**C&IT Big Data Environment**

**Cluster Details**

There are 13 hosts in the 12 - physical and 1 – virtual. The 12 physical hosts have Linux CentOS 7.5 Operating System installed, the virtual host has RHEL 7.5. Only virtual host has public IP address and it is dedicated for users interactive login via ssh shell to work with cluster. The cluster has following hardware specification:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Hostname** | **Manufacturer/Product** | **CPU Model** | **#CPUs/Cores** | **RAM Type** | **RAM Volume** | **Network** | **HDD (Hard Drives)** |
| **cehpn** | DELL PowerEdge R740 | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 8X2TB 7.2K RPM NLSAS |
| **cehsn** | DELL PowerEdge R740 | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 8X2TB 7.2K RPM NLSAS |
| **cehcd01** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd02** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd03** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd04** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd05** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd06** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd07** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd08** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd09** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehcd10** | DELL PowerEdge R740xd | Intel(R) Xeon(R)Gold 6126 2.60GHz | 2 X 12 = 24 TC | DDR4 | 192GB | Intel X520DP 10Gb DA/SFP+ | 12X4TB 7.2K RPM NLSAS |
| **cehg** | VMware Virtual Platform | Intel(R) Xeon(R) CPU 2.30GHz | Total – 8 Cores | DDR4 | 16GB | 10GB | 100GB |

More detailed hardware specification can be seen if to use web link <http://ohpc.grid.wayne.edu/srvspec.html> hostname – parameter. Please note that link works only from WSU network 141.217.X.X, or from cehg VM – gateway, or from private grid network. Following table shows role of each server in the cluster and gives the information regarding Hadoop ecosystem installed on them.

|  |  |  |
| --- | --- | --- |
| **Hostname** | **Host Role in the Cluster** | **Hadoop Ecosystems on the Host** |
| **cehpn** | HDFS Primary Name Node  YARN Resource Manager  Job History Server | HDFS, YARN, Spark, Kafka, Zookeeper, MySQL server (Hive Metastore Database, Oozie Database). |
| **cehsn** | HDFS Secondary Name Node | HDFS, Spark, Hive Metastore Service, HiveServer2, Oozie Server, Flume |
| **cehcd01** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd02** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd03** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd04** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd05** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd06** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd07** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd08** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd09** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehcd10** | HDFS DataNode Host  YARN NodeManager  Compute Host | HDFS, YARN, Spark |
| **cehg** | Gateway/Edge/Frontend Server.  Users must use the server as the ssh login host, to use Hadoop ecosystems and Spark. | HDFS client, MySQL client, Pig, Sqoop, Spark client (spark-shell, pyspark, spark-submit) |

Following links can be used for monitoring and cluster control and ecosystems:

<http://cehpn.nas.grid.wayne.edu:50070/>  HDFS, basic File system Status

[http://cehpn.nas.grid.wayne.edu:19888/](http://cehpn:19888/)  Job History for MapReduce

[http://cehpn.nas.grid.wayne.edu:8088](http://cehpn:8088)    All applications YARN/Spark

<http://cehsn.nas.grid.wayne.edu:11000/oozie/> Oozie Server

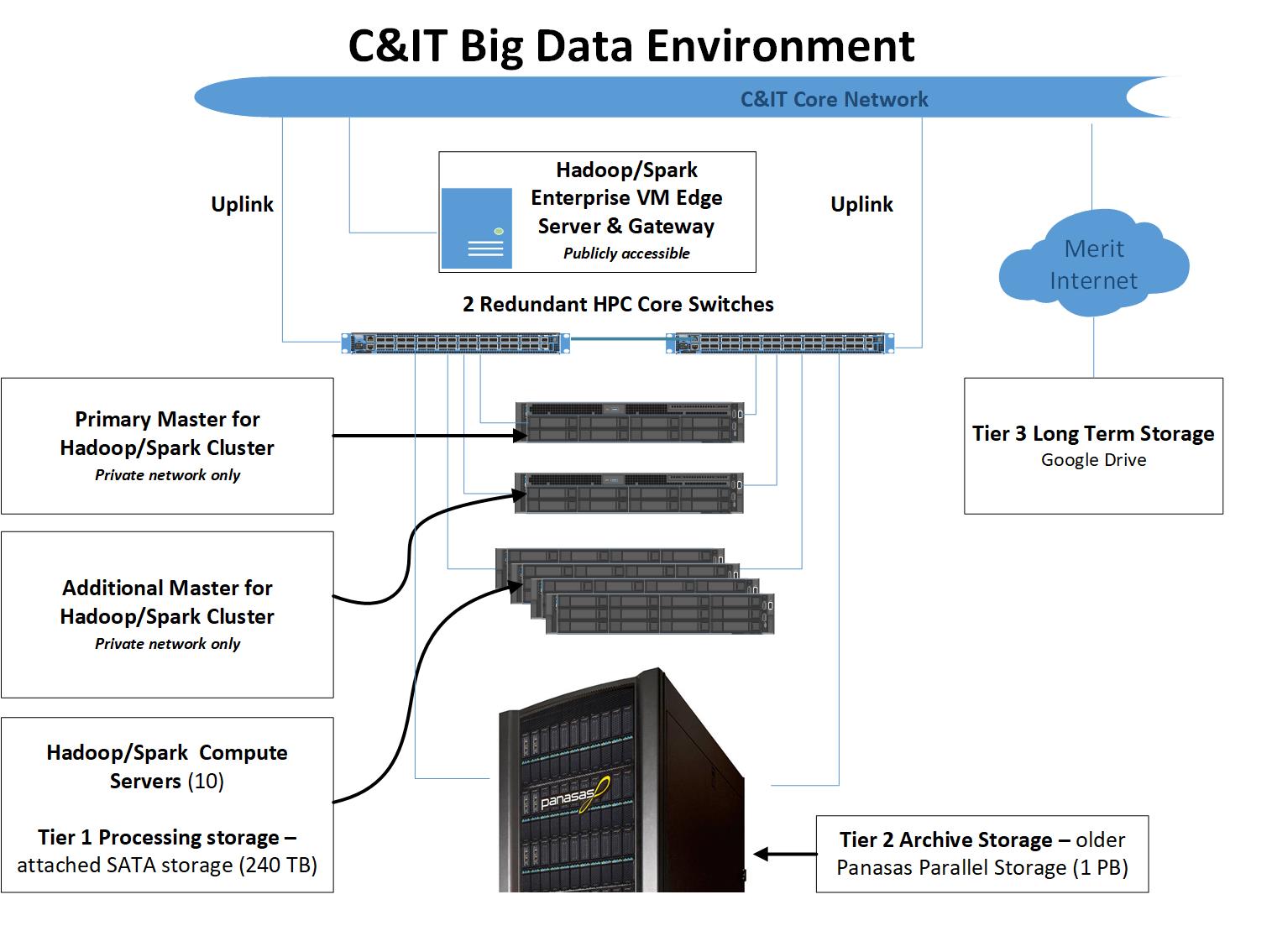
[http://cehcd01.nas.grid.wayne.edu:8042](http://cehcd01:8042)    Individual HOST

Please note that above links work only from cehg VM – gateway, or from private grid network.

The next page shows network diagram of the cluster environment.

<http://xymon.grid.wayne.edu/xymon/HP/CH/CH.html> - Cluster monitoring page, can be used from any WSU server.

<http://xymon.grid.wayne.edu/xymon-cgi/svcstatus.sh?HOST=cehpn&SERVICE=hadoop> - HDFS status



**Account/Access:**

The procedure of getting user account in the cluster consists of two steps

1. Getting Grid account: <https://www.grid.wayne.edu/application/index.html> (This step is required)
2. Sending the request for the account in cluster this can be done by advisor of the calls through the e-mail with list of students AccessID or each user can do it individually through WSU Cherwell web system <http://csm.wayne.edu> using following path: Forms -> High Performance Computing -> Big Data.

There is no priority or different level of rights and authorization for any cluster users. All have the same rights and belong to the same Linux group hadoop. There is no storage quotes for users. The general storage with current replication factor 2 ~ 240TB. Since user gets the account he/she can login to gateway/edge server cehg (virtual machine) using ssh and use Hadoop HDFS commands to get an access to Hadoop data, for Spark it can be pyspark, spark-shell (Scala) or spark-submit. Only limited operations can be done through the web basically for monitoring and controlling purposes, see web links on the second page of this document.

Free ssh clients can be downloaded from websites:

<https://www.putty.org/>

<https://www.bitvise.com/>

User’s account and data in the Hadoop file system will be available 6 months after the last login and usage of cehg server, after that the account will be removed with HDFS data. The automatic procedure of that is in the stage of development, since it is finalized and developed it will be implemented. But now it can be done manually.

The cluster has Hortonworks Hadoop implementation I would recommend some of the useful links :

<https://hadoop.apache.org/> ; <https://spark.apache.org/> ; <https://hive.apache.org/> ; <https://oozie.apache.org/>

<https://hortonworks.com/tutorial/manage-files-on-hdfs-via-cli-ambari-files-view/section/1/>

<https://docs.hortonworks.com/HDPDocuments/HDP2/HDP-2.6.5/bk_spark-component-guide/bk_spark-component-guide.pdf>

<https://hortonworks.com/tutorial/manage-files-on-hdfs-via-cli-ambari-files-view/section/1/>

**Examples/Smoke Tests:**

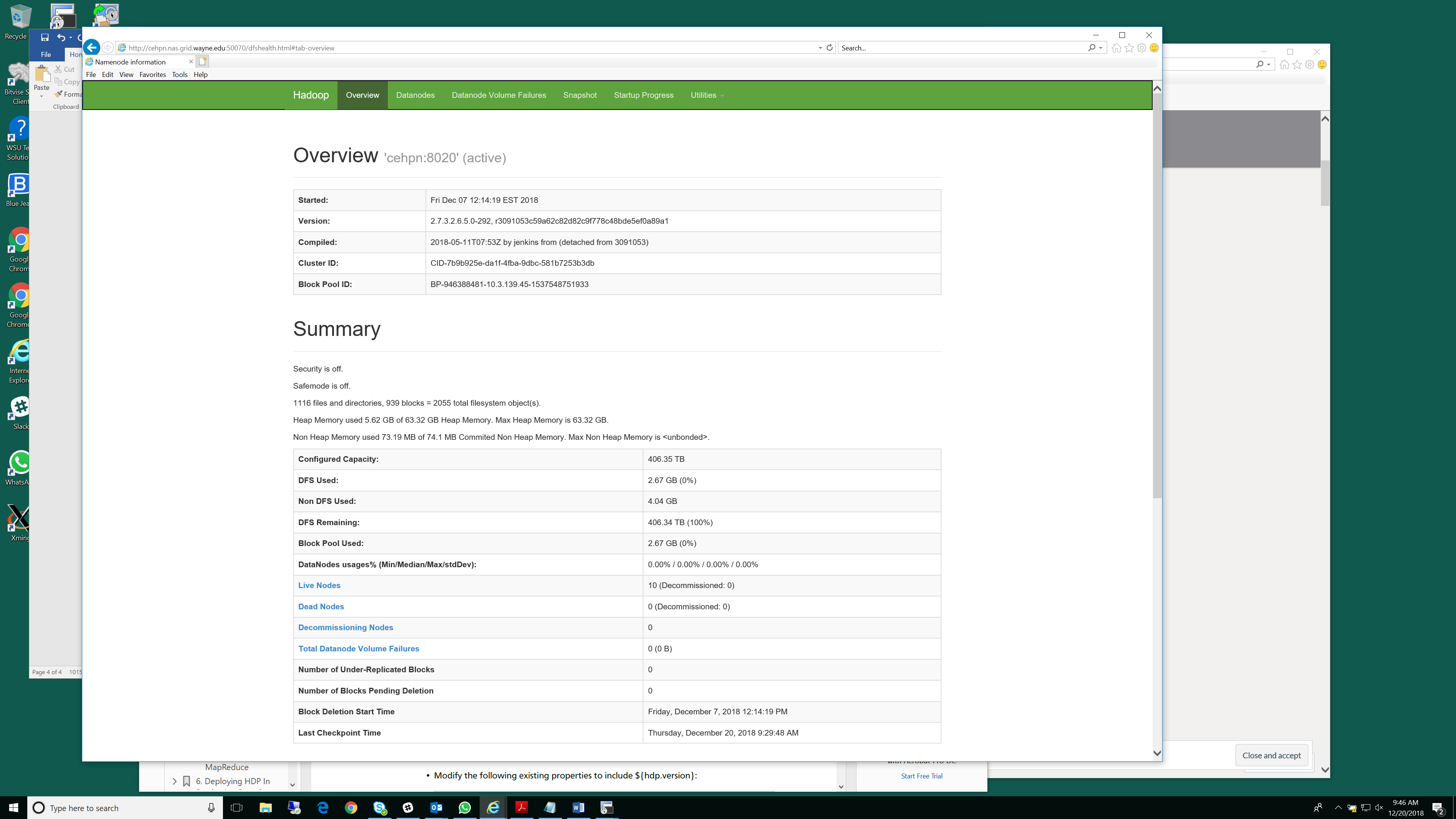
***HDFS:***

From cehg.grid.wayne.edu or any server in the grid network (warrior for example),

check that you can reach NameNode with your browser:

#firefox

<http://cehpn.nas.grid.wayne.edu:50070>



hdfs dfs -mkdir -p /user/<username>/test

hdfs dfs -copyFromLocal /user/<username>/test/

hdfs dfs –ls /user/<username>/test/

hdfs dfs –cat /user/<username>/test/passwd

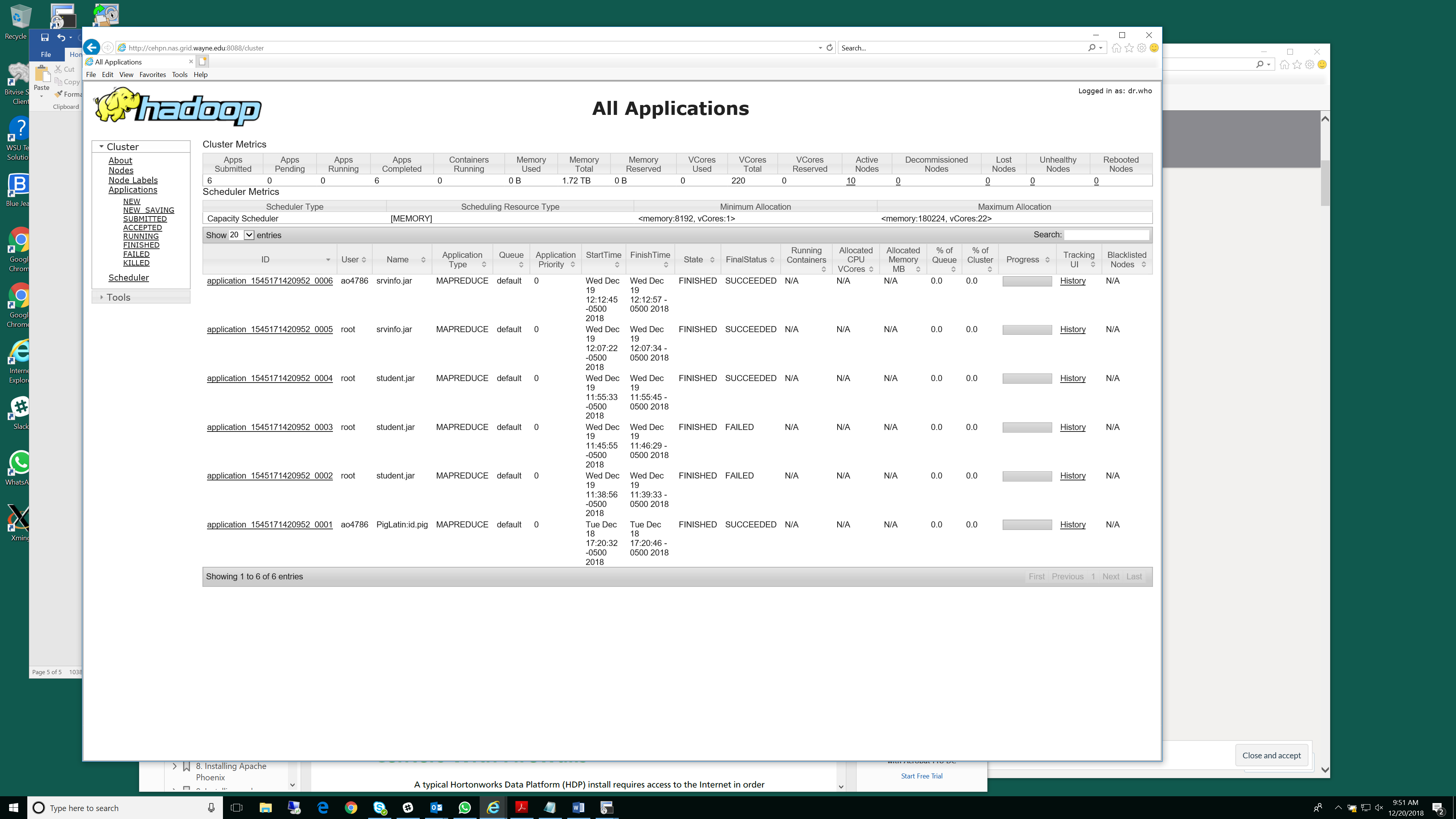
***MapReduce:***

From cehg.grid.wayne.edu or any server in the grid network (warrior for example),

check that you can reach ResourceManager with your browser:

#firefox

<http://cehpn.nas.grid.wayne.edu:8088>



Using Terasort, sort 10GB of data.

hdfs dfs –mkdir /user/<username>/tmp

/usr/hdp/current/hadoop-client/bin/hadoop jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples-\*.jar teragen 10000 /user/<username>/tmp/teragenout

**[ao4786@cehg ~]$ /usr/hdp/current/hadoop-client/bin/hadoop jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples-\*.jar teragen 10000 /user/ao4786/tmp/teragenout**

18/12/20 10:00:21 INFO client.RMProxy: Connecting to ResourceManager at cehpn.local/10.3.139.45:8050

18/12/20 10:00:21 INFO terasort.TeraSort: Generating 10000 using 2

18/12/20 10:00:21 INFO mapreduce.JobSubmitter: number of splits:2

18/12/20 10:00:21 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1545171420952\_0007

18/12/20 10:00:21 WARN mapred.YARNRunner: Usage of -Djava.library.path in mapreduce.admin.map.child.java.opts can cause programs to no longer function if hadoop native libraries are used. These values should be set as part of the LD\_LIBRARY\_PATH in the map JVM env using mapreduce.admin.user.env config settings.

18/12/20 10:00:21 WARN mapred.YARNRunner: Usage of -Djava.library.path in mapreduce.admin.reduce.child.java.opts can cause programs to no longer function if hadoop native libraries are used. These values should be set as part of the LD\_LIBRARY\_PATH in the reduce JVM env using mapreduce.admin.user.env config settings.

18/12/20 10:00:21 INFO impl.YarnClientImpl: Submitted application application\_1545171420952\_0007

18/12/20 10:00:21 INFO mapreduce.Job: The url to track the job: http://cehpn:8088/proxy/application\_1545171420952\_0007/

18/12/20 10:00:21 INFO mapreduce.Job: Running job: job\_1545171420952\_0007

18/12/20 10:00:26 INFO mapreduce.Job: Job job\_1545171420952\_0007 running in uber mode : false

18/12/20 10:00:26 INFO mapreduce.Job: map 0% reduce 0%

18/12/20 10:00:31 INFO mapreduce.Job: map 50% reduce 0%

18/12/20 10:00:34 INFO mapreduce.Job: map 100% reduce 0%

18/12/20 10:00:35 INFO mapreduce.Job: Job job\_1545171420952\_0007 completed successfully

18/12/20 10:00:35 INFO mapreduce.Job: Counters: 31

File System Counters

FILE: Number of bytes read=0

FILE: Number of bytes written=286870

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=164

HDFS: Number of bytes written=1000000

HDFS: Number of read operations=8

HDFS: Number of large read operations=0

HDFS: Number of write operations=4

Job Counters

Launched map tasks=2

Other local map tasks=2

Total time spent by all maps in occupied slots (ms)=7451

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=7451

Total vcore-milliseconds taken by all map tasks=7451

Total megabyte-milliseconds taken by all map tasks=61038592

Map-Reduce Framework

Map input records=10000

Map output records=10000

Input split bytes=164

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=111

CPU time spent (ms)=2150

Physical memory (bytes) snapshot=698765312

Virtual memory (bytes) snapshot=18258096128

Total committed heap usage (bytes)=1180172288

org.apache.hadoop.examples.terasort.TeraGen$Counters

CHECKSUM=21555350172850

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=1000000

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/tmp

Found 1 items

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 10:00 /user/ao4786/tmp/teragenout

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/tmp/teragenout

Found 3 items

-rw-r--r-- 2 ao4786 hadoop 0 2018-12-20 10:00 /user/ao4786/tmp/teragenout/\_SUCCESS

-rw-r--r-- 2 ao4786 hadoop 500000 2018-12-20 10:00 /user/ao4786/tmp/teragenout/part-m-00000

-rw-r--r-- 2 ao4786 hadoop 500000 2018-12-20 10:00 /user/ao4786/tmp/teragenout/part-m-00001

/usr/hdp/current/hadoop-client/bin/hadoop jar /usr/hdp/current/hadoopmapreduce-

client/hadoop-mapreduce-examples-\*.jar terasort /user/<username>/tmp/teragenout

/user/<username>/tmp/terasortout

**[ao4786@cehg ~]$ /usr/hdp/current/hadoop-client/bin/hadoop jar /usr/hdp/current/hadoop-mapreduce-client/hadoop-mapreduce-examples-\*.jar terasort /user/ao4786/tmp/teragenout /user/ao4786/tmp/terasortout**

18/12/20 10:12:48 INFO terasort.TeraSort: starting

18/12/20 10:12:49 INFO input.FileInputFormat: Total input paths to process : 2

Spent 149ms computing base-splits.

Spent 2ms computing TeraScheduler splits.

Computing input splits took 152ms

Sampling 2 splits of 2

Making 1 from 10000 sampled records

Computing parititions took 250ms

Spent 403ms computing partitions.

18/12/20 10:12:49 INFO client.RMProxy: Connecting to ResourceManager at cehpn.local/10.3.139.45:8050

18/12/20 10:12:51 INFO mapreduce.JobSubmitter: number of splits:2

18/12/20 10:12:51 INFO mapreduce.JobSubmitter: Submitting tokens for job: job\_1545171420952\_0008

18/12/20 10:12:51 WARN mapred.YARNRunner: Usage of -Djava.library.path in mapreduce.admin.map.child.java.opts can cause programs to no longer function if hadoop native libraries are used. These values should be set as part of the LD\_LIBRARY\_PATH in the map JVM env using mapreduce.admin.user.env config settings.

18/12/20 10:12:51 WARN mapred.YARNRunner: Usage of -Djava.library.path in mapreduce.admin.reduce.child.java.opts can cause programs to no longer function if hadoop native libraries are used. These values should be set as part of the LD\_LIBRARY\_PATH in the reduce JVM env using mapreduce.admin.user.env config settings.

18/12/20 10:12:51 INFO impl.YarnClientImpl: Submitted application application\_1545171420952\_0008

18/12/20 10:12:51 INFO mapreduce.Job: The url to track the job: http://cehpn:8088/proxy/application\_1545171420952\_0008/

18/12/20 10:12:51 INFO mapreduce.Job: Running job: job\_1545171420952\_0008

18/12/20 10:12:56 INFO mapreduce.Job: Job job\_1545171420952\_0008 running in uber mode : false

18/12/20 10:12:56 INFO mapreduce.Job: map 0% reduce 0%

18/12/20 10:13:02 INFO mapreduce.Job: map 100% reduce 0%

18/12/20 10:13:08 INFO mapreduce.Job: map 100% reduce 100%

18/12/20 10:13:08 INFO mapreduce.Job: Job job\_1545171420952\_0008 completed successfully

18/12/20 10:13:08 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=1040006

FILE: Number of bytes written=2514537

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=1000256

HDFS: Number of bytes written=1000000

HDFS: Number of read operations=9

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

Job Counters

Launched map tasks=2

Launched reduce tasks=1

Rack-local map tasks=2

Total time spent by all maps in occupied slots (ms)=7386

Total time spent by all reduces in occupied slots (ms)=3405

Total time spent by all map tasks (ms)=7386

Total time spent by all reduce tasks (ms)=3405

Total vcore-milliseconds taken by all map tasks=7386

Total vcore-milliseconds taken by all reduce tasks=3405

Total megabyte-milliseconds taken by all map tasks=60506112

Total megabyte-milliseconds taken by all reduce tasks=27893760

Map-Reduce Framework

Map input records=10000

Map output records=10000

Map output bytes=1020000

Map output materialized bytes=1040012

Input split bytes=256

Combine input records=0

Combine output records=0

Reduce input groups=10000

Reduce shuffle bytes=1040012

Reduce input records=10000

Reduce output records=10000

Spilled Records=20000

Shuffled Maps =2

Failed Shuffles=0

Merged Map outputs=2

GC time elapsed (ms)=152

CPU time spent (ms)=4620

Physical memory (bytes) snapshot=5370384384

Virtual memory (bytes) snapshot=27413204992

Total committed heap usage (bytes)=5697437696

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=1000000

File Output Format Counters

Bytes Written=1000000

18/12/20 10:13:08 INFO terasort.TeraSort: done

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/tmp/terasortout

Found 3 items

-rw-r--r-- 1 ao4786 hadoop 0 2018-12-20 10:13 /user/ao4786/tmp/terasortout/\_SUCCESS

-rw-r--r-- 10 ao4786 hadoop 0 2018-12-20 10:12 /user/ao4786/tmp/terasortout/\_partition.lst

-rw-r--r-- 1 ao4786 hadoop 1000000 2018-12-20 10:13 /user/ao4786/tmp/terasortout/part-r-00000

***Spark:***

/usr/hdp/current/spark2-client

The result approximates pi.

**./bin/spark-submit --class org.apache.spark.examples.SparkPi --master yarn-client --num-executors 1 --driver-memory 512m --executor-memory 512m --executor-cores 1 examples/jars/spark-examples\*.jar 10**

------

18/12/20 14:25:17 INFO YarnScheduler: Removed TaskSet 0.0, whose tasks have all completed, from pool

18/12/20 14:25:17 INFO DAGScheduler: ResultStage 0 (reduce at SparkPi.scala:38) finished in 0.748 s

18/12/20 14:25:17 INFO DAGScheduler: Job 0 finished: reduce at SparkPi.scala:38, took 0.797881 s

**Pi is roughly 3.1414991414991413**

18/12/20 14:25:17 INFO SparkUI: Stopped Spark web UI at http://cehg:4040

18/12/20 14:25:18 INFO YarnClientSchedulerBackend: Interrupting monitor thread

Compute the frequncy of words in the novel of Tolstoy's "War and Peace".

spark-shell --master yarn-client --driver-memory 512m --executor-memory 512m

scala> val file = sc.textFile("/public/warandpeace.txt")

file: org.apache.spark.rdd.RDD[String] = /public/warandpeace.txt MapPartitionsRDD[1] at textFile at <console>:24

scala> val counts = file.flatMap(line => line.split(" ")).map(word => (word, 1)).reduceByKey(\_ + \_)

counts: org.apache.spark.rdd.RDD[(String, Int)] = ShuffledRDD[4] at reduceByKey at <console>:25

scala> counts.saveAsTextFile("/user/ao4786/tmp/wordcount")

scala> counts.count()

res1: Long = 41992

Exit Scala

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/tmp/

Found 3 items

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 10:00 /user/ao4786/tmp/teragenout

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 10:13 /user/ao4786/tmp/terasortout

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 14:55 /user/ao4786/tmp/wordcount

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/tmp/wordcount

Found 3 items

-rw-r--r-- 2 ao4786 hadoop 0 2018-12-20 14:55 /user/ao4786/tmp/wordcount/\_SUCCESS

-rw-r--r-- 2 ao4786 hadoop 283660 2018-12-20 14:55 /user/ao4786/tmp/wordcount/part-00000

-rw-r--r-- 2 ao4786 hadoop 287621 2018-12-20 14:55 /user/ao4786/tmp/wordcount/part-00001

[ao4786@cehg ~]$ hdfs dfs -cat /user/ao4786/tmp/wordcount/part-00000

*Please note you should use your username instead of ao4786*

***Sqoop:***

Check that Sqoop is available

[ao4786@cehg ~]$ sqoop version | grep 'Sqoop [0-9].\*'

18/12/20 15:39:58 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6.2.6.5.0-292

Sqoop 1.4.6.2.6.5.0-292

***$ mysql -u <username> -p***

***Enter password:***

***mysql> create database testDb;***

***mysql> use testDb;***

***mysql> create table student(id integer,name char(20));***

***mysql> insert into student values(1,'Archana');***

***mysql> insert into student values(2,'XYZ');***

***mysql> exit;***

***sqoop import --connect jdbc:mysql://<hostname>/testDb --username <username> --password <password> --table student --m 1***

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786

Found 7 items

drwx------ - ao4786 hadoop 0 2018-11-20 01:00 /user/ao4786/.Trash

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 14:56 /user/ao4786/.sparkStaging

drwx------ - ao4786 hadoop 0 2018-12-20 15:49 /user/ao4786/.staging

-rw-r--r-- 2 ao4786 hadoop 2402 2018-12-18 15:07 /user/ao4786/passwd

drwxr-xr-x - ao4786 hadoop 0 2018-12-19 12:12 /user/ao4786/srvinfo

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 15:49 /user/ao4786/student

drwxr-xr-x - ao4786 hadoop 0 2018-12-20 14:55 /user/ao4786/tmp

[ao4786@cehg ~]$ hdfs dfs -ls /user/ao4786/student

Found 2 items

-rw-r--r-- 2 ao4786 hadoop 0 2018-12-20 15:49 /user/ao4786/student/\_SUCCESS

-rw-r--r-- 2 ao4786 hadoop 16 2018-12-20 15:49 /user/ao4786/student/part-m-00000

[ao4786@cehg ~]$ hdfs dfs -cat /user/ao4786/student/part-m-00000

1,Archana

2,XYZ

**Support**

To request support for troubleshooting jobs, software, or hardware, fill out form **Report an Issue**, under Cherwell <http://csm.wayne.edu> -> High Performance Computing -> Big Data