CSP554—Big Data Technologies

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

1. (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 /

   Take a screen snapshot of names of the files or directories that are listed and include it in your assignment submission.
   Command: hadoop fs -ls /

   ```
   [hadoop@ip-172-31-14-138 ~]$ ls
   yashpatel.txt
   [hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /
   Found 4 items
   drwxr-xr-x   - hdfs hadoop          0 2021-02-04 04:03 /apps
   drwxrwxrwt   - hdfs hadoop          0 2021-02-04 04:05 /tmp
   drwxr-xr-x   - hdfs hadoop          0 2021-02-04 04:03 /user
   drwxr-xr-x   - hdfs hadoop          0 2021-02-04 04:03 /var
   [hadoop@ip-172-31-14-138 ~]$
   ```

2. (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

   Write down the command you executed and also take a screen snapshot of names of the files or directories that are listed and include it in your assignment submission.
   Command: hadoop fs -ls /user

   ```
   [hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /user
   Found 6 items
   drwxrwxrwx   - hadoop hadoop        0 2021-02-04 04:03 /user/hadoop
   drwxr-xr-x   - mapred mapred        0 2021-02-04 04:03 /user/history
   drwxrwxrwx   - hdfs   hadoop        0 2021-02-04 04:03 /user/hive
   drwxrwxrwx   - hue    hue           0 2021-02-04 04:03 /user/hue
   drwxrwxrwx   - oozie  oozie         0 2021-02-04 04:05 /user/oozie
   drwxrwxrwx   - root   hadoop        0 2021-02-04 04:03 /user/root
   [hadoop@ip-172-31-14-138 ~]$ |
   ```

3. (2 points) Execute a command to create the following HDFS directory:

   /user/csp554

   Record the command you executed and include it in your assignment submission.

Command: hadoop fs -mkdir /user/csp554

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -mkdir /user/csp554
[hadoop@ip-172-31-14-138 ~]$ |
```

4. (2 points) Execute a command to create the following HDFS directory:

    /user/csp554-2

    Record the command you executed and include it in your assignment submission.
    Command: hadoop fs -mkdir /user/csp554-2

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -mkdir /user/csp554-2
[hadoop@ip-172-31-14-138 ~]$ |
```

5. (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

    Record the command you executed and include it in your assignment submission.
    Command1: hadoop fs -put /home/hadoop/yashpatel.txt /user/csp554
    Command2: hadoop fs -ls /user/csp554

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -put /home/hadoop/yashpatel.txt /user/csp554
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /user/csp554
Found 1 items
-rw-r--r--   1 hadoop hadoop         23 2021-02-04 04:59 /user/csp554/yashpatel.txt
[hadoop@ip-172-31-14-138 ~]$ |
```

6. (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

    Record the command you executed and include it in your assignment submission.
    Command1: hadoop fs -cp /user/csp554/yashpatel.txt /user/csp554-2
    Command2: hadoop fs -ls /user/csp554-2

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -cp /user/csp554/yashpatel.txt /user/csp554-2
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /user/csp554-2
Found 1 items
-rw-r--r--   1 hadoop hadoop         23 2021-02-04 05:07 /user/csp554-2/yashpatel.txt
[hadoop@ip-172-31-14-138 ~]$ |
```

7. (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.

The way you do this would be as follows to copy an object from an S3 bucket to the Linux file system of the Hadoop master node.

aws s3 cp s3://mybucket/myid.txt   /home/hadoop/myid.txt

After you executed the above command perform an "ls /home/hadoop" and take a screen snapshot of names of the files or directories that are listed and include it in your assignment submission.
Command1: aws s3 cp s3://a20451170-csp554-bdt-a2/A20451170.txt
/home/hadoop/A20451170.txt
Command2: ls/home/hadoop

```
[hadoop@ip-172-31-14-138 ~]$ aws s3 cp s3://a20451170-csp554-bdt-a2/A20451170.txt /home/hadoop/A20451170.txt
download: s3://a20451170-csp554-bdt-a2/A20451170.txt to ./A20451170.txt
[hadoop@ip-172-31-14-138 ~]$ ls /home/hadoop
A20451170.txt   yashpatel.txt
[hadoop@ip-172-31-14-138 ~]$
```

8. (2 points) Copy the same object myid.txt you created in an S3 bucket into HDFS into the directory /users/csp554

hadoop fs -cp s3://mybucket/myid.txt   hdfs:///user/csp554-2
Note, the three slashes after the "hdfs:"

After you executed the above command, execute another command (you needed to figure out which one) to list the files and directories under the hdfs directory listed below:

/user/csp554-2

Write down the command you executed and also take a screen snapshot of names of the files or directories that are listed and include it in your assignment submission.
Command1: hadoop fs -cp s3://a20451170-csp554-bdt-a2/A20451170.txt
hdfs:///user/csp554-2
Command2: hadoop fs -ls /user/csp554-2

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -cp s3://a20451170-csp554-bdt-a2/A20451170.txt hdfs:///user/csp554-2
21/02/04 05:21:19 INFO s3n.S3NativeFileSystem: Opening 's3://a20451170-csp554-bdt-a2/A20451170.txt' for reading
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /user/csp554-2
Found 2 items
-rw-r--r--   1 hadoop hadoop         21 2021-02-04 05:21 /user/csp554-2/A20451170.txt
-rw-r--r--   1 hadoop hadoop         23 2021-02-04 05:07 /user/csp554-2/yashpatel.txt
[hadoop@ip-172-31-14-138 ~]$
```

9. (2 points) Execute a command to show the contents of the myid.txt file in the hdfs directory /user/csp554-2

Clue: look up about how to use the "cat" command in the file system shell document.

Write down the command you executed and also take a screen snapshot of the listed content of the file and include it in your assignment submission.
Command: hadoop fs -cat /user/csp554-2/A20451170.txt

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -cat /user/csp554-2/A20451170.txt
this is the id file
[hadoop@ip-172-31-14-138 ~]$
```

10. (2 points) Execute a command to remove the myid.txt file in the hdfs directory /user/csp554-2

    Clue: look up about how to use the "rm" command in the file system shell document.

    Write down the command you executed, then list the content of the /user/csp554-2 HDFS
    directory and take a screen snapshot of the listed content of the directory and include it in your
    assignment submission.
    Command1: hadoop fs -rm /user/csp554-2/A20451170.txt
    Command2: hadoop fs -ls /user/csp554-2

```
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -rm /user/csp554-2/A20451170.txt
Deleted /user/csp554-2/A20451170.txt
[hadoop@ip-172-31-14-138 ~]$ hadoop fs -ls /user/csp554-2
Found 1 items
-rw-r--r--   1 hadoop hadoop         23 2021-02-04 05:07 /user/csp554-2/yashpatel.txt
[hadoop@ip-172-31-14-138 ~]$
```