HDFS Commands

commonly used commands

mkdir

hadoop fs -mkdir <paths>
Ex- Hadoop fs -mkdir /mpr

List the file

```
hadoop fs -ls <args>
```

View the contents

```
hadoop fs -cat <path[filename]>
```

Put

hadoop fs -put <source:localFile> <destination>

Get

```
hadoop fs -get   <source> <dest:localFileSystem>
```

copy from local

hadoop fs -copyFromLocal <src:localFileSystem>
<dest:Hdfs>

copy to local

hadoop fs -copyToLocal <src:Hdfs> <dest:localFileSystem>

move

```
hadoop fs -mv <src> <dest>
```

remove

hadoop fs -rm <arg>

Other commands

```
# 1. Print the Hadoop version
hadoop version

# 2. List the contents of the root directory in HDFS
#
hadoop fs -ls /

# 3. Report the amount of space used and
# available on currently mounted filesystem
#
hadoop dfsadmin -report

# 4. Count the number of directories,files and bytes under
# the paths that match the specified file pattern
#

  hadoop dfsadmin -report

# 5. Run a DFS filesystem checking utility
#
hadoop fsck   /

# 6. Run a cluster balancing utility
#
hadoop balancer

# 7. Create a new directory named “hadoop” below the
# /user/training directory in HDFS. Since you’re
# currently logged in with the “training” user ID,
# /user/training is your home directory in HDFS.
#
hadoop fs -mkdir /user/training/hadoop

# 8. Add a sample text file from the local directory
# named “data” to the new directory you created in HDFS
```

```
# during the previous step.
#
hadoop fs -put data/sample.txt /user/training/hadoop

# 9. List the contents of this new directory in HDFS.
#
hadoop fs -ls /user/training/hadoop

# 10. Add the entire local directory called “retail” to the
# /user/training directory in HDFS.
#
hadoop fs -put data/retail /user/training/hadoop

# 11. Since /user/training is your home directory in HDFS,
# any command that does not have an absolute path is
# interpreted as relative to that directory. The next
# command will therefore list your home directory, and
# should show the items you’ve just added there.
#
hadoop fs -ls /

# 12. See how much space this directory occupies in HDFS.
#
hadoop fs -du /hadoop/retail

Displays the number of bytes occupied

# 13. Delete a file ‘customers’ from the “retail” directory.
#
hadoop fs -rm /hadoop/retail/customers

hadoop fs -rmr   /hadoop/retail – for directory removal

# 14. Ensure this file is no longer in HDFS.
#
hadoop fs -ls hadoop/retail/customers

# 15. Delete all files from the “retail” directory using a wildcard.
#
hadoop fs -rm hadoop/retail/*

# 16. To empty the trash
#
hadoop fs -expunge
```

18. List the hadoop directory again
#
hadoop fs -ls /hadoop

19. Add the purchases.txt file from the local directory
named “/home/training/” to the hadoop directory you created in HDFS
#
hadoop fs -copyFromLocal /home/training/purchases.txt hadoop/

20. To view the contents of your text file purchases.txt
which is present in your hadoop directory.
#
hadoop fs -cat hadoop/purchases.txt

21. Add the purchases.txt file from “hadoop” directory which is present in HDFS directory
to the directory “data” which is present in your local directory
#
hadoop fs -copyToLocal hadoop/purchases.txt /home/training/data

22. cp is used to copy files between directories present in HDFS or local
#
hadoop fs -cp /user/training/*.txt /user/training/hadoop

cp

Hadoop fs -cp source destination

hadoop fs -cp hdfs:///wc1 /mpr

Transfer files within hdfs

hadoop fs -cp hdfs:///wco2/part-r-00000 file:///home/mahi/mpr

Transfer files from hdfs to local

23. ‘-get’ command can be used alternaively to ‘-copyToLocal’ command
#
hadoop fs -get hadoop/sample.txt /home/training/

24. Display last kilobyte of the file “purchases.txt” to stdout.
#
hadoop fs -tail hadoop/purchases.txt

25. Move a directory from one location to other
#
hadoop fs -mv hadoop apache_hadoop

26 Command to make the name node leave safe mode
#
sudo -u hdfs dfsadmin -safemode leave

27. List all the hadoop file system shell commands
#
hadoop fs

28. Last but not least, always ask for help!
#
hadoop fs -help

29. hadoop fs touchz

The hadoop touchz command creates a zero byte file. This is similar to the touch command in unix. The syntax is shown below:

hadoop fs -touchz /user/hadoop/filename