## **DSC 650**

Week 2 Assignment

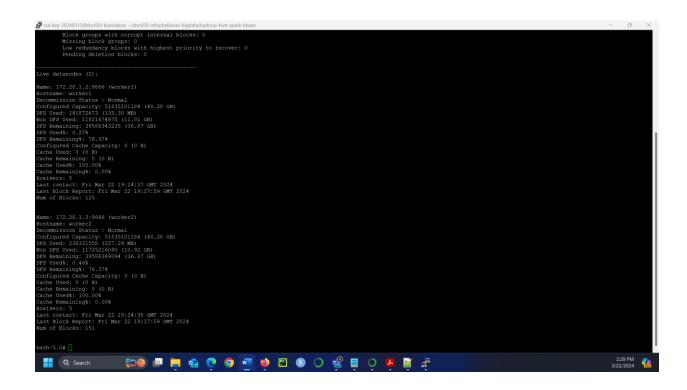
Eyram Kueviakoe

March 22, 2024

# Deep Dive into HDFS

Screenshot 1: hdfs dfsadmin -report

```
Assistance of the control of the con
```



# Screenshot 2: Screenshot proving the data has been loaded hdfs dfs -Ls /

# Screenshot:

```
docker-compose exec worker1 bash
hdfs dfs -ls /
and
docker-compose exec worker2 bash
hdfs dfs -ls /
```

```
The skey 2003/1586cc650-burdabec -/dec650-infra/bellevue-bigdata/hadocy-hive-spark-thase$ sudo docker-compose exec worker1 bash bash-5.04 hdfs dfs -1s /
skr4: class path contains multiple SLAW1 bindings.
SLRAS: Pound binding in [sar:file/usr/program/hadocy/common/lib/slf4]-log6312-1.7.25.jar!/org/slf4/impl/staticLoggerBinder.class]
SLRAS: Found binding in [sar:file/usr/program/hadocy/common/lib/slf4]-log6312-1.7.25.jar!/org/slf4/impl/staticLoggerBinder.class]
SLRAS: Found binding in [sar:file/usr/program/hadocy/common/lib/slf4]-log6312-1.7.25.jar!/org/slf4/impl/staticLoggerBinder.class]
SLRAS: Status binding is of type [org.slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4]-slf4
```

Screenshots of the three chosen HDFS command outputs.

## Command 1

hdfs dfs -mkdir assignment2

and

hdfs dfs -ls

## to show the folder was created

## Command 2

## hdfs dfs -getfacl assignment2

```
bash-5.0# hdfs dfs -getfacl assignment2
SLF40: Class path contains multiple SLF40 bindings.
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/had/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF41: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF41: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
2024-03-22 20:30:22.904 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable # file: assignment2
# owner: root
# group: supergroup
user::rwx
group::r-x
other::r-x
bash-5.0# [
```

## Command 3

hdfs dfs -rm -r assignment2

and

hdfs dfs -ls

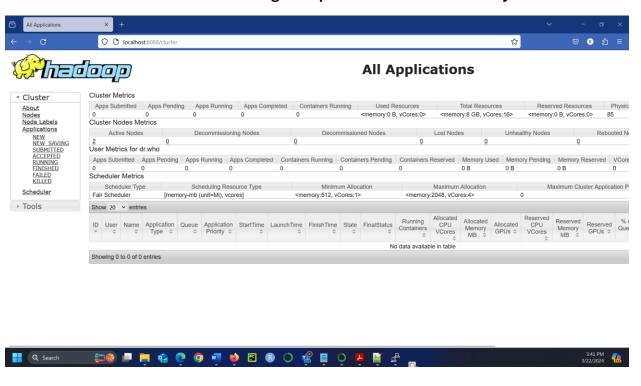
to show the folder is deleted

```
SLF40: Class path contains multiple SLF4J bindings.
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hive/lib/log4j-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF40: Actual binding is of type [org.slf4j.impl.Log4jloggerEactory]
2024-03-22 20:31:08,013 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable beleted assignment2
bash-5.0# hdfs dfs -ls
SLF40: Class path contains multiple SLF4D bindings.
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hadoop/share/hadoop/common/lib/slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF40: Found binding in [jar:file:/usr/program/hive/lib/slf4j-log4j12-1.7.10.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF
```

## Screenshot

yarn node -list

# Screenshot from the YARN UI showing the updated maximum memory



**Experimenting with MapReduce** 

```
### Case Local map taskes!

Rack-local map taskes!

Rack-local map taskes!

Total time spent by all maps in occupied slots (ms)=22494

Total time spent by all reduces in occupied slots (ms)=1716

Total time spent by all reduces in occupied slots (ms)=1716

Total time spent by all reduces taskes (ms)=429

Total voor-milliseconds taken by all map taskes=1129

Total voor-milliseconds taken by all map taskes=1129

Total mapshyte-milliseconds taken by all map taskes=1129

Total mapshyte-milliseconds taken by all map taskes=1129

Rop-index particles and the state of the state of taskes of taskes
```

# **Summary and significance of the result:**

The estimated Pi value of 3.8 from the MapReduce job is higher than the actual Pi value. This means that the calculation wasn't very accurate.

There are many reasons that can explain the inaccuracy:

- We are not using enough samples or data points
- The random numbers are not properly generated.

To get an accurate result, we will need a larger sample. This will reduce the estimation error. Also the random number generation should be tuned to generate points that are evenly distributed.