- SSH command to connect to Hadoop EMR-
  ssh -i /c/Users/Ankan\ Mazumdar/Downloads/emr-key-pair.pem
  hadoop@ec2-18-220-141-223.us-east-2.compute.amazonaws.com

```
Ankan Mazumdar@DESKTOP-CMULEBA MINGW64 ~/Downloads
$ ssh -i emr_key_pair.pem hadoop@ec2-18-220-141-223.us-east-2.compute.amazonaws.com
Last login: Fri Sep 15 01:21:45 2023

       __|  __|_  )
       _|  (     /    Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
23 package(s) needed for security, out of 57 available
Run "sudo yum update" to apply all updates.

EEEEEEEEEEEEEEEEEEEE MMMMMMMM           MMMMMMMM RRRRRRRRRRRRRRRR
E::::::::::::::::::::E M:::::::M         M:::::::M R:::::::::::::::R
EE:::::EEEEEEEEE:::::E M:::::::::M       M:::::::M R:::::RRRRRR:::::R
  E::::E       EEEEE M::::::::::M       M:::::::::M RR::::R      R::::R
  E::::E             M:::::M:::M       M:::M:::::M   R:::R      R::::R
  E::::EEEEEEEEEE     M:::::M M:::M   M:::M M:::::M   R:::RRRRRR:::::R
  E::::::::::::::E     M:::::M  M:::M:::M  M:::::M   R:::::::::::RR
  E::::EEEEEEEEEE     M:::::M   M:::::M   M:::::M   R:::RRRRRR:::::R
  E::::E             M:::::M    M:::M    M:::::M   R:::R      R::::R
  E::::E       EEEEE M:::::M     MMM     M:::::M   R:::R      R::::R
EE:::::EEEEEEEEE::::E M:::::M             M:::::M   R:::R      R::::R
E::::::::::::::::::::E M:::::M             M:::::M RR::::R      R::::R
EEEEEEEEEEEEEEEEEEEE MMMMMMM             MMMMMMM RRRRRRR        RRRRRR
```
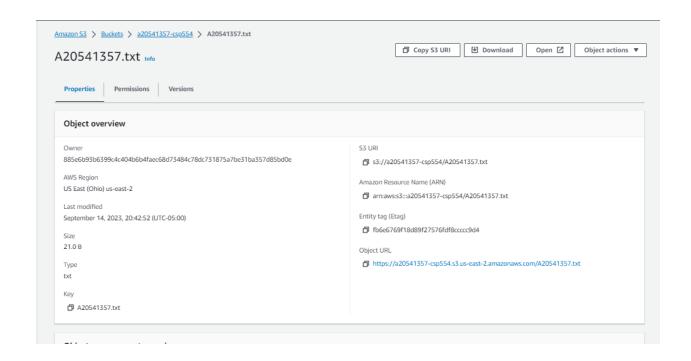
- SCP file from local PC to the Hadoop master node-

  Ankan Mazumdar@DESKTOP-CMULEBA MINGW64 ~/Downloads
  scp -i emr-key-pair.pem ankanmazumdar.txt
  hadoop@ec2-18-220-141-223.us-east-2.compute.amazonaws.com:/home/hadoop

```
Ankan Mazumdar@DESKTOP-CMULEBA MINGW64 ~/Downloads
$ scp -i emr_key_pair.pem ankanmazumdar.txt hadoop@ec2-18-220-141-223.us-east-2.compute.amazonaws.com:/home/hadoop/
ankanmazumdar.txt                                                                          100%   21     0.7KB/s   00:00
```

- Upload "myid.txt" in S3 bucket-

A20541357.txt Info

| Copy S3 URI | Download | Open ☑ | Object actions ▼ |

Properties | Permissions | Versions

Object overview

Owner
885e6b93b6399c4c404b6b4faec68d73484c78dc731875a7be31ba357d85bd0e

AWS Region
US East (Ohio) us-east-2

Last modified
September 14, 2023, 20:42:52 (UTC-05:00)

Size
21.0 B

Type
txt

Key
A20541357.txt

S3 URI
s3://a20541357-csp554/A20541357.txt

Amazon Resource Name (ARN)
arn:aws:s3:::a20541357-csp554/A20541357.txt

Entity tag (Etag)
fb6e6769f18d89f27576fdf8ccccc9d4

Object URL
https://a20541357-csp554.s3.us-east-2.amazonaws.com/A20541357.txt

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 dfs -ls /

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 5 items
drwxr-xr-x   - hdfs    hdfsadmingroup          0 2023-09-14 23:05 /apps
drwxr-xr-x   - hadoop  hdfsadmingroup          0 2023-09-15 01:45 /test
drwxrwxrwt   - hdfs    hdfsadmingroup          0 2023-09-14 23:09 /tmp
drwxr-xr-x   - hdfs    hdfsadmingroup          0 2023-09-14 23:05 /user
drwxr-xr-x   - hdfs    hdfsadmingroup          0 2023-09-14 23:05 /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 dfs -ls /user

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 6 items
drwxrwxrwx   - hadoop hdfsadmingroup          0 2023-09-14 23:05 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2023-09-14 23:05 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2023-09-14 23:05 /user/hive
drwxrwxrwx   - hue    hue                     0 2023-09-14 23:05 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2023-09-14 23:09 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2023-09-14 23:05 /user/root
```

11. (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.

hadoop dfs -mkdir -p /user/csp554

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -mkdir -p /user/csp554/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/csp554/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 7 items
drwxr-xr-x   - hadoop hdfsadmingroup          0 2023-09-15 02:23 /user/csp554
drwxrwxrwx   - hadoop hdfsadmingroup          0 2023-09-14 23:05 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2023-09-14 23:05 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2023-09-14 23:05 /user/hive
drwxrwxrwx   - hue    hue                     0 2023-09-14 23:05 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2023-09-14 23:09 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2023-09-14 23:05 /user/root
```

12. (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.

hadoop dfs -mkdir -p /user/csp554-2/

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -mkdir -p /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 8 items
drwxr-xr-x   - hadoop hdfsadmingroup          0 2023-09-15 02:23 /user/csp554
drwxr-xr-x   - hadoop hdfsadmingroup          0 2023-09-15 02:24 /user/csp554-2
drwxrwxrwx   - hadoop hdfsadmingroup          0 2023-09-14 23:05 /user/hadoop
drwxr-xr-x   - mapred mapred                  0 2023-09-14 23:05 /user/history
drwxrwxrwx   - hdfs   hdfsadmingroup          0 2023-09-14 23:05 /user/hive
drwxrwxrwx   - hue    hue                     0 2023-09-14 23:05 /user/hue
drwxrwxrwx   - oozie  oozie                   0 2023-09-14 23:09 /user/oozie
drwxrwxrwx   - root   hdfsadmingroup          0 2023-09-14 23:05 /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 Record the command you executed and include it in your assignment submission.

hadoop dfs -copyFromLocal -f   /home/hadoop/ankanmazumdar.txt /user/csp554/

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -copyFromLocal -f /home/hadoop/ankanmazumdar.txt /user/csp554/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/csp554/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup         21 2023-09-15 02:35 /user/csp554/ankanmazumdar.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 Record the command you executed and include it in your assignment submission.

hadoop dfs -cp /user/csp554/ankanmazumdar.txt /user/csp554-2/

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -cp /user/csp554/ankanmazumdar.txt /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup         21 2023-09-15 02:36 /user/csp554-2/ankanmazumdar.txt
```

15. (2 points) Copy the object myid.txt you uploaded to an S3 bucket into the Hadoop master node Linux file system.

aws s3 cp s3://a20541357-csp554/A20541357.txt /home/hadoop/A20541357.txt

```
[hadoop@ip-172-31-35-150 ~]$ aws s3 cp help
[hadoop@ip-172-31-35-150 ~]$ aws s3 cp s3://a20541357-csp554/A20541357.txt /home/hadoop/A20541357.txt
download: s3://a20541357-csp554/A20541357.txt to ./A20541357.txt
[hadoop@ip-172-31-35-150 ~]$ ls /home/hadoop
A20541357.txt  ankanmazumdar.txt
[hadoop@ip-172-31-35-150 ~]$ ls -l  /home/hadoop
total 8
-rw-rw-r-- 1 hadoop hadoop 21 Sep 15 01:42 A20541357.txt
-rw-r--r-- 1 hadoop hadoop 21 Sep 15 01:34 ankanmazumdar.txt
```

16. (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.

hadoop fs -cp s3://a20541357-csp554/A20541357.txt hdfs:///user/csp554-2/

```
[hadoop@ip-172-31-35-150 ~]$ hadoop fs -cp s3://a20541357-csp554/A20541357.txt hdfs:///user/csp554-2/
2023-09-15 02:45:20,699 INFO s3n.S3NativeFileSystem: Opening 's3://a20541357-csp554/A20541357.txt' for reading
[hadoop@ip-172-31-35-150 ~]$  hadoop dfs -ls /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 2 items
-rw-r--r--   1 hadoop hdfsadmingroup         21 2023-09-15 02:45 /user/csp554-2/A20541357.txt
-rw-r--r--   1 hadoop hdfsadmingroup         21 2023-09-15 02:36 /user/csp554-2/ankanmazumdar.txt
```

17. (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.

hadoop dfs -cat /user/csp554-2/A20541357.txt

```
[hadoop@ip-172-31-35-150 ~]$  hadoop dfs -cat /user/csp554-2/A20541357.txt
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

this is the name file[hadoop@ip-172-31-35-150 ~]$
[hadoop@ip-172-31-35-150 ~]$
```

18. (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.

hadoop dfs -rm -f /user/csp554-2/A20541357.txt

```
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 2 items
-rw-r--r--   1 hadoop hdfsadmingroup          21 2023-09-15 02:45 /user/csp554-2/A20541357.txt
-rw-r--r--   1 hadoop hdfsadmingroup          21 2023-09-15 02:36 /user/csp554-2/ankanmazumdar.txt
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -rm -f /user/csp554-2/A20541357.txt
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Deleted /user/csp554-2/A20541357.txt
[hadoop@ip-172-31-35-150 ~]$ hadoop dfs -ls /user/csp554-2/
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

Found 1 items
-rw-r--r--   1 hadoop hdfsadmingroup          21 2023-09-15 02:36 /user/csp554-2/ankanmazumdar.txt
[hadoop@ip-172-31-35-150 ~]$
```