

Sean Hinchee
2/15/19

Lab 1

Exercise 1	1
Contents of afile.txt:	1
WinSCP:	1
scp(1) from hpc-class to hadoop:	1
Observing the results of the upload:	2
Observing the contents of the file created by HDFSWrite.jar:	2
Exercise 2	2
Refined source for Lab1Ex2.java:	2
Source of Makefile:	3
Building, and output of Lab1Ex2.jar:	4

Exercise 1

The file afile.txt was created on the Windows lab machine, copied to hpc-class, then copied to the hadoop machine. The contents of afile.txt and the operations performed are provided below.

The HDFSWrite.java program opens the designed file for writing, then writes the designated string to the file.

Contents of afile.txt:

this is experiment 1's afile

WinSCP:

X:\se419\lab1\					/home/seh/**				
Name	Size	Type	Changed		Name	Size	Changed	Rights	Ow
..		Parent directory	2/8/2019 11:40:57 AM		..		2/8/2019 12:22:54 PM	rwxt--xt-x	roo
afile.txt	1 KB	Text Document	2/8/2019 12:24:12 PM		afile.txt	1 KB	2/8/2019 12:24:12 PM	rw-r--r--	seh
ex1ls.PNG	16 KB	PNG File	2/8/2019 12:29:06 PM						
ex1scp.PNG	4 KB	PNG File	2/8/2019 12:29:49 PM						
hadoopjarfiles.zip	4,037 KB	zip Archive	2/8/2019 11:40:34 AM						
Handout-Lab1-Sprin...	1,642 KB	Chrome HTML Do...	2/8/2019 11:40:32 AM						

scp(1) from hpc-class to hadoop:

```
[seh@hpc-class ~]$ scp afile.txt hadoop:~/
Password:
Password:
afile.txt                                     100% 30    1.3KB/s   00:00
[seh@hpc-class ~]$
```

Observing the results of the upload:

```
[seh@hadoop000 ~]$ hdfs dfs -ls /user/seh/lab1/
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hadoop/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Found 1 items
-rw-r--r--  2 seh supergroup      30 2019-02-08 12:27 /user/seh/lab1/afile.txt
[seh@hadoop000 ~]$
```

Observing the contents of the file created by HDFSWrite.jar:

```
[seh@hadoop000 ~]$ hadoop dfs -cat lab1/newfile
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hadoop/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
The first Hadoop Program![seh@hadoop000 ~]$
[seh@hadoop000 ~]$
```

Exercise 2

I went back and edited the source for Lab1Ex2.java so that it would be readable and accurate in its comments. The updated source is provided below.

The Lab1Ex2.java program opens the designated file for reading, then reads the 1000 byte block requested, and finally calculates an XOR checksum of the 1000 byte block, outputting the results.

Refined source for Lab1Ex2.java:

```
import org.apache.hadoop.conf.*;
import org.apache.hadoop.fs.*;

public class Lab1Ex2 {
    public static void main ( String [] args ) throws Exception {
        // The system configuration
        Configuration conf = new Configuration();
```

```

// Get an instance of the Filesystem
FileSystem fs = FileSystem.get(conf);

// Set and build path to file to read from
String path_name = "/cpre419/bigdata";

Path path = new Path(path_name);

// The Output Data Stream to write into
FSDataInputStream file = fs.open(path);

// Read 1000 byte block from file of [1000000000, 1000000999]
byte[] buf = new byte[1000];
file.readFully(1000000000, buf, 0, 1000);

// Calculate XOR checksum
byte xor = 0;
for(int i = 0; i < buf.length; i++)
    xor ^= buf[i];

System.out.println("Checksum: " + xor);

// Close the file and the file system instance
file.close();
fs.close();
}
}

```

Source of Makefile:

```

all:
    # Compile Lab1Ex2.class
    javac Lab1Ex2.java
    # Create Lab1Ex2.jar
    jar -cvf Lab1Ex2.jar Lab1Ex2.class

run:
    # Runs the jar in hadoop
    hadoop jar Lab1Ex2.jar Lab1Ex2

```

Building, and output of Lab1Ex2.jar:

```
[seh@hadoop000 ~]$ cat Lab1Ex2.java
import org.apache.hadoop.conf.*;
import org.apache.hadoop.fs.*;
//import

public class Lab1Ex2 {
    public static void main ( String [] args ) throws Exception {
        // The system configuration
        Configuration conf = new Configuration();
        // Get an instance of the Filesystem
        FileSystem fs = FileSystem.get(conf);

        String path_name = "/cpre419/bigdata";

        Path path = new Path(path_name);

        // The Output Data Stream to write into
        FSDataInputStream file = fs.open(path);

        // seek to first offset
        //file.seek(0);

        byte[] buf = new byte[1000];
        file.readFully(1000000000, buf, 0, 1000);

        // calculate checksum

        byte xor = 0;
        for(int i = 0; i < buf.length; i++)
            xor ^= buf[i];

        System.out.println("Checksum: " + xor);

        // Close the file and the file system instance
        file.close();
        fs.close();
    }
}

[seh@hadoop000 ~]$ make && make run
javac Lab1Ex2.java
jar -cvf Lab1Ex2.jar Lab1Ex2.class
added manifest
adding: Lab1Ex2.class(in = 1378) (out= 790) (deflated 42%)
hadoop jar Lab1Ex2.jar Lab1Ex2
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/opt/hadoop/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/opt/hadoop/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
Checksum: 90
[seh@hadoop000 ~]$
```