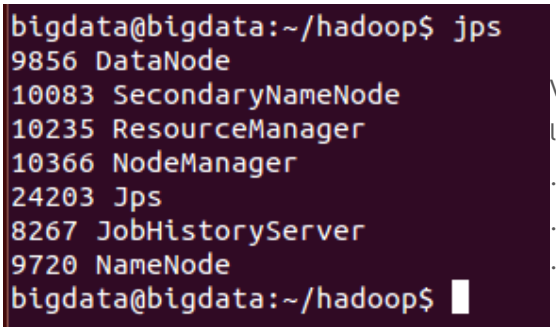
EJERCICIOS DE HDFS

PI

En primer lugar compruebo si están arrancados los demonios:



Veo que están arrancados. Si no estuvieran tendría que arrancarlos con los siguientes comandos:

./sbin/start-dfs.sh

./sbin/start-yarn.sh

./sbin/mr-jobhistory-daemon.sh start historyserver

Me encuentro en la carpeta home/bigdata/hadoop. Para acceder más rápido he creado una variable de entorno:

HADOOP_HOME=/home/bigdata/hadoop

Con el comando cd HADOOP_HOME puedo acceder a la carpeta directamente.

Ejecutar el programa "pi" con los siguientes parámetros de entrada e indicar el valor devuelto.

- Programa: pi

- Nº de mapeos: 2

- Nº de muestras por mapeo: 4

Ejecuto el programa pi con los parámetros indicados:

hadoop jar /home/bigdata/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar pi 2 4

```
@bigdata: ~
bigdata@bigdata:~/EjercicioSudoku$ cd ..
bigdata@bigdata:~$ hadoop jar /home/bigdata/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar pi 2 4
Number of Maps  = 2
Samples per Map = 4
16/08/03 00:29:32 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Wrote input for Map #0
Wrote input for Map #1
Starting Job
16/08/03 00:29:34 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
16/08/03 00:29:35 INFO input.FileInputFormat: Total input paths to process : 2
16/08/03 00:29:35 INFO mapreduce.JobSubmitter: number of splits:2
16/08/03 00:29:35 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1469204298658_0011
16/08/03 00:29:36 INFO impl.YarnClientImpl: Submitted application application_1469204298658_0011
16/08/03 00:29:36 INFO mapreduce.Job: The url to track the job: http://bigdata:8088/proxy/application_1469204298658_0011/
16/08/03 00:29:36 INFO mapreduce.Job: Running job: job_1469204298658_0011
16/08/03 00:29:45 INFO mapreduce.Job: Job job_1469204298658_0011 running in uber mode : false
16/08/03 00:29:45 INFO mapreduce.Job:  map 0% reduce 0%
16/08/03 00:29:58 INFO mapreduce.Job:  map 100% reduce 0%
16/08/03 00:30:06 INFO mapreduce.Job:  map 100% reduce 100%
16/08/03 00:30:07 INFO mapreduce.Job: Job job_1469204298658_0011 completed successfully
16/08/03 00:30:08 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=50
                FILE: Number of bytes written=354033
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=534
                HDFS: Number of bytes written=215
                HDFS: Number of read operations=11
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=3
        Job Counters
                Launched map tasks=2
                Launched reduce tasks=1
                Data-local map tasks=2
                Total time spent by all maps in occupied slots (ms)=22515
                Total time spent by all reduces in occupied slots (ms)=5515
                Total time spent by all map tasks (ms)=22515
                Total time spent by all reduce tasks (ms)=5515
                Total vcore-milliseconds taken by all map tasks=22515
                Total vcore-milliseconds taken by all reduce tasks=5515
                Total megabyte-milliseconds taken by all map tasks=23055360
                Total megabyte-milliseconds taken by all reduce tasks=5647360
        Map-Reduce Framework
                Map input records=2
                Map output records=4
                Map output bytes=36
                Map output materialized bytes=56
                Input split bytes=298
                Combine input records=0
                HDFS: Number of bytes written=215
                HDFS: Number of read operations=11
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=3
        Job Counters
                Launched map tasks=2
                Launched reduce tasks=1
                Data-local map tasks=2
                Total time spent by all maps in occupied slots (ms)=22515
                Total time spent by all reduces in occupied slots (ms)=5515
                Total time spent by all map tasks (ms)=22515
                Total time spent by all reduce tasks (ms)=5515
                Total vcore-milliseconds taken by all map tasks=22515
                Total vcore-milliseconds taken by all reduce tasks=5515
                Total megabyte-milliseconds taken by all map tasks=23055360
                Total megabyte-milliseconds taken by all reduce tasks=5647360
        Map-Reduce Framework
                Map input records=2
                Map output records=4
                Map output bytes=36
                Map output materialized bytes=56
                Input split bytes=298
                Combine input records=0
                Combine output records=0
                Reduce input groups=2
                Reduce shuffle bytes=56
                Reduce input records=4
                Reduce output records=0
                Spilled Records=8
                Shuffled Maps =2
                Failed Shuffles=0
                Merged Map outputs=2
                GC time elapsed (ms)=320
                CPU time spent (ms)=1490
                Physical memory (bytes) snapshot=621404160
                Virtual memory (bytes) snapshot=2400448512
                Total committed heap usage (bytes)=485031936
        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
ob Finished in 33.598 seconds
stimated value of Pi is 3.5000000000000000000
igdata@bigdata:~$
```

Se ha calculado el valor aproximado de PI = 3.5

Ejecutar el programa "pi" con los siguientes parámetros de entrada e indicar el valor devuelto.

- Programa: pi

- Nº de mapeos: 5

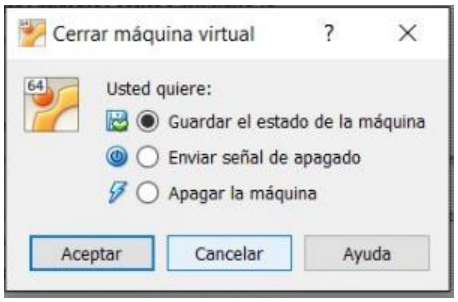
- Nº de muestras por mapeo: 10

hadoop jar /home/bigdata/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar pi 5 10

```
bigdata@bigdata:~$ hadoop jar /home/bigdata/hadoop/share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.2.jar pi 5 10
Number of Maps  = 5
Samples per Map = 10
16/08/03 00:37:53 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
Wrote input for Map #0
Wrote input for Map #1
Wrote input for Map #2
Wrote input for Map #3
Wrote input for Map #4
Starting Job
16/08/03 00:37:55 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
16/08/03 00:37:55 INFO input.FileInputFormat: Total input paths to process : 5
16/08/03 00:37:55 INFO mapreduce.JobSubmitter: number of splits:5
16/08/03 00:37:56 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1469204298658_0012
16/08/03 00:37:56 INFO impl.YarnClientImpl: Submitted application application_1469204298658_0012
16/08/03 00:37:56 INFO mapreduce.Job: The url to track the job: http://bigdata:8088/proxy/application_1469204298658_0012/
16/08/03 00:37:56 INFO mapreduce.Job: Running job: job_1469204298658_0012
16/08/03 00:38:06 INFO mapreduce.Job: Job job_1469204298658_0012 running in uber mode : false
16/08/03 00:38:06 INFO mapreduce.Job:  map 0% reduce 0%
16/08/03 00:38:36 INFO mapreduce.Job:  map 80% reduce 0%
16/08/03 00:38:37 INFO mapreduce.Job:  map 100% reduce 0%
16/08/03 00:38:45 INFO mapreduce.Job:  map 100% reduce 100%
16/08/03 00:38:45 INFO mapreduce.Job: Job job_1469204298658_0012 completed successfully
16/08/03 00:38:45 INFO mapreduce.Job: Counters: 49
        File System Counters
                FILE: Number of bytes read=116
                FILE: Number of bytes written=708165
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=1335
                HDFS: Number of bytes written=215
                HDFS: Number of read operations=23
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=3
        Job Counters
                Launched map tasks=5
                Launched reduce tasks=1
                Data-local map tasks=5
                Total time spent by all maps in occupied slots (ms)=145641
                Total time spent by all reduces in occupied slots (ms)=5413
                Total time spent by all map tasks (ms)=145641
                Total time spent by all reduce tasks (ms)=5413
                Total vcore-milliseconds taken by all map tasks=145641
                Total vcore-milliseconds taken by all reduce tasks=5413
                Total megabyte-milliseconds taken by all map tasks=149136384
                Total megabyte-milliseconds taken by all reduce tasks=5542912
        Map-Reduce Framework
                Map input records=5
                Map output records=10
                Map output bytes=90
                Map output materialized bytes=140
                Input split bytes=745
                Combine input records=0
                Combine output records=0
                Reduce input groups=2
                Reduce shuffle bytes=140
                Reduce input records=10
                Reduce output records=0
                Spilled Records=20
                Shuffled Maps =5
                Failed Shuffles=0
                Merged Map outputs=5
                GC time elapsed (ms)=1867
                CPU time spent (ms)=3160
                Physical memory (bytes) snapshot=1340383232
                Virtual memory (bytes) snapshot=4794564608
                Total committed heap usage (bytes)=1022754816
        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=590
        File Output Format Counters
                Bytes Written=97
Job Finished in 50.572 seconds
Estimated value of Pi is 3.28000000000000000000
bigdata@bigdata:~$
```

Con más mapeos y muestras por mapeo el valor es 3.28 y se aproxima más a PI.

Para finalizar cierro y guardo la máquina Archivo/cerrar y guardar el estado de la máquina.



Si no hago eso tendía que parar los demonios:

cd $HADOOP_HOME

./sbin/stop-dfs.sh
./sbin/stop-yarn.sh
./sbin/mr-jobhistory-daemon.sh stop historyserver