

CPR E 419 - Lab 1

Experiment 1:

1. Created a directory called /lab1 under HDFS.
For this, I issued the command, `hadoop fs -mkdir /lab1`.
This would create the directory lab1 under /.

Output:

```
cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -mkdir /lab1
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$

cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -ls /lab1
Found 3 items
-rw-r--r-- 1 cpre419 supergroup 64 2021-02-17 14:28 /lab1/afile.txt
drwxr-xr-x - cpre419 supergroup 0 2021-02-18 00:02 /lab1/experiment2
-rw-r--r-- 1 cpre419 supergroup 50 2021-02-17 22:45 /lab1/newfile
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
```

2. I created a file called afile.txt. And moved it to the Hadoop FS using the following:

```
cpre419@cpre419-VirtualBox:/$ cd $HADOOP_HOME/sbin
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -copyFromLocal /home/cpre419/Downloads/afile.txt /lab1
2021-02-17 14:28:53,109 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
```

The output of the file under /lab1 called afile.txt which I filled with random text is as follows:

```
cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -cat /lab1/afile.txt
2021-02-17 14:38:27,327 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
This is the first file I am writing to test Hadoop FS commands.
```

3. I created a new Maven project, and followed the instructions as specified in the lab to generate a .jar package. The Java file would write a new file called newfile under /lab1.

I got the result by running

`hadoop jar ".jar file location" "Java class name under file"`

Command run as follows:

```

cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop jar /home/cpre419/Downloads/exp3.jar lab1
2021-02-17 22:45:51,005 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$

```

```

cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -cat /lab1/newfile
2021-02-17 22:51:48,777 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
The first Hadoop Program!cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$

```

Experiment 2:

1. Created a new directory called Experiment2 under /lab1. Downloaded the bigdata file into Downloads. Copied the bigdata file to the Hadoop file system, using the following commands.

```

cpre419@cpre419-VirtualBox:~/hadoop/sbin$
cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -mkdir /lab1/experiment2
cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop fs -copyFromLocal /home/cpre419/Downloads/bigdata /lab1/experiment2
2021-02-18 00:01:57,382 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:01:59,253 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:00,957 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:02,568 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:04,179 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:05,641 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:06,877 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:08,173 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:09,339 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:10,529 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:11,670 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:12,834 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:14,353 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:15,721 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2021-02-18 00:02:17,026 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
cpre419@cpre419-VirtualBox:~/hadoop/sbin$

```

2. Created a Maven project following the instructions specified in the Lab tutorial. Attaching the project in submissions, for reference.
3. Created a Java class under the default package called FindChecksum that computes the checksum for the specified range.
4. Exported the packer to a .jar under Downloads. Ran the .jar file using hadoop jar command as displayed below to get the results.

```

cpre419@cpre419-VirtualBox:~/hadoop/sbin$ hadoop jar /home/cpre419/Downloads/chksum.jar FindChecksum
2021-02-18 00:37:04,964 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
The checksum obtained is : -4
cpre419@cpre419-VirtualBox:~/hadoop/sbin$

```

In the submissions, attaching the Java code for experiment 2, and also the Project folder created for experiment 2. For running code in the first experiment, used the same Java code that was provided in the Lab Manual.