs dfsadmin -report

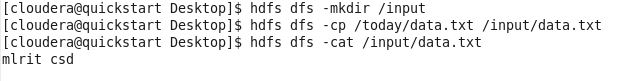
**cat , ls :**

**Du,rm:**

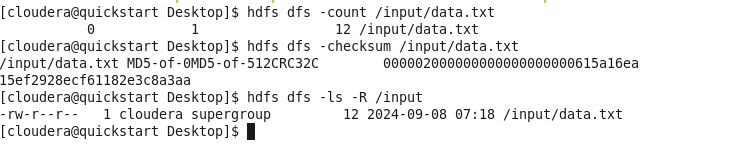




**Mkdir ,cp,cat:**

****

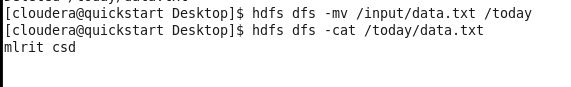
**Get,count,checksum,:**

****

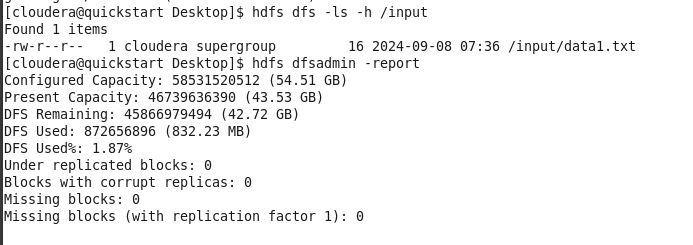
**Stat,tail,mv:**

****

****

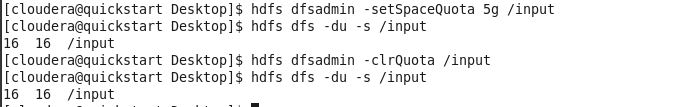
****

**H,getconf:**

****

****

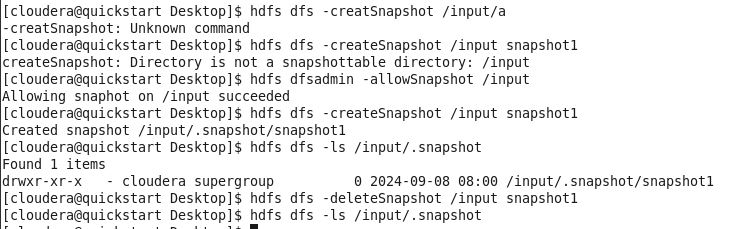
**spaceQuota:**

****

**Safemod:**

****

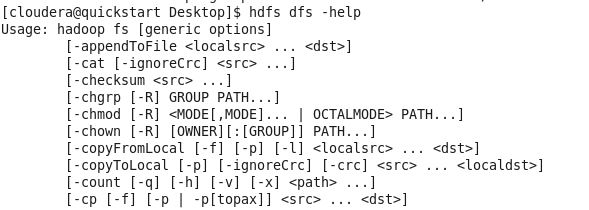
**Snapshot:**

****

**-rm:**

****

**-help:**

****