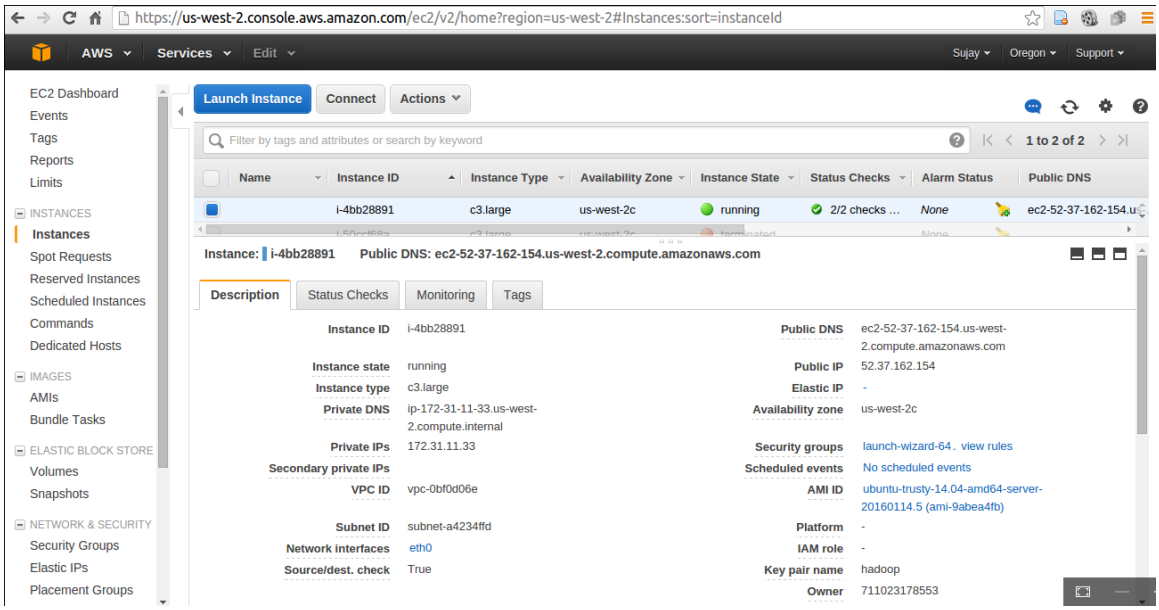
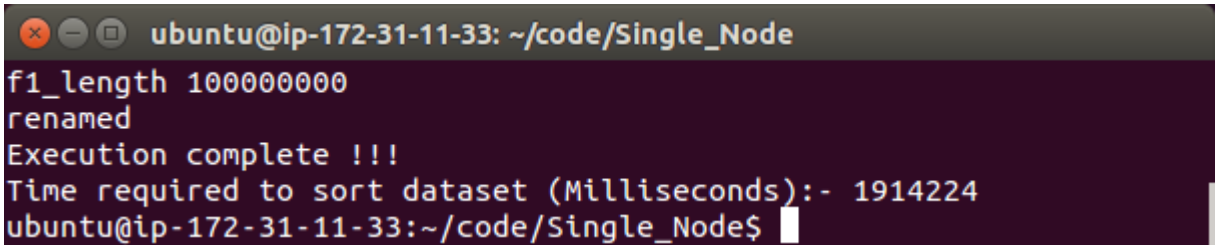
CS 553 Cloud Computing
Programming Assignment 2
Sujay Gunjal (CWID: A20351746)
Screenshots

Screenshots of Shared-Memory, Hadoop and spark execution:-

Aws screenshot of Shared memory 10GB sort execution:-



4 threads Execution screenshot:-



Time taken by 4 Thread to sort 10Gb file is: - 1914224 milliseconds

31.90375 minutes

Valsort output of shared memory sort:-

```
ubuntu@ip-172-31-11-33: ~/64
ubuntu@ip-172-31-11-33:~/code/Single_Node$ pwd
/home/ubuntu/code/Single_Node
ubuntu@ip-172-31-11-33:~/code/Single_Node$ ls -lrt
total 9765696
-rw------- 1 ubuntu ubuntu        3085 Mar 26 18:53 Terasort.class
-rw------- 1 ubuntu ubuntu        3480 Mar 26 18:53 Terasort.java~
-rw------- 1 ubuntu ubuntu        8901 Mar 26 18:53 File_Sort.java
-rw------- 1 ubuntu ubuntu        1904 Mar 26 18:53 MergeSort.class
-rw------- 1 ubuntu ubuntu         724 Mar 26 18:53 File_Collector.java
-rw------- 1 ubuntu ubuntu         724 Mar 26 18:53 File_Collector.java~
-rw------- 1 ubuntu ubuntu        8901 Mar 26 18:53 File_Sort.java~
-rw------- 1 ubuntu ubuntu        6637 Mar 26 18:53 File_Sort.class
-rw------- 1 ubuntu ubuntu        2707 Mar 26 18:53 MergeSort.java
-rw------- 1 ubuntu ubuntu        3480 Mar 26 18:53 Terasort.java
-rw------- 1 ubuntu ubuntu         876 Mar 26 18:53 File_Collector.class
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 26 19:42 final_sorted_dataset
ubuntu@ip-172-31-11-33:~/code/Single_Node$ cd /home/ubuntu/64
ubuntu@ip-172-31-11-33:~/64$ ./valsort /home/ubuntu/code/Single_Node/final_sorted_dataset
Records: 100000000
Checksum: 2faf0ab746e89a8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-11-33:~/64$
```

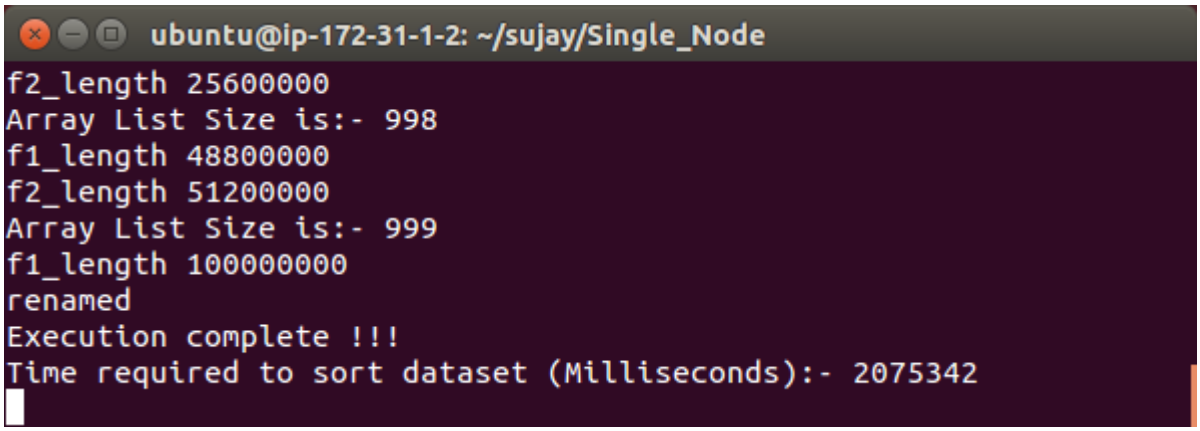
Head output of shared memory sort:-

```
ubuntu@ip-172-31-11-33: ~/code/Single_Node
ubuntu@ip-172-31-11-33:~/code/Single_Node$ head -10 final_sorted_dataset
    "O!uve  00000000000000000000000001228D4  77778888000022224444DDDDDDDDEEEE00000000CCCC7777DDDD
    PMd32=  00000000000000000000000003440CC1  FFFFEEEE6666CCCCBBBB9999333335555DDDDDDDD777788886666
    ^3CO],  00000000000000000000000158C5C5  5555AAAA9999EEEE888822229999CCCCDDDD6666555544442222
   !&S3/]]  00000000000000000000000002145D78  8888BBBBDDDD1111CCCC55556666BBBB1111EEEEDDDD22229999
   !,=U#,9  00000000000000000000000019072E3  33332222FFFFBBBB0000FFFFAAAA6666555553333DDDD3333CCCC
   !0f[ITd  00000000000000000000000003CAAB4B  9999FFFF555533337777CCCC4444BBBB7777EEEEBBBBDDDD4444
   !f6Suy2  00000000000000000000000003ABFD84  EEEE555555556666AAAA5555BBBBDDDD0000111166660000DDDD
   #%NIpq.  00000000000000000000000003B36FB9  1111000033334444111166666666AAAAAAAA00001111CCCCEEEE
   #'^cl'~  00000000000000000000000002EDC5C8  8888AAAA11114444FFFF77773333EEEE44440000FFFF99999999
   $"-'Q)]  00000000000000000000000005F1265D  CCCC6666EEEE22220000DDDDAAAA88886666BBBB00006666AAAA
ubuntu@ip-172-31-11-33:~/code/Single_Node$
```

Tail output of shared memory sort:-

```
ubuntu@ip-172-31-11-33: ~/code/Single_Node
ubuntu@ip-172-31-11-33:~/code/Single_Node$ tail -10 final_sorted_dataset
~~~uq2k#=U  00000000000000000000000002C06745  99991111DDDD222211110000FFFFEEEEFFFF33337777CCCC2222
~~~v/0&Qnm  00000000000000000000000004709701  CCCC88883333FFFF000000000000099991111FFFF777744446666
~~~yKOl:gE  00000000000000000000000002048B4F  CCCC11114444888822226666BBBB888855557777EEEEBBBB0000
~~~yK^H.il  00000000000000000000000463D004  44440000FFFF3333999944447777DDDDFFFFAAAA11118888DDDD
~~~yL;C'XE  00000000000000000000000005B0D211  2222EEEE33330000222211110CCCCFFFF5555777744448BBB6666
~~~zbA_ Tt  00000000000000000000000007F9F4F  BBBBCCCC666655559999FFFF8888AAAA11116666AAAABBBB0000
~~~zeO^FEg  00000000000000000000000001E06130  4444CCCCBBBB99992222888855558888CCCCFFFF000011111111
~~~}GxjWHI  00000000000000000000000000CA1345  777711118888AAAAAAAA22221111BBBB00002222BBBBCCCC2222
~~~}P;]g0g  00000000000000000000000040DA3E4  4444FFFF444466663333EEEE88888888DDDDEEEE44442222DDDD
~~~}kU|K<p  00000000000000000000000005E4A0AA  0000666655551111BBBB88889999AAAA55550000333355557777
ubuntu@ip-172-31-11-33:~/code/Single_Node$
```

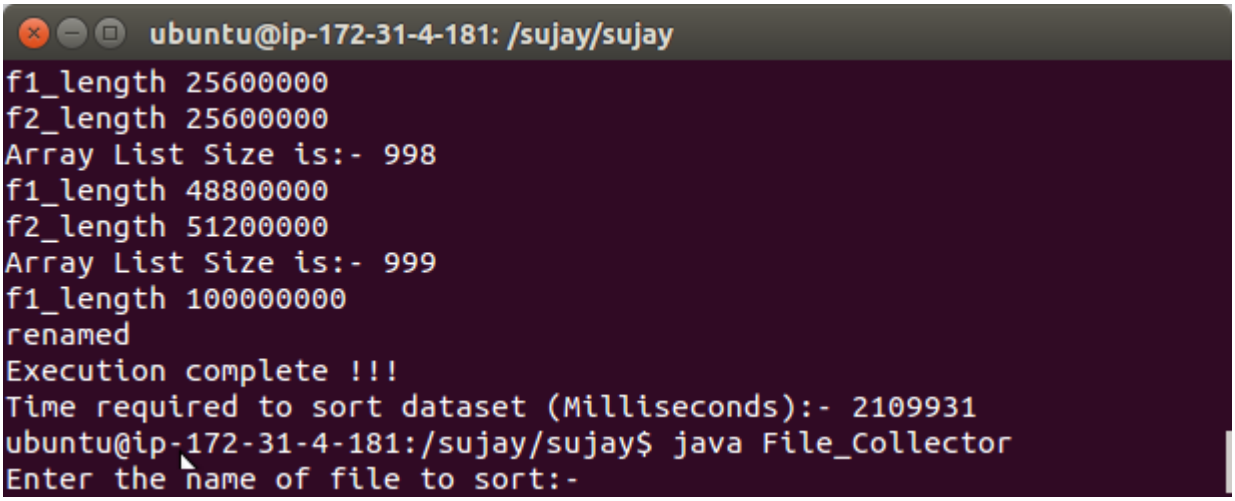
8 threads Execution screenshot:-



Time taken by 8 Thread to sort 10Gb file is :- 2075342 milliseconds

34.58903 minutes

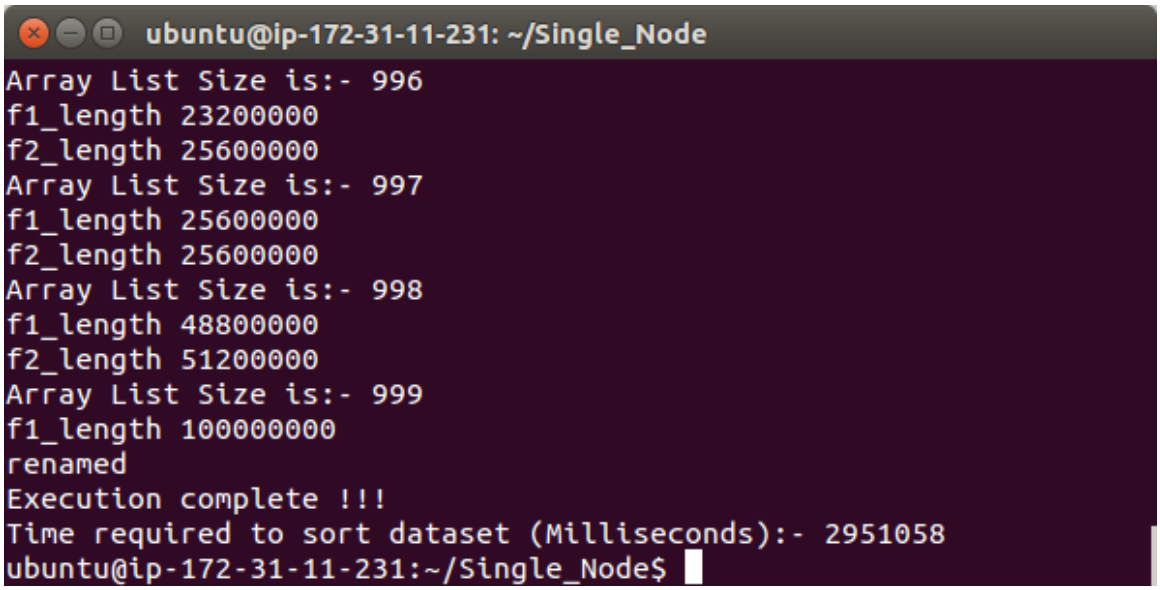
2 threads Execution screenshot:-



Time taken by 2 Thread to sort 10Gb file is :- 2109931 milliseconds

35.165516 minutes

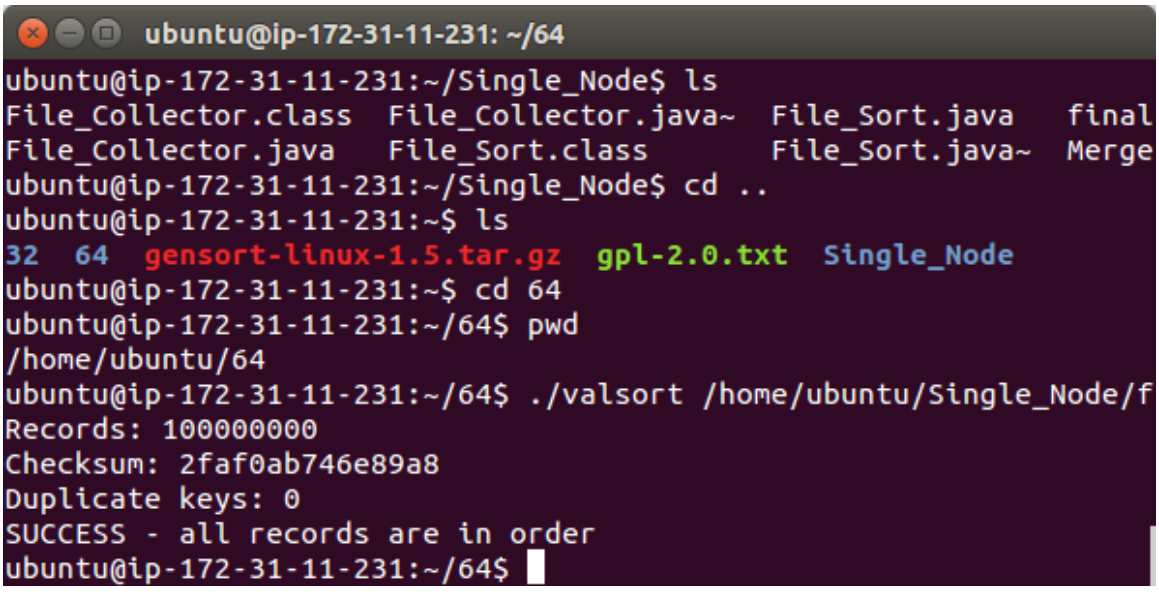
1 thread Execution screenshot:-



Time taken by 1 Thread to sort 10 Gb file is: - 2951058 milliseconds

49.1843 minutes

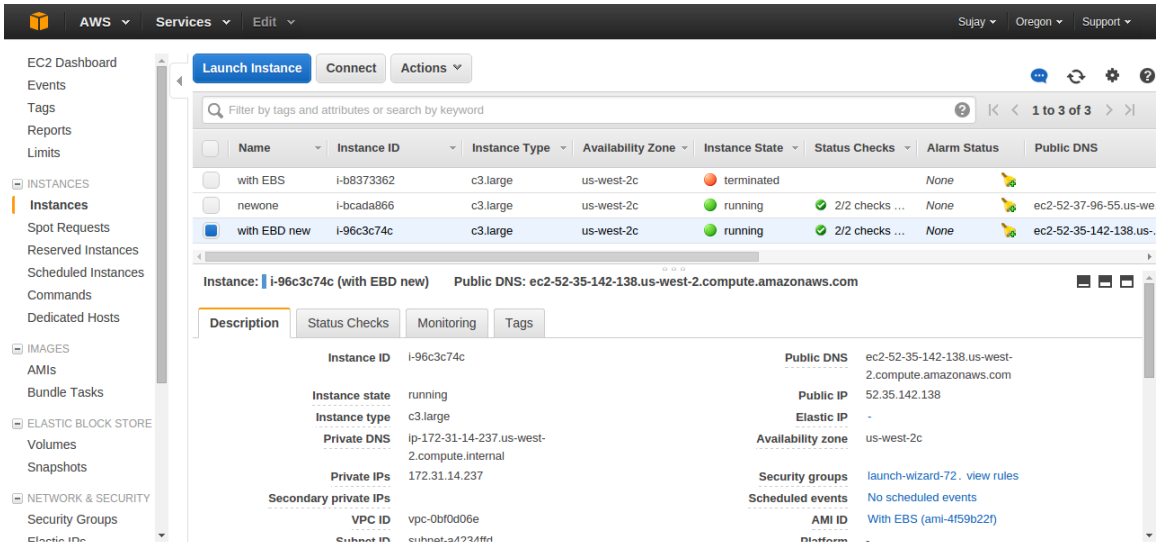
Valsort output of Shared memory application:-



Hadoop Execution:-

10 Gb Single Node Hadoop executions:-

Aws Screenshot:-



10Gb Hadoop output:-

```
ubuntu@ip-172-31-14-237:~/hadoop-2.7.2$ bin/hadoop jar TeraSort.jar
TeraSort /hdfs_sujay/10GB_input_file /hdfs_sujay/output
16/03/28 07:06:37 INFO client.RMProxy: Connecting to ResourceManager
at ec2-52-35-142-138.us-west-
2.compute.amazonaws.com/172.31.14.237:9050
16/03/28 07:06:37 WARN mapreduce.JobResourceUploader: Hadoop
command-line option parsing not performed. Implement the Tool
interface and execute your application with ToolRunner to remedy
this.
16/03/28 07:06:38 INFO input.FileInputFormat: Total input paths to
process : 1
16/03/28 07:06:38 INFO mapreduce.JobSubmitter: number of splits:75
16/03/28 07:06:38 INFO mapreduce.JobSubmitter: Submitting tokens for
job: job_1459147758523_0001
16/03/28 07:06:39 INFO impl.YarnClientImpl: Submitted application
application_1459147758523_0001
16/03/28 07:06:39 INFO mapreduce.Job: The url to track the job:
http://ec2-52-35-142-138.us-west-
2.compute.amazonaws.com:9006/proxy/application_1459147758523_0001/
16/03/28 07:06:39 INFO mapreduce.Job: Running job:
job_1459147758523_0001
16/03/28 07:06:48 INFO mapreduce.Job: Job job_1459147758523_0001
running in uber mode : false
16/03/28 07:06:48 INFO mapreduce.Job:  map 0% reduce 0%
16/03/28 07:07:12 INFO mapreduce.Job:  map 1% reduce 0%
16/03/28 07:07:16 INFO mapreduce.Job:  map 2% reduce 0%
```

```
16/03/28 07:07:20 INFO mapreduce.Job:  map 3% reduce 0%
16/03/28 07:07:23 INFO mapreduce.Job:  map 4% reduce 0%
16/03/28 07:07:36 INFO mapreduce.Job:  map 5% reduce 0%
16/03/28 07:07:48 INFO mapreduce.Job:  map 6% reduce 0%
16/03/28 07:07:51 INFO mapreduce.Job:  map 7% reduce 0%
16/03/28 07:07:57 INFO mapreduce.Job:  map 8% reduce 0%
16/03/28 07:08:19 INFO mapreduce.Job:  map 9% reduce 0%
16/03/28 07:08:22 INFO mapreduce.Job:  map 10% reduce 0%
16/03/28 07:08:24 INFO mapreduce.Job:  map 11% reduce 0%
16/03/28 07:08:27 INFO mapreduce.Job:  map 12% reduce 0%
16/03/28 07:08:38 INFO mapreduce.Job:  map 13% reduce 0%
16/03/28 07:08:50 INFO mapreduce.Job:  map 14% reduce 0%
16/03/28 07:08:56 INFO mapreduce.Job:  map 15% reduce 0%
16/03/28 07:09:00 INFO mapreduce.Job:  map 16% reduce 0%
16/03/28 07:09:25 INFO mapreduce.Job:  map 17% reduce 0%
16/03/28 07:09:28 INFO mapreduce.Job:  map 19% reduce 0%
16/03/28 07:09:32 INFO mapreduce.Job:  map 20% reduce 0%
16/03/28 07:09:47 INFO mapreduce.Job:  map 21% reduce 0%
16/03/28 07:09:59 INFO mapreduce.Job:  map 22% reduce 0%
16/03/28 07:10:02 INFO mapreduce.Job:  map 23% reduce 0%
16/03/28 07:10:08 INFO mapreduce.Job:  map 24% reduce 0%
16/03/28 07:10:32 INFO mapreduce.Job:  map 25% reduce 0%
16/03/28 07:10:35 INFO mapreduce.Job:  map 26% reduce 0%
16/03/28 07:10:38 INFO mapreduce.Job:  map 27% reduce 0%
16/03/28 07:10:41 INFO mapreduce.Job:  map 27% reduce 1%
16/03/28 07:10:44 INFO mapreduce.Job:  map 27% reduce 2%
16/03/28 07:10:51 INFO mapreduce.Job:  map 28% reduce 2%
16/03/28 07:10:53 INFO mapreduce.Job:  map 28% reduce 3%
16/03/28 07:11:00 INFO mapreduce.Job:  map 29% reduce 3%
16/03/28 07:11:06 INFO mapreduce.Job:  map 30% reduce 3%
16/03/28 07:11:11 INFO mapreduce.Job:  map 31% reduce 3%
16/03/28 07:11:14 INFO mapreduce.Job:  map 31% reduce 4%
16/03/28 07:11:20 INFO mapreduce.Job:  map 31% reduce 5%
16/03/28 07:11:23 INFO mapreduce.Job:  map 31% reduce 6%
16/03/28 07:11:29 INFO mapreduce.Job:  map 31% reduce 7%
16/03/28 07:11:35 INFO mapreduce.Job:  map 31% reduce 8%
16/03/28 07:11:42 INFO mapreduce.Job:  map 32% reduce 8%
16/03/28 07:11:45 INFO mapreduce.Job:  map 33% reduce 8%
16/03/28 07:11:51 INFO mapreduce.Job:  map 34% reduce 8%
16/03/28 07:11:53 INFO mapreduce.Job:  map 34% reduce 9%
16/03/28 07:12:04 INFO mapreduce.Job:  map 35% reduce 9%
16/03/28 07:12:06 INFO mapreduce.Job:  map 35% reduce 10%
16/03/28 07:12:16 INFO mapreduce.Job:  map 36% reduce 10%
16/03/28 07:12:22 INFO mapreduce.Job:  map 37% reduce 10%
16/03/28 07:12:40 INFO mapreduce.Job:  map 37% reduce 11%
16/03/28 07:12:46 INFO mapreduce.Job:  map 37% reduce 12%
16/03/28 07:13:01 INFO mapreduce.Job:  map 39% reduce 12%
16/03/28 07:13:04 INFO mapreduce.Job:  map 40% reduce 12%
16/03/28 07:13:19 INFO mapreduce.Job:  map 41% reduce 12%
16/03/28 07:13:23 INFO mapreduce.Job:  map 42% reduce 12%
16/03/28 07:13:35 INFO mapreduce.Job:  map 43% reduce 12%
16/03/28 07:13:41 INFO mapreduce.Job:  map 44% reduce 12%
16/03/28 07:13:56 INFO mapreduce.Job:  map 44% reduce 13%
16/03/28 07:14:05 INFO mapreduce.Job:  map 44% reduce 14%
16/03/28 07:14:11 INFO mapreduce.Job:  map 44% reduce 15%
16/03/28 07:14:17 INFO mapreduce.Job:  map 45% reduce 15%
```

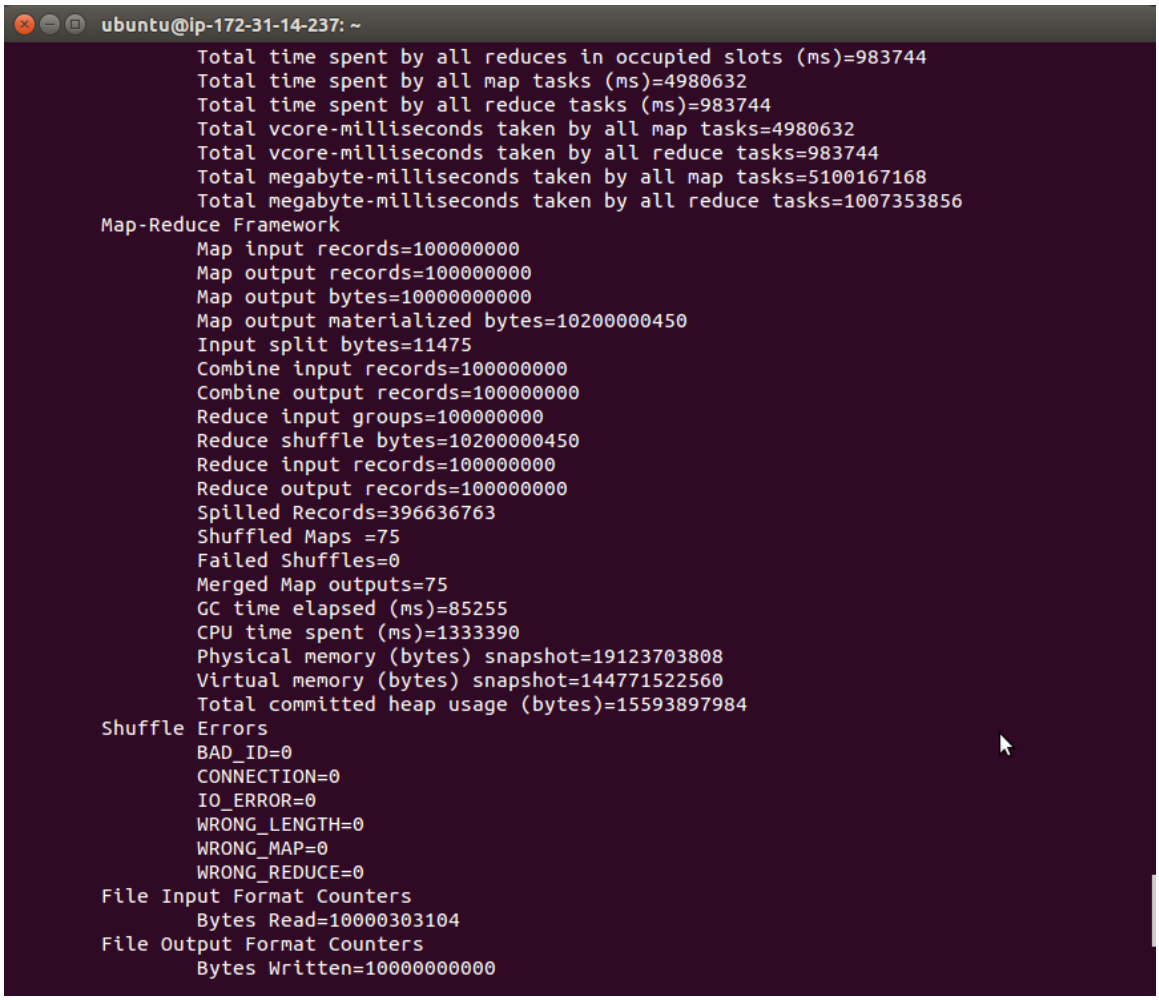
```
16/03/28 07:14:20 INFO mapreduce.Job:  map 46% reduce 15%
16/03/28 07:14:21 INFO mapreduce.Job:  map 47% reduce 15%
16/03/28 07:14:33 INFO mapreduce.Job:  map 48% reduce 15%
16/03/28 07:14:45 INFO mapreduce.Job:  map 49% reduce 15%
16/03/28 07:14:48 INFO mapreduce.Job:  map 50% reduce 15%
16/03/28 07:14:52 INFO mapreduce.Job:  map 51% reduce 15%
16/03/28 07:14:57 INFO mapreduce.Job:  map 51% reduce 16%
16/03/28 07:15:15 INFO mapreduce.Job:  map 51% reduce 17%
16/03/28 07:15:29 INFO mapreduce.Job:  map 52% reduce 17%
16/03/28 07:15:32 INFO mapreduce.Job:  map 53% reduce 17%
16/03/28 07:15:35 INFO mapreduce.Job:  map 54% reduce 17%
16/03/28 07:15:51 INFO mapreduce.Job:  map 55% reduce 17%
16/03/28 07:16:01 INFO mapreduce.Job:  map 56% reduce 17%
16/03/28 07:16:04 INFO mapreduce.Job:  map 57% reduce 17%
16/03/28 07:16:13 INFO mapreduce.Job:  map 57% reduce 18%
16/03/28 07:16:19 INFO mapreduce.Job:  map 57% reduce 19%
16/03/28 07:16:35 INFO mapreduce.Job:  map 58% reduce 19%
16/03/28 07:16:36 INFO mapreduce.Job:  map 59% reduce 19%
16/03/28 07:16:40 INFO mapreduce.Job:  map 60% reduce 19%
16/03/28 07:16:52 INFO mapreduce.Job:  map 62% reduce 19%
16/03/28 07:17:04 INFO mapreduce.Job:  map 63% reduce 19%
16/03/28 07:17:08 INFO mapreduce.Job:  map 64% reduce 19%
16/03/28 07:17:14 INFO mapreduce.Job:  map 64% reduce 20%
16/03/28 07:17:20 INFO mapreduce.Job:  map 64% reduce 21%
16/03/28 07:17:44 INFO mapreduce.Job:  map 65% reduce 21%
16/03/28 07:17:47 INFO mapreduce.Job:  map 66% reduce 21%
16/03/28 07:17:50 INFO mapreduce.Job:  map 67% reduce 21%
16/03/28 07:18:03 INFO mapreduce.Job:  map 68% reduce 21%
16/03/28 07:18:15 INFO mapreduce.Job:  map 69% reduce 21%
16/03/28 07:18:18 INFO mapreduce.Job:  map 70% reduce 21%
16/03/28 07:18:23 INFO mapreduce.Job:  map 71% reduce 21%
16/03/28 07:18:27 INFO mapreduce.Job:  map 71% reduce 22%
16/03/28 07:18:30 INFO mapreduce.Job:  map 71% reduce 23%
16/03/28 07:18:33 INFO mapreduce.Job:  map 71% reduce 24%
16/03/28 07:18:48 INFO mapreduce.Job:  map 72% reduce 24%
16/03/28 07:18:51 INFO mapreduce.Job:  map 73% reduce 24%
16/03/28 07:18:54 INFO mapreduce.Job:  map 74% reduce 24%
16/03/28 07:19:07 INFO mapreduce.Job:  map 75% reduce 24%
16/03/28 07:19:16 INFO mapreduce.Job:  map 76% reduce 24%
16/03/28 07:19:19 INFO mapreduce.Job:  map 77% reduce 24%
16/03/28 07:19:34 INFO mapreduce.Job:  map 77% reduce 25%
16/03/28 07:19:40 INFO mapreduce.Job:  map 77% reduce 26%
16/03/28 07:20:02 INFO mapreduce.Job:  map 78% reduce 26%
16/03/28 07:20:03 INFO mapreduce.Job:  map 79% reduce 26%
16/03/28 07:20:06 INFO mapreduce.Job:  map 80% reduce 26%
16/03/28 07:20:18 INFO mapreduce.Job:  map 81% reduce 26%
16/03/28 07:20:21 INFO mapreduce.Job:  map 82% reduce 26%
16/03/28 07:20:30 INFO mapreduce.Job:  map 83% reduce 26%
16/03/28 07:20:33 INFO mapreduce.Job:  map 84% reduce 26%
16/03/28 07:20:43 INFO mapreduce.Job:  map 84% reduce 27%
16/03/28 07:20:46 INFO mapreduce.Job:  map 84% reduce 28%
16/03/28 07:21:01 INFO mapreduce.Job:  map 85% reduce 28%
16/03/28 07:21:04 INFO mapreduce.Job:  map 86% reduce 28%
16/03/28 07:21:05 INFO mapreduce.Job:  map 87% reduce 28%
16/03/28 07:21:17 INFO mapreduce.Job:  map 88% reduce 28%
16/03/28 07:21:27 INFO mapreduce.Job:  map 89% reduce 28%
```

```
16/03/28 07:21:30 INFO mapreduce.Job:  map 90% reduce 28%
16/03/28 07:21:36 INFO mapreduce.Job:  map 91% reduce 28%
16/03/28 07:21:42 INFO mapreduce.Job:  map 91% reduce 29%
16/03/28 07:21:48 INFO mapreduce.Job:  map 91% reduce 30%
16/03/28 07:22:15 INFO mapreduce.Job:  map 92% reduce 30%
16/03/28 07:22:18 INFO mapreduce.Job:  map 93% reduce 30%
16/03/28 07:22:22 INFO mapreduce.Job:  map 94% reduce 30%
16/03/28 07:22:31 INFO mapreduce.Job:  map 95% reduce 30%
16/03/28 07:22:43 INFO mapreduce.Job:  map 96% reduce 30%
16/03/28 07:22:46 INFO mapreduce.Job:  map 97% reduce 30%
16/03/28 07:22:54 INFO mapreduce.Job:  map 97% reduce 31%
16/03/28 07:22:57 INFO mapreduce.Job:  map 97% reduce 32%
16/03/28 07:23:13 INFO mapreduce.Job:  map 98% reduce 32%
16/03/28 07:23:14 INFO mapreduce.Job:  map 99% reduce 32%
16/03/28 07:23:25 INFO mapreduce.Job:  map 99% reduce 33%
16/03/28 07:23:28 INFO mapreduce.Job:  map 100% reduce 33%
16/03/28 07:23:37 INFO mapreduce.Job:  map 100% reduce 41%
16/03/28 07:23:40 INFO mapreduce.Job:  map 100% reduce 61%
16/03/28 07:23:43 INFO mapreduce.Job:  map 100% reduce 67%
16/03/28 07:23:49 INFO mapreduce.Job:  map 100% reduce 68%
16/03/28 07:23:55 INFO mapreduce.Job:  map 100% reduce 69%
16/03/28 07:23:58 INFO mapreduce.Job:  map 100% reduce 70%
16/03/28 07:24:04 INFO mapreduce.Job:  map 100% reduce 71%
16/03/28 07:24:07 INFO mapreduce.Job:  map 100% reduce 72%
16/03/28 07:24:13 INFO mapreduce.Job:  map 100% reduce 73%
16/03/28 07:24:19 INFO mapreduce.Job:  map 100% reduce 74%
16/03/28 07:24:22 INFO mapreduce.Job:  map 100% reduce 75%
16/03/28 07:24:28 INFO mapreduce.Job:  map 100% reduce 76%
16/03/28 07:24:31 INFO mapreduce.Job:  map 100% reduce 77%
16/03/28 07:24:37 INFO mapreduce.Job:  map 100% reduce 78%
16/03/28 07:24:43 INFO mapreduce.Job:  map 100% reduce 79%
16/03/28 07:24:46 INFO mapreduce.Job:  map 100% reduce 80%
16/03/28 07:24:52 INFO mapreduce.Job:  map 100% reduce 81%
16/03/28 07:24:58 INFO mapreduce.Job:  map 100% reduce 82%
16/03/28 07:25:04 INFO mapreduce.Job:  map 100% reduce 83%
16/03/28 07:25:07 INFO mapreduce.Job:  map 100% reduce 84%
16/03/28 07:25:13 INFO mapreduce.Job:  map 100% reduce 85%
16/03/28 07:25:19 INFO mapreduce.Job:  map 100% reduce 86%
16/03/28 07:25:22 INFO mapreduce.Job:  map 100% reduce 87%
16/03/28 07:25:28 INFO mapreduce.Job:  map 100% reduce 88%
16/03/28 07:25:34 INFO mapreduce.Job:  map 100% reduce 89%
16/03/28 07:25:37 INFO mapreduce.Job:  map 100% reduce 90%
16/03/28 07:25:43 INFO mapreduce.Job:  map 100% reduce 91%
16/03/28 07:25:49 INFO mapreduce.Job:  map 100% reduce 92%
16/03/28 07:25:52 INFO mapreduce.Job:  map 100% reduce 93%
16/03/28 07:25:58 INFO mapreduce.Job:  map 100% reduce 94%
16/03/28 07:26:04 INFO mapreduce.Job:  map 100% reduce 95%
16/03/28 07:26:07 INFO mapreduce.Job:  map 100% reduce 96%
16/03/28 07:26:13 INFO mapreduce.Job:  map 100% reduce 97%
16/03/28 07:26:19 INFO mapreduce.Job:  map 100% reduce 98%
16/03/28 07:26:22 INFO mapreduce.Job:  map 100% reduce 99%
16/03/28 07:26:28 INFO mapreduce.Job:  map 100% reduce 100%
16/03/28 07:26:30 INFO mapreduce.Job: Job job_1459147758523_0001
completed successfully
16/03/28 07:26:31 INFO mapreduce.Job: Counters: 50
      File System Counters
```

```
            FILE: Number of bytes read=30256951212
            FILE: Number of bytes written=40465905460
            FILE: Number of read operations=0
            FILE: Number of large read operations=0
            FILE: Number of write operations=0
            HDFS: Number of bytes read=10000314579
            HDFS: Number of bytes written=10000000000
            HDFS: Number of read operations=228
            HDFS: Number of large read operations=0
            HDFS: Number of write operations=2
      Job Counters
            Killed map tasks=2
            Launched map tasks=77
            Launched reduce tasks=1
            Data-local map tasks=77
            Total time spent by all maps in occupied slots
(ms)=4980632
            Total time spent by all reduces in occupied slots
(ms)=983744
            Total time spent by all map tasks (ms)=4980632
            Total time spent by all reduce tasks (ms)=983744
            Total vcore-milliseconds taken by all map tasks=4980632
            Total vcore-milliseconds taken by all reduce tasks=983744
            Total megabyte-milliseconds taken by all map
tasks=5100167168
            Total megabyte-milliseconds taken by all reduce
tasks=1007353856
      Map-Reduce Framework
            Map input records=100000000
            Map output records=100000000
            Map output bytes=10000000000
            Map output materialized bytes=10200000450
            Input split bytes=11475
            Combine input records=100000000
            Combine output records=100000000
            Reduce input groups=100000000
            Reduce shuffle bytes=10200000450
            Reduce input records=100000000
            Reduce output records=100000000
            Spilled Records=396636763
            Shuffled Maps =75
            Failed Shuffles=0
            Merged Map outputs=75
            GC time elapsed (ms)=85255
            CPU time spent (ms)=1333390
            Physical memory (bytes) snapshot=19123703808
            Virtual memory (bytes) snapshot=144771522560
            Total committed heap usage (bytes)=15593897984
      Shuffle Errors
            BAD_ID=0
            CONNECTION=0
            IO_ERROR=0
            WRONG_LENGTH=0
            WRONG_MAP=0
            WRONG_REDUCE=0
      File Input Format Counters
```

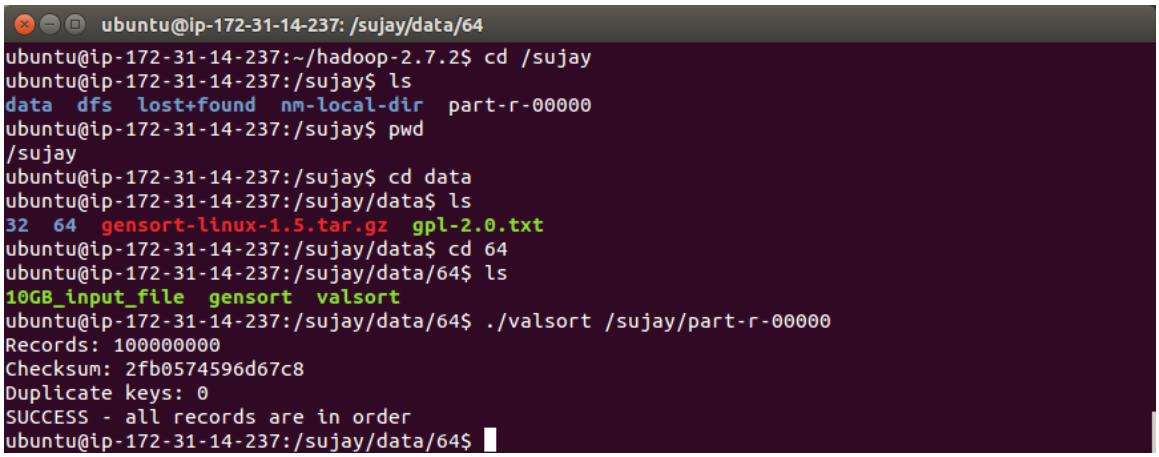
```
        Bytes Read=10000303104
File Output Format Counters
        Bytes Written=10000000000
```

Total Time taken for Execution: 20.18 minutes

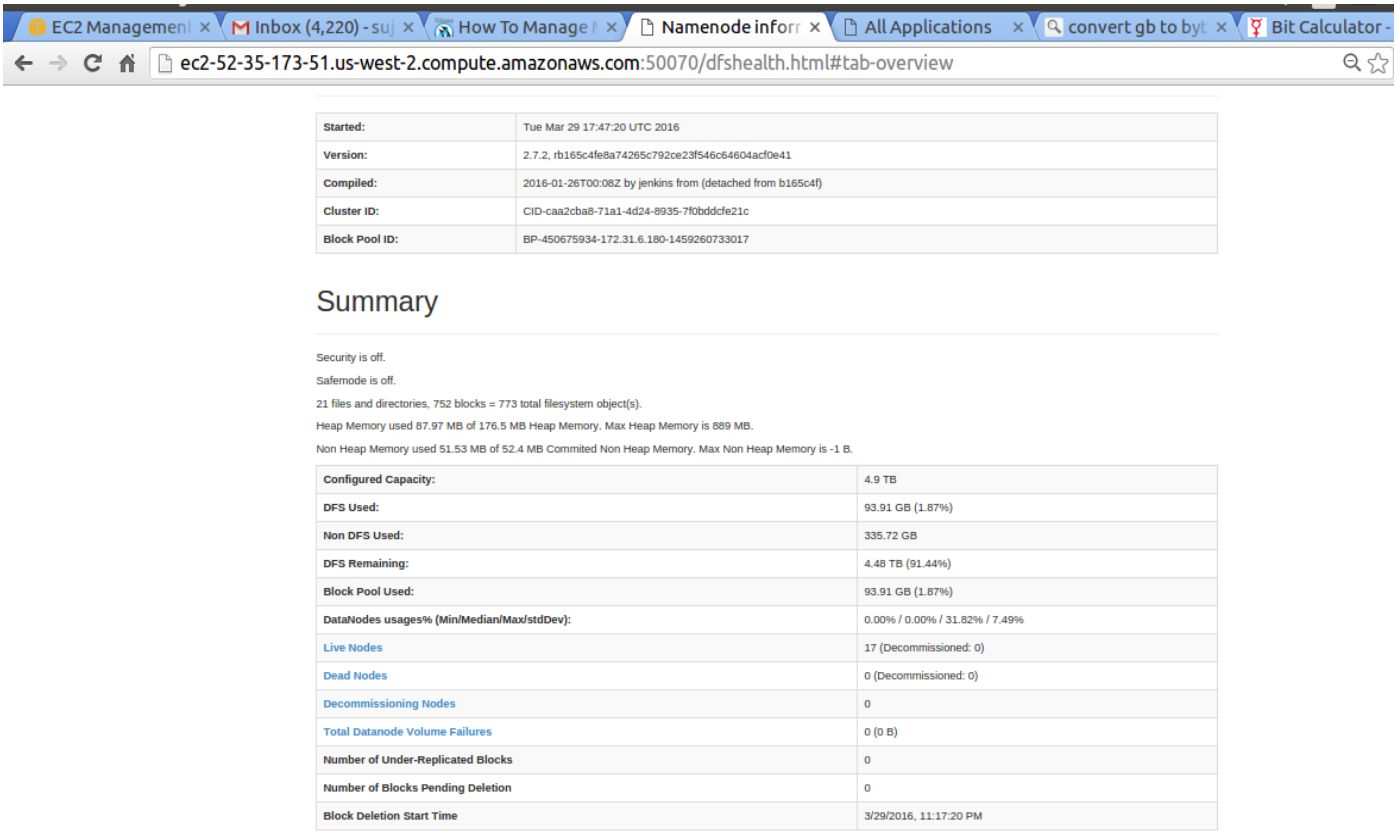


Final Output on terminal Hadoop single node:-

Valsort output Hadoop single node:-



Hadoop 100GB 16 Nodes Execution Name node:-



Started:	Tue Mar 29 17:47:20 UTC 2016
Version:	2.7.2, rb165c4fe8a74265c792ce23f546c64604acf0e41
Compiled:	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
Cluster ID:	CID-caa2cba8-71a1-4d24-8935-7f0bddcfe21c
Block Pool ID:	BP-450675934-172.31.6.180-1459260733017

Summary

Security is off.

Safemode is off.

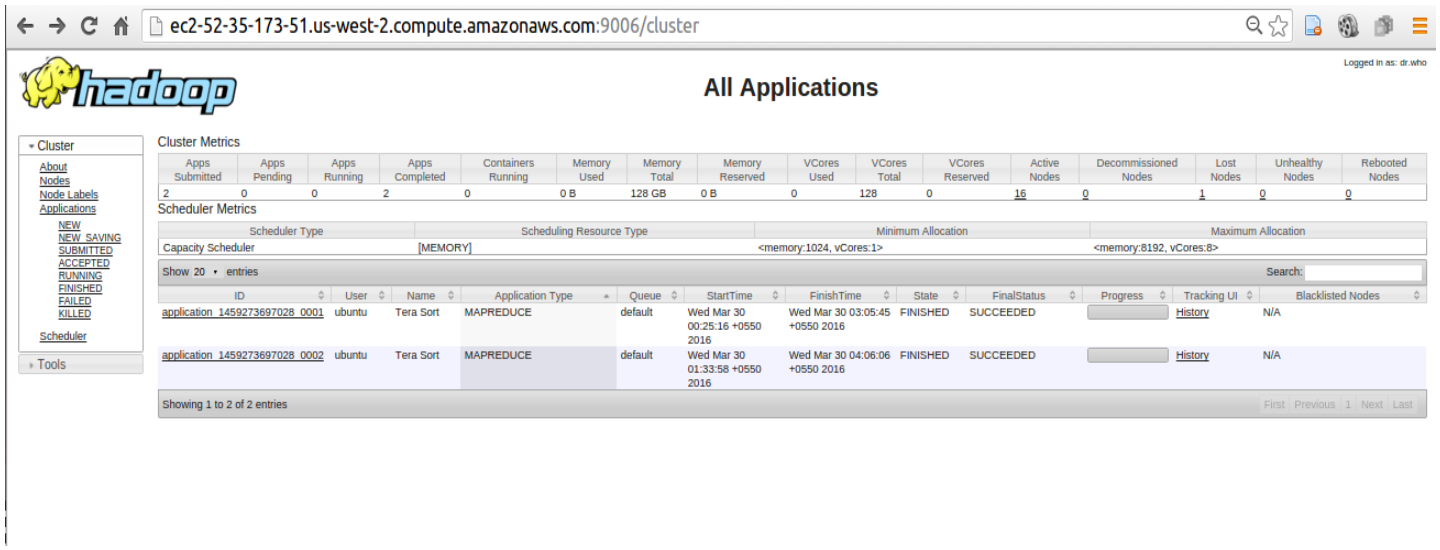
21 files and directories, 752 blocks = 773 total filesystem object(s).

Heap Memory used 87.97 MB of 176.5 MB Heap Memory. Max Heap Memory is 889 MB.

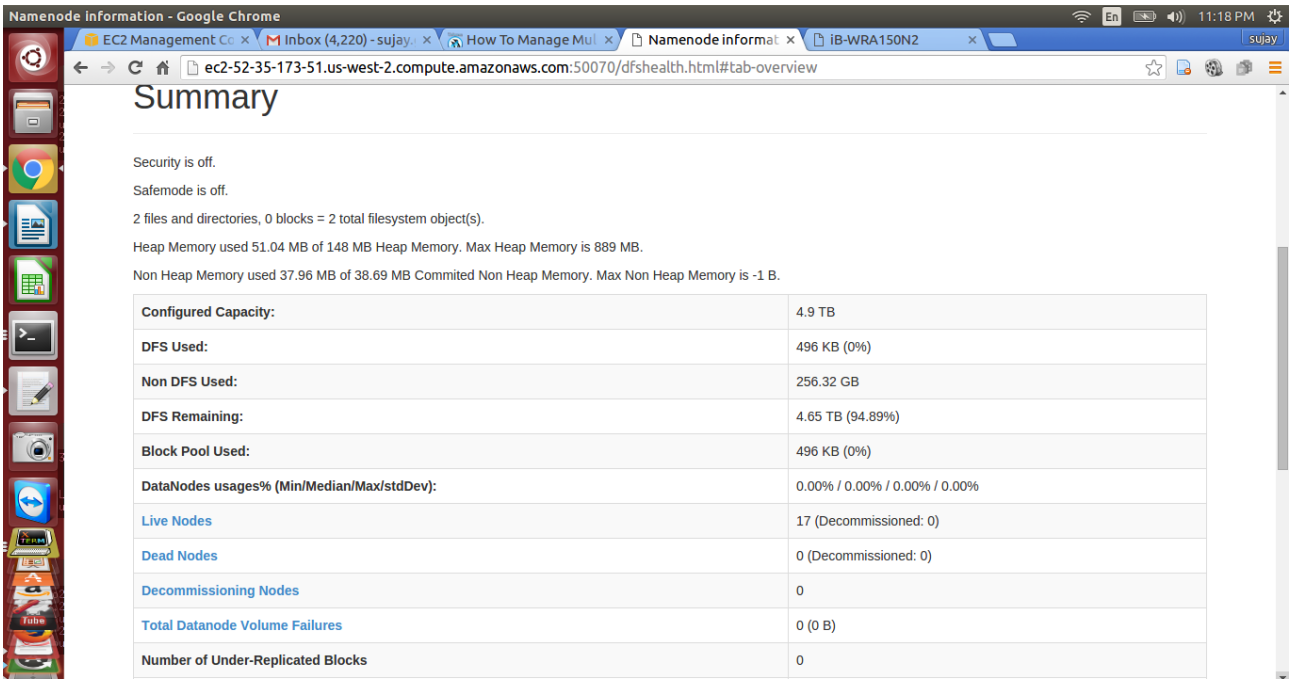
Non Heap Memory used 51.53 MB of 52.4 MB Commited Non Heap Memory. Max Non Heap Memory is -1 B.

Configured Capacity:	4.9 TB
DFS Used:	93.91 GB (1.87%)
Non DFS Used:	335.72 GB
DFS Remaining:	4.48 TB (91.44%)
Block Pool Used:	93.91 GB (1.87%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 31.82% / 7.49%
Live Nodes	17 (Decommissioned: 0)
Dead Nodes	0 (Decommissioned: 0)
Decommissioning Nodes	0
Total Datanode Volume Failures	0 (0 B)
Number of Under-Replicated Blocks	0
Number of Blocks Pending Deletion	0
Block Deletion Start Time	3/29/2016, 11:17:20 PM

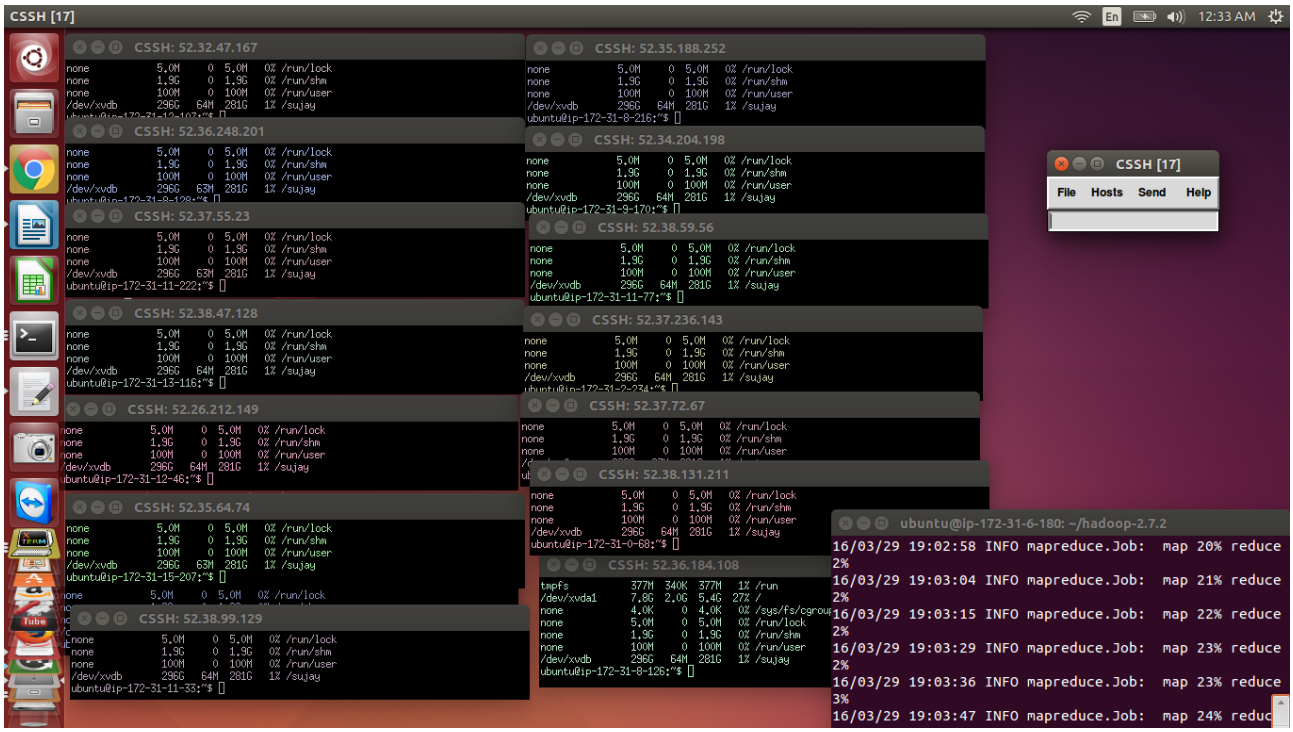
Hadoop 16 Nodes execution completion Screenshot:-



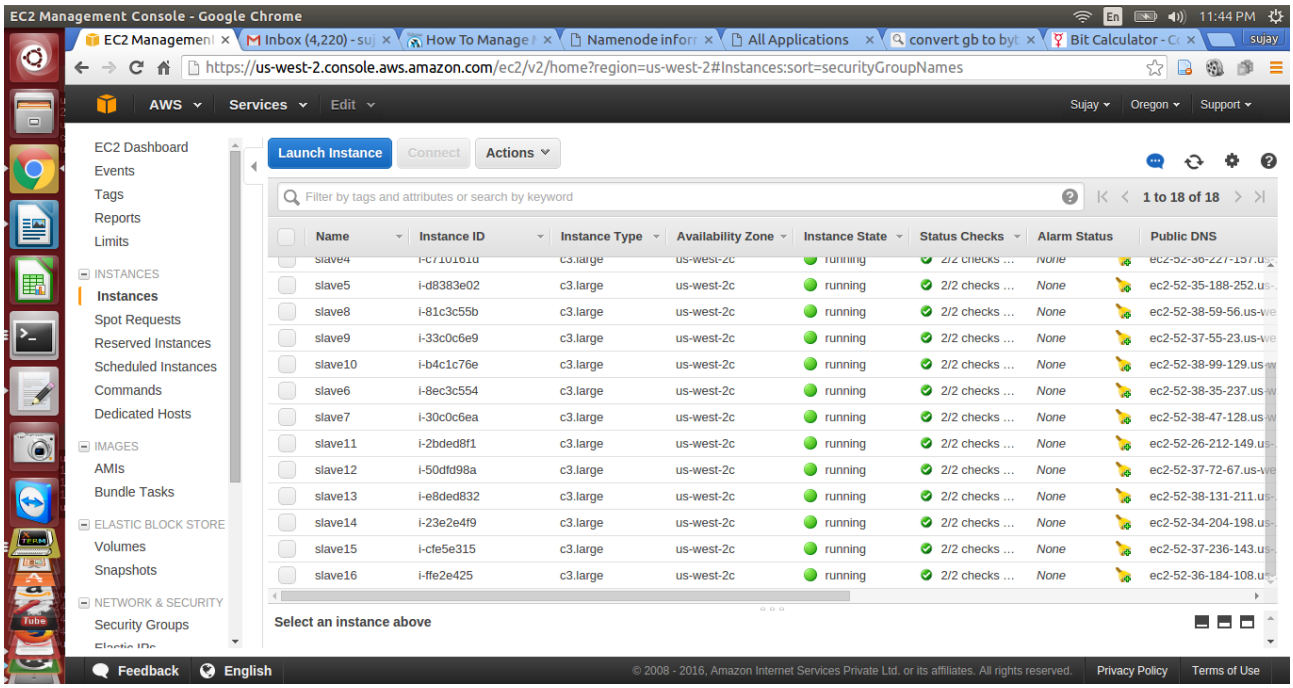
Name node one more screenshot:-



Hadoop 16 Nodes execution on Ubuntu screenshot:-



Hadoop 16 Nodes execution screen shot of AWS:-



Hadoop 100GB tail output:-

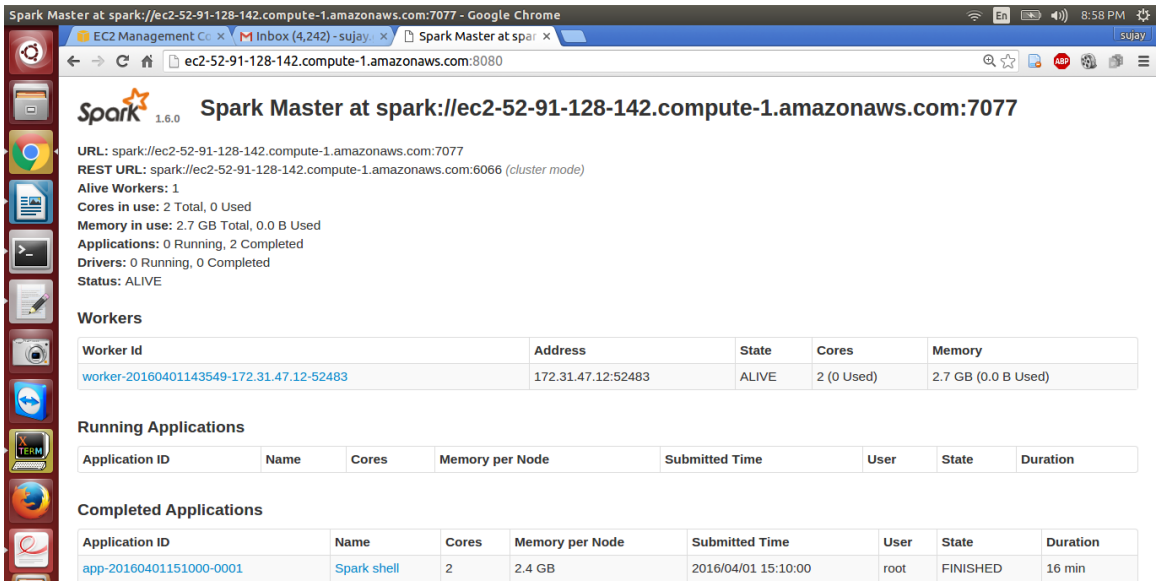
```
ubuntu@ip-172-31-6-180: /sujay
ubuntu@ip-172-31-6-180:/sujay$ tail -10 part-r-00000
~~~~#iay1X      00000000000000000000000025D35EDF  6666AAAA555599997777000022223333888 8FFFF999922220000
~~~~+@p){@      00000000000000000000000085426F4   77773333555511111111000 0CCCC55559999AAAA7777DDDDDDDD
~~~~,R^_?n      00000000000000000000001034E347    11111111999900001111888 8AAAA55554444EEEE999933338888
~~~~.Ey`^)      00000000000000000000016F0E66B     CCCC6666DDDD2222DDDD111188889999EEEEEEEEEEEEBBBB4444
~~~~4!kA7x      00000000000000000000001F1A1E26    EEEE777711117777BBBB1111EEEE88884444DDDDDDDDEEEEBBBB
~~~~8Ii/!@      00000000000000000000001F05932F    11119999BBBB444477770000111144444CCCCAAAA6666DDDD0000
~~~~<I'5>F      00000000000000000000008CB2293     88883333BBBB1111666699998888555588888888222288888CCCC
~~~~G-)m^)      00000000000000000000013397F73     DDDDFFFFBBBBCCCCFFFF44446666AAAA111133333333AAAACCCC
~~~~c+I&cP      00000000000000000000074BDF64      888800005555000ODDDD22227777AAAA000033332222AAAADDDD
~~~~hb&5X*      00000000000000000000032C0E06B     7777BBBBBBBB9999EEEEAAAAAAAA0000CCCCDDDD4444BBBB4444
ubuntu@ip-172-31-6-180:/sujay$
```

Hadoop 100GB head Output:-

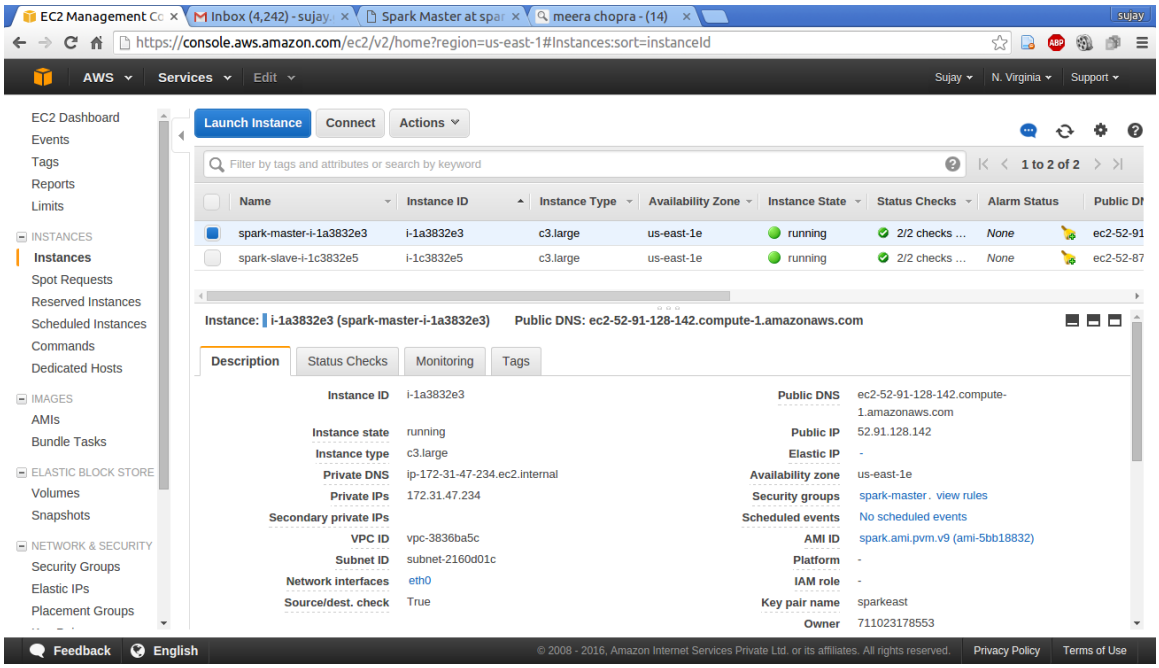
```
ubuntu@ip-172-31-6-180: /sujay
ubuntu@ip-172-31-6-180:/sujay$ head -10 part-r-00000
    !4+ABv      00000000000000000000000017F7E829  EEEE333344441111222288883333444466663333222 2DDDDEEEE
    "O!uve      00000000000000000000000001228D4   777788880000222244444DDDDDDDEEEE00000000CCCC7777DDDD
    %!$sU(      00000000000000000000000002E6C821C 22223333777744445555111119999CCCC4444EEEEFFFF11115555
    &5rX|X      00000000000000000000000399BC288   5555CCCCBBBB99999999DDDD111100001111EEEE7777DDDD9999
    'ic%So      00000000000000000000000031F06B7D  EEEEBBBBAAAA8888DDDDDDDD777722224444111166664444AAAA
    *0G1Io      00000000000000000000000003B5E85A1 1111AAAA9999CCCCBBBB11119999111133339999111 1AAAA6666
    ,(GhT_      00000000000000000000000002D0172DC 1111CCCC1111DDDDCCCCEEEE9999CCCC8888CCCCFFFF55555555
    0*vYm3      00000000000000000000000026D61578  DDDD7777AAAAEEEEEEEE6666AAAA2222CCCC5555555522229999
    2C>)8d      00000000000000000000000026C79E66  444400001111CCCC6666BBBB555577776666CCCC2222AAAABBBB
    PMd32=      00000000000000000000000003440CC1  FFFFEEEE6666CCCCBBBB999933335555DDDDDDDD777788886666
ubuntu@ip-172-31-6-180:/sujay$
```

Spark execution screen-shots:-

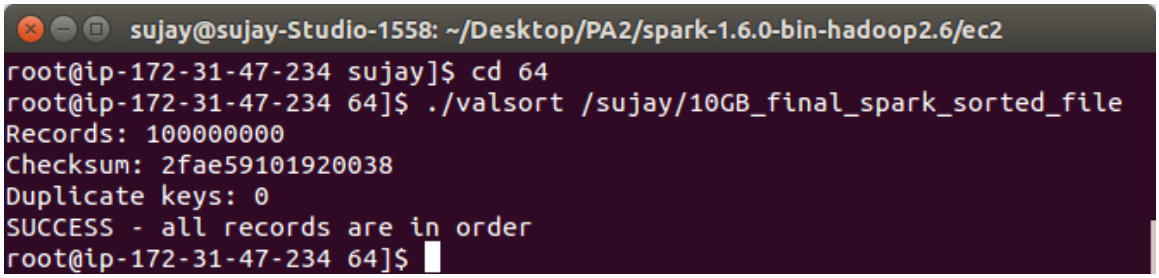
Spark 10Gb Execution report:-



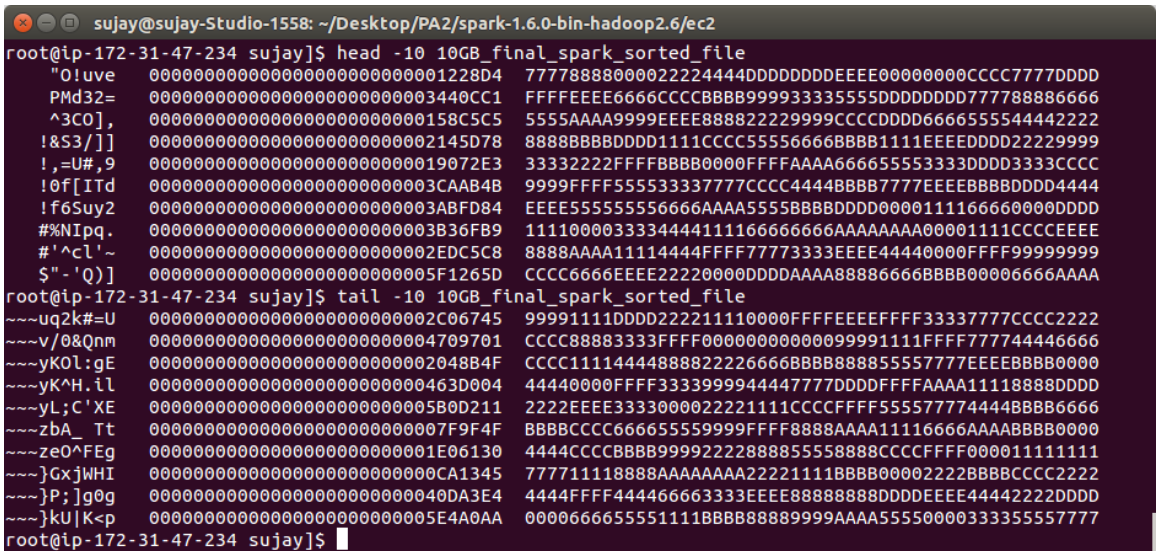
Spark Single node AWS screenshot:-



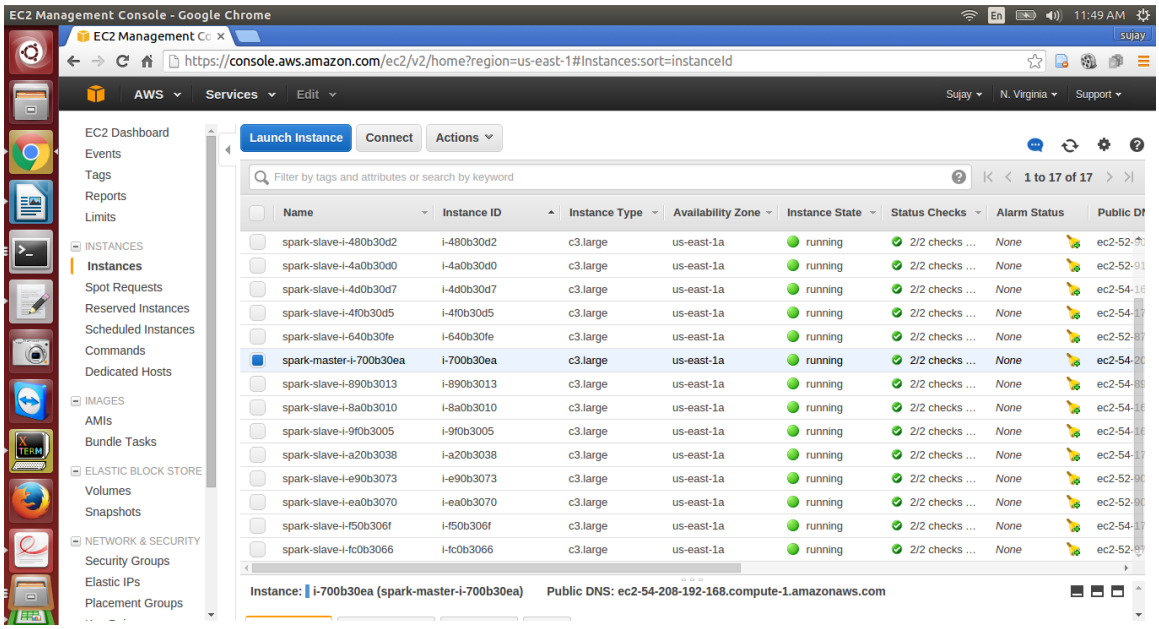
Spark 10GB Valsort output:-



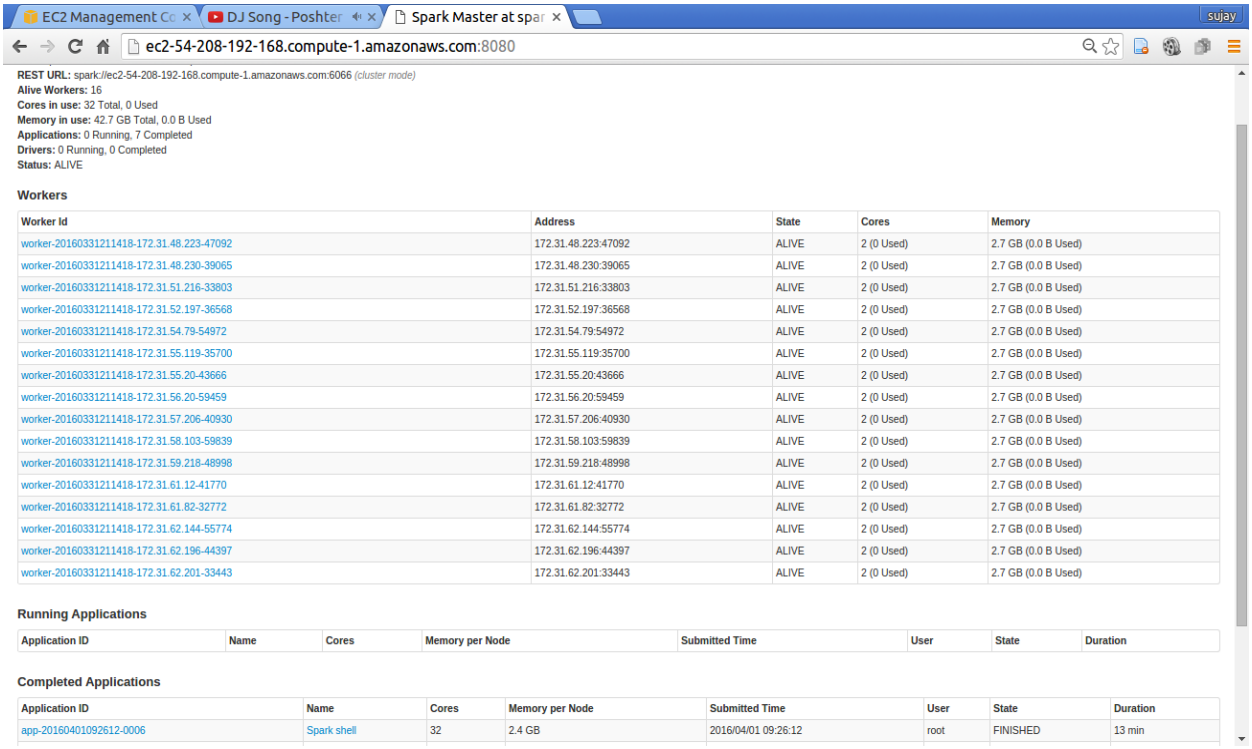
Spark 10GB head and tail screenshot:-



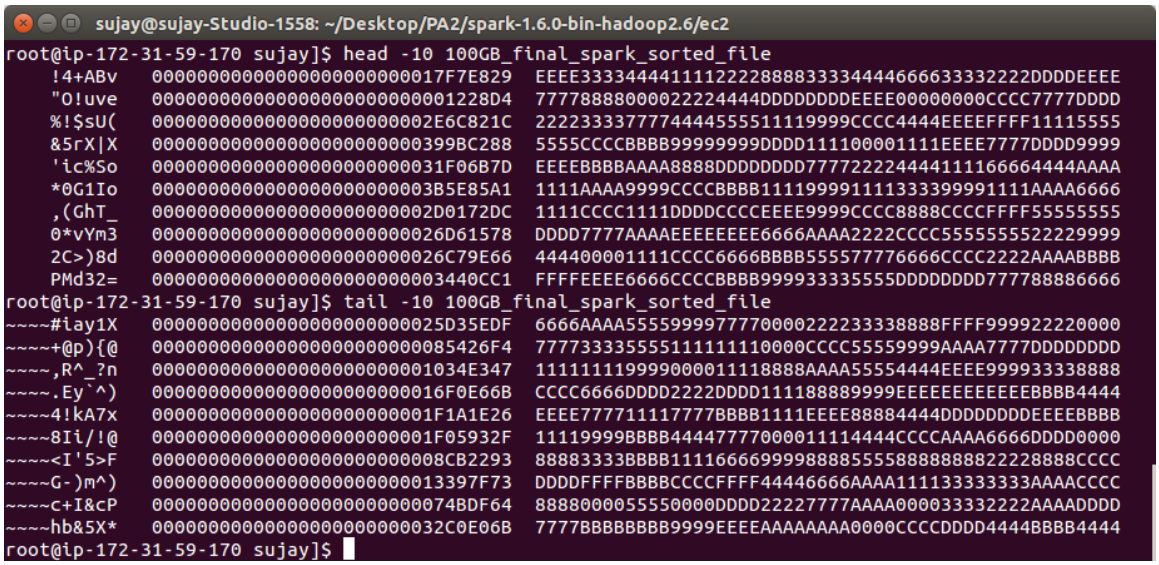
Spark 100 Gb 16 nodes AWS screen-shot:-



Spark 100 Gb 16 nodes completion screen-shot:-



Spark 100Gb 16 nodes head tail output:-



Spark 100GB 16 nodes valsort output:-