Week 2

Felipe Rodriguez

Bellevue University

DSC 650 Big Data

Professor Nasheb Ismaily

December 10, 2023

HDFS Report

Loaded Data

```
[bash-5.0# hdfs dfs -put /data/grades.csv /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-08 23:03:20,522 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[bash-5.0# hdfs dfs -ls /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-08 23:04:04,295 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 7 items
-rw-r--r--   1 root supergroup        747 2023-12-08 23:03 /grades.csv
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /hbase
drwxr-xr-x   - root supergroup          0 2023-12-08 22:54 /log
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /spark-jars
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /tez
drwxr-xr-x   - root supergroup          0 2023-12-08 22:53 /tmp
drwxrwx---   - root supergroup          0 2023-12-08 22:55 /user
bash-5.0# 
```

Loaded Data Worker 1

```
[feliperodriguez@bigdata-week1:~/dsc650-infra/bellevue-bigdata/hadoop-hive-spark-hbase$ docker-compose exec worker1 bash
[bash-5.0# hdfs dfs -ls /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-08 23:10:29,107 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 7 items
-rw-r--r--   1 root supergroup        747 2023-12-08 23:03 /grades.csv
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /hbase
drwxr-xr-x   - root supergroup          0 2023-12-08 22:54 /log
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /spark-jars
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /tez
drwxr-xr-x   - root supergroup          0 2023-12-08 22:53 /tmp
drwxrwx---   - root supergroup          0 2023-12-08 22:55 /user
```

Loaded Data Worker 2

```
[feliperodriguez@bigdata-week1:~/dsc650-infra/bellevue-bigdata/hadoop-hive-spark-hbase$ docker-compose exec worker2 bash
[bash-5.0# hdfs dfs -ls /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-08 23:13:04,254 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 7 items
-rw-r--r--   1 root supergroup        747 2023-12-08 23:03 /grades.csv
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /hbase
drwxr-xr-x   - root supergroup          0 2023-12-08 22:54 /log
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /spark-jars
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /tez
drwxr-xr-x   - root supergroup          0 2023-12-08 22:53 /tmp
drwxrwx---   - root supergroup          0 2023-12-08 22:55 /user
```

HDFS Command 1 – Count

```
bash-5.0# hdfs dfs -count /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-08 23:19:15,463 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
          72          289        377268168 /
bash-5.0#
```

HDFS Command 2 – Mkdir

```
[bash-5.0# hdfs dfs -mkdir /test
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-09 00:01:03,772 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
[bash-5.0# hdfs dfs -ls /
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-09 00:01:15,330 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Found 8 items
-rw-r--r--   1 root supergroup        747 2023-12-08 23:03 /grades.csv
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /hbase
drwxr-xr-x   - root supergroup          0 2023-12-08 22:54 /log
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /spark-jars
drwxr-xr-x   - root supergroup          0 2023-12-09 00:01 /test
drwxr-xr-x   - root supergroup          0 2023-12-08 22:55 /tez
drwxr-xr-x   - root supergroup          0 2023-12-08 22:53 /tmp
drwxrwx---   - root supergroup          0 2023-12-08 22:55 /user
bash-5.0#
```
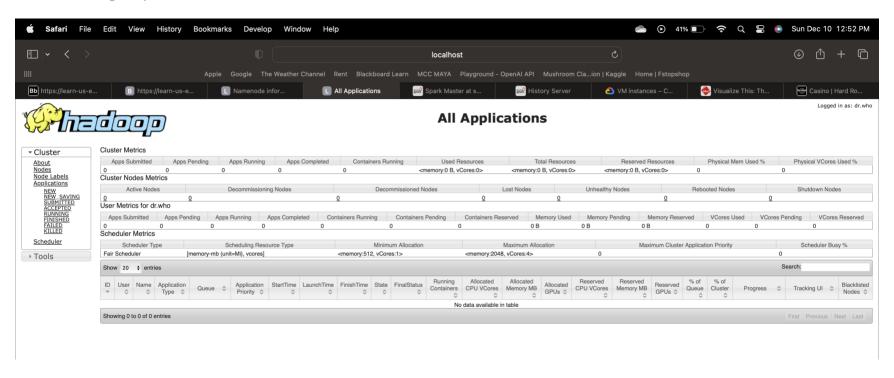
HDFS Command 3 – Stat

```
bash-5.0# hdfs dfs -stat "%b %F"  /grades.csv
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-10 22:02:19,143 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
747 regular file
bash-5.0#
```

Yard Node –List

```
bash-5.0# yarn node -list
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/tez/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.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]
2023-12-09 01:26:23,797 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-12-09 01:26:23,897 INFO client.RMProxy: Connecting to ResourceManager at master/172.28.1.1:8032
Total Nodes:2
         Node-Id             Node-State Node-Http-Address       Number-of-Running-Containers
   worker1:45583               RUNNING      worker1:8042                                  0
   worker2:33977               RUNNING      worker2:8042                                  0
bash-5.0#
```

Maximum Capacity Increased

MapReduce Results

Summary of Results

The method used for calculating Pi in this assignment is the Monte Carlo Method. This method takes the number of samples, in this case 10, randomly scatter points inside a square and counts the number of points that fall inside a circle inscribed in the square. This number counted is the estimate of Pi. Pi value is 3.1415~, however in the test ran, the Pi value is 3.8. This is likely because of the small sample size. Because the Monte Carlo method is looking at the amount of scatter points within the inscribed circle, the more samples that are observed, the closer the value will be to Pi.