Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

mdyaseen@test:~$ su - hdoop
Password:
hdoop@test:~$ cd $HADOOP_HOME
hdoop@test:~/hadoop-3.2.1$ cd ./sbin
hdoop@test:~/hadoop-3.2.1/sbin$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hdoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [test]

2021-05-11 12:48:32,204 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers

hdoop@test:~/hadoop-3.2.1/sbin$ jps
2418 DataNode
2611 SecondaryNameNode
2281 NameNode
3501 Jps
2813 ResourceManager
2943 NodeManager

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -mkdir -p lab6
2021-05-11 12:55:59,418 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -ls
2021-05-11 12:56:08,256 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x   - hdoop supergroup   0 2021-05-11 12:56 lab6
drwxr-xr-x   - hdoop supergroup   0 2021-05-11 06:29 sampleDir

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -put /home/mdyaseen/Desktop/wc.txt lab6/input.txt
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:04:06,952 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:04:08,096 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -cat lab6/input.txt

Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

2021-05-11 13:04:52,787 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:04:53,886 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
Deer Bear River
Car Car River
Deer Car Bear

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -ls lab6
2021-05-11 13:05:08,237 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 hdoop supergroup   44 2021-05-11 13:04 lab6/input.txt

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop jar /home/mdyaseen/Desktop/wordcount.jar wordcount lab6/ /lab6_output
Exception in thread "main" java.lang.ClassNotFoundException: wordcount
        at java.net.URLClassLoader.findClass(URLClassLoader.java:382)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:418)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:351)
        at java.lang.Class.forName0(Native Method)
        at java.lang.Class.forName(Class.java:348)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:316)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:236)

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop jar /home/mdyaseen/Desktop/wordcount.jar wordcount.wordcount lab6/ /lab6_output
Exception in thread "main" java.lang.ClassNotFoundException: wordcount.wordcount
        at java.net.URLClassLoader.findClass(URLClassLoader.java:382)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:418)
        at java.lang.ClassLoader.loadClass(ClassLoader.java:351)
        at java.lang.Class.forName0(Native Method)
        at java.lang.Class.forName(Class.java:348)
        at org.apache.hadoop.util.RunJar.run(RunJar.java:316)
        at org.apache.hadoop.util.RunJar.main(RunJar.java:236)

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop jar /home/mdyaseen/Desktop/wordcount.jar WordCount lab6/ /lab6_output
2021-05-11 13:16:27,874 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:16:28,911 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2021-05-11 13:16:29,553 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_1620717524992_0001

Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

2021-05-11 13:16:29,752 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:29,987 INFO input.FileInputFormat: Total input files to process : 1
2021-05-11 13:16:30,057 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,112 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,145 INFO mapreduce.JobSubmitter: number of splits:1
2021-05-11 13:16:30,436 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,496 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1620717524992_0001
2021-05-11 13:16:30,496 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-05-11 13:16:30,833 INFO conf.Configuration: resource-types.xml not found
2021-05-11 13:16:30,833 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2021-05-11 13:16:31,178 INFO impl.YarnClientImpl: Submitted application application_1620717524992_0001
2021-05-11 13:16:31,317 INFO mapreduce.Job: The url to track the job: http://test:8088/proxy/application_1620717524992_0001/
2021-05-11 13:16:31,318 INFO mapreduce.Job: Running job: job_1620717524992_0001
2021-05-11 13:16:43,642 INFO mapreduce.Job: Job job_1620717524992_0001 running in uber mode : false
2021-05-11 13:16:43,648 INFO mapreduce.Job: map 0% reduce 0%
2021-05-11 13:16:49,771 INFO mapreduce.Job: map 100% reduce 0%
2021-05-11 13:16:55,843 INFO mapreduce.Job: map 100% reduce 100%
2021-05-11 13:16:56,887 INFO mapreduce.Job: Job job_1620717524992_0001 completed successfully
2021-05-11 13:16:57,020 INFO mapreduce.Job: Counters: 54
        File System Counters
                FILE: Number of bytes read=50
                FILE: Number of bytes written=451185
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=156
                HDFS: Number of bytes written=28
                HDFS: Number of read operations=8
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=3750

Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

            Total time spent by all reduces in occupied slots (ms)=3327
            Total time spent by all map tasks (ms)=3750
            Total time spent by all reduce tasks (ms)=3327
            Total vcore-milliseconds taken by all map tasks=3750
            Total vcore-milliseconds taken by all reduce tasks=3327
            Total megabyte-milliseconds taken by all map tasks=3840000
            Total megabyte-milliseconds taken by all reduce tasks=3406848
        Map-Reduce Framework
            Map input records=3
            Map output records=9
            Map output bytes=80
            Map output materialized bytes=50
            Input split bytes=112
            Combine input records=9
            Combine output records=4
            Reduce input groups=4
            Reduce shuffle bytes=50
            Reduce input records=4
            Reduce output records=4
            Spilled Records=8
            Shuffled Maps =1
            Failed Shuffles=0
            Merged Map outputs=1
            GC time elapsed (ms)=120
            CPU time spent (ms)=1090
            Physical memory (bytes) snapshot=346484736
            Virtual memory (bytes) snapshot=4967919616
            Total committed heap usage (bytes)=230821888
            Peak Map Physical memory (bytes)=224583680
            Peak Map Virtual memory (bytes)=2479607808
            Peak Reduce Physical memory (bytes)=121901056
            Peak Reduce Virtual memory (bytes)=2488311808
        Shuffle Errors
            BAD_ID=0
            CONNECTION=0
            IO_ERROR=0
            WRONG_LENGTH=0
            WRONG_MAP=0
            WRONG_REDUCE=0
        File Input Format Counters
            Bytes Read=44
        File Output Format Counters
            Bytes Written=28

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -ls /lab6_output

Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:17:19,627 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 hdoop supergroup   0 2021-05-11 13:16 /lab6_output/_SUCCESS
-rw-r--r--   1 hdoop supergroup   28 2021-05-11 13:16 /lab6_output/part-r-00000

hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -cat /lab6_output/part-r-00000
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:18:08,517 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:18:09,598 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
Bear   2
Car   3
Deer   2
River   2
hdoop@test:~/hadoop-3.2.1/sbin$

```
mdyaseen@test:~/Desktop$ cd
mdyaseen@test:~$ su - hdoop
Password:
hdoop@test:~$ cd $HADOOP_HOME
hdoop@test:~/hadoop-3.2.1$ cd ./sbin
hdoop@test:~/hadoop-3.2.1/sbin$ ./start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hdoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [test]
2021-05-11 12:48:32,204 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Starting resourcemanager
Starting nodemanagers
hdoop@test:~/hadoop-3.2.1/sbin$ jps
2418 DataNode
2611 SecondaryNameNode
2281 NameNode
3501 Jps
2813 ResourceManager
2943 NodeManager
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -mkdir -p lab6
2021-05-11 12:55:59,418 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -ls
2021-05-11 12:56:08,256 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
drwxr-xr-x   - hdoop supergroup        0 2021-05-11 12:56 lab6
drwxr-xr-x   - hdoop supergroup        0 2021-05-11 06:29 sampleDir
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -put /home/mdyaseen/Desktop/wc.txt lab6/input.txt
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:04:06,952 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:04:08,096 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -cat lab6/input.txt
2021-05-11 13:04:52,787 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:04:53,886 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
Deer Bear River
Car Car River
Deer Car Bear
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop fs -ls lab6
2021-05-11 13:05:08,237 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 1 items
-rw-r--r--   1 hdoop supergroup       44 2021-05-11 13:04 lab6/input.txt
```

Lab – 06: WordCount.jar

Md Yaseen Ahmed
1BM19CS404

```
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop jar /home/mdyaseen/Desktop/wordcount.jar WordCount lab6/ /lab6_output
2021-05-11 13:16:27,874 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:16:28,911 INFO client.RMProxy: Connecting to ResourceManager at /127.0.0.1:8032
2021-05-11 13:16:29,553 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/hdoop/.staging/job_1620717524992_0001
2021-05-11 13:16:29,752 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:29,987 INFO input.FileInputFormat: Total input files to process : 1
2021-05-11 13:16:30,057 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,112 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,145 INFO mapreduce.JobSubmitter: number of splits:1
2021-05-11 13:16:30,436 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
2021-05-11 13:16:30,496 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1620717524992_0001
2021-05-11 13:16:30,496 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-05-11 13:16:30,833 INFO conf.Configuration: resource-types.xml not found
2021-05-11 13:16:30,833 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2021-05-11 13:16:31,178 INFO impl.YarnClientImpl: Submitted application application_1620717524992_0001
2021-05-11 13:16:31,317 INFO mapreduce.Job: The url to track the job: http://test:8088/proxy/application_1620717524992_0001/
2021-05-11 13:16:31,318 INFO mapreduce.Job: Running job: job_1620717524992_0001
2021-05-11 13:16:43,642 INFO mapreduce.Job: Job job_1620717524992_0001 running in uber mode : false
2021-05-11 13:16:43,648 INFO mapreduce.Job:  map 0% reduce 0%
2021-05-11 13:16:49,771 INFO mapreduce.Job:  map 100% reduce 0%
2021-05-11 13:16:55,843 INFO mapreduce.Job:  map 100% reduce 100%
2021-05-11 13:16:56,887 INFO mapreduce.Job: Job job_1620717524992_0001 completed successfully
2021-05-11 13:16:57,020 INFO mapreduce.Job: Counters: 54
        File System Counters
                FILE: Number of bytes read=50
                FILE: Number of bytes written=451185
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=156
                HDFS: Number of bytes written=28
                HDFS: Number of read operations=8
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=2
                HDFS: Number of bytes read erasure-coded=0
        Job Counters
                Launched map tasks=1
                Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=3750
                Total time spent by all reduces in occupied slots (ms)=3327
                Total time spent by all map tasks (ms)=3750
                Total time spent by all reduce tasks (ms)=3327
                Total vcore-milliseconds taken by all map tasks=3750
                Total vcore-milliseconds taken by all reduce tasks=3327
                Total megabyte-milliseconds taken by all map tasks=3840000
                Total megabyte-milliseconds taken by all reduce tasks=3406848
        Map-Reduce Framework
```

```
                Combine output records=4
                Reduce input groups=4
                Reduce shuffle bytes=50
                Reduce input records=4
                Reduce output records=4
                Spilled Records=8
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=120
                CPU time spent (ms)=1090
                Physical memory (bytes) snapshot=346484736
                Virtual memory (bytes) snapshot=4967919616
                Total committed heap usage (bytes)=230821888
                Peak Map Physical memory (bytes)=224583680
                Peak Map Virtual memory (bytes)=2479607808
                Peak Reduce Physical memory (bytes)=121901056
                Peak Reduce Virtual memory (bytes)=2488311808
        Shuffle Errors
                BAD_ID=0
                CONNECTION=0
                IO_ERROR=0
                WRONG_LENGTH=0
                WRONG_MAP=0
                WRONG_REDUCE=0
        File Input Format Counters
                Bytes Read=44
        File Output Format Counters
                Bytes Written=28
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -ls /lab6_output
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:17:19,627 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 2 items
-rw-r--r--   1 hdoop supergroup          0 2021-05-11 13:16 /lab6_output/_SUCCESS
-rw-r--r--   1 hdoop supergroup         28 2021-05-11 13:16 /lab6_output/part-r-00000
hdoop@test:~/hadoop-3.2.1/sbin$ hadoop dfs -cat /lab6_output/part-r-00000
WARNING: Use of this script to execute dfs is deprecated.
WARNING: Attempting to execute replacement "hdfs dfs" instead.

2021-05-11 13:18:08,517 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2021-05-11 13:18:09,598 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localHostTrusted = false, remoteHostTrusted = false
Bear    2
Car     3
Deer    2
River   2
hdoop@test:~/hadoop-3.2.1/sbin$ s
```