Calculating Pi using Map Reduce

Norina Akhtar 19643

Table Of Contents

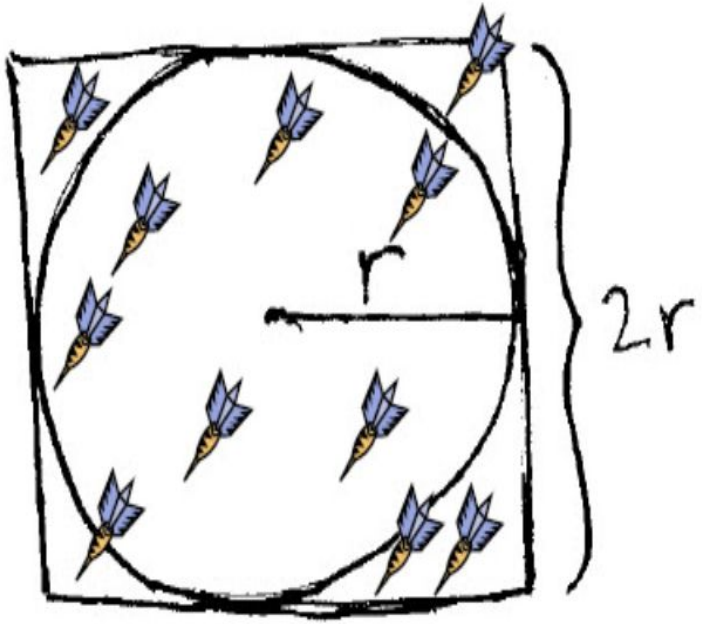
Introduction

MapReduce is a programming model and an associated implementation for processing and generating big data sets with a parallel, distributed algorithm on a cluster.

The number π is a mathematical constant that is the ratio of a circle's circumference to its diameter, approximately equal to 3.14159.

Overview of pi calculation using map reduce

- Throw N darts on the board . Each dart lands at a random position (x,y) on the board.
- If each dart landed inside the circle or not:
- Check if **x^2 + y^2 < r**
- Take the total number of darts that landed in the circle as s:
  - 4(S/N) = π



How to find inside and outside

$(x - center\_x)^2 + (x - center\_x)^2$ compare r^2

- Inside: if compare is <
- Outside: if compare is >
- On the circle: if compare is =

Explanation

If radius is 5, then based on the input values in Map, we can calculate
pi = 4 * (S / N)

= 4 * (Insde / (Inside + Outside))

= 4 * (5 / (5 + 7))

= 4 * (5 / 12)

= 1.66

Note:

- To get more accurate pi value you can
  - increase the value of the radius, and
  - create much much more input values (e.g., a million values)

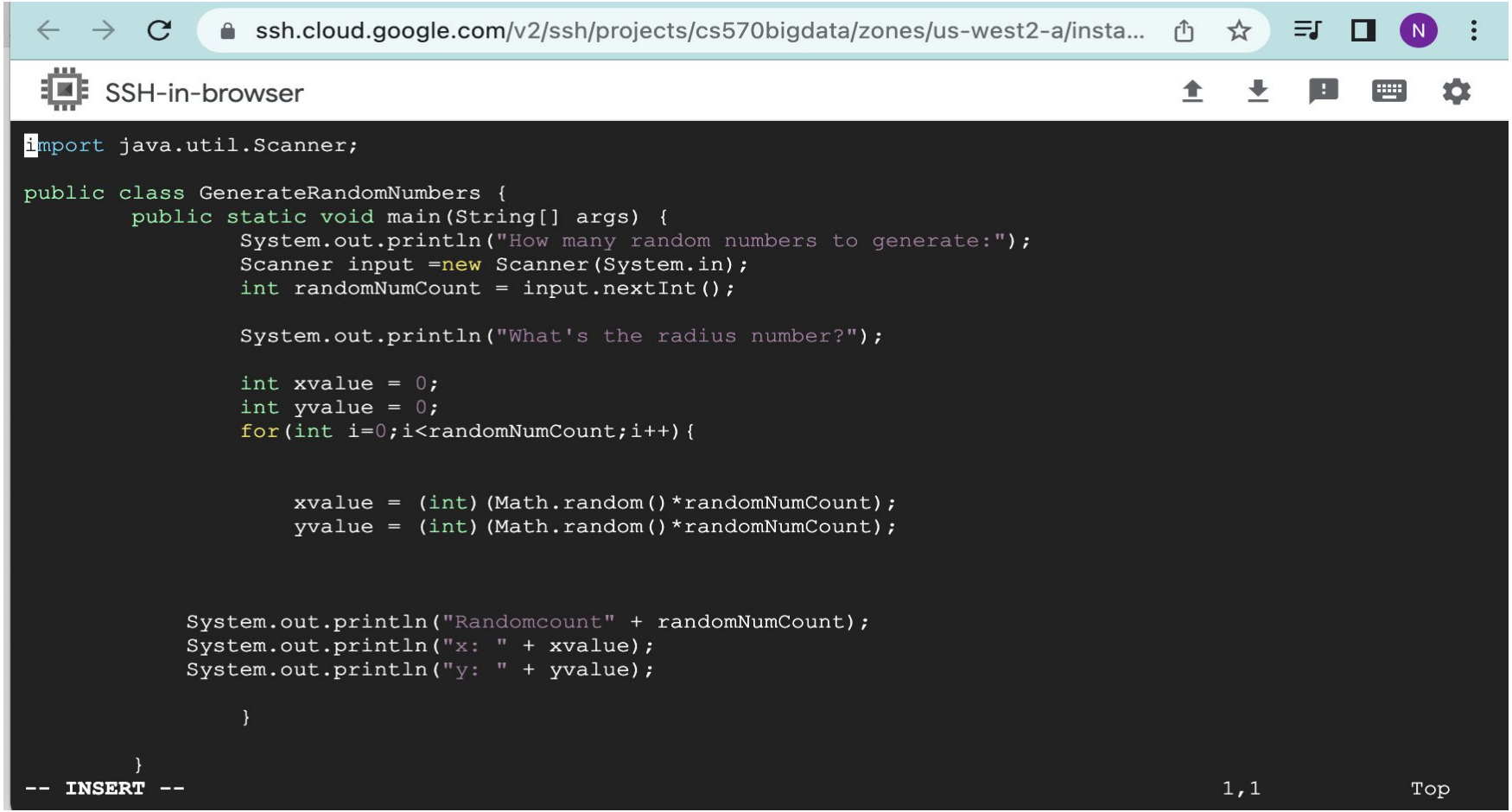
Design

Job: Pi										
Map Task								Reduce Task		
map()				combine()				reduce()		
Input (Given)		Output (Program)		Input (Given)		Output (Program)		Input (Given)		Output (Program)
Key	Value (radius=2)	Key	Value (radius=2)	Key	Values	Key	Value	Key	Values	
file1	(0, 1)	Outside	1	Inside	[1]	Inside	1	Inside	[1, 3, 1]	Inside 5
	(1, 3)	Inside	1	Outside	[1, 1]	Outside	2	Outside	[2, 1, 4]	Outside 7
	(4, 3)	Outside	1							
file2	(2, 3)	Inside	1	Inside	[1, 1, 1]	Inside	3			
	(1, 3)	Inside	1	Outside	[1]	Outside	1			
	(1, 4)	Outside	1							
	(3, 2)	Inside	1							
file3	(3, 0)	Outside	1	Inside	[1]	Inside	1			
	(3, 3)	Inside	1	Outside	[1, 1, 1, 1]	Outside	4			
	(3, 4)	Outside	1							
	(0, 0)	Outside	1							
	(4, 4)	Outside	1							

Implementation

Java program to generate random numbers Using hadoop cluster

```
nakhtar@vm:~/hadoop-3.3.4$ cd ..
nakhtar@vm:~$ ls
Pi  hadoop-3.3.4  hadoop-3.3.4.tar.gz
nakhtar@vm:~$ ls Pi
input
nakhtar@vm:~$ cd Pi
nakhtar@vm:~/Pi$ vi GenerateRandomNumbers.java
```


ssh.cloud.google.com/v2/ssh/projects/cs570bigdata/zones/us-west2-a/insta...
SSH-in-browser
import java.util.Scanner;
public class GenerateRandomNumbers {
public static void main(String[] args) {
System.out.println("How many random numbers to generate:");
Scanner input =new Scanner(System.in);
int randomNumCount = input.nextInt();
System.out.println("What's the radius number?");
int xvalue = 0;
int yvalue = 0;
for(int i=0;i<randomNumCount;i++){
xvalue = (int)(Math.random()*randomNumCount);
yvalue = (int)(Math.random()*randomNumCount);
System.out.println("Randomcount" + randomNumCount);
System.out.println("x: " + xvalue);
System.out.println("y: " + yvalue);
}
}
-- INSERT --
1,1
Top

Made directories , adjust files and connect to ssh

SSH-in-browser

```
nakhtar@vm:~$ cd Pi
nakhtar@vm:~/Pi$ cd input
nakhtar@vm:~/Pi/input$ ls
nakhtar@vm:~/Pi/input$ cd ..
nakhtar@vm:~/Pi$ ls
GenerateRandomNumbers.class  GenerateRandomNumbers.java  PiCalculation.java  input
nakhtar@vm:~/Pi$ vi PiCalculation.java
nakhtar@vm:~/Pi$ vi PiCalculation.java
nakhtar@vm:~/Pi$ rm -r Picalculation.java
rm: cannot remove 'Picalculation.java': No such file or directory
nakhtar@vm:~/Pi$ rm -r PiCalculation.java
nakhtar@vm:~/Pi$ ls
GenerateRandomNumbers.class  GenerateRandomNumbers.java  input
nakhtar@vm:~/Pi$ vi PiCalculation.java
nakhtar@vm:~/Pi$ vi PiCalculation.java
nakhtar@vm:~/Pi$ rm -r PiCalculation.java
nakhtar@vm:~/Pi$ ls
GenerateRandomNumbers.class  GenerateRandomNumbers.java  input
nakhtar@vm:~/Pi$ vi PiCalculation.java
nakhtar@vm:~/Pi$ ls
GenerateRandomNumbers.class  GenerateRandomNumbers.java  PiCalculation.java  input
nakhtar@vm:~/Pi$ cd input
nakhtar@vm:~/Pi/input$ vi file01
nakhtar@vm:~/Pi/input$ cd ..
nakhtar@vm:~/Pi$ cd ..
nakhtar@vm:~$ vi ~/.bashrc
nakhtar@vm:~$ ssh-keygen -t rsa -P '' -f ~/.ssh/id_rsa
Generating public/private rsa key pair.
/home/nakhtar/.ssh/id_rsa already exists.
Overwrite (y/n)? y
Your identification has been saved in /home/nakhtar/.ssh/id_rsa
Your public key has been saved in /home/nakhtar/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:iZRyiLXMvQV67gR+3g9lG52vxkZKDeK/vzup98zie4o nakhtar@vm
The key's randomart image is:
+---[RSA 3072]----+
|    . .          |
|   = = o         |
|  . O * .        |
|   . B +.... .   |
|    . *.S.+oo    |
|     = ..o.oo.   |
|      o oo.+ ..  |
|         oo.O+.  |
|          EOOX+  |
+----[SHA256]-----+
nakhtar@vm:~$ cat ~/.ssh/id_rsa.pub >> ~/.ssh/authorized_keys
nakhtar@vm:~$ chmod 0600 ~/.ssh/authorized_keys
nakhtar@vm:~$ ssh localhost
Welcome to Ubuntu 20.04.5 LTS (GNU/Linux 5.15.0-1018-gcp x86_64)
```

```
nakhtar@vm:~/Pi$ javac GenerateRandomNumbers.java
nakhtar@vm:~/Pi$ java GenerateRandomNumbers
How many random numbers to generate:
10
What's the radius number?
Randomcount10
x: 1
y: 6
Randomcount10
x: 3
y: 8
Randomcount10
x: 3
y: 6
Randomcount10
x: 1
y: 0
Randomcount10
x: 1
y: 0
Randomcount10
x: 5
y: 7
Randomcount10
x: 6
y: 5
Randomcount10
x: 0
y: 9
Randomcount10
x: 2
y: 2
Randomcount10
x: 6
y: 7
nakhtar@vm:~/Pi$
```

Create a MapReduce program to calculate the numbers of inside darts and outside darts

```java
import java.io.*;
import java.util.*;
import java.lang.Object;
import java.net.URI;

import org.apache.hadoop.fs.Path;
import org.apache.hadoop.conf.*;
import org.apache.hadoop.io.*;
import org.apache.hadoop.mapreduce.*;
import org.apache.hadoop.mapreduce.Mapper.Context;
import org.apache.hadoop.mapreduce.lib.input.FileInputFormat;
import org.apache.hadoop.mapreduce.lib.input.TextInputFormat;
import org.apache.hadoop.mapreduce.lib.output.FileOutputFormat;
import org.apache.hadoop.mapreduce.lib.output.TextOutputFormat;
import org.apache.hadoop.fs.*;


public class PiCalculation {

  public static class TokenizerMapper
       extends Mapper<Object, Text, Text, IntWritable>{

    private final static IntWritable one = new IntWritable(1);
    private Text word = new Text();


    public void map(Object key, Text value, Context context
                    ) throws IOException, InterruptedException {

      String line = value.toString();
      line = line.replace("(","");
      line = line.replace(")","");
      line = line.replace(","," ");

      StringTokenizer itr = new StringTokenizer(line);
      int radius = 20;
      while (itr.hasMoreTokens()) {
        String x, y;
        x = itr.nextToken();
        if (itr.hasMoreTokens()) {
            y = itr.nextToken();
        }else {
            y = "0";
        }
        int xvalue = (int)(Integer.parseInt(x));
        int yvalue = (int)(Integer.parseInt(y));
        double check = Math.sqrt(Math.pow((radius-xvalue), 2) + Math.pow((radius-yvalue), 2));

        if (check < radius) {
            word.set("inside");
```

```
        x = itr.nextToken();
        if (itr.hasMoreTokens()) {
            y = itr.nextToken();
        }else {
            y = "0";
        }
        int xvalue = (int)(Integer.parseInt(x));
        int yvalue = (int)(Integer.parseInt(y));
        double check = Math.sqrt(Math.pow((radius-xvalue), 2) + Math.pow((radius-yvalue), 2));

        if (check < radius) {
            word.set("inside");
        } else {
            word.set("outside");
        }
        context.write(word, one);
      }
    }
  }

  public static class IntSumReducer
       extends Reducer<Text,IntWritable,Text,IntWritable> {
    private IntWritable result = new IntWritable();

    public void reduce(Text key, Iterable<IntWritable> values,
                       Context context
                       ) throws IOException, InterruptedException {
      int sum = 0;
      for (IntWritable val : values) {
        sum += val.get();
      }
      result.set(sum);
      context.write(key, result);
    }
  }

  public static void main(String[] args) throws Exception {
    Configuration conf = new Configuration();
    Job job = Job.getInstance(conf, "pi calculation");
    job.setJarByClass(PiCalculation.class);
    job.setMapperClass(TokenizerMapper.class);
    job.setCombinerClass(IntSumReducer.class);
    job.setReducerClass(IntSumReducer.class);
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    //System.exit(job.waitForCompletion(true) ? 0 : 1);
    job.waitForCompletion(true);
    String filePath = args[1] + "/" + "part-r-00000";
```

```
        sum += val.get();
      }
      result.set(sum);
      context.write(key, result);
    }
  }

  public static void main(String[] args) throws Exception {
    Configuration conf = new Configuration();
    Job job = Job.getInstance(conf, "pi calculation");
    job.setJarByClass(PiCalculation.class);
    job.setMapperClass(TokenizerMapper.class);
    job.setCombinerClass(IntSumReducer.class);
    job.setReducerClass(IntSumReducer.class);
    job.setOutputKeyClass(Text.class);
    job.setOutputValueClass(IntWritable.class);
    FileInputFormat.addInputPath(job, new Path(args[0]));
    FileOutputFormat.setOutputPath(job, new Path(args[1]));
    //System.exit(job.waitForCompletion(true) ? 0 : 1);
    job.waitForCompletion(true);
    String filePath = args[1] + "/" + "part-r-00000";
    Path path = new Path(filePath);
    FileSystem fs = FileSystem.get(path.toUri(), conf);

    BufferedReader br=new BufferedReader(new InputStreamReader(fs.open(path)));

    String z, inside= null, outside= null;

    String line1,line2;

    line1=br.readLine();
    System.out.println(line1);
    line2=br.readLine();
    System.out.println(line2);

    line1 = line1.replace("inside","").trim();
    line2 = line2.replace("outside","").trim();

    System.out.println("Inside:"+line1+", Outside:"+line2);

    if (line1 != null && line2 != null) {
       double invalue = Double.valueOf(line1);
       double outvalue = Double.valueOf(line2);
       double pi =4*( invalue /(invalue+outvalue));
       System.out.println("PI:"+pi);
    }

    fs.close();
  }
}
```

Format the file system

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs namenode -format
2022-10-11 17:33:26,830 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = vm.us-west2-a.c.cs570bigdata.internal/10.168.0.2
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 3.3.4
STARTUP_MSG:   classpath = /home/nakhtar/hadoop-3.3.4/etc/hadoop:/home/nakhtar/hadoop-3.3.4/share/hadoop/commo
hadoop/common/lib/kerby-config-1.0.1.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jaxb-api-2.2.11.ja
.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/kerb-server-1.0.1.jar:/home/nakhtar/hadoop-3.3.4/share/
are/hadoop/common/lib/commons-beanutils-1.9.4.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jcip-anno
eload4j-1.2.22.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/listenablefuture-9999.0-empty-to-avoid-c
lib/gson-2.8.9.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jetty-util-9.4.43.v20210629.jar:/home/na
akhtar/hadoop-3.3.4/share/hadoop/common/lib/slf4j-api-1.7.36.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/commo
ib/commons-text-1.4.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/kerby-util-1.0.1.jar:/home/nakhtar/
adoop-3.3.4/share/hadoop/common/lib/checker-qual-2.5.2.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/
b/jetty-servlet-9.4.43.v20210629.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jersey-server-1.19.jar
ar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/kerby-xdr-1.0.1.jar:/home/nakhtar/hadoop-3.3.4/share/had
are/hadoop/common/lib/slf4j-reload4j-1.7.36.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/netty-3.10.
ypto-1.0.1.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/zookeeper-3.5.6.jar:/home/nakhtar/hadoop-3.3
/hadoop-3.3.4/share/hadoop/common/lib/curator-recipes-4.2.0.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/
mon/lib/curator-framework-4.2.0.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jul-to-slf4j-1.7.36.jar
3.v20210629.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/curator-client-4.2.0.jar:/home/nakhtar/hado
/hadoop-3.3.4/share/hadoop/common/lib/javax.servlet-api-3.1.0.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/comm
re/hadoop/common/lib/commons-cli-1.2.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jaxb-impl-2.2.3-1.
7.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jackson-annotations-2.12.7.jar:/home/nakhtar/hadoop-3
4/share/hadoop/common/lib/hadoop-annotations-3.3.4.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/comm
ommon/lib/woodstox-core-5.3.0.jar:/home/nakhtar/hadoop-3.3.4/share/hadoop/common/lib/jetty-webapp-9.4.43.v2021
```

Start NameNode daemon and DataNode daemon

```
SHUTDOWN_MSG: Shutting down NameNode at vm.us-west2-a.c.cs570bigdata.internal/10.168.0.2
************************************************************/
nakhtar@vm:~/hadoop-3.3.4$ sbin/start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [vm]
nakhtar@vm:~/hadoop-3.3.4$ wget http://localhost:9870/
--2022-10-11 17:34:00--  http://localhost:9870/
Resolving localhost (localhost)... 127.0.0.1
Connecting to localhost (localhost)|127.0.0.1|:9870... connected.
HTTP request sent, awaiting response... 302 Found
Location: http://localhost:9870/index.html [following]
--2022-10-11 17:34:01--  http://localhost:9870/index.html
Reusing existing connection to localhost:9870.
HTTP request sent, awaiting response... 200 OK
Length: 1079 (1.1K) [text/html]
Saving to: 'index.html.5'

index.html.5              100%[=====================================>]   1.05K  --.-KB/s    in 0s

2022-10-11 17:34:01 (125 MB/s) - 'index.html.5' saved [1079/1079]
```

Make the HDFS directories required to execute MapReduce jobs:

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -mkdir /user
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -mkdir /user/nakhtar
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -mkdir input
nakhtar@vm:~/hadoop-3.3.4$  bin/hdfs dfs -put etc/hadoop/*.xml input
nakhtar@vm:~/hadoop-3.3.4$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.3.4.jar grep input output 'dfs
2022-10-11 17:23:13,580 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2022-10-11 17:23:13,713 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2022-10-11 17:23:13,714 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2022-10-11 17:23:14,079 INFO input.FileInputFormat: Total input files to process : 10
2022-10-11 17:23:14,117 INFO mapreduce.JobSubmitter: number of splits:10
2022-10-11 17:23:14,310 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1734921756_0001
2022-10-11 17:23:14,311 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-10-11 17:23:14,488 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2022-10-11 17:23:14,489 INFO mapreduce.Job: Running job: job_local1734921756_0001
2022-10-11 17:23:14,498 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2022-10-11 17:23:14,508 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2022-10-11 17:23:14,509 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output
2022-10-11 17:23:14,510 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.FileOutputCo
2022-10-11 17:23:14,578 INFO mapred.LocalJobRunner: Waiting for map tasks
2022-10-11 17:23:14,579 INFO mapred.LocalJobRunner: Starting task: attempt_local1734921756_0001_m_000000_0
```

Copy the input files into the distributed file system

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -put ../Pi/input/* pi/input
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -ls /user/nakhtar/pi/input
Found 1 items
-rw-r--r--   1 nakhtar supergroup        619 2022-10-11 17:38 /user/nakhtar/pi/input/file01
nakhtar@vm:~/hadoop-3.3.4$ bin/hadoop com.sun.tools.javac.Main ../Pi/PiCalculation.java
nakhtar@vm:~/hadoop-3.3.4$ cp ../Pi/*.class .
nakhtar@vm:~/hadoop-3.3.4$ cp ../Pi/*.java .
nakhtar@vm:~/hadoop-3.3.4$ jar cf wc.jar PiCalculation*.class
nakhtar@vm:~/hadoop-3.3.4$ ls
```

Compile PiCalculation.java and create a jar

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -put ../Pi/input/* pi/input
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -ls /user/nakhtar/pi/input
Found 1 items
-rw-r--r--   1 nakhtar supergroup        619 2022-10-11 17:38 /user/nakhtar/pi/input/file01
nakhtar@vm:~/hadoop-3.3.4$ bin/hadoop com.sun.tools.javac.Main ../Pi/PiCalculation.java
nakhtar@vm:~/hadoop-3.3.4$ cp ../Pi/*.class .
nakhtar@vm:~/hadoop-3.3.4$ cp ../Pi/*.java .
nakhtar@vm:~/hadoop-3.3.4$ jar cf wc.jar PiCalculation*.class
nakhtar@vm:~/hadoop-3.3.4$ ls
```

Run Program

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hadoop jar wc.jar PiCalculation /user/nakhtar/pi/input /user/nakhtar/pi/output
2022-10-11 17:43:42,681 INFO impl.MetricsConfig: Loaded properties from hadoop-metrics2.properties
2022-10-11 17:43:42,770 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot period at 10 second(s).
2022-10-11 17:43:42,770 INFO impl.MetricsSystemImpl: JobTracker metrics system started
2022-10-11 17:43:43,046 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. I
ner to remedy this.
2022-10-11 17:43:43,222 INFO input.FileInputFormat: Total input files to process : 1
2022-10-11 17:43:43,249 INFO mapreduce.JobSubmitter: number of splits:1
2022-10-11 17:43:43,417 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1419692810_0001
2022-10-11 17:43:43,417 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-10-11 17:43:43,607 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
2022-10-11 17:43:43,609 INFO mapreduce.Job: Running job: job_local1419692810_0001
2022-10-11 17:43:43,618 INFO mapred.LocalJobRunner: OutputCommitter set in config null
2022-10-11 17:43:43,630 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2022-10-11 17:43:43,630 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders u
2022-10-11 17:43:43,631 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapreduce.lib.output.
2022-10-11 17:43:43,711 INFO mapred.LocalJobRunner: Waiting for map tasks
2022-10-11 17:43:43,712 INFO mapred.LocalJobRunner: Starting task: attempt_local1419692810_0001_m_000000_0
2022-10-11 17:43:43,749 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 2
2022-10-11 17:43:43,750 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders u
2022-10-11 17:43:43,785 INFO mapred.Task:  Using ResourceCalculatorProcessTree : [ ]
2022-10-11 17:43:43,790 INFO mapred.MapTask: Processing split: hdfs://localhost:9000/user/nakhtar/pi/input/fil
```

Output

```
                HDFS: Number of bytes read=2158
                HDFS: Number of bytes written=22
                HDFS: Number of read operations=15
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=4
                HDFS: Number of bytes read erasure-coded=0
        Map-Reduce Framework
                Map input records=150
                Map output records=150
                Map output bytes=1694
                Map output materialized bytes=33
                Input split bytes=115
                Combine input records=150
                Combine output records=2
                Reduce input groups=2
                Reduce shuffle bytes=33
                Reduce input records=2
                Reduce output records=2
                Spilled Records=4
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=26
                Total committed heap usage (bytes)=759169024
        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=1079
        File Output Format Counters
                Bytes Written=22
inside  106
outside 44
Inside:106, Outside:44
PI:2.8266666666666667
```

```
nakhtar@vm:~/hadoop-3.3.4$ bin/hdfs dfs -get pi/output output
nakhtar@vm:~/hadoop-3.3.4$ cat output/*
cat: output/output: Is a directory
nakhtar@vm:~/hadoop-3.3.4$ cd output
nakhtar@vm:~/hadoop-3.3.4/output$ ls
output
nakhtar@vm:~/hadoop-3.3.4/output$ cd output
nakhtar@vm:~/hadoop-3.3.4/output/output$ ls
_SUCCESS  part-r-00000
nakhtar@vm:~/hadoop-3.3.4/output/output$ cat part-r-00000
inside  106
outside 44
nakhtar@vm:~/hadoop-3.3.4/output/output$
```

Enhancement Ideas

- If you want to get an accurate estimate of Pi, you need a large number of random samples.
- Notice that each dart can be thrown at any time and its position can be evaluated independently.
- With one person throwing all the darst, it will take a long time to finish.
- If we had N people throwing a dart each , this would be much faster.

Conclusion

The Pi algorithm is concise and clear that,

- N should be large
- Points should be chosen uniformly at random

Hence, we should increase the the number of random values to get accurate result of Pi.

References

Time for action – using Hadoop to calculate Pi

MapReduce calculation process

Overview of Pi calculation using MapReduce

Pi Computation With MapReduce

MapReduce Pi