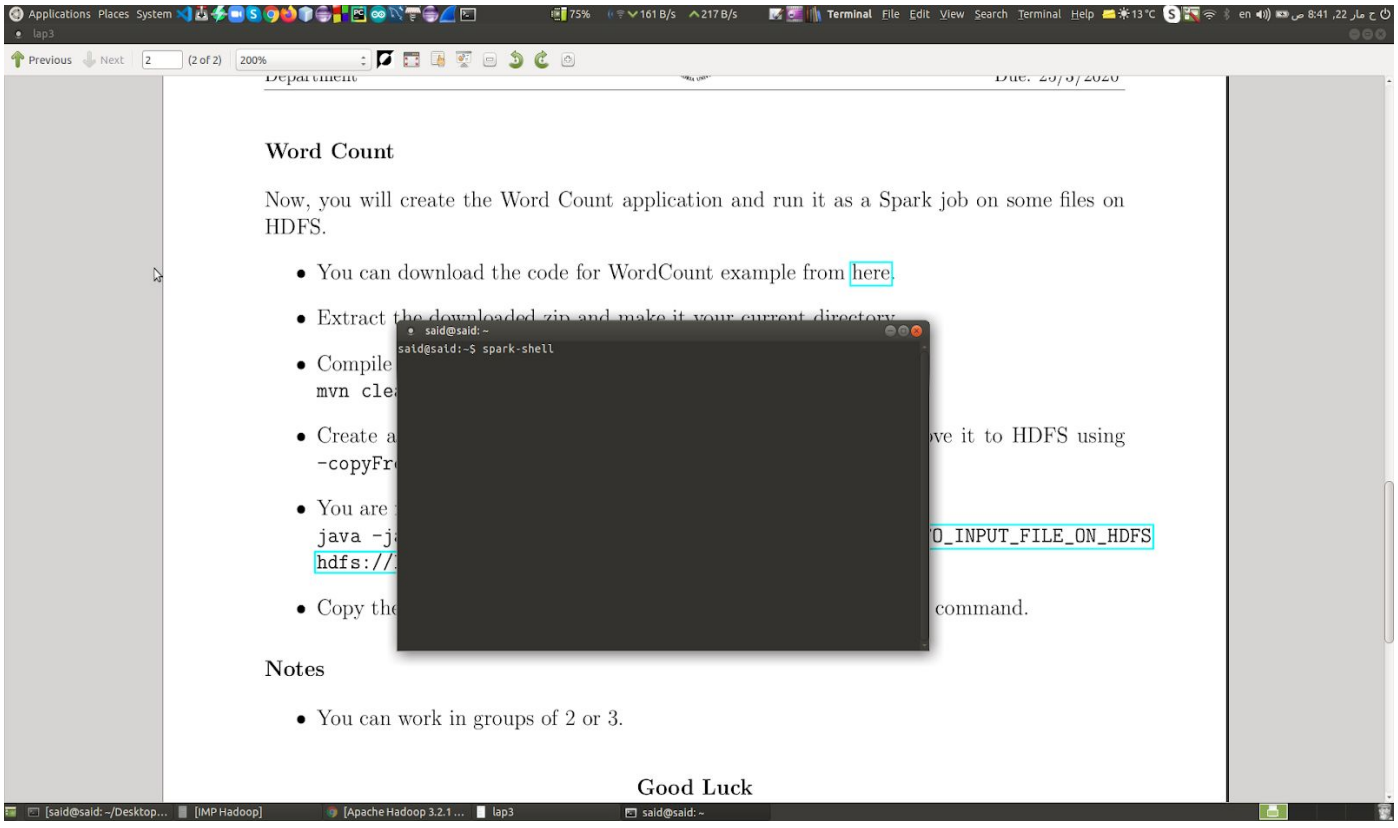
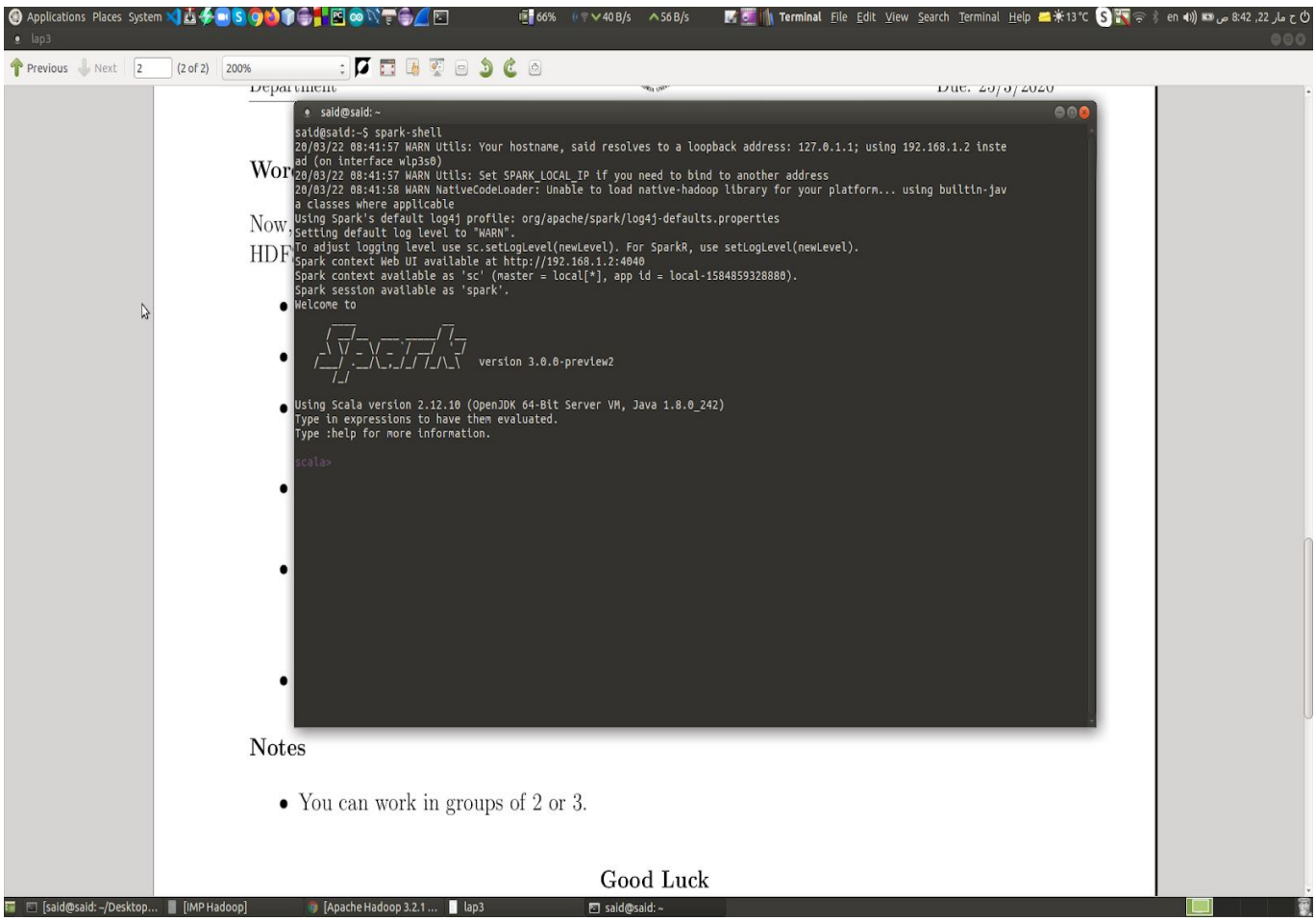
Lab 3 Spark,

By:

Mohamed Said	**57**
Ahmed El Bawab	**8**
Khalil Ismail	**23**

Screen Shots:

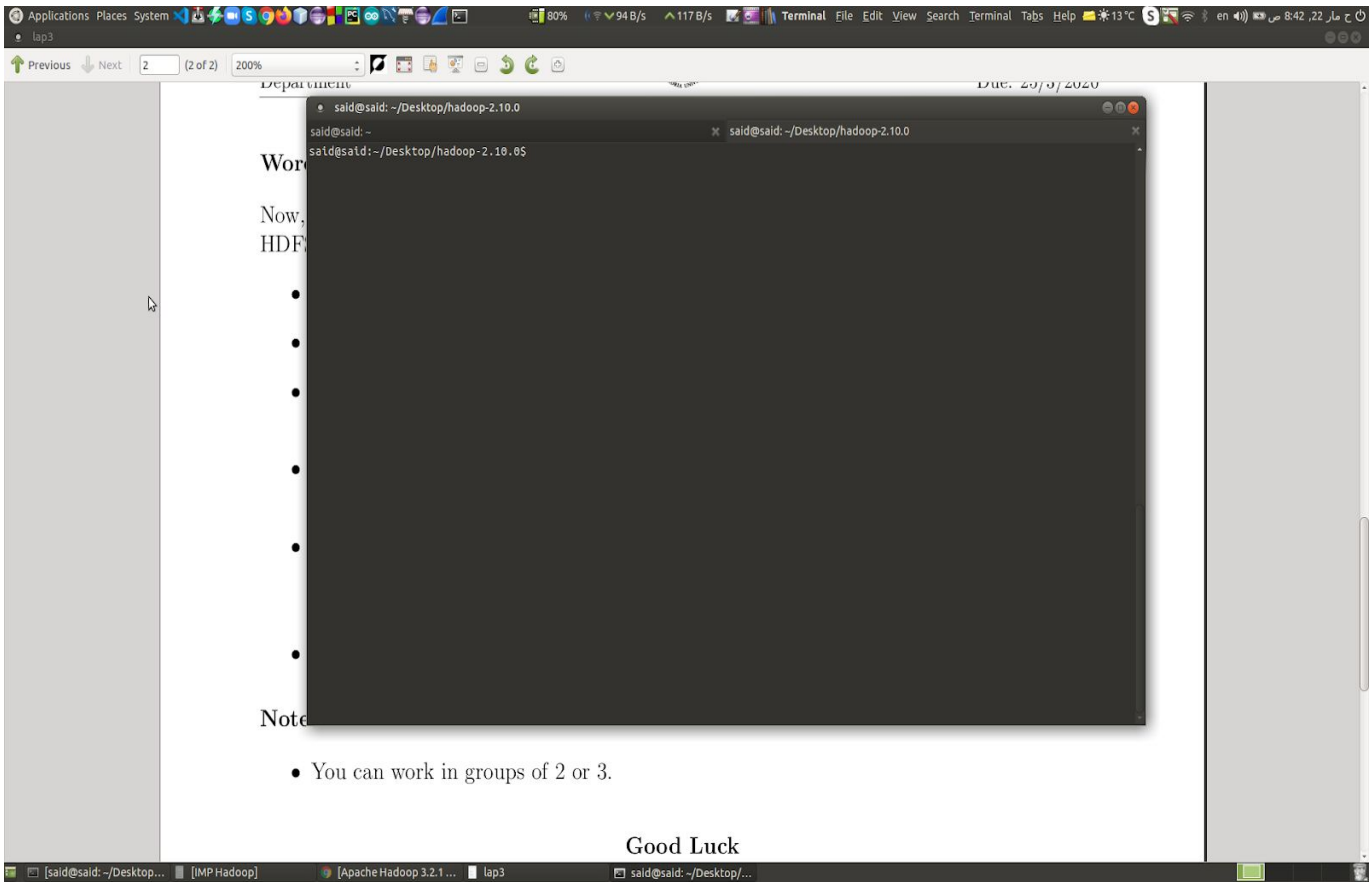




Notes

- You can work in groups of 2 or 3.

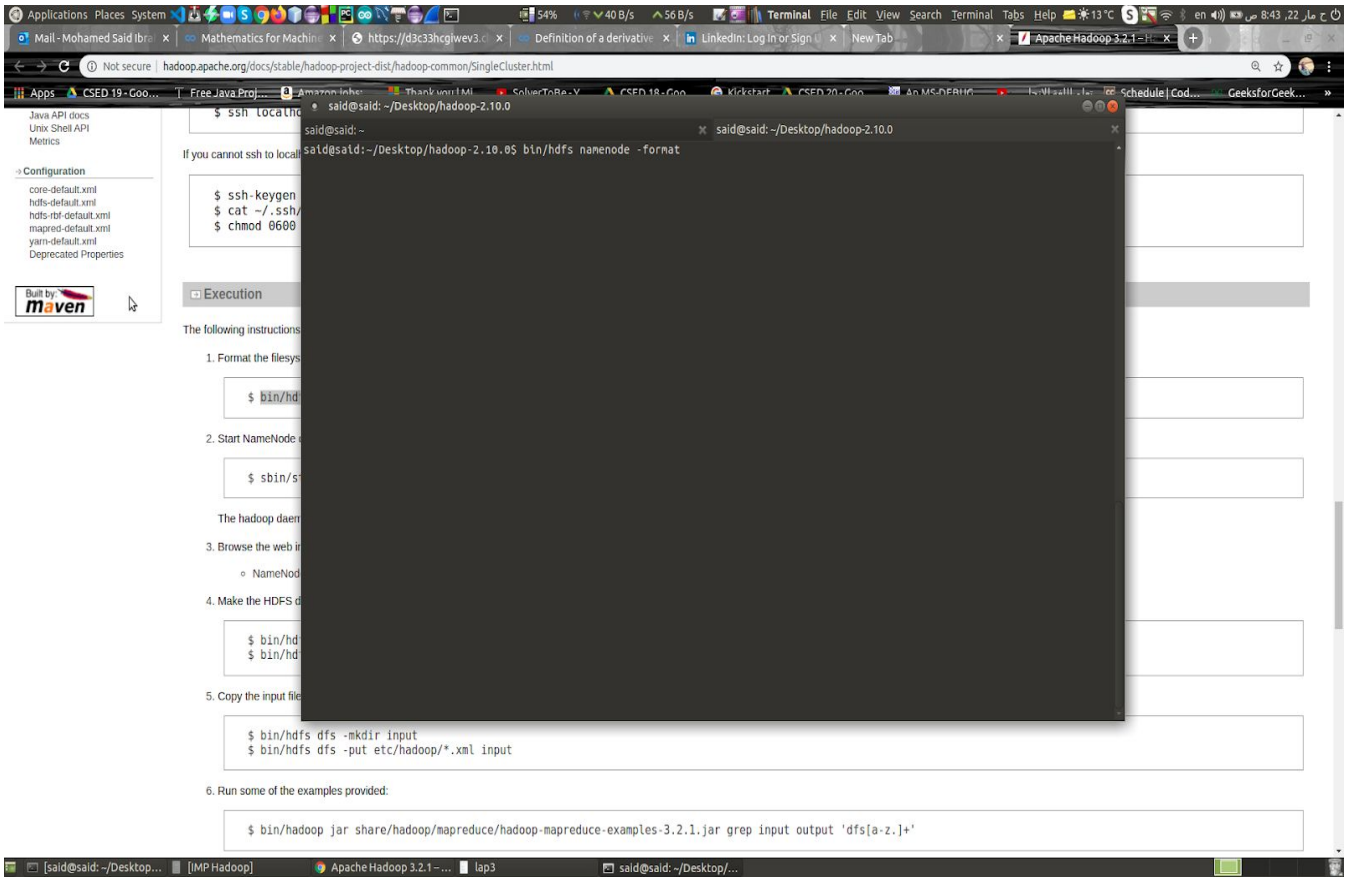
Good Luck

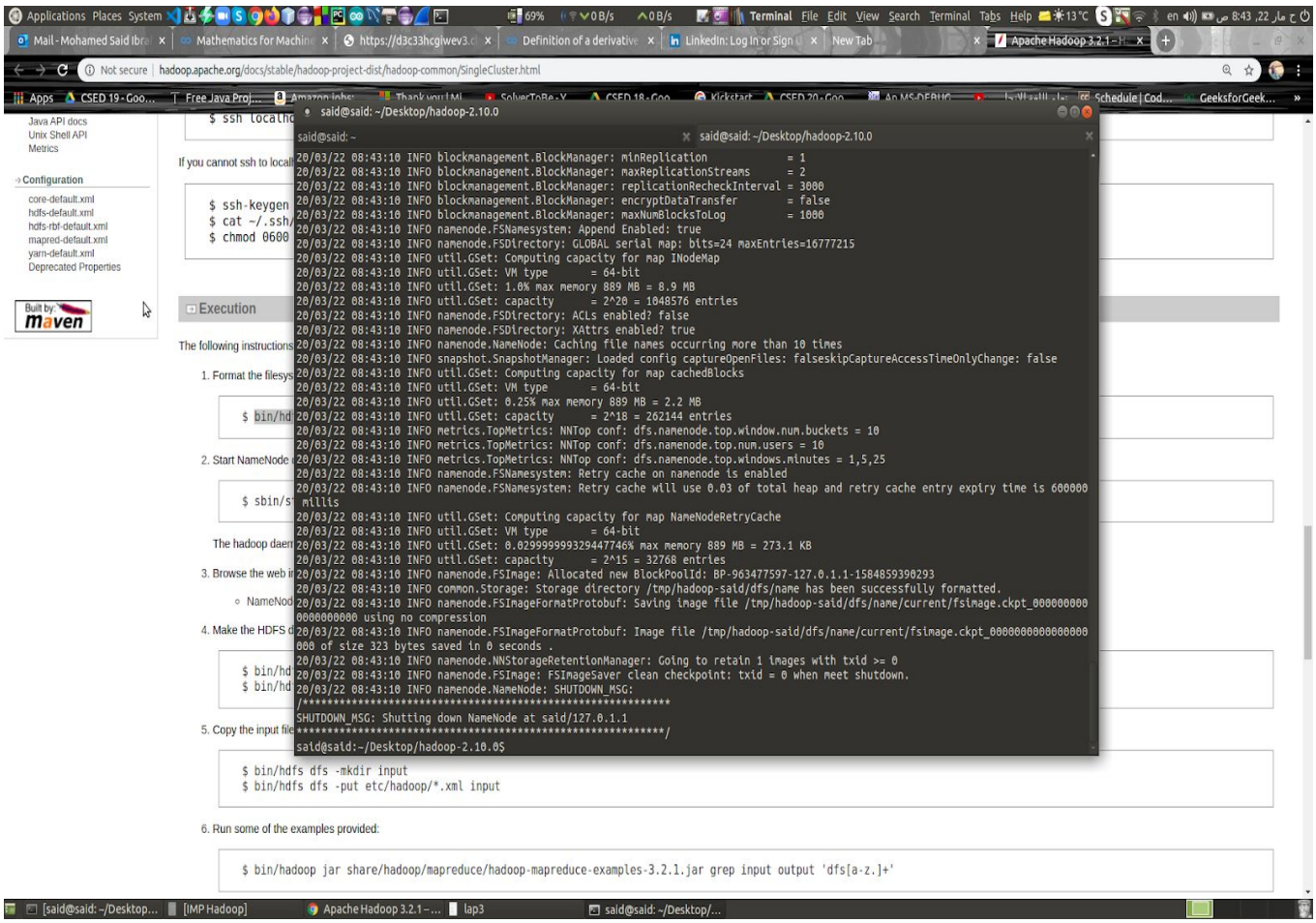


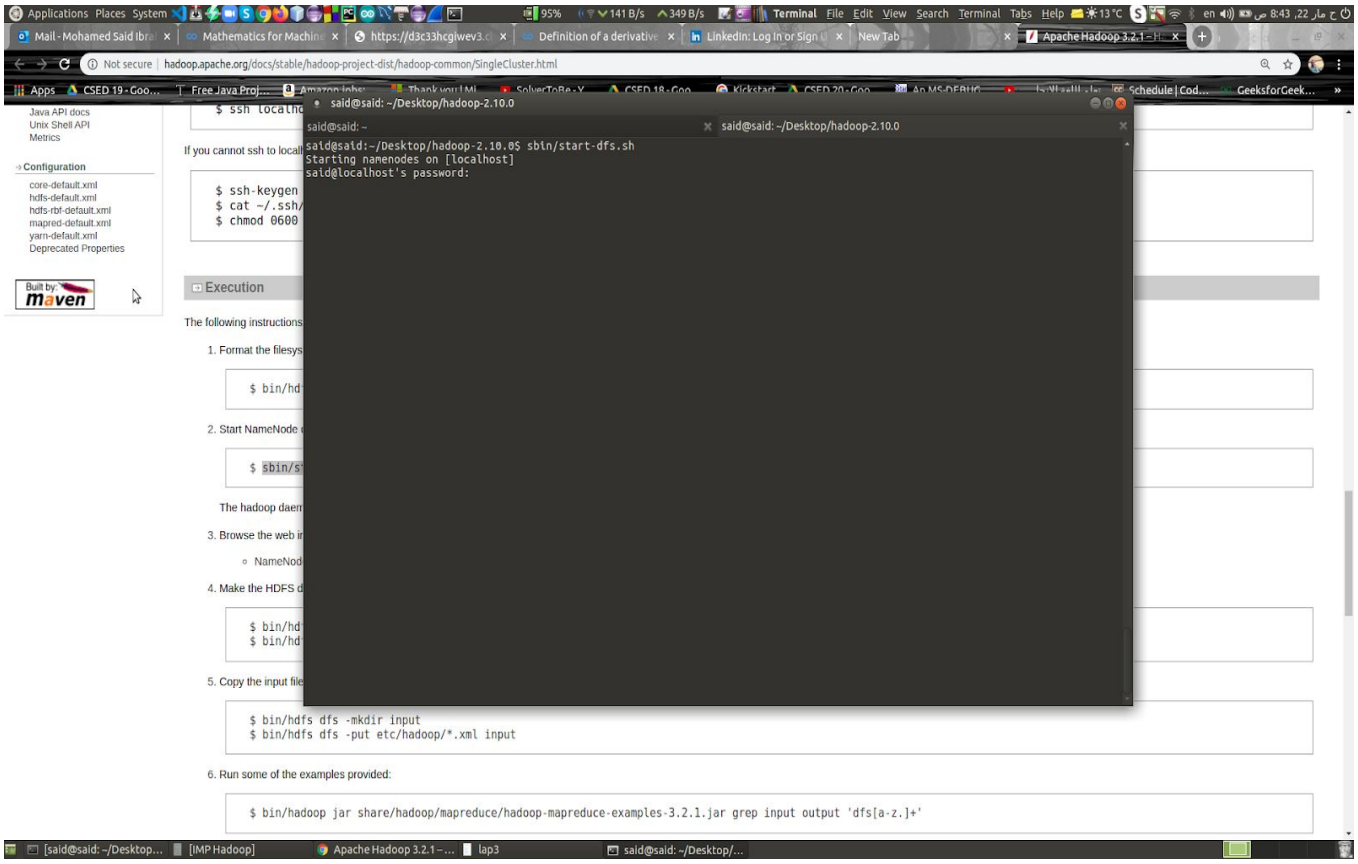
Notes

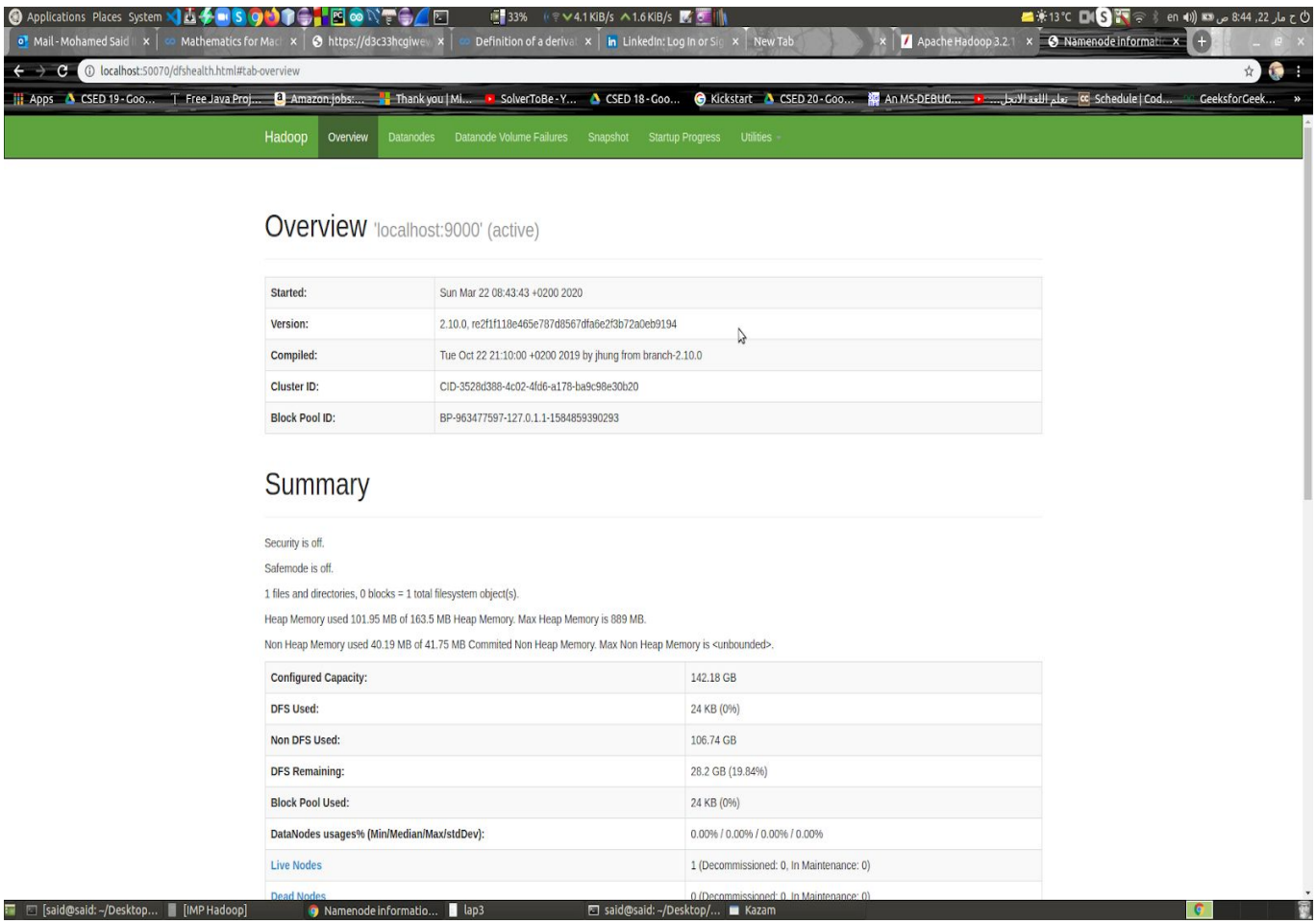
- You can work in groups of 2 or 3.

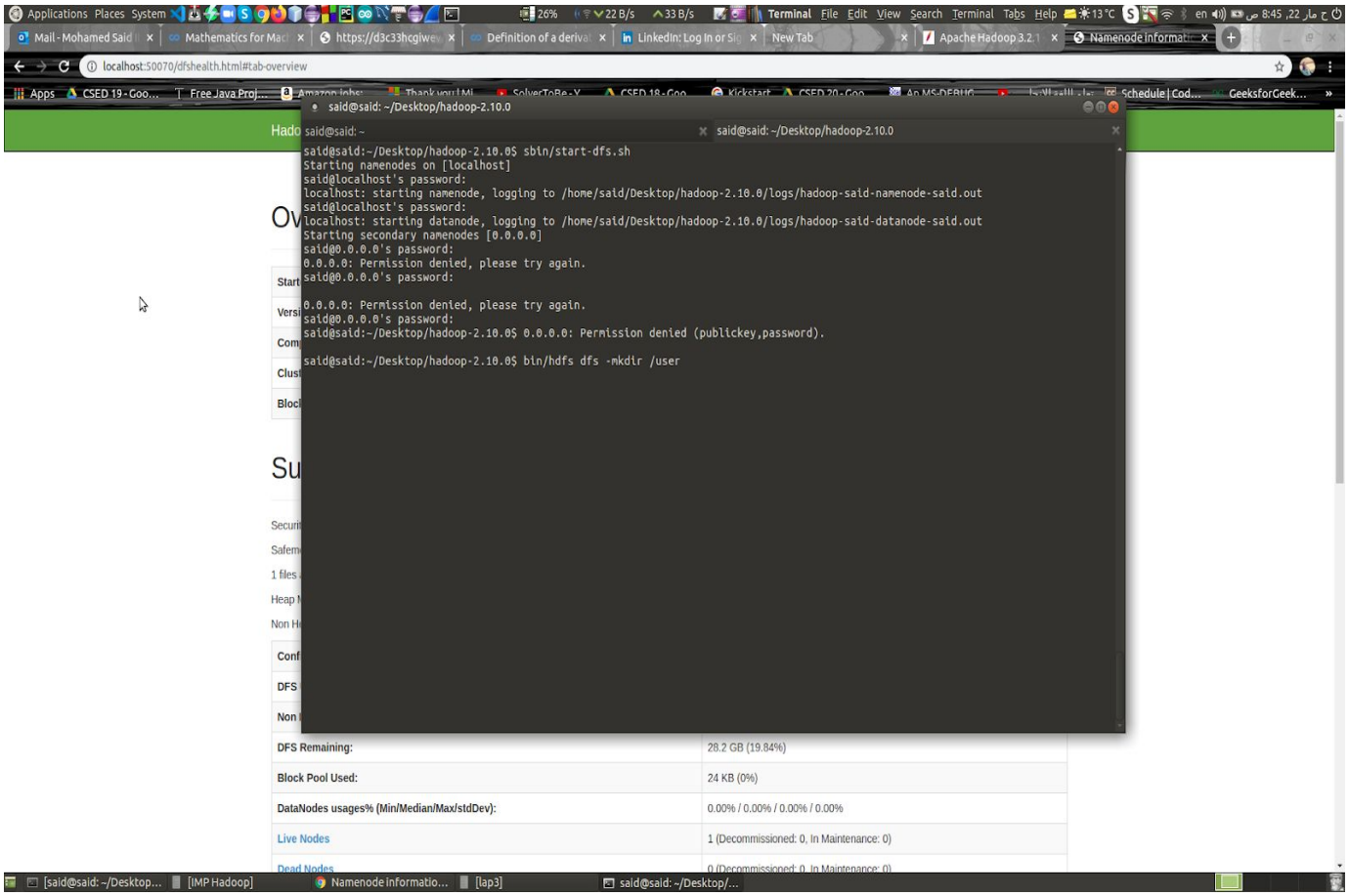
Good Luck

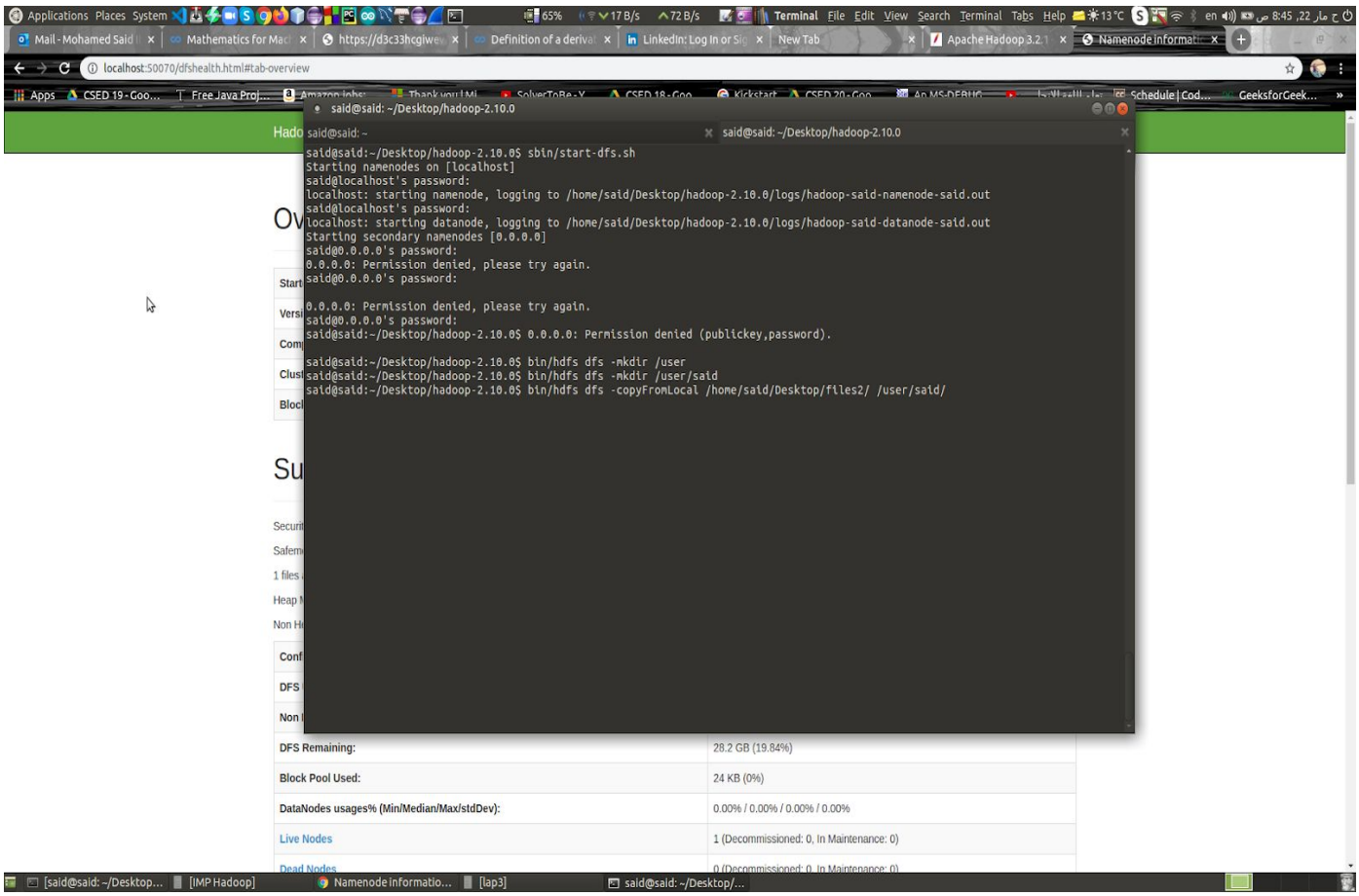

said@said: ~/Desktop/hadoop-2.10.0
said@said: ~
said@said: ~/Desktop/hadoop-2.10.0
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs namenode -format

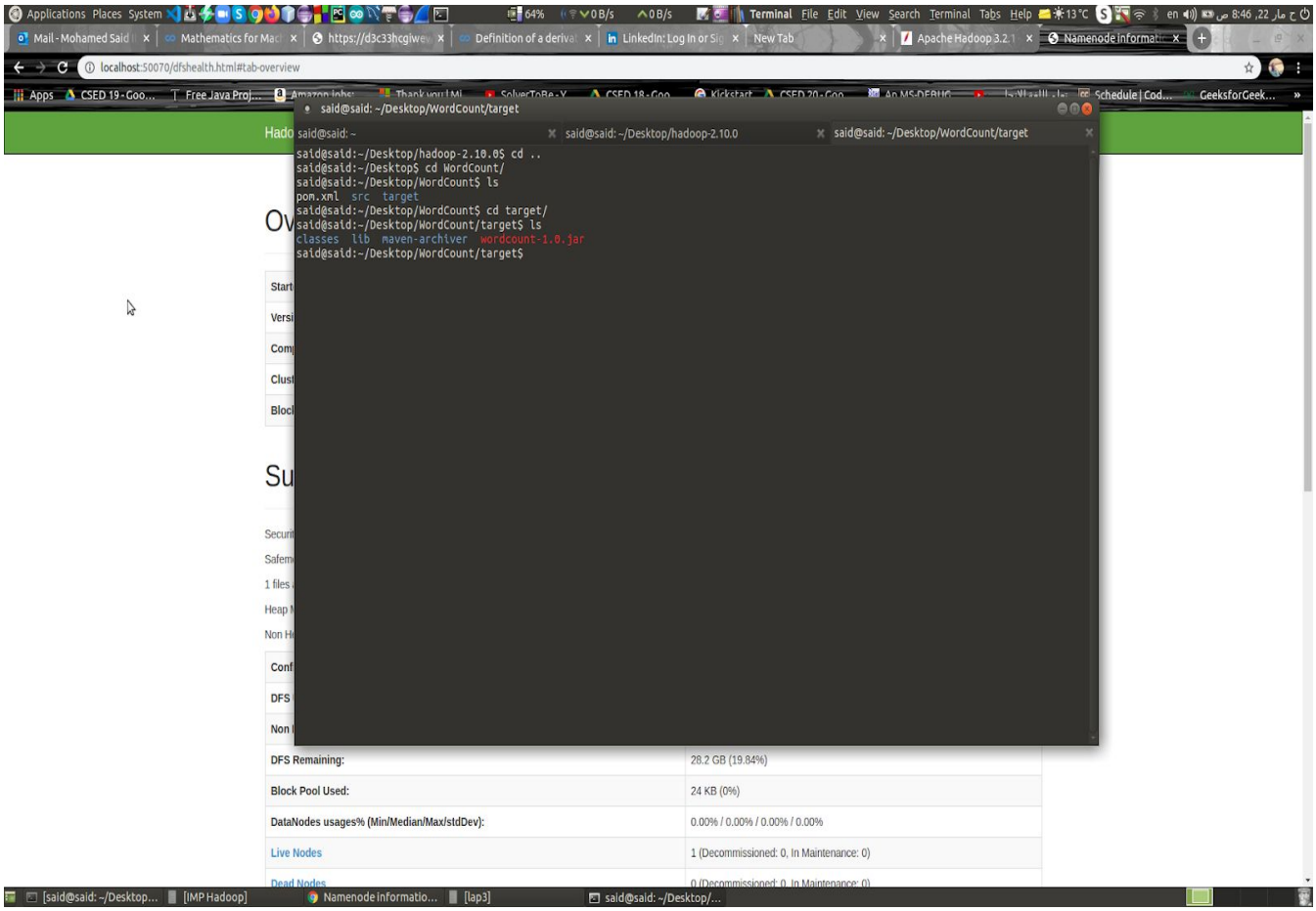

said@said: ~/Desktop/hadoop-2.10.0
said@said: ~
said@said: ~/Desktop/hadoop-2.10.0
20/03/22 08:43:10 INFO blockmanagement.BlockManager: minReplication = 1
20/03/22 08:43:10 INFO blockmanagement.BlockManager: maxReplicationStreams = 2
20/03/22 08:43:10 INFO blockmanagement.BlockManager: replicationRecheckInterval = 3000
20/03/22 08:43:10 INFO blockmanagement.BlockManager: encryptDataTransfer = false
20/03/22 08:43:10 INFO blockmanagement.BlockManager: maxNumBlocksToLog = 1000
20/03/22 08:43:10 INFO namenode.FSNamesystem: Append Enabled: true
20/03/22 08:43:10 INFO namenode.FSDirectory: GLOBAL serial map: bits=24 maxEntries=16777215
20/03/22 08:43:10 INFO util.GSet: Computing capacity for map INodeMap
20/03/22 08:43:10 INFO util.GSet: VM type = 64-bit
20/03/22 08:43:10 INFO util.GSet: 1.0% max memory 889 MB = 8.9 MB
20/03/22 08:43:10 INFO util.GSet: capacity = 2^20 = 1048576 entries
20/03/22 08:43:10 INFO namenode.FSDirectory: ACLs enabled? false
20/03/22 08:43:10 INFO namenode.FSDirectory: XAttrs enabled? true
20/03/22 08:43:10 INFO namenode.NameNode: Caching file names occurring more than 10 times
20/03/22 08:43:10 INFO snapshot.SnapshotManager: Loaded config captureOpenFiles: falseskipCaptureAccessTimeOnlyChange: false
20/03/22 08:43:10 INFO util.GSet: Computing capacity for map cachedBlocks
20/03/22 08:43:10 INFO util.GSet: VM type = 64-bit
20/03/22 08:43:10 INFO util.GSet: 0.25% max memory 889 MB = 2.2 MB
20/03/22 08:43:10 INFO util.GSet: capacity = 2^18 = 262144 entries
20/03/22 08:43:10 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.window.num.buckets = 10
20/03/22 08:43:10 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.num.users = 10
20/03/22 08:43:10 INFO metrics.TopMetrics: NNTop conf: dfs.namenode.top.windows.minutes = 1,5,25
20/03/22 08:43:10 INFO namenode.FSNamesystem: Retry cache on namenode is enabled
20/03/22 08:43:10 INFO namenode.FSNamesystem: Retry cache will use 0.03 of total heap and retry cache entry expiry time is 600000 millis
20/03/22 08:43:10 INFO util.GSet: Computing capacity for map NameNodeRetryCache
20/03/22 08:43:10 INFO util.GSet: VM type = 64-bit
20/03/22 08:43:10 INFO util.GSet: 0.029999999329447746% max memory 889 MB = 273.1 KB
20/03/22 08:43:10 INFO util.GSet: capacity = 2^15 = 32768 entries
20/03/22 08:43:10 INFO namenode.FSImage: Allocated new BlockPoolId: BP-963477597-127.0.1.1-1584859390293
20/03/22 08:43:10 INFO common.Storage: Storage directory /tmp/hadoop-said/dfs/name has been successfully formatted.
20/03/22 08:43:10 INFO namenode.FSImageFormatProtobuf: Saving image file /tmp/hadoop-said/dfs/name/current/fsimage.ckpt_0000000000000000000 using no compression
20/03/22 08:43:10 INFO namenode.FSImageFormatProtobuf: Image file /tmp/hadoop-said/dfs/name/current/fsimage.ckpt_0000000000000000000 of size 323 bytes saved in 0 seconds .
20/03/22 08:43:10 INFO namenode.NNStorageRetentionManager: Going to retain 1 images with txid >= 0
20/03/22 08:43:10 INFO namenode.FSImage: FSImageSaver clean checkpoint: txid = 0 when meet shutdown.
20/03/22 08:43:10 INFO namenode.NameNode: SHUTDOWN_MSG:
/**
SHUTDOWN_MSG: Shutting down NameNode at said/127.0.1.1
**/
said@said:~/Desktop/hadoop-2.10.0$

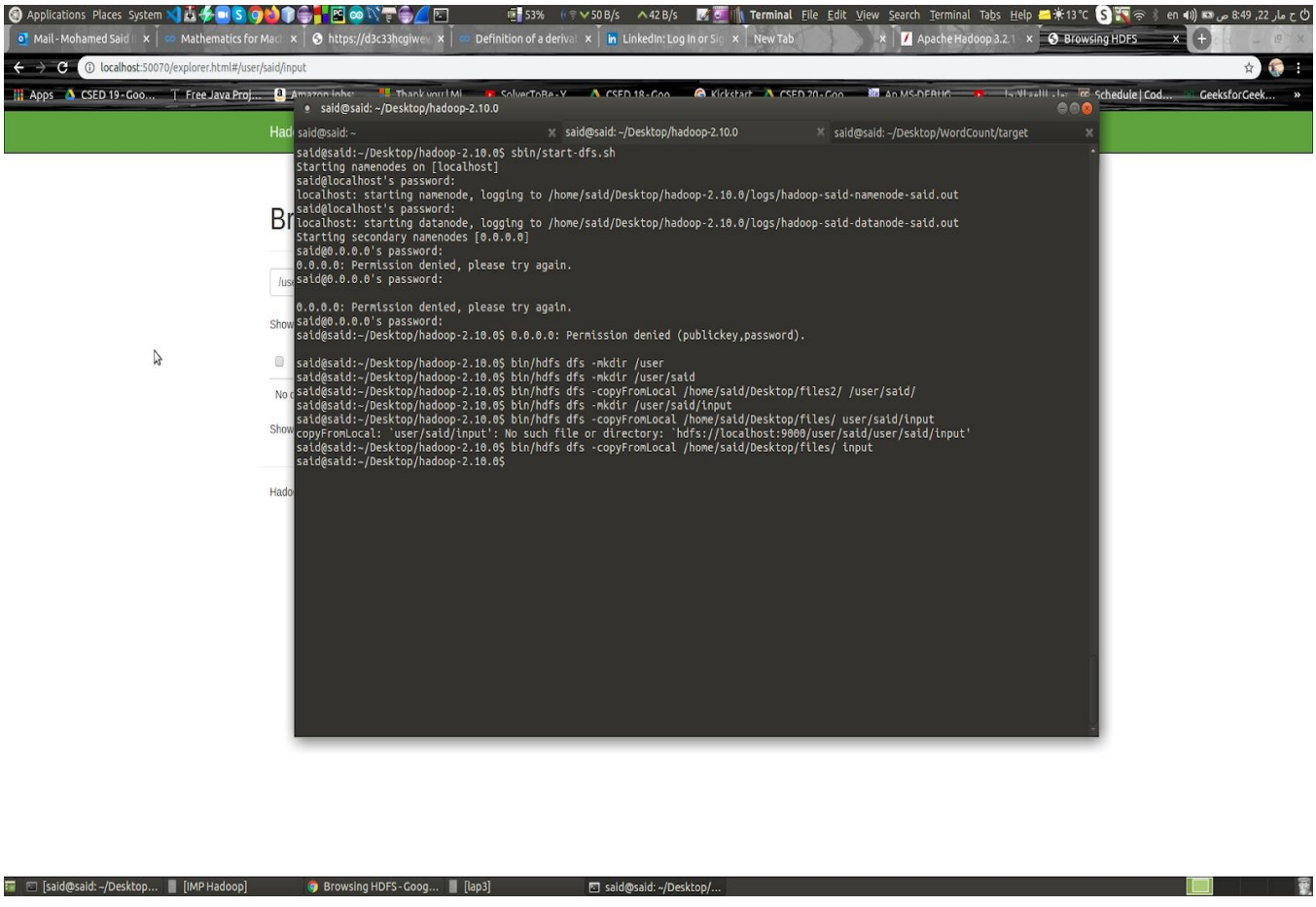

said@said: ~/Desktop/hadoop-2.10.0
said@said: ~
said@said: ~/Desktop/hadoop-2.10.0
said@said:~/Desktop/hadoop-2.10.0$ sbin/start-dfs.sh
Starting namenodes on [localhost]
said@localhost's password:
Execution
$ bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-3.2.1.jar grep input output 'dfs[a-z.]+'

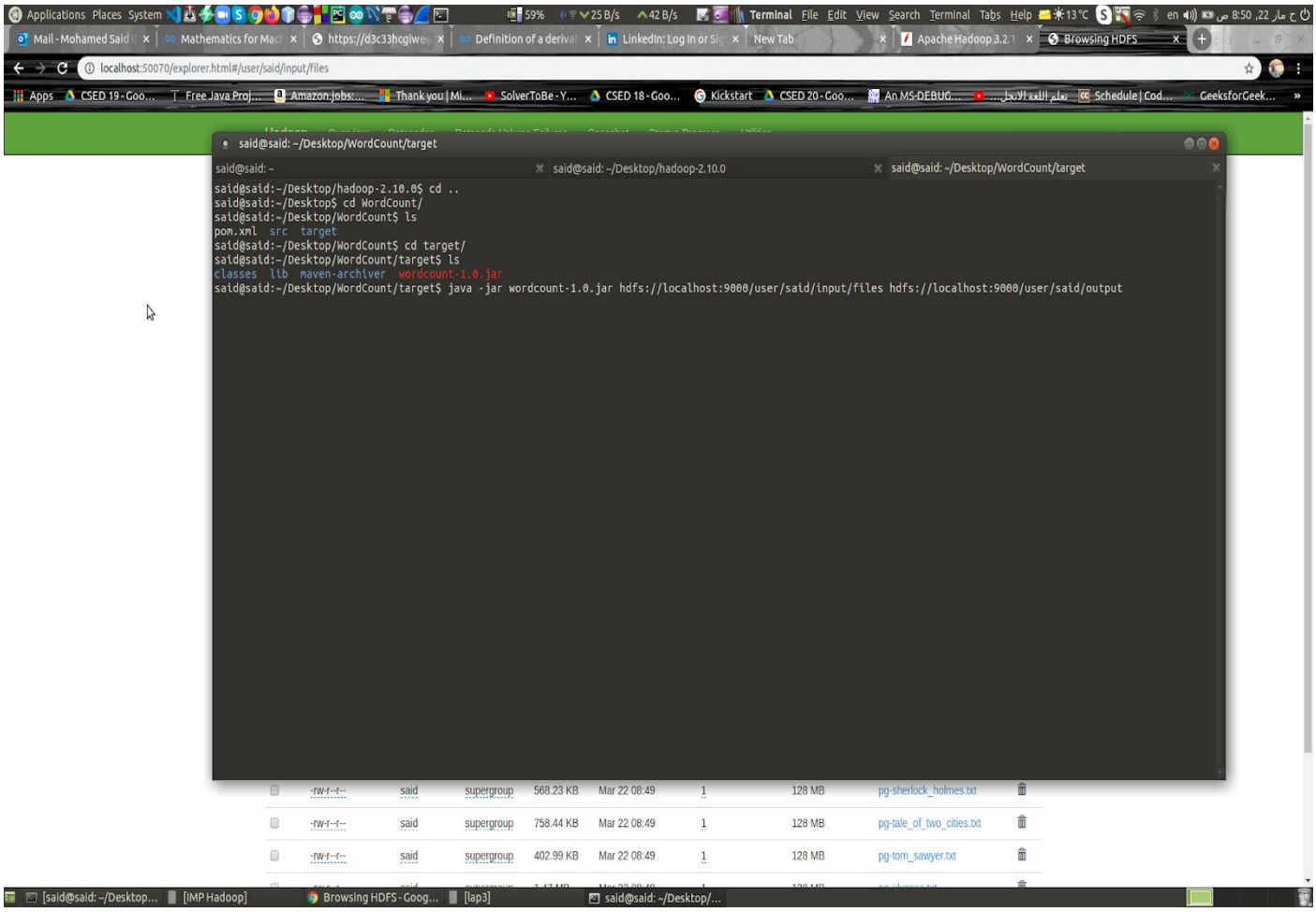

Hadoop
Overview
Datanodes
Datanode Volume Failures
Snapshot
Startup Progress
Utilities
Overview 'localhost:9000' (active)
Started: Sun Mar 22 08:43:43 +0200 2020
Version: 2.10.0, re2f1f118e465e787d8567dfa6e2f3b72a0eb9194
Compiled: Tue Oct 22 21:10:00 +0200 2019 by jhung from branch-2.10.0
Cluster ID: CID-3528d388-4c02-4fd6-a178-ba9c98e30b20
Block Pool ID: BP-963477597-127.0.1.1-1584859390293
Summary
Security is off.
Safemode is off.
1 files and directories, 0 blocks = 1 total filesystem object(s).
Heap Memory used 101.95 MB of 163.5 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 40.19 MB of 41.75 MB Commited Non Heap Memory. Max Non Heap Memory is <unbounded>.
Configured Capacity: 142.18 GB
DFS Used: 24 KB (0%)
Non DFS Used: 106.74 GB
DFS Remaining: 28.2 GB (19.84%)
Block Pool Used: 24 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev): 0.00% / 0.00% / 0.00% / 0.00%
Live Nodes 1 (Decommissioned: 0, In Maintenance: 0)

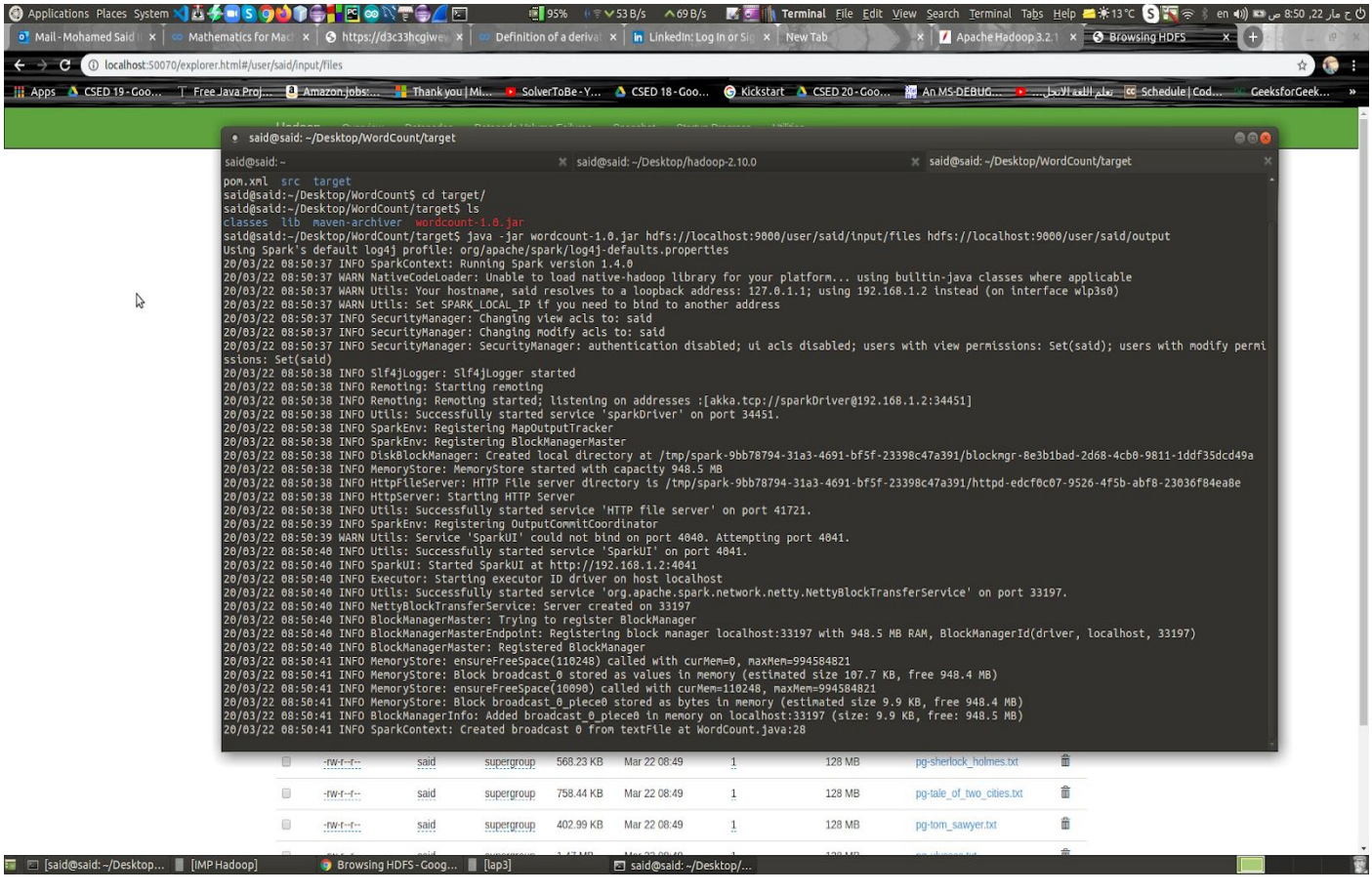

said@said: ~/Desktop/hadoop-2.10.0
said@said:~/Desktop/hadoop-2.10.0$ sbin/start-dfs.sh
Starting namenodes on [localhost]
said@localhost's password:
localhost: starting namenode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-namenode-said.out
said@localhost's password:
localhost: starting datanode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-datanode-said.out
Starting secondary namenodes [0.0.0.0]
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
said@said:~/Desktop/hadoop-2.10.0$ 0.0.0.0: Permission denied (publickey,password).
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user
DFS Remaining: 28.2 GB (19.84%)
Block Pool Used: 24 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev): 0.00% / 0.00% / 0.00% / 0.00%
Live Nodes 1 (Decommissioned: 0, In Maintenance: 0)

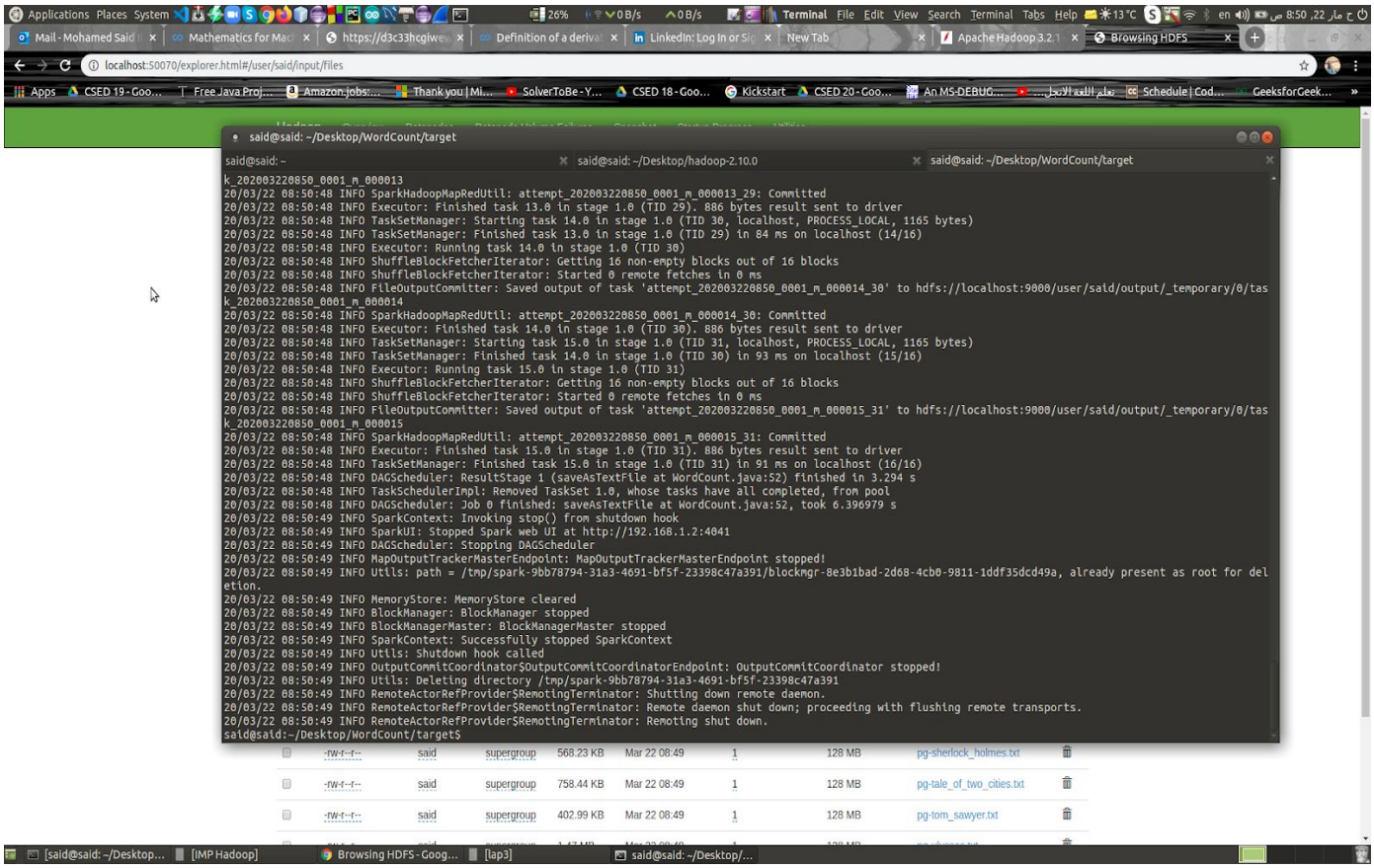

said@said: ~/Desktop/hadoop-2.10.0
said@said:~/Desktop/hadoop-2.10.0$ sbin/start-dfs.sh
Starting namenodes on [localhost]
said@localhost's password:
localhost: starting namenode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-namenode-said.out
said@localhost's password:
localhost: starting datanode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-datanode-said.out
Starting secondary namenodes [0.0.0.0]
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
said@said:~/Desktop/hadoop-2.10.0$ 0.0.0.0: Permission denied (publickey,password).
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user/said
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -copyFromLocal /home/said/Desktop/files2/ /user/said/
DFS Remaining: 28.2 GB (19.84%)
Block Pool Used: 24 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev): 0.00% / 0.00% / 0.00% / 0.00%
Live Nodes 1 (Decommissioned: 0, In Maintenance: 0)

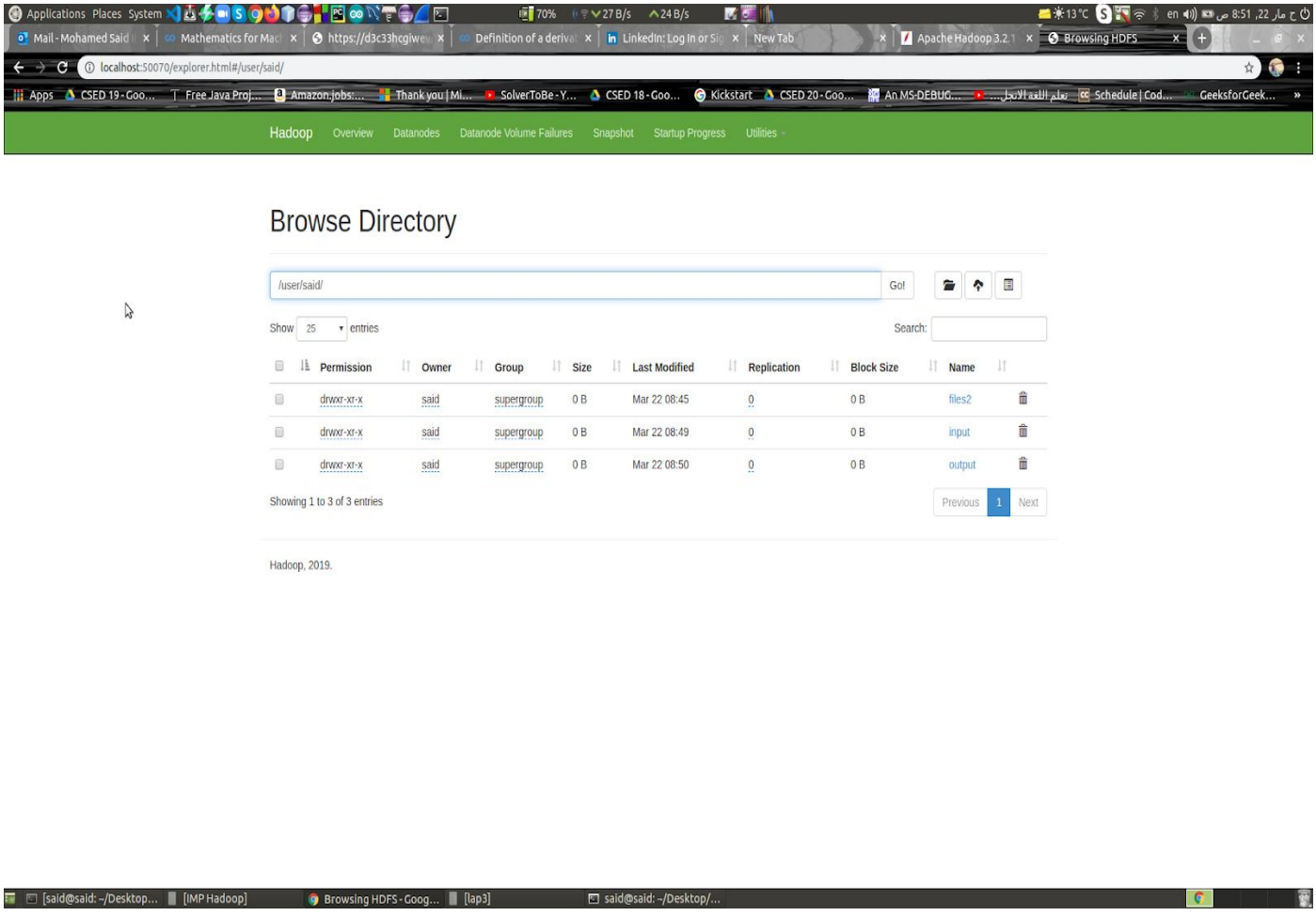

said@said: ~/Desktop/WordCount/target
said@said:~/Desktop/hadoop-2.10.0$ cd ..
said@said:~/Desktop$ cd WordCount/
said@said:~/Desktop/WordCount$ ls
pom.xml src target
said@said:~/Desktop/WordCount$ cd target/
said@said:~/Desktop/WordCount/target$ ls
classes lib maven-archiver wordcount-1.0.jar
said@said:~/Desktop/WordCount/target$
DFS Remaining: 28.2 GB (19.84%)
Block Pool Used: 24 KB (0%)
DataNodes usages% (Min/Median/Max/stdDev): 0.00% / 0.00% / 0.00% / 0.00%
Live Nodes 1 (Decommissioned: 0, In Maintenance: 0)

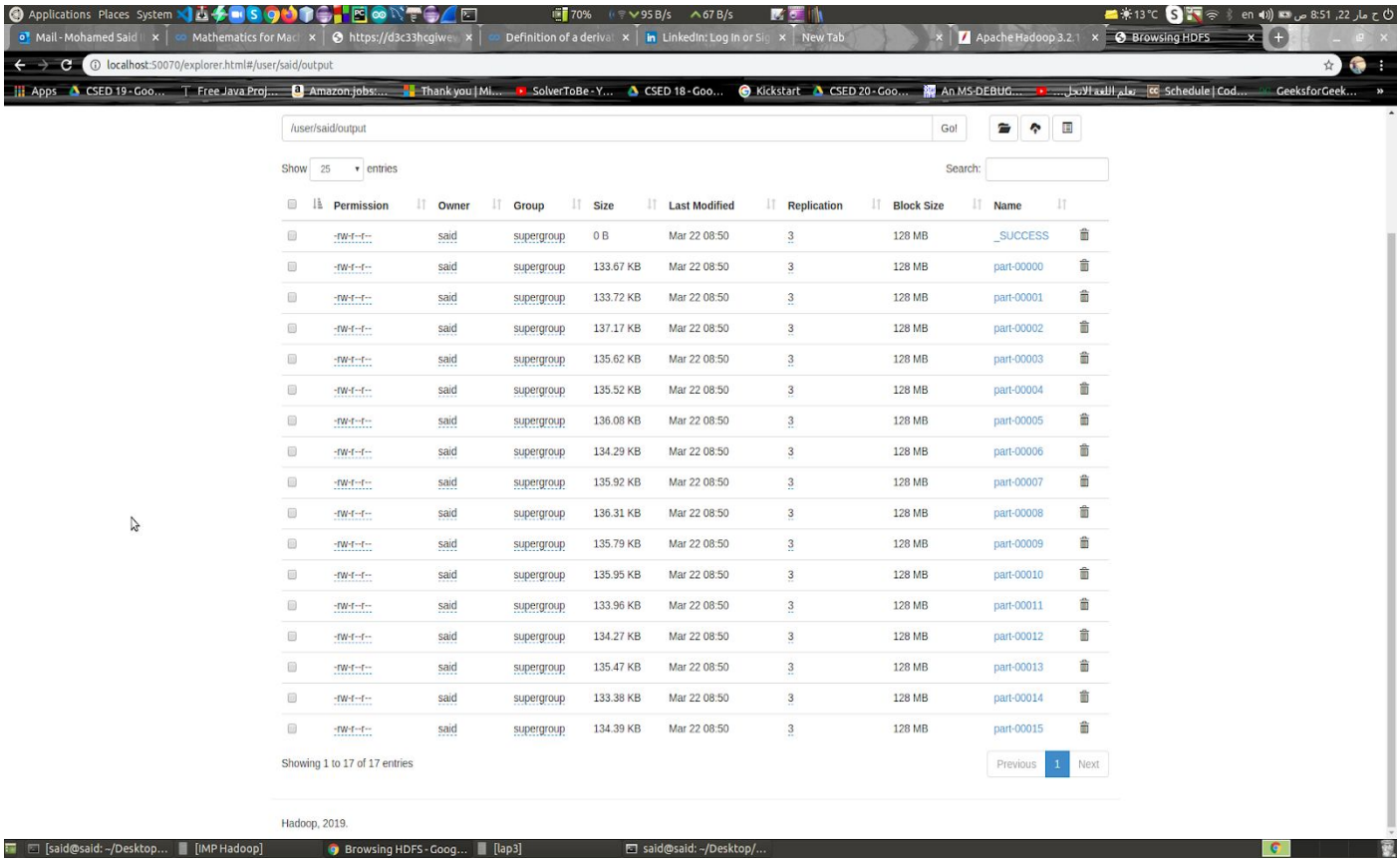

said@said: ~/Desktop/hadoop-2.10.0
said@said:~/Desktop/hadoop-2.10.0$ sbin/start-dfs.sh
Starting namenodes on [localhost]
said@localhost's password:
localhost: starting namenode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-namenode-said.out
said@localhost's password:
localhost: starting datanode, logging to /home/said/Desktop/hadoop-2.10.0/logs/hadoop-said-datanode-said.out
Starting secondary namenodes [0.0.0.0]
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
0.0.0.0: Permission denied, please try again.
said@0.0.0.0's password:
said@said:~/Desktop/hadoop-2.10.0$ 0.0.0.0: Permission denied (publickey,password).
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user/said
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -copyFromLocal /home/said/Desktop/files2/ /user/said/
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -mkdir /user/said/input
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -copyFromLocal /home/said/Desktop/files/ user/said/input
copyFromLocal: `user/said/input': No such file or directory: `hdfs://localhost:9000/user/said/user/said/input'
said@said:~/Desktop/hadoop-2.10.0$ bin/hdfs dfs -copyFromLocal /home/said/Desktop/files/ input
said@said:~/Desktop/hadoop-2.10.0$

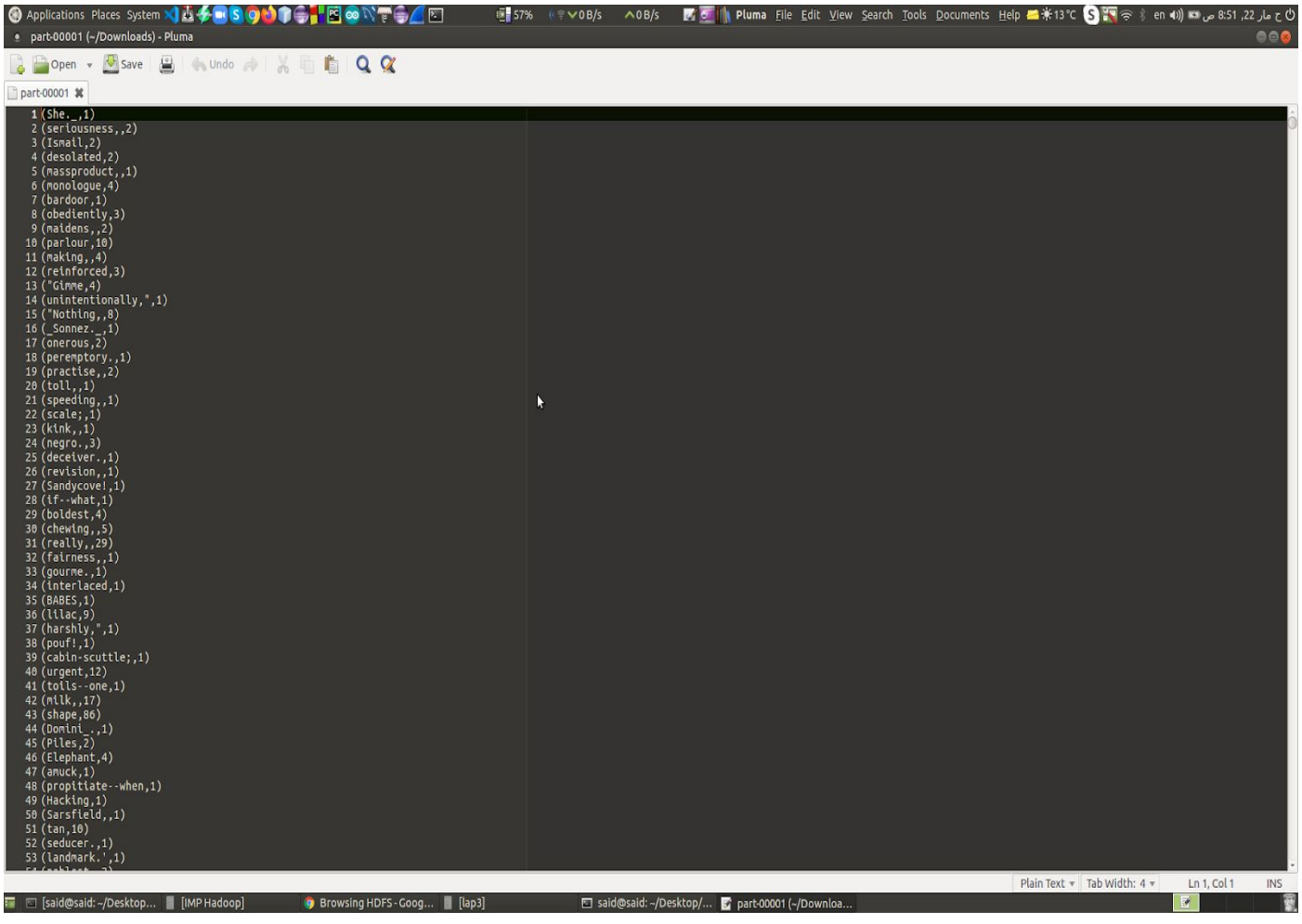

said@said: ~/Desktop/WordCount/target
said@said:~/Desktop/hadoop-2.10.0$ cd ..
said@said:~/Desktop$ cd WordCount/
said@said:~/Desktop/WordCount$ ls
pom.xml src target
said@said:~/Desktop/WordCount$ cd target/
said@said:~/Desktop/WordCount/target$ ls
classes lib maven-archiver wordcount-1.0.jar
said@said:~/Desktop/WordCount/target$ java -jar wordcount-1.0.jar hdfs://localhost:9000/user/said/input/files hdfs://localhost:9000/user/said/output

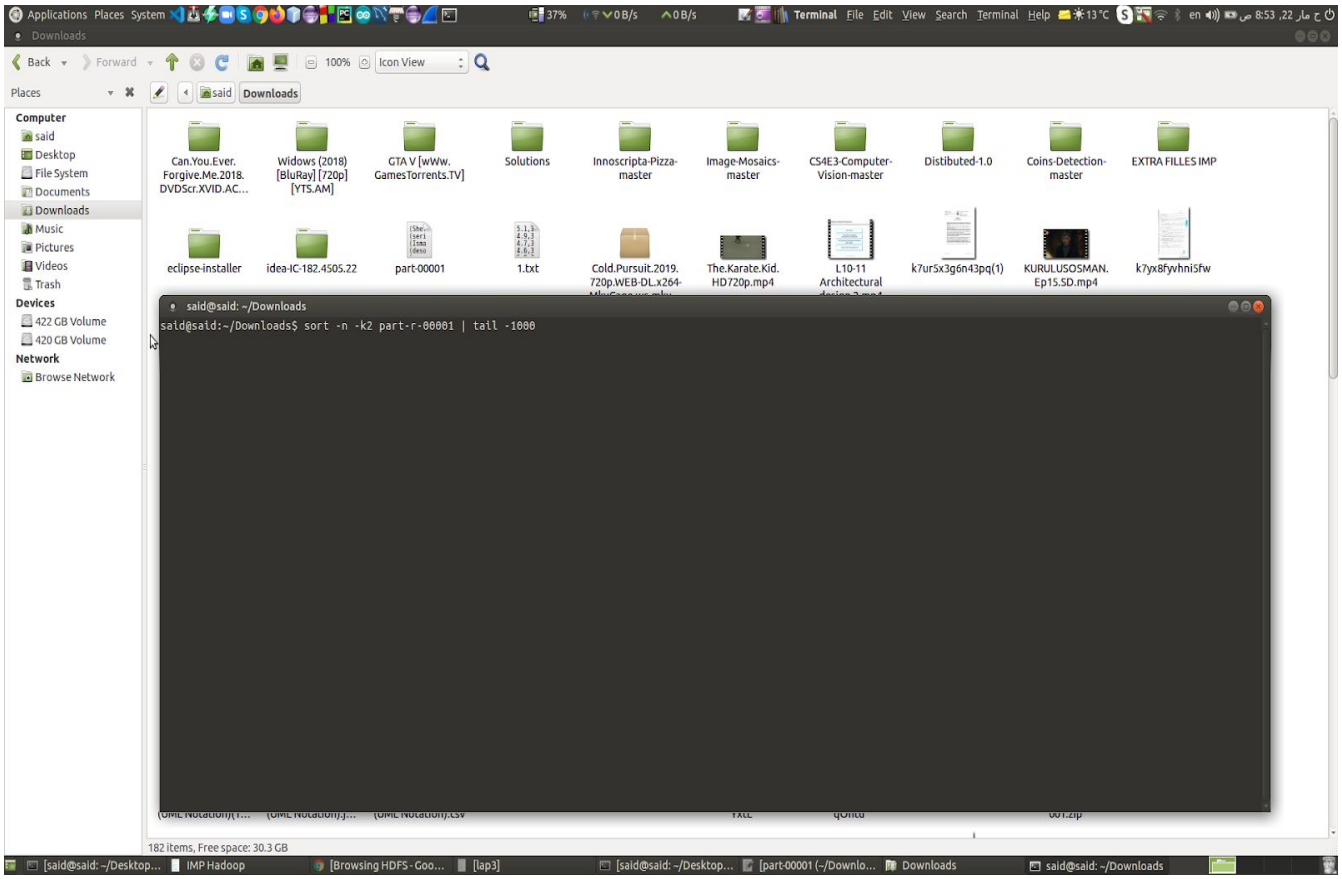

said@said: ~/Desktop/WordCount/target
pom.xml src target
said@said:~/Desktop/WordCount$ cd target/
said@said:~/Desktop/WordCount/target$ ls
classes lib maven-archiver wordcount-1.0.jar
said@said:~/Desktop/WordCount/target$ java -jar wordcount-1.0.jar hdfs://localhost:9000/user/said/input/files hdfs://localhost:9000/user/said/output
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
20/03/22 08:50:37 INFO SparkContext: Running Spark version 1.4.0
20/03/22 08:50:37 WARN NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
20/03/22 08:50:37 WARN Utils: Your hostname, said resolves to a loopback address: 127.0.1.1; using 192.168.1.2 instead (on interface wlp3s0)
20/03/22 08:50:37 WARN Utils: Set SPARK_LOCAL_IP if you need to bind to another address
20/03/22 08:50:37 INFO SecurityManager: Changing view acls to: said
20/03/22 08:50:37 INFO SecurityManager: Changing modify acls to: said
20/03/22 08:50:37 INFO SecurityManager: SecurityManager: authentication disabled; ui acls disabled; users with view permissions: Set(said); users with modify permissions: Set(said)
20/03/22 08:50:38 INFO Slf4jLogger: Slf4jLogger started
20/03/22 08:50:38 INFO Remoting: Starting remoting
20/03/22 08:50:38 INFO Remoting: Remoting started; listening on addresses :[akka.tcp://sparkDriver@192.168.1.2:34451]
20/03/22 08:50:38 INFO Utils: Successfully started service 'sparkDriver' on port 34451.
20/03/22 08:50:38 INFO SparkEnv: Registering MapOutputTracker
20/03/22 08:50:38 INFO SparkEnv: Registering BlockManagerMaster
20/03/22 08:50:38 INFO DiskBlockManager: Created local directory at /tmp/spark-9bb78794-31a3-4691-bf5f-23398c47a391/blockmgr-8e3b1bad-2d68-4cb0-9811-1ddf35dcd49a
20/03/22 08:50:38 INFO MemoryStore: MemoryStore started with capacity 948.5 MB
20/03/22 08:50:38 INFO HttpFileServer: HTTP File server directory is /tmp/spark-9bb78794-31a3-4691-bf5f-23398c47a391/httpd-edcf0c07-9526-4f5b-abf8-23036f84ea8e
20/03/22 08:50:38 INFO HttpServer: Starting HTTP Server
20/03/22 08:50:38 INFO Utils: Successfully started service 'HTTP file server' on port 41721.
20/03/22 08:50:39 INFO SparkEnv: Registering OutputCommitCoordinator
20/03/22 08:50:39 WARN Utils: Service 'SparkUI' could not bind on port 4040. Attempting port 4041.
20/03/22 08:50:40 INFO Utils: Successfully started service 'SparkUI' on port 4041.
20/03/22 08:50:40 INFO SparkUI: Started SparkUI at http://192.168.1.2:4041
20/03/22 08:50:40 INFO Executor: Starting executor ID driver on host localhost
20/03/22 08:50:40 INFO Utils: Successfully started service 'org.apache.spark.network.netty.NettyBlockTransferService' on port 33197.
20/03/22 08:50:40 INFO NettyBlockTransferService: Server created on 33197
20/03/22 08:50:40 INFO BlockManagerMaster: Trying to register BlockManager
20/03/22 08:50:40 INFO BlockManagerMasterEndpoint: Registering block manager localhost:33197 with 948.5 MB RAM, BlockManagerId(driver, localhost, 33197)
20/03/22 08:50:40 INFO BlockManagerMaster: Registered BlockManager
20/03/22 08:50:41 INFO MemoryStore: ensureFreeSpace(110248) called with curMem=0, maxMem=994584821
20/03/22 08:50:41 INFO MemoryStore: Block broadcast_0 stored as values in memory (estimated size 107.7 KB, free 948.4 MB)
20/03/22 08:50:41 INFO MemoryStore: ensureFreeSpace(10090) called with curMem=110248, maxMem=994584821
20/03/22 08:50:41 INFO MemoryStore: Block broadcast_0_piece0 stored as bytes in memory (estimated size 9.9 KB, free 948.4 MB)
20/03/22 08:50:41 INFO BlockManagerInfo: Added broadcast_0_piece0 in memory on localhost:33197 (size: 9.9 KB, free: 948.5 MB)
20/03/22 08:50:41 INFO SparkContext: Created broadcast 0 from textFile at WordCount.java:28

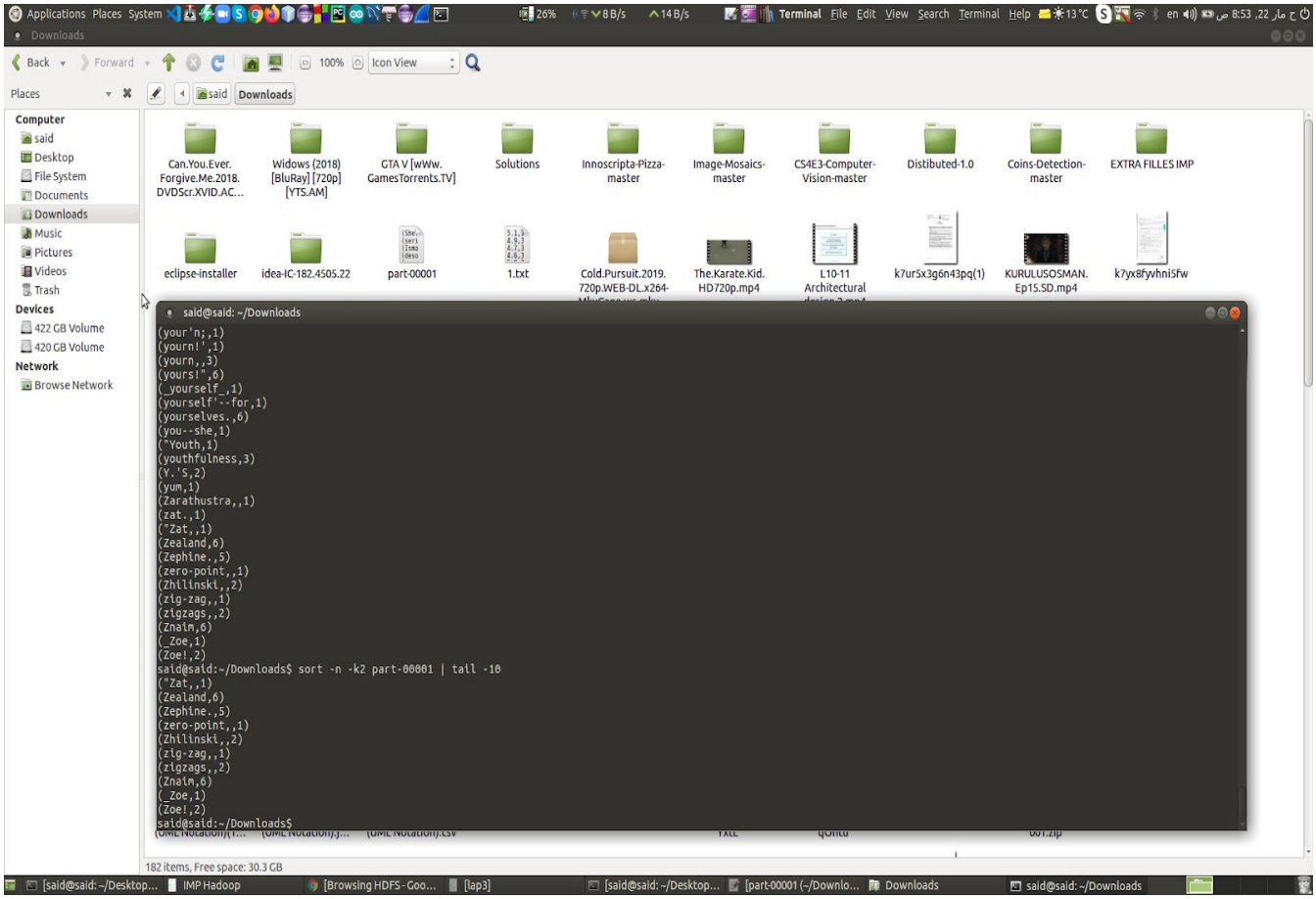

said@said: ~/Desktop/WordCount/target
said@said: ~
said@said: ~/Desktop/hadoop-2.10.0
said@said: ~/Desktop/WordCount/target
20/03/22 08:50:48 INFO SparkHadoopMapRedUtil: attempt_202003220850_0001_m_000013_29: Committed
20/03/22 08:50:48 INFO Executor: Finished task 13.0 in stage 1.0 (TID 29). 886 bytes result sent to driver
20/03/22 08:50:48 INFO TaskSetManager: Starting task 14.0 in stage 1.0 (TID 30, localhost, PROCESS_LOCAL, 1165 bytes)
20/03/22 08:50:48 INFO TaskSetManager: Finished task 13.0 in stage 1.0 (TID 29) in 84 ms on localhost (14/16)
20/03/22 08:50:48 INFO Executor: Running task 14.0 in stage 1.0 (TID 30)
20/03/22 08:50:48 INFO ShuffleBlockFetcherIterator: Getting 16 non-empty blocks out of 16 blocks
20/03/22 08:50:48 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 0 ms
20/03/22 08:50:48 INFO FileOutputCommitter: Saved output of task 'attempt_202003220850_0001_m_000014_30' to hdfs://localhost:9000/user/said/output/_temporary/0/tas
k_202003220850_0001_m_000014
20/03/22 08:50:48 INFO SparkHadoopMapRedUtil: attempt_202003220850_0001_m_000014_30: Committed
20/03/22 08:50:48 INFO Executor: Finished task 14.0 in stage 1.0 (TID 30). 886 bytes result sent to driver
20/03/22 08:50:48 INFO TaskSetManager: Starting task 15.0 in stage 1.0 (TID 31, localhost, PROCESS_LOCAL, 1165 bytes)
20/03/22 08:50:48 INFO TaskSetManager: Finished task 14.0 in stage 1.0 (TID 30) in 93 ms on localhost (15/16)
20/03/22 08:50:48 INFO Executor: Running task 15.0 in stage 1.0 (TID 31)
20/03/22 08:50:48 INFO ShuffleBlockFetcherIterator: Getting 16 non-empty blocks out of 16 blocks
20/03/22 08:50:48 INFO ShuffleBlockFetcherIterator: Started 0 remote fetches in 0 ms
20/03/22 08:50:48 INFO FileOutputCommitter: Saved output of task 'attempt_202003220850_0001_m_000015_31' to hdfs://localhost:9000/user/said/output/_temporary/0/tas
k_202003220850_0001_m_000015
20/03/22 08:50:48 INFO SparkHadoopMapRedUtil: attempt_202003220850_0001_m_000015_31: Committed
20/03/22 08:50:48 INFO Executor: Finished task 15.0 in stage 1.0 (TID 31). 886 bytes result sent to driver
20/03/22 08:50:48 INFO TaskSetManager: Finished task 15.0 in stage 1.0 (TID 31) in 91 ms on localhost (16/16)
20/03/22 08:50:48 INFO DAGScheduler: ResultStage 1 (saveAsTextFile at WordCount.java:52) finished in 3.294 s
20/03/22 08:50:48 INFO TaskSchedulerImpl: Removed TaskSet 1.0, whose tasks have all completed, from pool
20/03/22 08:50:48 INFO DAGScheduler: Job 0 finished: saveAsTextFile at WordCount.java:52, took 6.396979 s
20/03/22 08:50:49 INFO SparkContext: Invoking stop() from shutdown hook
20/03/22 08:50:49 INFO SparkUI: Stopped Spark web UI at http://192.168.1.2:4041
20/03/22 08:50:49 INFO DAGScheduler: Stopping DAGScheduler
20/03/22 08:50:49 INFO MapOutputTrackerMasterEndpoint: MapOutputTrackerMasterEndpoint stopped!
20/03/22 08:50:49 INFO Utils: path = /tmp/spark-9bb78794-31a3-4691-bf5f-23398c47a391/blockmgr-8e3b1bad-2d68-4cb0-9811-1ddf35dcd49a, already present as root for del
etion.
20/03/22 08:50:49 INFO MemoryStore: MemoryStore cleared
20/03/22 08:50:49 INFO BlockManager: BlockManager stopped
20/03/22 08:50:49 INFO BlockManagerMaster: BlockManagerMaster stopped
20/03/22 08:50:49 INFO SparkContext: Successfully stopped SparkContext
20/03/22 08:50:49 INFO Utils: Shutdown hook called
20/03/22 08:50:49 INFO OutputCommitCoordinator$OutputCommitCoordinatorEndpoint: OutputCommitCoordinator stopped!
20/03/22 08:50:49 INFO Utils: Deleting directory /tmp/spark-9bb78794-31a3-4691-bf5f-23398c47a391
20/03/22 08:50:49 INFO RemoteActorRefProvider$RemotingTerminator: Shutting down remote daemon.
20/03/22 08:50:49 INFO RemoteActorRefProvider$RemotingTerminator: Remote daemon shut down; proceeding with flushing remote transports.
20/03/22 08:50:49 INFO RemoteActorRefProvider$RemotingTerminator: Remoting shut down.
said@said:~/Desktop/WordCount/target$
-rw-r--r-- said supergroup 568.23 KB Mar 22 08:49 1 128 MB pg-sherlock_holmes.txt
-rw-r--r-- said supergroup 758.44 KB Mar 22 08:49 1 128 MB pg-tale_of_two_cities.txt
-rw-r--r-- said supergroup 402.99 KB Mar 22 08:49 1 128 MB pg-tom_sawyer.txt


localhost:50070/explorer.html#/user/said/
Hadoop Overview Datanodes Datanode Volume Failures Snapshot Startup Progress Utilities
Browse Directory
/user/said/
Go!
Show 25 entries
Search:
Permission Owner Group Size Last Modified Replication Block Size Name
drwxr-xr-x said supergroup 0 B Mar 22 08:45 0 0 B files2
drwxr-xr-x said supergroup 0 B Mar 22 08:49 0 0 B input
drwxr-xr-x said supergroup 0 B Mar 22 08:50 0 0 B output
Showing 1 to 3 of 3 entries
Previous 1 Next
Hadoop, 2019.


Browsing HDFS
localhost:50070/explorer.html#/user/said/output
/user/said/output
Go!
Show 25 entries
Search:
Permission Owner Group Size Last Modified Replication Block Size Name
-rw-r--r-- said supergroup 0 B Mar 22 08:50 3 128 MB _SUCCESS
-rw-r--r-- said supergroup 133.67 KB Mar 22 08:50 3 128 MB part-00000
-rw-r--r-- said supergroup 133.72 KB Mar 22 08:50 3 128 MB part-00001
-rw-r--r-- said supergroup 137.17 KB Mar 22 08:50 3 128 MB part-00002
-rw-r--r-- said supergroup 135.62 KB Mar 22 08:50 3 128 MB part-00003
-rw-r--r-- said supergroup 135.52 KB Mar 22 08:50 3 128 MB part-00004
-rw-r--r-- said supergroup 136.08 KB Mar 22 08:50 3 128 MB part-00005
-rw-r--r-- said supergroup 134.29 KB Mar 22 08:50 3 128 MB part-00006
-rw-r--r-- said supergroup 135.92 KB Mar 22 08:50 3 128 MB part-00007
-rw-r--r-- said supergroup 136.31 KB Mar 22 08:50 3 128 MB part-00008
-rw-r--r-- said supergroup 135.79 KB Mar 22 08:50 3 128 MB part-00009
-rw-r--r-- said supergroup 135.95 KB Mar 22 08:50 3 128 MB part-00010
-rw-r--r-- said supergroup 133.96 KB Mar 22 08:50 3 128 MB part-00011
-rw-r--r-- said supergroup 134.27 KB Mar 22 08:50 3 128 MB part-00012
-rw-r--r-- said supergroup 135.47 KB Mar 22 08:50 3 128 MB part-00013
-rw-r--r-- said supergroup 133.38 KB Mar 22 08:50 3 128 MB part-00014
-rw-r--r-- said supergroup 134.39 KB Mar 22 08:50 3 128 MB part-00015
Showing 1 to 17 of 17 entries
Previous 1 Next
Hadoop, 2019.


part-00001 (~/Downloads) - Pluma
part-00001
(She._,1)
(seriousness,,2)
(Ismail,2)
(desolated,2)
(massproduct,,1)
(monologue,4)
(bardoor,1)
(obediently,3)
(maidens,,2)
(parlour,10)
(making,,4)
(reinforced,3)
("Gimme,4)
(unintentionally,",1)
("Nothing,,8)
(_Sonnez._,1)
(onerous,2)
(peremptory.,1)
(practise,,2)
(toll,,1)
(speeding,,1)
(scale;,1)
(kink,,1)
(negro.,3)
(deceiver.,1)
(revision,,1)
(Sandycove!,1)
(if--what,1)
(boldest,4)
(chewing,,5)
(really,,29)
(fairness,,1)
(gourme.,1)
(interlaced,1)
(BABES,1)
(lilac,9)
(harshly,",1)
(pouf!,1)
(cabin-scuttle;,1)
(urgent,12)
(toils--one,1)
(milk,,17)
(shape,86)
(Domini_.,1)
(Piles,2)
(Elephant,4)
(amuck,1)
(propitiate--when,1)
(Hacking,1)
(Sarsfield,,1)
(tan,10)
(seducer.,1)
(landmark.',1)


said@said: ~/Downloads
said@said:~/Downloads$ sort -n -k2 part-r-00001 | tail -1000
182 items, Free space: 30.3 GB


said@said: ~/Downloads
(your'n;,1)
(yourn!',1)
(yourn,,3)
(yours!",6)
(_yourself_,1)
(yourself'--for,1)
(yourselves.,6)
(you--she,1)
("Youth,1)
(youthfulness,3)
(Y.'S,2)
(yum,1)
(Zarathustra,,1)
(zat.,1)
("Zat,,1)
(Zealand,6)
(Zephine.,5)
(zero-point,,1)
(Zhilinski,,2)
(zig-zag,,1)
(zigzags,,2)
(Znaim,6)
(_Zoe,1)
(Zoe!,2)
said@said:~/Downloads$ sort -n -k2 part-00001 | tail -10
("Zat,,1)
(Zealand,6)
(Zephine.,5)
(zero-point,,1)
(Zhilinski,,2)
(zig-zag,,1)
(zigzags,,2)
(Znaim,6)
(_Zoe,1)
(Zoe!,2)
said@said:~/Downloads$
182 items, Free space: 30.3 GB