Jamie Tran

Professor Nasheb Ismaily

DSC650 Big Data

12 June 2024

Week 2 Screenshots

**Screenshot of HDFS report output:**

**A screenshot of a computer program

Description automatically generated** **A screenshot of a computer

Description automatically generated**

**Screenshot proving the data has been loaded:**

**A screenshot of a computer

Description automatically generated**

**Screenshots of 3 chosen HDFS command outputs:**

* **Command #1**

A computer screen shot of a black screen

Description automatically generated

A screenshot of a computer program

Description automatically generated

* **Command #2**

**A screenshot of a computer program

Description automatically generated**

**A screenshot of a computer screen

Description automatically generated**

* **Command #3**

**A screenshot of a computer program

Description automatically generated**

**A computer screen shot of a black screen

Description automatically generated**

**Screenshot of YARN nodes:**

**A computer screen with white text

Description automatically generated**

**Screenshot from YARN UI with 2048 MB:**

**A screenshot of a computer

Description automatically generated**

**Screenshot of MapReduce Pi job:**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated** A screenshot of a computer

Description automatically generated

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

The MapReduce Pi job is utilizing the monte carlo method to estimate the value of Pi. This method produces its results by creating random samples to get our end value. The two parameters that are needed are the number of maps that the program is going to create and the number of points that will be generated for each map. From our input of ‘hadoop jar /usr/program/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.3.jar pi -libjars ${libjars} 2 10’ we are creating 2 random maps with 10 points each. Based on these random generated maps our program has predicted that the estimated value of Pi is 3.80000. Since we know that the value of Pi is 3.14159 we can infer that our program would probably benefit from a few more samples.