WEEK 7 HADOOP

HDFS commands are used to access the Hadoop File System. HDFS stands for 'Hadoop Distributed File System'. The HDFS is a sub-project of the Apache Hadoop project. This Apache Software Foundation project is designed to provide a fault-tolerant file system designed to run on commodity hardware. HDFS is accessed through a set of shell commands.

1. To Check the hadoop version

[cloudera@quickstart Desktop]$ hadoop version

2. To list the files and subdirectories present in HDFS : This command will list all the available files and subdirectories under default directory. For instance, the **default directory for Cloudera VM is /user/cloudera**

[cloudera@quickstart Desktop]$ hadoop fs -ls

3. Return all the directories under root directory

[cloudera@quickstart Desktop]$ hadoop fs -ls /

4. Copy a file from local to HDFS?

Usage: hadoop fs -copyFromLocal <localsrc> <dest>

First create a file in local named student.txt

[cloudera@quickstart Desktop]$ gedit students.txt

[cloudera@quickstart Desktop]$ hadoop fs -copyFromLocal students.txt /user/cloudera

(or)

Usage: hadoop fs -put <localsrc> <dest>

[cloudera@quickstart Desktop]$ hadoop fs -put students.txt /user/cloudera/studentscopied.txt

[cloudera@quickstart Desktop]$ hadoop fs -ls

5. Check the contents of file that you copied in HDFS?

Usage: hadoop fs –cat /path_to_file_in_hdfs

[cloudera@quickstart Desktop]$ hadoop fs -cat students.txt

6. Copy any file from HDFS to Local File System

Usage: hadoop fs -copyToLocal <hdfs src> < local dest>

[cloudera@quickstart Desktop]$ hadoop fs -copyToLocal students.txt studentscopiedtolocal.txt

[cloudera@quickstart Desktop]$ ls

7. Check the health of the Hadoop file system.

Command: hdfs fsck /

8. Create a directory in HDFS

Usage: hadoop fs –mkdir /directory_name

[cloudera@quickstart Desktop]$ hadoop fs –mkdir kriti2018

[cloudera@quickstart Desktop]$ hadoop fs -ls

9. Copy any file present in HDFS to a directory which is also present in HDFS

[cloudera@quickstart Desktop]$ hadoop fs –cp kriti.txt kriti2018

10. Create an empty file in HDFS?

Usage : hadoop fs –touchz /directory/filename

[cloudera@quickstart Desktop]$ hadoop fs –touchz empty.txt

[cloudera@quickstart Desktop]$ hadoop fs -ls

11. Check the file size of any file in HDFS?

Usage: hadoop fs –du –s /directory/filename

[cloudera@quickstart Desktop]$ hadoop fs –du students.txt

12. Print the contents of a file stored in HDFS?

Usage: hadoop fs –cat /path/to/file_in_hdfs

[cloudera@quickstart Desktop]$ hadoop fs -cat students.txt

13. Count the number of directories and files inside a directory in HDFS? : HDFS Command to count the number of directories, files, and bytes under the paths that match the specified file pattern.

Usage: hadoop fs -count <path>

[cloudera@quickstart Desktop]$ hadoop fs –count kriti2018

14. Remove a file from HDFS?

Usage: hadoop fs dfs –rm <path>

[cloudera@quickstart Desktop]$ hadoop fs -rm empty.txt

15. Delete a directory completely in HDFS?

Usage: hadoop fs -rm -r <path>

[cloudera@quickstart Desktop]$ hadoop fs -rm –r kriti2018

16. Move a file into a directory in HDFS?

Usage: hadoop fs -mv <src> <dest>

[cloudera@quickstart Desktop]$ hadoop fs –mv emptyfile.txt kriti2018

17. Find a help for an individual command?

Usage: hadoop dfs -usage <command>

[cloudera@quickstart Desktop]$ hdfs dfs -usage mkdir

18. Find the help for a given or all commands?

Command: hadoop fs -help

19. Check the memory status?

Usage: hadoop fs -df hdfs :/

[cloudera@quickstart Desktop]$ hadoop fs -df

20. Display the last kilobyte of the particular file?

[cloudera@quickstart Desktop]$ hadoop fs -tail /user/cloudera/n.txt

21. Append the contents of a file present in local to a file present in HDFS

[cloudera@quickstart Desktop]$ gedit first.txt

[cloudera@quickstart Desktop]$ gedit second.txt

[cloudera@quickstart Desktop]$ hadoop fs -copyFromLocal second.txt /user/cloudera/

[cloudera@quickstart Desktop]$ hadoop fs -appendToFile /home/cloudera/Desktop/first.txt /user/cloudera/second.txt

[cloudera@quickstart Desktop]$ hadoop fs -cat /user/cloudera/second.txt