

#screenshot of "hdfs dfsadmin -report" Deliverable: Screenshot of the output.

```
Decommission Status : Normal
Configured Capacity: 51835101184 (48.28 GB)
DFS Used: 205692476 (196.16 MB)
Non DFS Used: 11747856836 (10.94 GB)
DFS Remaining: 39596341073 (36.88 GB)
DFS Used%: 0.40%
DFS Remaining%: 76.39%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 5
Last contact: Mon Dec 04 22:06:02 GMT 2023
Last Block Report: Mon Dec 04 22:03:50 GMT 2023
Num of Blocks: 127
```

```
Name: 172.28.1.3:9866 (worker2)
Hostname: worker2
Decommission Status : Normal
Configured Capacity: 51835101184 (48.28 GB)
DFS Used: 174511755 (166.43 MB)
Non DFS Used: 11779037557 (10.97 GB)
DFS Remaining: 39596367181 (36.88 GB)
DFS Used%: 0.34%
DFS Remaining%: 76.39%
Configured Cache Capacity: 0 (0 B)
Cache Used: 0 (0 B)
Cache Remaining: 0 (0 B)
Cache Used%: 100.00%
Cache Remaining%: 0.00%
Xceivers: 7
Last contact: Mon Dec 04 22:06:02 GMT 2023
Last Block Report: Mon Dec 04 22:03:50 GMT 2023
Num of Blocks: 149
```

```
bash-5.0#
```

“hdfs dfs -ls / Deliverable: Screenshot proving the data has been loaded.

```
bash-5.0# ls /data
batches grades.csv input.txt output.txt ssn-address.tsv
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]
2023-12-04 22:09:03,886 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]
2023-12-04 22:09:22,402 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 2023-12-04 22:09 /grades.csv
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /hbase
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /log
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /spark-jars
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /tez
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /tmp
drwxrwx--- - root supergroup 0 2023-12-04 22:05 /user
bash-5.0#
```

```
br39@dc650-brennaparker:~/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]
2023-12-04 22:12:05,066 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 2023-12-04 22:09 /grades.csv
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /hbase
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /log
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /spark-jars
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /tez
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /tmp
drwxrwx--- - root supergroup 0 2023-12-04 22:05 /user
bash-5.0# exit
exit
br39@dc650-brennaparker:~/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]
2023-12-04 22:13:05,767 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 2023-12-04 22:09 /grades.csv
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /hbase
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /log
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /spark-jars
drwxr-xr-x - root supergroup 0 2023-12-04 22:05 /tez
drwxr-xr-x - root supergroup 0 2023-12-04 22:04 /tmp
drwxrwx--- - root supergroup 0 2023-12-04 22:05 /user
bash-5.0#
```

Deliverable: Screenshots of the three chosen HDFS command outputs.

```
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]
2023-12-04 22:57:13,023 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
bash-5.0#
```

```

bash-5.0# hdfs dfs -count /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]
2023-12-04 23:25:57,254 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
      1          0          0 /test
bash-5.0#

```

```

bash-5.0# hdfs dfs -rmkdir /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]
2023-12-04 23:26:52,942 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
bash-5.0#

```

Deliverable: Screenshot of the results.

```

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]
2023-12-04 23:30:45,061 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
2023-12-04 23:30:45,166 INFO client.RMPProxy: Connecting to ResourceManager at master/172.28.1.1:8032
Total Nodes:2
      Node-Id          Node-State Node-Http-Address      Number-of-Running-Containers
      worker2:33557      RUNNING      worker2:8042          0
      worker1:38741      RUNNING      worker1:8042          0
bash-5.0#

```

Deliverable: Screenshot from the YARN UI showing the updated maximum memory (2048 MB).



All Applications

Cluster

About
Nodes
Node Labels
Applications
NEW
NEW SAVING
SUBMITTED
ACCEPTED
RUNNING
FINISHED
FAILED
KILLED
Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Used Resources	Total Resources
0	0	0	0	0	<memory:0 B, vCores:0>	<memory:8 GB, vCores:1>

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes
2	0	0	0

User Metrics for dr.who

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Containers Pending	Containers Reserved	Memory Used
0	0	0	0	0	0	0	0 B

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Fair Scheduler	[memory-mb (unit=Mi), vcores]	<memory:512, vCores:1>	<memory:2048, vCores:4>

Show 20 entries

ID	User	Name	Application Type	Queue	Application Priority	StartTime	LaunchTime	FinishTime	State	FinalStatus	Running Containers	Allocated CPU VCores	Allocated Memory MB	Allo GP
No data available in table														

Showing 0 to 0 of 0 entries

Summary:

Pi: a map/reduce program that estimates Pi using a quasi-Monte Carlo method

Number of maps: 2

Samples per map: 10

So, the quasi-Monte Carlo method is a program that estimates Pi by generating random points in a unit square and counts the number of points that is inside a circle within the square. This is a form of map reduction in that we're trying to simplify the points on the map into a single number, preferably as close to Pi as possible. In this instance we only used 2 maps with 10 points per map. Some of these points are inside and some are outside of the circle. Now, even though the quasi-Monte Carlo method removes all points outside of the circle, if there are only a few points inside of the circle, it still won't equal pi. The more samples that we use, the closer the final number will be to pi.

```
GC time elapsed (ms)=174
CPU time spent (ms)=1560
Physical memory (bytes) snapshot=739356672
Virtual memory (bytes) snapshot=7060795392
Total committed heap usage (bytes)=663748608
Peak Map Physical memory (bytes)=282963968
Peak Map Virtual memory (bytes)=2354556928
Peak Reduce Physical memory (bytes)=180117504
Peak Reduce Virtual memory (bytes)=2354200576

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=236

File Output Format Counters
Bytes Written=97

Job Finished in 25.344 seconds
Estimated value of Pi is 3.800000000000000000000000
bash-5.0#
```