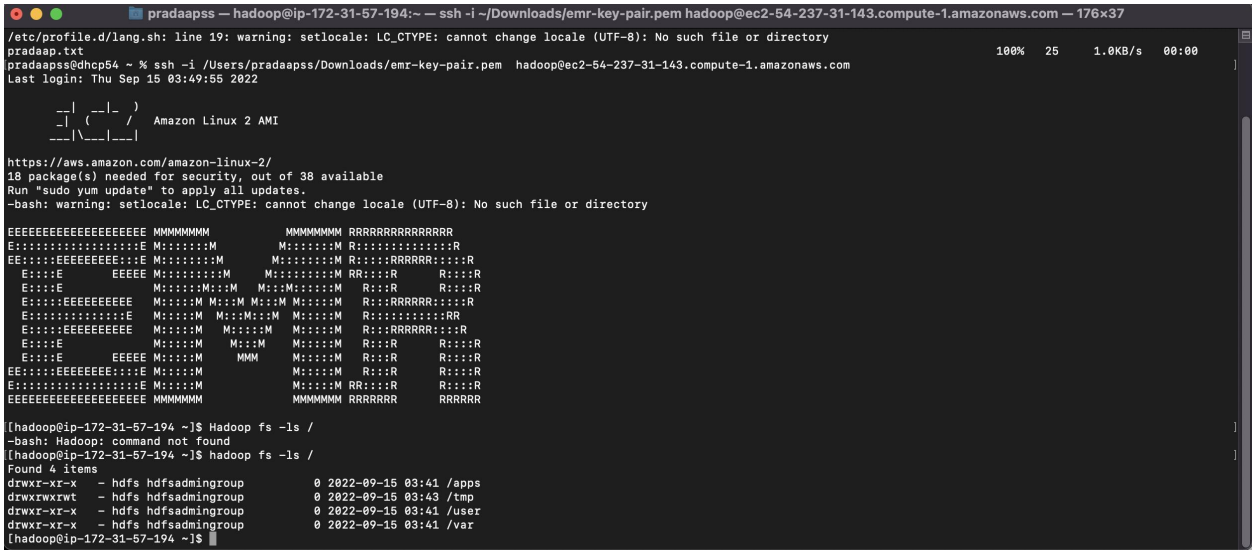
Assignment #2 (Modules 02a & 02b, 20 points) A20512400

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):

Take a screen snapshot of names of the files or directories that are listed and include it in your assignment submission.

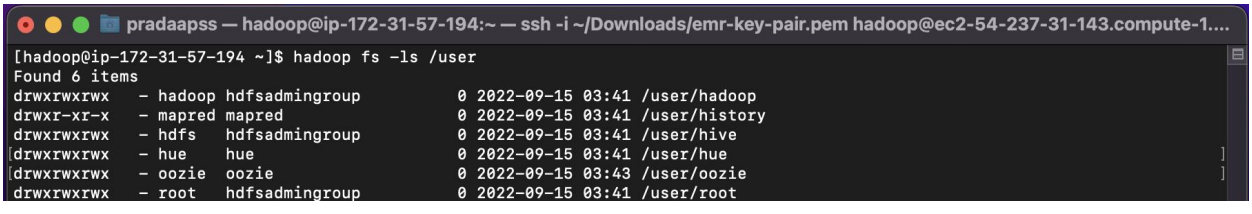
hadoop fs –ls /



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:

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.

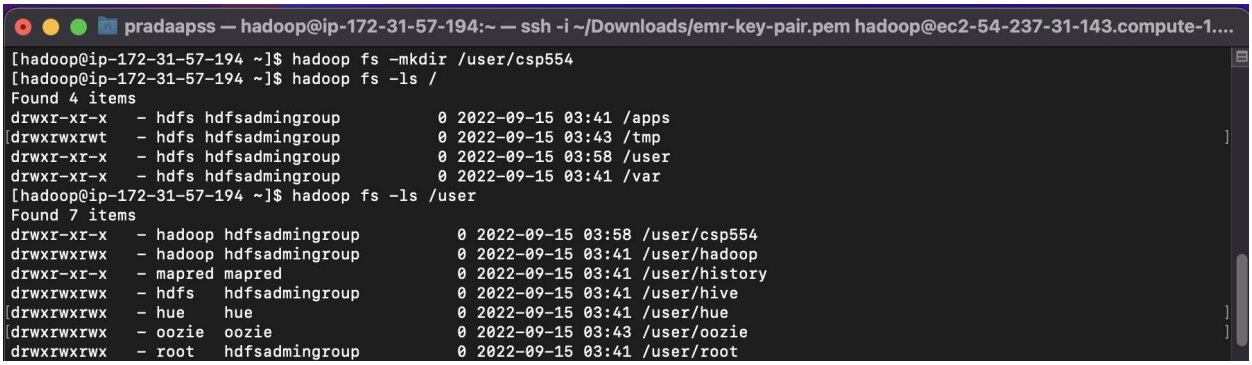
hadoop fs -ls /user



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

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

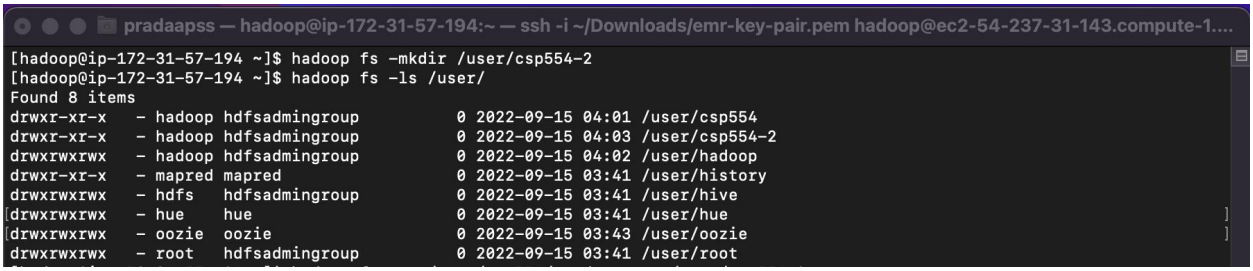
hadoop fs -mkdir /user/csp554



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

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

hadoop fs -mkdir /user/csp554-2



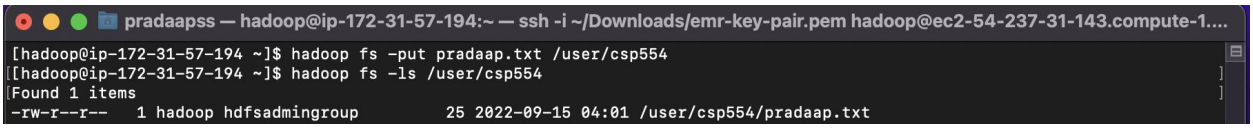
13. (2 points) Execute a command that copies a given local file to the given hdfs directory :

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 fs -put pradaap.txt /user/csp554



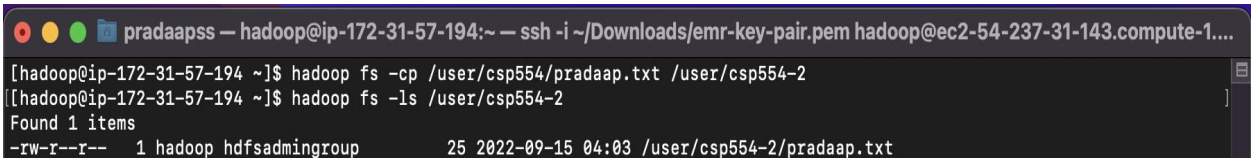
14. (2 points) Copy a file from one hdfs directory to another hdfs directory and write down the command

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 fs -cp /user/csp554/pradaap.txt /user/csp554-2

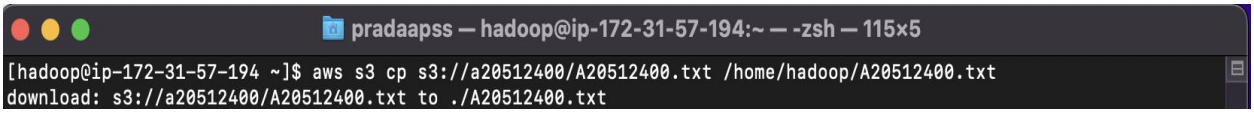


15. (2 points) Copy the object myid.txt you uploaded to an S3 bucket into the Hadoop master node Linux file system. The actual object includes your student id as above.

Note, Amazon EMR and Hadoop provide a variety of file systems that you can use with EMR. You specify which file system to use with a file system prefix. For example, s3://myawsbucket/path references an Amazon S3 bucket using EMRFS (EMR file system). See: https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-file-systems.html

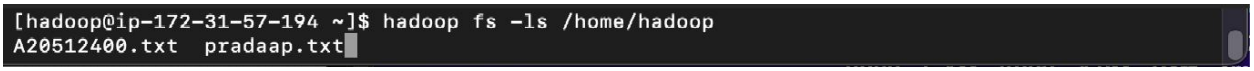
The way you do this would be as follows to copy an object from an S3 bucket to the Linux file system of the Hadoop master node.

aws s3 cp s3://a20512400/A20512400.txt /home/hadoop/A20512400.txt



The above is an AWS CLI (command line interpreter) command. For more information about how to use the CLI to manipulate S3 buckets see:
https://docs.aws.amazon.com/cli/latest/reference/s3/index.html

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.



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

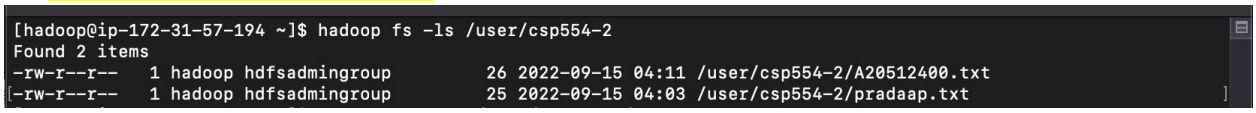
    hadoop fs -cp s3://mybucket/myid.txt hdfs:///user/csp554-2
    Note, the three slashes after the “hdfs:”

    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:

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://a20512400/A20512400.txt hdfs:///user/csp554-2

hadoop fs -ls /user/csp554-2

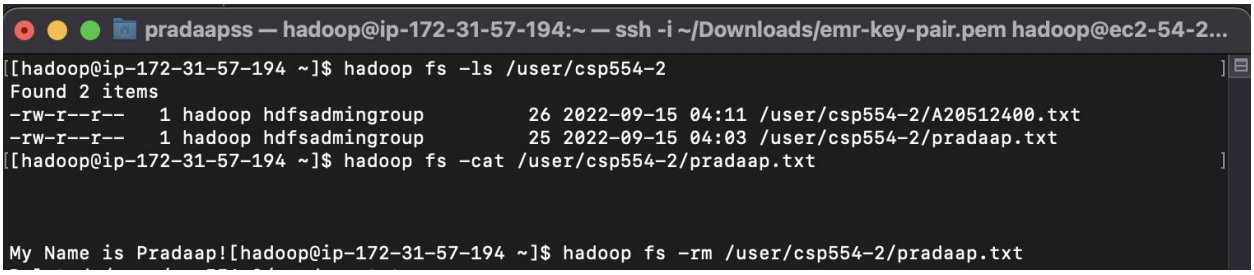


17. (2 points) Execute a command to show the contents of the myid.txt file in the hdfs directory /user/csp554-2

    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 fs -ls /user/csp554-2



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 fs -rm /user/csp554-2/pradaap.txt