CSP554—Big Data Technologies

Assignment #2 (Modules 02a & 02b, 20 points)

9. **(2 points) Execute the following hdfs command to list the files or directories that are listed (also indicating which is a file and which a directory): hadoop fs –ls /**

```
[hadoop@ip-172-31-12-51 /]$ hadoop fs -ls /
Found 4 items
drwxr-xr-x   - hdfs hdfsadmingroup          0 2024-02-01 22:32 /apps
drwxrwxrwt   - hdfs hdfsadmingroup          0 2024-02-01 22:35 /tmp
drwxr-xr-x   - hdfs hdfsadmingroup          0 2024-02-01 23:44 /user
drwxr-xr-x   - hdfs hdfsadmingroup          0 2024-02-01 22:32 /var
```

10. **(2 points) Execute a command (you needed to figure out which one) to list the files and directories under the hdfs directory listed below:**

/user

```
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -ls /user
Found 6 items
drwxrwxrwx   - hadoop hdfsadmingroup          0 2024-02-01 22:32 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2024-02-01 22:32 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2024-02-01 22:32 /user/hive
drwxrwxrwx   - hue    hue                     0 2024-02-01 22:32 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2024-02-01 22:34 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2024-02-01 22:32 /user/root
```

11. **(2 points) Execute a command to create the following HDFS directory:**

/user/csp554

```
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -mkdir /user/csp554
mkdir: `/user/csp554': File exists
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -ls /user
Found 7 items
drwxr-xr-x   - hadoop hdfsadmingroup          0 2024-02-01 23:41 /user/csp554
drwxrwxrwx   - hadoop hdfsadmingroup          0 2024-02-01 22:32 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2024-02-01 22:32 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2024-02-01 22:32 /user/hive
drwxrwxrwx   - hue    hue                     0 2024-02-01 22:32 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2024-02-01 22:34 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2024-02-01 22:32 /user/root
```

12. **(2 points) Execute a command to create the following HDFS directory:**

    /user/csp554-2

```
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -mkdir /user/csp554-2
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -mkdir /user
mkdir: `/user': File exists
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -ls /user/csp554-2
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -ls /user
Found 8 items
drwxr-xr-x   - hadoop hdfsadmingroup          0 2024-02-01 23:41 /user/csp554
drwxr-xr-x   - hadoop hdfsadmingroup          0 2024-02-01 23:44 /user/csp554-2
drwxrwxrwx   - hadoop hdfsadmingroup          0 2024-02-01 22:32 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2024-02-01 22:32 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2024-02-01 22:32 /user/hive
drwxrwxrwx   - hue    hue                     0 2024-02-01 22:32 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2024-02-01 22:34 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2024-02-01 22:32 /user/root
```

13. **(2 points) Execute a command that copies a given local file to the given hdfs directory :**

Source local file: /home/hadoop/myname.txt (where the actual name is your name as described above)

Destination HDFS directory: /user/csp554

```
[hadoop@ip-172-31-12-51 /]$ hadoop fs -copyFromLocal /home/hadoop/prachi.txt /user/csp554
```

```
Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup         23 2024-02-02 00:40 /user/csp554-2/prachi.txt
```

14. **(2 points) Copy a file from one hdfs directory to another hdfs directory and write down the command**

Source hdfs file: /user/csp554/myname.txt (where the actual name is your name as described above)

Destination HDFS directory: /user/csp554-2

```
[hadoop@ip-172-31-12-51 /]$ hadoop fs -cp /user/csp554-2/prachi.txt /user/csp554-2
```

15. **(2 points) Copy the object myid.txt you uploaded to an S3 bucket into the Hadoop master node Linux file system. The actual object includes your student id as above.**

```
[hadoop@ip-172-31-12-51 /]$ aws s3 cp s3://assign2prachia20549927/A20549927.txt   /home/hadoop/A20549927.txt
download: s3://assign2prachia20549927/A20549927.txt to home/hadoop/A20549927.txt
[hadoop@ip-172-31-12-51 /]$ ls /home/hadoop
A20549927.txt  prachi.txt
```

16. **(2 points) Copy the same object myid.txt you created in an S3 bucket into HDFS into the directory /users/csp554**

```
[hadoop@ip-172-31-12-51 /]$ hadoop fs -cp s3://assign2prachia20549927/A20549927.txt   hdfs:///user/csp554-2
2024-02-02 00:51:10,117 INFO s3n.S3NativeFileSystem: Opening 's3://assign2prachia20549927/A20549927.txt' for reading
[hadoop@ip-172-31-12-51 /]$ hadoop fs -ls /user/csp554-2
Found 2 items
-rw-r--r--   1 hadoop hdfsadmingroup         19 2024-02-02 00:51 /user/csp554-2/A20549927.txt
-rw-r--r--   1 hadoop hdfsadmingroup         23 2024-02-02 00:40 /user/csp554-2/prachi.txt
```

17. **(2 points) Execute a command to show the contents of the myid.txt file in the hdfs directory /user/csp554-2**

```
[hadoop@ip-172-31-12-51 ~]$ ls
A20549927.txt  prachi.txt
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -cat /user/csp554-2/A20549927.txt
this is the id file[hadoop@icd home^C
```

18. (2 points) Execute a command to remove the myid.txt file in the hdfs directory /user/csp554-2

```
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -rm /user/csp554-2/A20549927.txt
Deleted /user/csp554-2/A20549927.txt
[hadoop@ip-172-31-12-51 ~]$ hadoop fs -ls /user/csp554-2
Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup         23 2024-02-02 00:40 /user/csp554-2/prachi.txt
```

19. This might be very important to your wallet. Bucket deleting and termination section.