

Homework Assignment 1

1. Work *Example 3-3* on CSUEB Hadoop. Type out all the commands in each step of the process and print out a screenshot of the final results in CSUEB Hadoop.
2. Work *Example 3-4* on CSUEB Hadoop. Type out all the commands in each step of the process and print out a screenshot of the final results (showing the file directory) in CSUEB Hadoop.

Example 3-3. Displaying files from a Hadoop filesystem on standard output twice.

Run on Local disk

```
scp "/Users/amishaferhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileSystemDoubleCat.java"  
student39@134.154.190.204:/home/student39/
```

Command for copying a file from local disk to HDFS:

```
hdfs dfs -copyFromLocal "/home/student39/ FileSystemDoubleCat.java" /home/39student39/
```

Sets the HADOOP_CLASSPATH environment variable to /home/student39/ so Hadoop can find additional Java classes and JAR files.

```
export HADOOP_CLASSPATH=/home/student39
```

Command for compiling Java file:

```
javac -classpath /home/student9/hadoop-common-2.6.1.jar:/home/student9/hadoop-mapreduce-client-core-2.6.1.jar:/home/student9/commons-cli-2.0.jar -d /home/student9/ FileSystemDoubleCat.java
```

Command for running Java in Hadoop

```
hadoop FileSystemDoubleCat hdfs:///home/39student39/input39/quangle.txt
```

Local Disk:

```
Last login: Fri Feb 21 19:15:32 on ttys001
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % scp "/Users/amishafarhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileSystemDoubleCat.java" student39@134.154.190.204:/home/student39/
```

```
student39@134.154.190.204's password:
[FileSystemDoubleCat.java          100% 910   32.4KB/s  00:00
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ %
```

Hadoop Server:

```
zsh: command not found: hadoop
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % ssh student39@134.154.190.204
```

```
student39@134.154.190.204's password:
Last login: Fri Feb 21 13:22:58 2025 from 134.154.76.107
[student39@msba-hadoop-name ~]$ hadoop version
Hadoop 2.9.0
Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r 756ebc8394e473ac25feac05fa493f6d612e6c50
Compiled by aruresh on 2017-11-13T23:15Z
Compiled with protoc 2.5.0
From source with checksum 0a76a9a32a5257331741f8d5932f183
This command was run using /usr/local/hadoop-2.9.0/share/hadoop/common/hadoop-common-2.9.0.jar
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/FileSystemDoubleCat.java" /home/39student39
[student39@msba-hadoop-name ~]$ export HADOOP_CLASSPATH=/home/student39
[student39@msba-hadoop-name ~]$ hadoop FileSystemDoubleCat hdfs:///home/39student39/input39/quangle.txt
Error: Could not find or load main class FileSystemDoubleCat
[student39@msba-hadoop-name ~]$ javac -classpath "/home/student39/hadoop-common-2.6.1.jar:/home/student39/hadoop-mapreduce-client-core-2.6.1.jar:/home/student39/commons-cli-2.0.jar" -d /home/student39/ /home/student39/FileSystemDoubleCat.java
[student39@msba-hadoop-name ~]$ 
[student39@msba-hadoop-name ~]$ hadoop FileSystemDoubleCat hdfs:///home/39student39/input39/quangle.txt
On the top of the Crumpetty Tree
The Quangle Wangle sat,
But his face you could not see,
On account of his Beaver Hat,
On the top of the Crumpetty Tree
The Quangle Wangle sat,
But his face you could not see,
On account of his Beaver Hat.
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" /home/39student39/
```

Example 3. shows how to copy a local file to a Hadoop filesystem illustrating progress.

Run on Local disk

```
scp "/Users/amishafarhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileCopyWithProgress.java"  
student39@134.154.190.204:/home/student39/
```

Command for copying a file from local disk to HDFS:

```
hdfs dfs -copyFromLocal "/home/student39/FileCopyWithProgress.java" /home/39student39/
```

Command for compiling Java file:

```
javac -classpath /home/student39/hadoop-common-2.6.1.jar:/home/student39/hadoop-mapreduce-client-core-2.6.1.jar:/home/student39/commons-cli-2.0.jar -d /home/student39/ FileCopyWithProgress.java
```

Sets the HADOOP_CLASSPATH environment variable to /home/student39/ so Hadoop can find additional Java classes and JAR files.

```
export HADOOP_CLASSPATH=/home/student39
```

Command for running Java in Hadoop

```
hadoop FileCopyWithProgress /home/student39/1400-8.txt  
hdfs://home/39student39/input39/1400-8.txt
```

Local Disk:

```
Last login: Fri Feb 21 19:15:32 on ttys001
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % scp */Users/amishafarhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileSystemDoubleCat.java student39@134.154.190.204:/home
/base/student39/

student39@134.154.190.204's password:
FileSystemDoubleCat.java          100%  910   32.4KB/s  00:00
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % scp */Users/amishafarhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileCopyWithProgress.java student39@134.154.190.204:/home
/base/student39/
scp: stat local `?/200/234/Users/amishafarhanashaik/Desktop/MSBA/BAN': No such file or directory
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % scp */Users/amishafarhanashaik/Desktop/MSBA/BAN 632 - Big Data/hadoop-book-3e/ch03/src/main/java/FileCopyWithProgress.java student39@134.154.190.204:/home
/base/student39/

student39@134.154.190.204's password:
FileCopyWithProgress.java          100% 1039   38.3KB/s  00:00
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ %
```

Hadoop Server:

```
Last login: Fri Feb 21 18:58:01 on ttys001
(base) amishafarhanashaik@Amishas-MacBook-Air-3 ~ % ssh student39@134.154.190.204

student39@134.154.190.204's password:
Last login: Fri Feb 21 18:56:04 2025 from 134.154.76.107
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" "/home/39student39/
copyFromLocal: unexpected URISyntaxException
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" "/user/student39/
copyFromLocal: '/user/student39/': No such file or directory: 'hdfs://msba-hadoop-name:9000/user/student39'
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" "/user/student39/
copyFromLocal: '/user/student39/': No such file or directory: 'hdfs://msba-hadoop-name:9000/user/student39'
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" "/user/student39/
copyFromLocal: '/user/student39/': No such file or directory: 'hdfs://msba-hadoop-name:9000/user/student39'
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/1400-8.txt" "/home/39student39/
copyFromLocal: '/home/39student39/1400-8.txt': File exists
[student39@msba-hadoop-name ~]$ hdfs dfs -copyFromLocal "/home/student39/FileCopyWithProgress.java" "/home/39student39/
[student39@msba-hadoop-name ~]$ javac -classpath /home/student39/hadoop-common-2.6.1.jar:/home/student39/hadoop-mapreduce-client-core-2.6.1.jar:/home/student39/commons-cli-2.0.jar -d /home/student39/ FileCopyWithProgress.java
[student39@msba-hadoop-name ~]$ export HADOOP_CLASSPATH=/home/student39
[student39@msba-hadoop-name ~]$ hadoop FileCopyWithProgress /home/student39/1400-8.txt hdfs://home/39student39/input39/1400-8.txt
[student39@msba-hadoop-name ~]$ hadoop FileCopyWithProgress /home/student39/1400-8.txt hdfs://home/39student39/input39/1400-8.txt
[student39@msba-hadoop-name ~]$ hadoop FileCopyWithProgress /home/student39/1400-8.txt hdfs://home/39student39/input39/1400-8.txt
.....[student39@msba-hadoop-name ~]$
```