

Shared-Memory

1 Thread

Total time taken by Shared-Memory with 1 thread.

```
ubuntu@ip-172-31-1-196:~$ mke2fs 1.42.9 (4-Feb-2014)
Filesystem label=
OS type: Linux
Block size=4096 (Log=2)
Fragment size=4096 (Log=2)
Stride=0 blocks, Stripe width=0 blocks
13107200 inodes, 52428800 blocks
2621440 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=4294967296
1600 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98384, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

Buildfile: /data/SharedMemoryCode/build.xml

clean:
precompile:
[jmkdir] Created dir: /data/SharedMemoryCode/output

compile:
[mkdir] Created dir: /data/SharedMemoryCode/build/classes
[javac] /data/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' wa
s not set, defaulting to build.sysclasspath=last; set to false for repeatable bu
ilds
[javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes

jar:
[jkdir] Created dir: /data/SharedMemoryCode/build/jar
[jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar

run:
[java] numberOfFilestoCreate: 1000
[java]
[java]
[java]
[java] Total Time= 4782776.0 milliseconds
[java] Total Time= 4782.776 seconds

BUILD SUCCESSFUL
Total time: 79 minutes 44 seconds
ubuntu@ip-172-31-1-196:~$ 
ubuntu@ip-172-31-1-196:~$ 
```

Output file generated of required size 10 GB

```
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
drwxrwxr-x 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
-rw-rw-r-- 1 ubuntu ubuntu 2320000000 Mar 28 02:44 output994
drwxrwxr-x 1 ubuntu ubuntu 2560000000 Mar 28 02:46 output995
-rw-rw-r-- 1 ubuntu ubuntu 2203369472 Mar 28 02:48 output996
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 11992500
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:46 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:41 output992
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
drwxrwxr-x 1 ubuntu ubuntu 2320000000 Mar 28 02:44 output994
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:46 output995
-rw-rw-r-- 1 ubuntu ubuntu 2280267776 Mar 28 02:48 output996
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 10598968
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:50 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:41 output992
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
drwxrwxr-x 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 853295104 Mar 28 02:51 output997
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ y
y: command not found
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ y
y: command not found
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 12129744
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:50 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:41 output992
drwxrwxr-x 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 2420809728 Mar 28 02:52 output997
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 17982804
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:54 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 28 02:54 output997
drwxrwxr-x 1 ubuntu ubuntu 3227213824 Mar 28 02:56 output998
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 17982804
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:54 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 28 02:54 output997
drwxrwxr-x 1 ubuntu ubuntu 8414347264 Mar 28 03:00 output998
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ ll
total 9765660
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 03:01 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 28 03:01 output999
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode/output$ 
```

Valsort on output file – all records are in order

```
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
total 10599868
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:50 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:41 output992
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
y: command not found
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
y: command not found
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
total 12129744
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:50 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:41 output992
-rw-rw-r-- 1 ubuntu ubuntu 2560000000 Mar 28 02:43 output993
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 2420000728 Mar 28 02:52 output997
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
total 12971244
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:54 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 28 02:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 327213824 Mar 28 02:56 output998
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
total 7932000
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 02:54 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 28 02:50 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 28 02:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 8414347264 Mar 28 03:00 output998
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
total 9765660
drwxrwxr-x 2 ubuntu ubuntu 24576 Mar 28 03:01 ./
drwxrwxr-x 9 ubuntu ubuntu 4096 Mar 28 01:41 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 28 03:01 output998
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ cd /home/ubuntu/gensorst-linux-1.5/64/
gensorst valsort
ubuntu@ip-172-31-1-196:~/gensorst-linux-1.5/64$ mv valsort /data/SharedMemoryCode
./output/
ubuntu@ip-172-31-1-196:~/gensorst-linux-1.5/64$ cd /data/SharedMemoryCode/output/
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$ ./valsort < /dev/null > /tmp/valsort.out
Records: 1000000000
Checksum: 2fa58ab746e89a8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-1-196:/data/SharedMemoryCode$
```

Head and tail command on output file, I have included these records in separate output file.

2 Thread

Total time taken by Shared-Memory with 2 thread.

```
ubuntu@ip-172-31-6-46:~  
mke2fs 1.42.9 (4-Feb-2014)  
Filesystem label=""  
Filesystem type: ext4  
Blocksize=4096 (log=2)  
Fragment size=4096 (log=2)  
Stride=0 blocks, Stripe width=0 blocks  
13107200 inodes, 52428800 blocks  
2621440 blocks (5.00%) reserved for the super user  
First data block=0  
Maximum filesystem blocks=4294967296  
1600 block groups  
32768 blocks per group, 32768 fragments per group  
8192 inodes per group  
Superblock backups stored on blocks:  
    32768, 98304, 163840, 229376, 294912, 192096, 884736, 1605632, 2654208,  
    4096000, 7962624, 11239424, 20480000, 23887872  
  
Allocating group tables: done  
Writing inode tables: done  
Creating journal (32768 blocks): done  
Writing superblocks and filesystem accounting information: done  
  
Buildfile: /data/SharedMemoryCode/build.xml  
  
clean:  
  
precompile:  
    [mkdir] Created dir: /data/SharedMemoryCode/output  
  
compile:  
    [mkdir] Created dir: /data/SharedMemoryCode/build/classes  
    [javac] /data/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' wa  
s not set, defaulting to build.sysclasspath=last; set to false for repeatable bu  
ilds  
        [javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes  
  
jar:  
    [mkdir] Created dir: /data/SharedMemoryCode/build/jar  
    [jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar  
  
run:  
    [java] numberOfFilestoCreate: 1000  
    [java]  
    [java]  
    [java] Total Time= 4040103.0 milliseconds  
    [java] Total Time= 4040.103 seconds  
  
BUILD SUCCESSFUL  
Total time: 67 minutes 22 seconds  
ubuntu@ip-172-31-6-46:~$  
ubuntu@ip-172-31-6-46:~$  
ubuntu@ip-172-31-6-46:~$
```

Output file generated of required size 10 GB

```
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode/outputs  
output0 output177 output277 output368 output436 output515 output570 output98  
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode/outputs$ ls  
output0 output139 output172 output216 output255 output35 output338 output371 output409 output446 output480 output511 output535 output558 output58 output600 output82  
output10 output144 output174 output217 output259 output30 output34 output372 output410 output449 output481 output512 output536 output559 output580 output601 output83  
output100 output141 output176 output219 output261 output300 output340 output373 output411 output450 output482 output513 output537 output56 output581 output602 output86  
output103 output143 output177 output220 output262 output302 output342 output377 output412 output452 output486 output514 output538 output560 output582 output603 output87  
output104 output144 output179 output226 output263 output303 output343 output379 output413 output453 output487 output515 output539 output561 output583 output604 output9  
output107 output145 output181 output227 output264 output306 output344 output380 output414 output455 output488 output516 output54 output562 output584 output605 output90  
output108 output148 output182 output228 output265 output307 output345 output381 output415 output456 output489 output517 output540 output563 output585 output606 output92  
output109 output149 output183 output229 output266 output308 output346 output382 output416 output457 output490 output518 output541 output564 output586 output607 output94  
output110 output150 output187 output235 output271 output315 output353 output383 output417 output458 output491 output519 output542 output565 output587 output608 output95  
output112 output152 output188 output237 output272 output316 output356 output385 output419 output460 output492 output520 output543 output566 output588 output609 output96  
output113 output153 output189 output238 output273 output317 output357 output386 output421 output461 output493 output521 output544 output567 output589 output610 output97  
output116 output154 output191 output239 output279 output320 output359 output388 output422 output463 output494 output522 output545 output569 output590 output613 output98  
output117 output155 output194 output240 output281 output321 output360 output389 output424 output465 output495 output523 output546 output570 output593 output614 output99  
output120 output158 output195 output241 output285 output321 output361 output391 output425 output467 output496 output524 output547 output570 output592 output616 output100  
output121 output159 output196 output242 output288 output322 output359 output392 output427 output468 output497 output525 output549 output571 output593 output617 output101  
output122 output160 output197 output243 output289 output323 output362 output393 output428 output469 output498 output526 output550 output572 output594 output618 output102  
output124 output161 output198 output244 output290 output324 output363 output394 output429 output470 output499 output527 output551 output573 output595 output619 output103  
output126 output162 output199 output245 output291 output325 output364 output395 output430 output471 output500 output528 output552 output574 output596 output620 output104  
output127 output163 output200 output246 output292 output326 output365 output396 output431 output472 output501 output529 output553 output575 output597 output621 output105  
output131 output165 output206 output248 output293 output327 output367 output397 output432 output473 output502 output530 output554 output576 output598 output622 output106  
output132 output166 output207 output251 output301 output328 output368 output398 output433 output474 output503 output531 output555 output577 output599 output623 output107  
output134 output168 output209 output252 output302 output329 output369 output399 output434 output475 output504 output532 output556 output578 output604 output108 output108  
output137 output171 output214 output253 output303 output330 output370 output400 output444 output505 output533 output567 output590 output605 output109 output109  
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode/outputs$ ls  
output0 output139 output172 output216 output255 output35 output338 output371 output409 output446 output480 output511 output535 output558 output58 output600 output82  
output10 output144 output174 output217 output259 output30 output34 output372 output410 output449 output481 output512 output536 output559 output580 output601 output83  
output100 output141 output176 output219 output261 output300 output340 output373 output411 output450 output482 output513 output537 output56 output581 output602 output86  
output103 output143 output177 output220 output262 output302 output342 output377 output412 output452 output486 output514 output538 output560 output582 output603 output87  
output104 output144 output179 output226 output263 output303 output343 output379 output413 output453 output487 output515 output539 output561 output583 output604 output9  
output107 output145 output181 output227 output264 output306 output344 output380 output414 output455 output488 output516 output54 output562 output584 output605 output90  
output108 output148 output182 output228 output265 output307 output345 output381 output415 output456 output489 output517 output540 output563 output585 output606 output92  
output109 output149 output183 output229 output266 output308 output346 output382 output416 output457 output490 output518 output541 output564 output586 output607 output94  
output110 output150 output187 output235 output271 output315 output353 output383 output417 output458 output491 output519 output542 output565 output587 output608 output95  
output112 output152 output188 output237 output272 output316 output356 output385 output419 output460 output494 output520 output543 output566 output588 output609 output96  
output113 output153 output189 output238 output273 output317 output357 output386 output421 output461 output493 output521 output544 output567 output589 output610 output97  
output116 output154 output191 output239 output279 output320 output359 output388 output422 output463 output494 output522 output545 output569 output590 output613 output98  
output117 output155 output194 output240 output281 output321 output360 output389 output424 output465 output495 output523 output546 output570 output593 output614 output99  
output120 output158 output195 output241 output285 output321 output361 output391 output425 output467 output496 output524 output547 output570 output592 output616 output100  
output121 output159 output196 output242 output288 output322 output359 output392 output427 output468 output497 output525 output549 output571 output593 output617 output101  
output122 output160 output197 output243 output289 output323 output362 output393 output428 output470 output500 output526 output550 output572 output594 output618 output102  
output124 output161 output198 output244 output290 output324 output363 output394 output429 output471 output501 output529 output551 output573 output595 output620 output103  
output126 output162 output199 output245 output291 output325 output364 output395 output430 output472 output502 output530 output552 output574 output596 output621 output104  
output127 output163 output200 output246 output292 output326 output365 output396 output431 output473 output503 output531 output553 output575 output597 output622 output105  
output131 output165 output206 output248 output293 output327 output367 output397 output432 output474 output504 output532 output554 output576 output598 output623 output106  
output132 output166 output207 output251 output301 output328 output368 output398 output433 output475 output505 output531 output555 output577 output599 output624 output107  
output134 output168 output209 output252 output302 output329 output369 output399 output434 output476 output506 output533 output556 output578 output604 output108 output108  
output137 output171 output214 output253 output303 output330 output370 output406 output444 output507 output534 output567 output590 output605 output109 output109  
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode/outputs$
```

Valsort on output file – all records are in order

```
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./output
output655 output768 output787 output806 output825 output844 output863
output750 output769 output788 output807 output826 output845 output864
output751 output770 output789 output808 output827 output846 output865
output752 output771 output790 output809 output828 output847 output866
output753 output772 output791 output810 output829 output848 output867
output754 output773 output792 output811 output830 output849 output868
output755 output774 output793 output812 output831 output850 output869
output756 output775 output794 output813 output832 output851 output870
output757 output776 output795 output814 output833 output852 output871
output758 output777 output796 output815 output834 output853 output872
output759 output778 output797 output816 output835 output854 output873
output760 output779 output798 output817 output836 output855 output874
output761 output780 output799 output818 output837 output856
output762 output781 output800 output819 output838 output857
output763 output782 output801 output820 output839 output858
output764 output783 output802 output821 output840 output859
output765 output784 output803 output822 output841 output860
output766 output785 output804 output823 output842 output861
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./output$ ls
output840
output841 output883 output894 output896 output916 output926 output936
output842 output884 output895 output995 output937 output928 output937
output843 output885 output996 output938 output938 output938
output844 output886 output997 output998 output918 output929 output938
output845 output887 output999 output999 output919 output930 output939
output846 output888 output999 output999 output919 output926 output931
output847 output889 output999 output999 output919 output926 output940
output848 output890 output999 output999 output919 output921 output932
output849 output891 output999 output999 output919 output922 output933
output850 output892 output999 output999 output919 output923 output934
output851 output893 output999 output999 output919 output924 output935
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./output$ ls
output998
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ cd ..
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ls
build.xml input output src
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ cd /home
ubuntu@ip-172-31-6-46:~/home$ cd ubuntu/
ubuntu@ip-172-31-6-46:~/home/ubuntu$ ls
abhijeetkedar1.pem gensort-linux-1.5.tar.gz SharedMemoryCode
gensort-linux-1.5 InstallationScript.sh SharedMemoryCode.tar.gz
ubuntu@ip-172-31-6-46:~/home/ubuntu$ cd gensort-linux-1.5/
ubuntu@ip-172-31-6-46:~/gensort-linux-1.5$ cd 64
ubuntu@ip-172-31-6-46:~/gensort-linux-1.5/64$ ls
gensort valsart
ubuntu@ip-172-31-6-46:~/gensort-linux-1.5/64$ mv valsart /data/SharedMemoryCode/
output/
ubuntu@ip-172-31-6-46:~/gensort-linux-1.5/64$ cd /data/SharedMemoryCode/output/
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./valsart
output998 valsart
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./valsart output998
Records: 00000000000000000000000000000000
Checksum: 262898746e89a8
Duplicate keys: 0
SUCFSS - all records are in order
ubuntu@ip-172-31-6-46:~/data/SharedMemoryCode$ ./output$
```

Head and tail command on output file, I have included these records in separate output file.

4 Thread

Total time taken by Shared-Memory with 2 thread

```
ubuntu@ip-172-31-10-32: ~
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

mke2fs 1.42.9 (4-Feb-2014)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=8 blocks, Stripe width=8 blocks
13107200 inodes, 52428800 blocks
2621440 blocks (5.00%) reserved for the super user
First data block=0
Block group size=4096 blocks
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

Buildfile: /data/SharedMemoryCode/build.xml

clean:

precompile:
[mkdir] Created dir: /data/SharedMemoryCode/output

compile:
[mkdir] Created dir: /data/SharedMemoryCode/build/classes
[javac] /data/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds
[javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes

jar:
[mkdir] Created dir: /data/SharedMemoryCode/build/jar
[jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar

run:
[java] numberOfFilestoCreate: 1000
[java]
[java]
[java]
[java] Total Time= 3310334.0 milliseconds
[java] Total Time= 3310.334 seconds

BUILD SUCCESSFUL
Total time: 55 minutes 12 seconds
ubuntu@ip-172-31-10-32:~$
```

Output file generated of required size 10 GB

```
ubuntu@ip-172-31-10-32: /data/SharedMemoryCode/output
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode/output$ ll
total 976564
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 25 19:58 ./
drwxrwxr-x 5 ubuntu ubuntu 4096 Mar 25 19:03 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 25 19:58 output998
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode/output$
```

Valsort on output file – all records are in order

```
ubuntu@ip-172-31-10-32:~/data/SharedMemoryCode/gensorT-linux-1.5/64
Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

Buildfile: /data/SharedMemoryCode/build.xml

clean:
[clean]

precompile:
  [mkdir] Created dir: /data/SharedMemoryCode/output

compile:
  [mkdir] Created dir: /data/SharedMemoryCode/build/classes
  [javac] /data/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds
  [javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes

jar:
  [mkdir] Created dir: /data/SharedMemoryCode/build/jar
  [jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar

run:
  [java] numberOffilestoCreate: 1000
  [java]
  [java]
  [java]
  [java] Total Time:   3310334.0 milliseconds
  [java] Total Time:   3310.334 seconds

BUILD SUCCESSFUL
Total time: 56 minutes 12 seconds
ubuntu@ip-172-31-10-32:~$ ls
abbijeekdarji.pem gensorT-linux-1.5 gensorT-linux-1.5.tar.gz InstallationScript.sh SharedMemoryCode SharedMemoryCode.tar.gz
ubuntu@ip-172-31-10-32:~$ cd /data/SharedMemoryCode/
ubuntu@ip-172-31-10-32:~/data/SharedMemoryCode$ ls
build build.xml input output src
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode$ cd ..
ubuntu@ip-172-31-10-32:/data$ cd /home/
ubuntu@ip-172-31-10-32:/home$ ls
ubuntu
ubuntu@ip-172-31-10-32:/home$ cd ubuntu/
ubuntu@ip-172-31-10-32:~/ubuntu$ gensorT-linux-1.5 gensorT-linux-1.5.tar.gz InstallationScript.sh SharedMemoryCode SharedMemoryCode.tar.gz
ubuntu@ip-172-31-10-32:~$ mv gensorT-linux-1.5 /data/SharedMemoryCode/
ubuntu@ip-172-31-10-32:~$ cd /data/SharedMemoryCode/gensorT-linux-1.5/64/
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode/gensorT-linux-1.5/64$ ls
gensorT valsort
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode/gensorT-linux-1.5/64$ ./valsort /data/SharedMemoryCode/output/output998
Records: 100000000
Checksum: 2faF0ab746e899a8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-10-32:/data/SharedMemoryCode/gensorT-linux-1.5/64$
```

Head and tail command on output file, I have included these records in separate output file

6 Thread

Total time taken by Shared-Memory with 2 thread

```
ubuntu@ip-172-31-8-231: ~
mke2fs 1.42.9 (4-Feb-2014)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
13107200 inodes, 52428800 blocks
2621440 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=4294967296
1600 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98384, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

Buildfile: /data/SharedMemoryCode/build.xml

clean:
[mkdir] Created dir: /data/SharedMemoryCode/output

compile:
[mkdir] Created dir: /data/SharedMemoryCode/build/classes
[javac] /data/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds
[javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes

jar:
[mkdir] Created dir: /data/SharedMemoryCode/build/jar
[jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar

run:
[java] numberOfFilestoCreate: 1000
[[[
[java]
[java]
[java]
[java] Total Time= 3165345.0 milliseconds
[java] Total Time= 3165.345 seconds

BUILD SUCCESSFUL
Total time: 52 minutes 47 seconds
ubuntu@ip-172-31-8-231:~$
```

Valsort on output file – all records are in order

```
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output
drwxrwxr-x 2 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 19:46 output996
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 19:46 output997
-rw-rw-r-- 1 ubuntu ubuntu 3744268288 Mar 27 19:49 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ll
total 14711312
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 27 19:46 ../
drwxrwxr-x 8 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 19:46 output996
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 19:46 output997
-rw-rw-r-- 1 ubuntu ubuntu 5864535360 Mar 27 19:50 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ll
total 15429576
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 27 19:46 ../
drwxrwxr-x 8 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 19:46 output996
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 19:46 output997
-rw-rw-r-- 1 ubuntu ubuntu 5799837696 Mar 27 19:50 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ll
total 16540296
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 27 19:46 ../
drwxrwxr-x 8 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 19:46 output996
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 19:46 output997
-rw-rw-r-- 1 ubuntu ubuntu 6937214976 Mar 27 19:51 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ll
total 18654528
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 27 19:46 ../
drwxrwxr-x 8 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 19:46 output996
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 19:46 output997
-rw-rw-r-- 1 ubuntu ubuntu 9102180352 Mar 27 19:53 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ll
total 9765664
drwxrwxr-x 2 ubuntu ubuntu 28672 Mar 27 19:53 ../
drwxrwxr-x 8 ubuntu ubuntu 4096 Mar 27 19:01 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 27 19:53 output998
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ cd /home/ubuntu/gensort-linux-1.5/64
gensort valsort
ubuntu@ip-172-31-8-231:~/gensort-linux-1.5/64$ ls
output998 valsor
ubuntu@ip-172-31-8-231:~/gensort-linux-1.5/64$ mv valsor /data/SharedMemoryCode/
output/
ubuntu@ip-172-31-8-231:~/gensort-linux-1.5/64$ cd /data/SharedMemoryCode/output/
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ls
output998 valsor
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$ ./valsor output998
Records: 100000000
Checksum: 2fa0ab746e89a8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-8-231:/data/SharedMemoryCode/output$
```

Head and tail command on output file, I have included these records in separate output file

8 Thread

Total time taken by Shared-Memory with 2 thread

```
ubuntu@ip-172-31-2-156:~  
Writing superblocks and filesystem accounting information: done  
  
mke2fs 1.42.9 (4-Feb-2014)  
Filesystem label=  
OS type: linux  
Block size=4096 (log=2)  
Fragment size=4096 (log=2)  
Stride=0 blocks, Stripe width=0 blocks  
13107200 inodes, 52428800 blocks  
2621440 blocks (5.00%) reserved for the super user  
First data block=0  
Maximum filesystem blocks=4294967296  
6400 blocks per group, 6400 fragments per group  
32768 blocks per group, 32768 fragments per group  
8192 inodes per group  
Superblock backups stored on blocks:  
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,  
    4096000, 7962624, 11239424, 20480000, 23887872  
  
Allocating group tables: done  
Writing inode tables: done  
Creating journal (32768 blocks): done  
Writing superblocks and filesystem accounting information: done  
  
Buildfile: /data/SharedMemoryCode/build.xml  
  
clean:  
  
precompile:  
    [mkdir] Created dir: /data/SharedMemoryCode/output  
  
compile:  
    [mkdir] Created dir: /data/SharedMemoryCode/build/classes  
    [javac] /data/SharedMemoryCode/build.xml:2: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds  
        [javac] Compiling 4 source files to /data/SharedMemoryCode/build/classes  
  
jar:  
    [mkdir] Created dir: /data/SharedMemoryCode/build/jar  
    [jar] Building jar: /data/SharedMemoryCode/build/jar/SharedMemory.jar  
  
run:  
    [java] numberOfFilestoCreate: 1000  
    [java]  
    [java]  
    [java]  
    [java] Total Time= 3081.287.0 milliseconds  
    [java] Total Time= 3081.287 seconds  
  
BUILD SUCCESSFUL  
Total time: 51 minutes 23 seconds  
ubuntu@ip-172-31-2-156:~$
```

Output file generated of required size 10 GB

```
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 17019452
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 7427874816 Mar 27 21:00 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18711428
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 7869112320 Mar 27 21:00 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18711428
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9160459240 Mar 27 21:01 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18745128
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9194962944 Mar 27 21:01 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 19346876
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9811156992 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 19518228
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9986613248 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 9765660
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 21:02 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 9765660
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 21:02 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
```

Valsort on output file – all records are in order

```
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18711428
-rw-rw-r-- 1 ubuntu ubuntu 7869112320 Mar 27 21:00 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18711428
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9160459240 Mar 27 21:01 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 18745128
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9194962944 Mar 27 21:01 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 19346876
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9811156992 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 19518228
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 20:54 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 4880000000 Mar 27 20:54 output996
-rw-rw-r-- 1 ubuntu ubuntu 5120000000 Mar 27 20:54 output997
-rw-rw-r-- 1 ubuntu ubuntu 9986613248 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 9765660
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 21:02 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ll
total 9765660
drwxrwxr-x 2 ubuntu ubuntu    24576 Mar 27 21:02 ./
drwxrwxr-x 9 ubuntu ubuntu    4096 Mar 27 20:10 ../
-rw-rw-r-- 1 ubuntu ubuntu 10000000000 Mar 27 21:02 output998
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ cd /home/ubuntu/gensort-l
nux-1.5/64/
ubuntu@ip-172-31-2-156:~/gensort-linux-1.5/64$ ls
gensort valsort
ubuntu@ip-172-31-2-156:~/gensort-linux-1.5/64$ mv valsort /data/SharedMemoryCode
/output/
ubuntu@ip-172-31-2-156:~/gensort-linux-1.5/64$ cd /data/SharedMemoryCode/output
/
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$ ./valsort output998
Records: 1000000000
Checksum: 2fa0ab746e89a8
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-2-156:/data/SharedMemoryCode/output$
```

Head and tail command on output file, I have included these records in separate output file

Extra Part – 1 GB – 1 NODE

Started Name node

Files created with size of 10 MB each file

```
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ls
-rw-rw-r- 1 ubuntu ubuntu     845 Mar 20 21:08 build.xml
-rw-rw-r- 1 ubuntu ubuntu   226 Mar 16 18:22 .classpath
-rwxrwxr-x 1 ubuntu ubuntu 1000000000 Mar 20 22:34 input*
drwxrwxr-x 2 ubuntu ubuntu    4096 Mar 20 21:53 output/
-rw-rw-r- 1 ubuntu ubuntu    371 Mar 20 18:22 .project
-rw-rw-r- 1 ubuntu ubuntu   1122 Mar 20 21:39 SCP_Commands
drwxrwxr-x 2 ubuntu ubuntu    4096 Mar 20 22:21 src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ rm -rf output/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ll
total 976604
drwxrwxr-x 5 ubuntu ubuntu    4096 Mar 20 22:37 ./
drwxrwxr-x 6 ubuntu ubuntu    4096 Mar 20 22:33 ../
drwxrwxr-x 2 ubuntu ubuntu    4096 Mar 20 22:21 bin/
drwxrwxr-x 4 ubuntu ubuntu    4096 Mar 20 21:52 build/
-rw-rw-r- 1 ubuntu ubuntu    845 Mar 20 21:08 build.xml
drwxrwxr-x 1 ubuntu ubuntu   226 Mar 16 18:22 .classpath
-rwxrwxr-x 1 ubuntu ubuntu 1000000000 Mar 20 22:34 input*
-rw-rw-r- 1 ubuntu ubuntu    371 Mar 16 18:22 .project
-rw-rw-r- 1 ubuntu ubuntu   1122 Mar 20 21:39 SCP_Commands
drwxrwxr-x 2 ubuntu ubuntu    4096 Mar 20 22:21 src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ant
Buildfile: /home/ubuntu/SharedMemoryCode/build.xml

clean:
    [delete] Deleting directory /home/ubuntu/SharedMemoryCode/build

precompile:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/output

compile:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/build/classes
    [javac] /home/ubuntu/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds
        [javac] Compiling 4 source files to /home/ubuntu/SharedMemoryCode/build/classes
jar:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/build/jar
    [jar] Building jar: /home/ubuntu/SharedMemoryCode/build/jar/SharedMemory.jar
jar:
    [jar]

run:
    [java] numberOfFilestoCreate: 100
    [java]
    [java]
    [java]
    [java]
    [java] Total Time= 152704.0 milliseconds
    [java] Total Time= 152.704 seconds

BUILD SUCCESSFUL
Total time: 2 minutes 40 seconds
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$
```

Time taken by Shared-Memory for sorting 1 GB.

```
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ls
drwxrwxr-x 5 ubuntu ubuntu 152704 Mar 20 22:37 bin/
drwxrwxr-x 6 ubuntu ubuntu 152704 Mar 20 22:33 build/
drwxrwxr-x 2 ubuntu ubuntu 152704 Mar 20 22:21 config/
drwxrwxr-x 4 ubuntu ubuntu 152704 Mar 20 21:52 classes/
-rw-rw-r- 1 ubuntu ubuntu 152704 Mar 20 21:08 build.xml
-rw-rw-r- 1 ubuntu ubuntu 152704 Mar 16 18:22 .classpath
-rwxrwxr-x 1 ubuntu ubuntu 152704 Mar 20 22:34 input*
drwxrwxr-x 2 ubuntu ubuntu 152704 Mar 20 22:21 output/
-rw-rw-r- 1 ubuntu ubuntu 152704 Mar 20 18:22 .project
-rw-rw-r- 1 ubuntu ubuntu 152704 Mar 20 21:39 SCP_Commands
drwxrwxr-x 2 ubuntu ubuntu 152704 Mar 20 22:21 src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ rm -rf output/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ll
total 976604
drwxrwxr-x 5 ubuntu ubuntu 152704 Mar 20 22:37 bin/
drwxrwxr-x 6 ubuntu ubuntu 152704 Mar 20 22:33 build/
drwxrwxr-x 2 ubuntu ubuntu 152704 Mar 20 22:21 config/
drwxrwxr-x 4 ubuntu ubuntu 152704 Mar 20 21:52 classes/
bin build.xml  input output SCP_Commands src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ cd src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode/src$ ls
config.properties MergeFiles.java SharedMemory.java SortFiles.java WriteFile.java
ubuntu@ip-172-31-1-35:~/SharedMemoryCode/src$ cd ..
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ls
bin build.xml  input output SCP_Commands src/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ rm -f output/
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$ ant
Buildfile: /home/ubuntu/SharedMemoryCode/build.xml

clean:
    [delete] Deleting directory /home/ubuntu/SharedMemoryCode/build

precompile:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/output

compile:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/build/classes
    [javac] /home/ubuntu/SharedMemoryCode/build.xml:12: warning: 'includeantruntime' was not set, defaulting to build.sysclasspath=last; set to false for repeatable builds
        [javac] Compiling 4 source files to /home/ubuntu/SharedMemoryCode/build/classes
jar:
    [mkdir] Created dir: /home/ubuntu/SharedMemoryCode/build/jar
    [jar] Building jar: /home/ubuntu/SharedMemoryCode/build/jar/SharedMemory.jar
run:
    [java] numberOfFilestoCreate: 10
    [java]
    [java]
    [java]
    [java]
    [java] Total Time= 132632.0 milliseconds
    [java] Total Time= 132.632 seconds

BUILD SUCCESSFUL
Total time: 2 minutes 14 seconds
ubuntu@ip-172-31-1-35:~/SharedMemoryCode$
```

Hadoop

A. 1 Node for 10 GB

Total Time taken

```
ubuntu@ip-172-31-6-158: ~
Killed map tasks=1
Launched map tasks=76
Launched reduce tasks=1
Data local tasks=76
Total time spent by all maps in occupied slots (ms)=4864779
Total time spent by all reduces in occupied slots (ms)=983532
Total time spent by all map tasks (ms)=4864779
Total time spent by all reduce tasks (ms)=983532
Total vcore-milliseconds taken by all map tasks=4864779
Total vcore-milliseconds taken by all reduce tasks=983532
Total megabyte-milliseconds taken by all map tasks=4981533696
Total megabyte-milliseconds taken by all reduce tasks=1007136768

Map-Reduce Framework
Map input records=100000000
Map output records=100000000
Map output bytes=10000000000
Map output materialized bytes=10200000450
Input split bytes=10050
Combine input records=0
Combine output records=0
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)=71024
CPU time spent (ms)=1463018
Physical memory (bytes) snapshot=19441934336
Virtual memory (bytes) snapshot=62900916224
Total committed heap usage (bytes)=15611723776

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= 1198136.0 milliseconds
Total Time= 1198.136 seconds
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
```

Validation sort

```
ubuntu@ip-172-31-6-158: ~
Total time spent by all maps in occupied slots (ms)=4864779
Total time spent by all reduces in occupied slots (ms)=983532
Total time spent by all map tasks (ms)=4864779
Total time spent by all reduce tasks (ms)=983532
Total vcore-milliseconds taken by all map tasks=4864779
Total vcore-milliseconds taken by all reduce tasks=983532
Total megabyte-milliseconds taken by all map tasks=4981533696
Total megabyte-milliseconds taken by all reduce tasks=1007136768

Map-Reduce Framework
Map input records=100000000
Map output records=100000000
Map output bytes=10000000000
Map output materialized bytes=10200000450
Input split bytes=10050
Combine input records=0
Combine output records=0
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)=71024
CPU time spent (ms)=1463018
Physical memory (bytes) snapshot=19441934336
Virtual memory (bytes) snapshot=62900916224
Total committed heap usage (bytes)=15611723776

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= 1198136.0 milliseconds
Total Time= 1198.136 seconds
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

Records: 100000000
Checksum: 2fafefb9aceec00
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-6-158:~$
```

Name Node

Overview 'ec2-54-172-180-220.compute-1.amazonaws.com:9000' (active)

Started:	Fri Mar 25 09:30:17 UTC 2016
Version:	2.7.2, r165c4fe8a74265c792e23f546c64604ac70e41
Compiled:	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
Cluster ID:	CID-2bd96c69-9e8b-45be-9180-b59200c00ff6
Block Pool ID:	BP-908280451-172.31.6.158-1458898213615

Summary

Security is off.
SafeMode is off.
17 files and directories, 153 blocks = 170 total filesystem object(s).
Heap Memory used 129.26 MB of 174.5 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 32.34 MB of 33.44 MB Committed Non Heap Memory. Max Non Heap Memory is 214 MB.

Configured Capacity	73.7 GB
DFS Used:	18.77 GB (25.47%)
Non DFS Used:	13.13 GB
DFS Remaining:	41.79 GB (56.71%)
Block Pool Used:	18.77 GB (25.47%)
DataNodes usages% (Min/Median/Max/stdDev):	25.47% / 25.47% / 25.47% / 0.00%
Live Nodes	1 (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	3/25/2016, 4:30:17 AM

Resource Node

FINISHED Applications

Cluster Metrics

Nodes	Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	Vcores Used	Vcores Total	Vcores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
1	0	0	1	0	0 B	8 GB	0 B	0	8	0	1	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:1024, vCores:1>	<memory:8192, vCores:8>

Tools

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
application_1458898240896_0001	ubuntu	TeraSort_AbhijeetKedari	MAPREDUCE	default	Fri Mar 25 04:36:34 -0500 2016	Fri Mar 25 04:56:27 -0500 2016	FINISHED	SUCCEEDED	N/A	History	N/A

Showing 1 to 1 of 1 entries

B. 16 Node for 100GB

Total time taken

```
ubuntu@ip-172-31-1-22:~/hadoop-2.7.2
Killed map tasks=4
Launched map tasks=749
Launched reduce tasks=1
Data local map tasks=20
Rack-local map tasks=710
Total time spent by all maps in occupied slots (ms)=243096590
Total time spent by all reduces in occupied slots (ms)=7003577
Total time spent by all map tasks (ms)=243096590
Total time spent by all reduce tasks (ms)=7003577
Total vcore-milliseconds taken by all map tasks=243096590
Total vcore-milliseconds taken by all reduce tasks=7003577
Total megabyte-milliseconds taken by all map tasks=24893098169
Total megabyte-milliseconds taken by all reduce tasks=7171662848
Map-Reduce Framework
  Map input records=10000000000
  Map output records=10000000000
  Map output bytes=100000000000
  Map output materialized bytes=102000004470
  Input split bytes=98340
  Combine input records=0
  Reduce input groups=10000000000
  Reduce shuffle bytes=102000004470
  Reduce input records=10000000000
  Reduce output records=10000000000
  Spilled Records=5038923127
  Shuffled Maps =745
  Failed Shuffles=0
  Merged Map outputs=745
  GC time elapsed (ms)=334349
  CPU time spent (ms)=14046850
  Physical memory (bytes) snapshot=189866491904
  Virtual memory (bytes) snapshot=618651140906
  Total committed heap usage (bytes)=131711631360
Shuffle Errors
  Bad Src Dst
  CONNECTION=0
  IO_ERROR=0
  WRONG_LENGTH=0
  WRONG_MAP=0
  WRONG_REDUCE=0
File Input Format Counters
  Bytes Read=100003047424
File Output Format Counters
  Bytes Written=1000000000000

Total Time= 745263.0 milliseconds
Total Time= 745.263 seconds
ubuntu@ip-172-31-1-22:~/hadoop-2.7.2$ 
ubuntu@ip-172-31-1-22:~/hadoop-2.7.2$ 
ubuntu@ip-172-31-1-22:~/hadoop-2.7.2$ 
```

Resource Node

The screenshot shows the EC2 Management console with the 'FINISHED Applications' tab selected. The main content area displays the 'FINISHED Applications' section, which includes a table of completed applications. The table has columns for ID, User, Name, Application Type, Queue, Start Time, Finish Time, State, Final Status, Progress, Tracking URL, and Blacklisted Nodes. One entry is visible:

ID	User	Name	Application Type	Queue	Start Time	Finish Time	State	Final Status	Progress	Tracking URL	Blacklisted Nodes
application_1459280939323_0001	ubuntu	TeraSort_AbhiheetKedari	MAPREDUCE	default	Tue Mar 29 15:53:13 -0500 2016	Tue Mar 29 17:57:20 -0500 2016	FINISHED	SUCCEEDED	N/A		

Below the table, it says 'Showing 1 to 1 of 1 entries'. The left sidebar contains navigation links for the cluster, nodes, and applications, along with a scheduler section and various tools.

Name Node

Overview 'ec2-52-91-226-46.compute-1.amazonaws.com:9000' (active)

Started:	Tue Mar 29 19:48:29 UTC 2016
Version:	2.7.2, rb165c4fe8a74265c792ce23f546c64604acf0e41
Compiled:	2016-01-26T00:08Z by jenkins from (detached from b165c4f)
Cluster ID:	CID-eb1ce23b-83ae-4b71-84e-b41b4ae3a9f1
Block Pool ID:	BP-723024956-172.31.1.22-1459280893876

Summary

Security is off.
SafeMode is off.
21 files and directories, 752 blocks = 773 total filesystem object(s).
Heap Memory used 33.24 MB of 210 MB Heap Memory, Max Heap Memory is 889 MB.
Non Heap Memory used 38.19 MB of 73.31 MB Committed Non Heap Memory, Max Non Heap Memory is 214 MB.

Configured Capacity:	11.44 TB
DFS Used:	93.92 GB (0.8%)
Non DFS Used:	784.31 GB
DFS Remaining:	10.58 TB (92.5%)
Block Pool Used:	93.92 GB (0.8%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 13.63% / 3.21%
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

C. Extra credit Part - 1 Node for 1 GB

```
ubuntu@ip-172-31-15-22:~/hadoop-2.7.2 abhijeet@abhijeet-Inspiron-5547: ~/Downloads/C5553/Assignment2
Job Counters
  Killed map tasks=1
  Launched map tasks=8
  Launched reduce tasks=1
  Data-local map tasks=8
  Total time spent by all maps in occupied slots (ms)=499952
  Total time spent by all reduces in occupied slots (ms)=186176
  Total time spent by all map tasks (ms)=249976
  Total time spent by all reduce tasks (ms)=46544
  Total vcore-milliseconds taken by all map tasks=249976
  Total vcore-milliseconds taken by all reduce tasks=46544
  Total megabyte-milliseconds taken by all map tasks=1023901696
  Total megabyte-milliseconds taken by all reduce tasks=381288448

Map-Reduce Framework
  Map input records=10000000
  Map output records=10000000
  Map output bytes=1000000000
  Map output materialized bytes=1020000048
  Input split bytes=1056
  Combine input records=0
  Combine output records=0
  Reduce input groups=10000000
  Reduce shuffle bytes=1020000048
  Reduce input records=10000000
  Reduce output records=10000000
  Spilled Records=9395241
  Shuffled Maps =8
  Failed Shuffles =0
  Merged Map outputs=8
  GC time elapsed (ms)=3977
  CPU time spent (ms)=125680
  Physical memory (bytes) snapshot=8360329216
  Virtual memory (bytes) snapshot=38937350144
  Total committed heap usage (bytes)=11781799936

Shuffle Errors
  BAD_ID=0
  CONNECTION=0
  IO_ERROR=0
  WRONG_HOSTNAME=0
  WRONG_MAP=0
  WRONG_REDUCE=0

File Input Format Counters
  Bytes Read=1000028672
File Output Format Counters
  Bytes Written=1000000000

Total Time= 94943.0 milliseconds
Total Time= 94.943 seconds
ubuntu@ip-172-31-15-22:~/hadoop-2.7.2$ cd hadoop-2.7.2/
```

Validation sort

Head and tail for 1 GB- 1 Node

Resource Node for 1GB – 1 Node

The screenshot shows the 'FINISHED Applications' section of the Hadoop interface. On the left, there's a sidebar with icons for various cluster metrics like 'About', 'Nodes', 'Node Labels', 'Applications' (with sub-options: NEW, NEW_SAVING, SUBMITTED, ACCEPTED, RUNNING, FINISHED, FAILED, KILLED), 'Scheduler', and 'Tools'. The main area has a title 'FINISHED Applications' and a sub-section 'Cluster Metrics' with a table showing 1 completed application. Below that is 'Scheduler Metrics' showing a Capacity Scheduler with [MEMORY] as the scheduling resource type, minimum allocation <memory:2048, vCores:1>, and maximum allocation <memory:8192, vCores:8>. A detailed table lists one finished job: 'application_1458447769777_0001' by user 'ubuntu' with name 'TeraSort_AbhijeetKedarl', type 'MAPREDUCE', queue 'default', started at 'Sat Mar 19 23:28:34 -0500 2016', finished at 'Sat Mar 19 23:30:05 -0500 2016', state 'FINISHED', final status 'SUCCEEDED', progress 'History N/A', and tracking UI 'N/A'. The bottom of the table shows 'Showing 1 to 1 of 1 entries' and navigation links.

Name Node for 1 GB – 1 Node

The screenshot shows the 'Name Node Information' page. The top navigation bar includes tabs for 'Hadoop', 'Overview', 'Datanodes', 'Datanode Volume Failures', 'Snapshot', 'Startup Progress', and 'Utilities'. The 'Overview' tab is active. Below it, the title is 'Overview 'ec2-52-201-255-9.compute-1.amazonaws.com:9000' (active)'. A table provides basic information: Started (Sun Mar 20 04:22:23 UTC 2016), Version (2.7.2, r1b165c4fe8a74265c792ce23f546c64604acf0e41), Compiled (2016-01-26T00:08Z by jenkins from (detached from b165c4f)), Cluster ID (CID-41e28670-7fe7-4c67-8542-e9c450bd0c0), and Block Pool ID (BP-1534490427-172.31.15.22-1458447191514). The 'Summary' section follows, containing text about security and safemode being off, and statistics on files, directories, and memory usage. Below this is a table with various metrics: Configured Capacity (7.74 GB), DFS Used (1.88 GB (24.26%)), Non DFS Used (3.12 GB), DFS Remaining (2.74 GB (35.37%)), Block Pool Used (1.88 GB (24.26%)), DataNodes usages% (Min/Median/Max/stdDev) (24.26% / 24.26% / 24.26% / 0.00%), Live Nodes (1 (Decommissioned: 0)), Dead Nodes (0 (Decommissioned: 0)), and Decommissioning Nodes (0).

D. Extra credit Part - 16 Node for 1 GB

Total time taken by Mapreduce

```
ubuntu@ip-172-31-15-150: ~/hadoop-2.7.2
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
  Killed map tasks=1
  Launched map tasks=9
  Launched reduce tasks=1
  Rack-local map tasks=9
  Total time spent by all maps in occupied slots (ms)=188636
  Total time spent by all reduces in occupied slots (ms)=120064
  Total time spent by all map tasks (ms)=94318
  Total time spent by all reduce tasks (ms)=30016
  Total vcore-milliseconds taken by all map tasks=94318
  Total vcore-milliseconds taken by all reduce tasks=30016
  Total negabyte-milliseconds taken by all map tasks=386326528
  Total negabyte-milliseconds taken by all reduce tasks=245891072
Map-Reduce Framework
  Map input records=100000000
  Map output records=100000000
  Map output bytes=1000000000
  Map output materialized bytes=10200000048
  Input split bytes=1072
  Combine input records=0
  Combine output records=0
  Reduce input groups=100000000
  Reduce shuffle bytes=1000000048
  Reduce input records=100000000
  Reduce output records=100000000
  Spilled Records=29395241
  Shuffled Maps =8
  Failed Shuffles=0
  Merged Map outputs=8
  GC time elapsed (ms)=659
  CPU time spent (ms)=95180
  Physical memory (bytes) snapshot=7644659712
  Virtual memory (bytes) snapshot=38989832192
  Total committed heap usage (bytes)=1176557056
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=1000028672
File Output Format Counters
  Bytes Written=10000000000
Total Time= 52891.0 milliseconds
Total Time= 52.891 seconds
ubuntu@ip-172-31-15-150:~/hadoop-2.7.2$
```

Validation result

```
ubuntu@ip-172-31-15-150: ~/gensort-linux-1.5/64
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done
ubuntu@ip-172-31-15-150:~$ sudo mke2fs -F -t ext4 /dev/md0
mke2fs 1.42.9 (4-Feb-2014)
Filesystem label=
OS type: Linux
Block size=4096 (Log=2)
Fragment size=4096 (Log=2)
Stride=128 blocks, Stripe width=384 blocks
1831432 inodes, 1465127040 blocks
7305632 (5.00%) reserved for the super user
First data block=9
Maximum filesystem blocks=4294967296
44713 block groups
32768 blocks per group, 32768 fragments per group
4096 inodes per group
Superblock backups stored on blocks:
    32768, 98384, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 2388782, 71663616, 78675968,
    102400000, 214990848, 512000000, 550731776, 644972544
Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done
ubuntu@ip-172-31-15-150:~$ sudo mkdir /data
ubuntu@ip-172-31-15-150:~$ sudo mount /dev/md0 /data
ubuntu@ip-172-31-15-150:~$ sudo chmod 777 /data
ubuntu@ip-172-31-15-150:~$ ls
abhijeetkedar1.gen  firstscript.sh  gensort-linux-1.5  hadoop-2.7.2  TeraSortJob.jar
ubuntu@ip-172-31-15-150:~$ cd hadoop-2.7.2
ubuntu@ip-172-31-15-150:~/hadoop-2.7.2$ bin/hadoop dfs -get /abhijeet/output/part-r-00000 /home/ubuntu
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
ubuntu@ip-172-31-15-150:~/hadoop-2.7.2$ ls
bin  etc  include  lib  libexec  LICENSE.txt  logs  NOTICE.txt  README.txt  sbin  share
ubuntu@ip-172-31-15-150:~/hadoop-2.7.2$ cd ..
ubuntu@ip-172-31-15-150:~$ ls
abhijeetkedar1.gen  firstscript.sh  gensort-linux-1.5  hadoop-2.7.2  part-r-00000  TeraSortJob.jar
ubuntu@ip-172-31-15-150:~/hadoop-2.7.2$ cd /home/ubuntu/gensort-linux-1.5/64/
ubuntu@ip-172-31-15-150:~/gensort-linux-1.5/64$ gensort inout part-r-00000 valsort
ubuntu@ip-172-31-15-150:~/gensort-linux-1.5/64$ ./valsort part-r-00000
Records: 10000000
Checksum: 4c49dd4547fe73
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-15-150:~/gensort-linux-1.5/64$
```

Name node

Summary

Security is off.
Safemode is off.
17 files and directories, 19 blocks = 36 total filesystem object(s).
Heap Memory used 55.44 MB of 459 MB Heap Memory. Max Heap Memory is 889 MB.
Non Heap Memory used 36.24 MB of 37.44 MB Committed Non Heap Memory. Max Non Heap Memory is 214 MB.

Configured Capacity:	123.83 GB
DFS Used:	1.88 GB (1.52%)
Non DFS Used:	32.84 GB
DFS Remaining:	89.11 GB (71.97%)
Block Pool Used:	1.88 GB (1.52%)
DataNodes usages% (Min/Median/Max/stdDev):	0.00% / 0.00% / 12.13% / 4.01%
Live Nodes	16 (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	3/21/2016, 4:35:09 AM

NameNode Journal Status

Mozilla Firefox seems slow... to... start. Learn How to Speed It Up | Don't Tell Me Again

Resource Node

FINISHED Applications - Mozilla Firefox

FINISHED Applications

Logged in as: dr.who

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
1	0	0	1	0	0 B	513.98 GB	0 B	0	136	0	17	0	0	0	0

Scheduler Metrics

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
application_1458552948440_0001	ubuntu	TeraSort_AbhijeetKedari	MAPREDUCE	default	Mon Mar 21 04:41:40 2016	Mon Mar 21 04:42:27 -0500 2016	FINISHED	SUCCEEDED	History	N/A	

Showing 1 to 1 of 1 entries First Previous 1 Next Last

Mozilla Firefox seems slow... to... start. Learn How to Speed It Up | Don't Tell Me Again

E. Extra credit Part - 1 Node for 1 TB

Time taken by MapReduce for 1TB dataset by 1 Node

```
ubuntu@ip-172-31-61-62:~/hadoop-2.7.2$ HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
  Killed map tasks=1
  Launched map tasks=7452
  Launched reduce tasks=1
  Data-local map tasks=7452
  Total time spent by all maps in occupied slots (ms)=468780208
  Total time spent by all reduces in occupied slots (ms)=293541984
  Total time spent by all map tasks (ms)=234390104
  Total time spent by all reduce tasks (ms)=73385496
  Total vcore-milliseconds taken by all map tasks=234390104
  Total vcore-milliseconds taken by all reduce tasks=73385496
  Total megabyte-milliseconds taken by all map tasks=960861865984
  Total megabyte-milliseconds taken by all reduce tasks=601173983232
Map-Reduce Framework
  Map input records=10000000000
  Map output records=10000000000
  Map output bytes=1000000000000
  Map output materialized bytes=10200000044706
  Input split bytes=999983
  Combine input records=0
  Combine output records=0
  Reduce input groups=10000000000
  Reduce shuffle bytes=10200000044706
  Reduce input records=10000000000
  Reduce output records=10000000000
  Spilled Records=49340996964
  Shuffled Maps =7451
  Failed Shuffles=0
  Merged Map outputs=7451
  GC time elapsed (ms)=5248742
  CPU time spent (ms)=171576270
  Physical memory (bytes) snapshot=6967979626496
  Virtual memory (bytes) snapshot=29467218808832
  Total committed heap usage (bytes)=10101884190720
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=1000030515200
File Output Format Counters
  Bytes Written=10000000000000
Total Time= 8.4242293E7 milliseconds
Total Time= 84242.293 seconds
ubuntu@ip-172-31-61-62:~/hadoop-2.7.2$
```

Valsort MapReduce for 1TB dataset by 1 Node

```
ubuntu@ip-172-31-61-62:~$ File Edit View Search Terminal Help
  Reduce output records=10000000000
  Spilled Records=49340996964
  Shuffled Maps =7451
  Failed Shuffles=0
  Merged Map outputs=7451
  GC time elapsed (ms)=5248742
  CPU time spent (ms)=171576270
  Physical memory (bytes) snapshot=6967979626496
  Virtual memory (bytes) snapshot=29467218808832
  Total committed heap usage (bytes)=10101884190720
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=1000030515200
File Output Format Counters
  Bytes Written=10000000000000
Total Time= 8.4242293E7 milliseconds
Total Time= 84242.293 seconds
ubuntu@ip-172-31-61-62:~/hadoop-2.7.2$ bin/hadoop dfs -get /abhijeet/output/part-r-00000 /data/
DEPRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.
ubuntu@ip-172-31-61-62:~/hadoop-2.7.2$ cd ..
ubuntu@ip-172-31-61-62:~$ ls
abhijeetkedar1.pem  gensorc-linux-1.5  gensorc-linux-1.5.tar.gz  hadoop-2.7.2
ubuntu@ip-172-31-61-62:~$ cd gensorc-linux-1.5/64/
gensorc valsor
ubuntu@ip-172-31-61-62:~/gensorc-linux-1.5/64$ cp valsor /data/
ubuntu@ip-172-31-61-62:~/gensorc-linux-1.5/64$ cd /data
ubuntu@ip-172-31-61-62:/data$ ls
dfs lost+found nn-local-dir part-r-00000  valsor
ubuntu@ip-172-31-61-62:/data$ ./valsor part-r-00000
Records: 10000000000
Checksum: 12a065cb70fe927fb
Duplicate keys: 0
SUCCESS - all records are in order
ubuntu@ip-172-31-61-62:/data$ Broadcast message from root@ip-172-31-61-62
(unknown) at 11:43 ...
The system is going down for power off NOW!
Connection to 54.173.56.162 closed by remote host.
Connection to 54.173.56.162 closed.
abhijeet@abhijeet-Inspiron-5547:~$
```

Head tail MapReduce for 1TB dataset by 1 Node

Resource Node for MapReduce 1 TB Dataset by 1 Node

RUNNING Applications - Mozilla Firefox

Billing Management C... | Inbox (261) - abhi... | FINISHED Applications | Namenode information | RUNNING Applications | Namenode information | The Winning Way: Lea... | +

ec2-54-173-56-162.compute-1.amazonaws.com:9006/cluster/apps/RUNNING

Search | Logged in as: dr who

RUNNING Applications

hadoop

Cluster Metrics

	Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
1	0	1	0	8	30 GB	30.23 GB	0 B	8	8	0	1	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:2048, vCores:1>	<memory:8192, vCores:8>

Show 20 entries

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
application_1458454242696_0001	ubuntu	TeraSort_Abhi...Kedari	MAPREDUCE	default	Sun Mar 20 03:40:35 -0500 2016	N/A	RUNNING	UNDEFINED		ApplicationMaster	0

Showing 1 to 1 of 1 entries

First Previous 1 Next Last

100% completed Resource Node

The screenshot shows the Hadoop FINISHED Applications interface. On the left, there's a sidebar with icons for Cluster, Applications (including NEW, SUBMITTED, ACCEPTED, RUNNING, PENDING, FAILED, KILLED), Scheduler, and Tools. The main area has tabs for FINISHED Applications, Namenode information, and WhatsApp Web. The FINISHED Applications tab is active, displaying a table of cluster metrics and a detailed view of a single application.

	Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers	Memory Used	Memory Total	Memory Reserved	Vcores Used	Vcores Total	Vcores Reserved	Active Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes
1	0	0	1	0	0 B	30.23 GB	0 B	0	8	0	1	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation
Capacity Scheduler	[MEMORY]	<memory:2048, vCores:1>	<memory:8192, vCores:8>

Applications

ID	User	Name	Application Type	Queue	StartTime	FinishTime	State	FinalStatus	Progress	Tracking UI	Blacklisted Nodes
application_1458454242696_0001	ubuntu	TeraSort_AbhijeetKedari	MARREDUCE	default	Sun Mar 20 03:40:35 -0500 2016	Mon Mar 21 03:04:32 -0500 2016	FINISHED	SUCCEEDED	History	N/A	

Showing 1 to 1 of 1 entries

Name Node for MapReduce 1 TB Dataset by 1 Node

The screenshot shows the Namenode Information page. It displays various metrics and configurations for the NameNode. At the top, it says "Security is off." and "Safemode is off." Below that, it lists file and directory counts, memory usage, and configured capacity (5.41 TB). A large table provides detailed information about DataNodes usage, live nodes, dead nodes, decommissioning nodes, and total datanode volume failures. The "Block Pool Used" section shows 1.83 TB (33.86%) used from a 5.41 TB pool. The "DataNodes usages%" section shows 33.86% usage across multiple nodes. The "NameNode Journal Status" section shows the current transaction ID (44879) and the state of the journal manager (EditLogFileOutputStream). The "NameNode Storage" section is partially visible at the bottom.

Configured Capacity:	5.41 TB
DFS Used:	1.83 TB (33.86%)
Non DFS Used:	279.53 GB
DFS Remaining:	3.31 TB (61.1%)
Block Pool Used:	1.83 TB (33.86%)
DataNodes usages% (Min/Median/Max/stdDev):	33.86% / 33.86% / 33.86% / 0.00%
Live Nodes	1 (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	3/20/2016, 1:10:09 AM

NameNode Journal Status

Current transaction ID: 44879

Journal Manager	State
FileJournalManager(root=/data/dfs/name)	EditLogFileOutputStream(/data/dfs/name/current_edits_inprogress_00000000000044879)

NameNode Storage

Spark

A. 1 Node for 10 GB

Total time taken

```
abhijeet@abhijeet-OptiPlex-5547:~/Downloads/$5533/Assignment2spark/spark-1.6-bin-hadoop2.6/ec2
16/03/03 08:23:13 INFO scheduler.TaskSetManager: Finished task 52.0 in stage 2.0 (TID 202) in 10097 ms on ip-172-31-23-225.ec2.internal (52/75)
16/03/03 08:23:14 INFO scheduler.TaskSetManager: Starting task 54.0 in stage 2.0 (TID 204, ip-172-31-23-225.ec2.internal, partition 54, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:14 INFO scheduler.TaskSetManager: Finished task 51.0 in stage 2.0 (TID 201) in 15084 ms on ip-172-31-23-225.ec2.internal (53/75)
16/03/03 08:23:24 INFO scheduler.TaskSetManager: Starting task 55.0 in stage 2.0 (TID 205, ip-172-31-23-225.ec2.internal, partition 55, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:24 INFO scheduler.TaskSetManager: Finished task 53.0 in stage 2.0 (TID 203) in 11559 ms on ip-172-31-23-225.ec2.internal (54/75)
16/03/03 08:23:25 INFO scheduler.TaskSetManager: Starting task 56.0 in stage 2.0 (TID 206, ip-172-31-23-225.ec2.internal, partition 56, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:25 INFO scheduler.TaskSetManager: Finished task 54.0 in stage 2.0 (TID 204) in 10599 ms on ip-172-31-23-225.ec2.internal (55/75)
16/03/03 08:23:37 INFO scheduler.TaskSetManager: Starting task 57.0 in stage 2.0 (TID 207, ip-172-31-23-225.ec2.internal, partition 57, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:37 INFO scheduler.TaskSetManager: Finished task 56.0 in stage 2.0 (TID 206) in 11698 ms on ip-172-31-23-225.ec2.internal (56/75)
16/03/03 08:23:37 INFO scheduler.TaskSetManager: Starting task 58.0 in stage 2.0 (TID 208, ip-172-31-23-225.ec2.internal, partition 58, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:37 INFO scheduler.TaskSetManager: Finished task 55.0 in stage 2.0 (TID 205) in 11559 ms on ip-172-31-23-225.ec2.internal (57/75)
16/03/03 08:23:47 INFO scheduler.TaskSetManager: Starting task 59.0 in stage 2.0 (TID 209, ip-172-31-23-225.ec2.internal, partition 59, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:47 INFO scheduler.TaskSetManager: Finished task 57.0 in stage 2.0 (TID 207) in 10809 ms on ip-172-31-23-225.ec2.internal (58/75)
16/03/03 08:23:48 INFO scheduler.TaskSetManager: Starting task 60.0 in stage 2.0 (TID 210, ip-172-31-23-225.ec2.internal, partition 60, NODE_LOCAL, 1894 bytes)
16/03/03 08:23:48 INFO scheduler.TaskSetManager: Finished task 58.0 in stage 2.0 (TID 208) in 11188 ms on ip-172-31-23-225.ec2.internal (59/75)
16/03/03 08:24:01 INFO scheduler.TaskSetManager: Starting task 61.0 in stage 2.0 (TID 211, ip-172-31-23-225.ec2.internal, partition 61, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:01 INFO scheduler.TaskSetManager: Finished task 59.0 in stage 2.0 (TID 209) in 13204 ms on ip-172-31-23-225.ec2.internal (60/75)
16/03/03 08:24:05 INFO scheduler.TaskSetManager: Starting task 62.0 in stage 2.0 (TID 212, ip-172-31-23-225.ec2.internal, partition 62, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:05 INFO scheduler.TaskSetManager: Finished task 60.0 in stage 2.0 (TID 210) in 16866 ms on ip-172-31-23-225.ec2.internal (61/75)
16/03/03 08:24:11 INFO scheduler.TaskSetManager: Starting task 63.0 in stage 2.0 (TID 213, ip-172-31-23-225.ec2.internal, partition 63, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:11 INFO scheduler.TaskSetManager: Finished task 61.0 in stage 2.0 (TID 211) in 13663 ms on ip-172-31-23-225.ec2.internal (62/75)
16/03/03 08:24:15 INFO scheduler.TaskSetManager: Starting task 64.0 in stage 2.0 (TID 214, ip-172-31-23-225.ec2.internal, partition 64, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:15 INFO scheduler.TaskSetManager: Finished task 62.0 in stage 2.0 (TID 212) in 9860 ms on ip-172-31-23-225.ec2.internal (63/75)
16/03/03 08:24:28 INFO scheduler.TaskSetManager: Starting task 65.0 in stage 2.0 (TID 215, ip-172-31-23-225.ec2.internal, partition 65, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:28 INFO scheduler.TaskSetManager: Finished task 64.0 in stage 2.0 (TID 214) in 12753 ms on ip-172-31-23-225.ec2.internal (64/75)
16/03/03 08:24:28 INFO scheduler.TaskSetManager: Starting task 66.0 in stage 2.0 (TID 216, ip-172-31-23-225.ec2.internal, partition 66, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:28 INFO scheduler.TaskSetManager: Finished task 63.0 in stage 2.0 (TID 213) in 15566 ms on ip-172-31-23-225.ec2.internal (65/75)
16/03/03 08:24:39 INFO scheduler.TaskSetManager: Starting task 67.0 in stage 2.0 (TID 217, ip-172-31-23-225.ec2.internal, partition 67, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:39 INFO scheduler.TaskSetManager: Finished task 65.0 in stage 2.0 (TID 215) in 11179 ms on ip-172-31-23-225.ec2.internal (66/75)
16/03/03 08:24:42 INFO scheduler.TaskSetManager: Starting task 66.0 in stage 2.0 (TID 216) in 12520 ms on ip-172-31-23-225.ec2.internal (67/75)
16/03/03 08:24:42 INFO scheduler.TaskSetManager: Starting task 68.0 in stage 2.0 (TID 218, ip-172-31-23-225.ec2.internal, partition 68, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:51 INFO scheduler.TaskSetManager: Starting task 69.0 in stage 2.0 (TID 219, ip-172-31-23-225.ec2.internal, partition 69, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:51 INFO scheduler.TaskSetManager: Finished task 67.0 in stage 2.0 (TID 217) in 10077 ms on ip-172-31-23-225.ec2.internal (68/75)
16/03/03 08:24:55 INFO scheduler.TaskSetManager: Finished task 68.0 in stage 2.0 (TID 220, ip-172-31-23-225.ec2.internal, partition 70, NODE_LOCAL, 1894 bytes)
16/03/03 08:24:55 INFO scheduler.TaskSetManager: Starting task 71.0 in stage 2.0 (TID 221, ip-172-31-23-225.ec2.internal, partition 71, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:06 INFO scheduler.TaskSetManager: Finished task 69.0 in stage 2.0 (TID 220, ip-172-31-23-225.ec2.internal, partition 70, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:06 INFO scheduler.TaskSetManager: Starting task 72.0 in stage 2.0 (TID 222, ip-172-31-23-225.ec2.internal, partition 72, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:11 INFO scheduler.TaskSetManager: Finished task 70.0 in stage 2.0 (TID 220) in 15476 ms on ip-172-31-23-225.ec2.internal (71/75)
16/03/03 08:25:22 INFO scheduler.TaskSetManager: Starting task 73.0 in stage 2.0 (TID 223, ip-172-31-23-225.ec2.internal, partition 73, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:22 INFO scheduler.TaskSetManager: Finished task 71.0 in stage 2.0 (TID 224, ip-172-31-23-225.ec2.internal, partition 74, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:23 INFO scheduler.TaskSetManager: Starting task 74.0 in stage 2.0 (TID 224, ip-172-31-23-225.ec2.internal, partition 75, NODE_LOCAL, 1894 bytes)
16/03/03 08:25:23 INFO scheduler.TaskSetManager: Finished task 72.0 in stage 2.0 (TID 222) in 10077 ms on ip-172-31-23-225.ec2.internal (73/75)
16/03/03 08:25:33 INFO scheduler.TaskSetManager: Finished task 74.0 in stage 2.0 (TID 224) in 9826 ms on ip-172-31-23-225.ec2.internal (74/75)
16/03/03 08:25:33 INFO scheduler.TaskSetManager: Finished task 73.0 in stage 2.0 (TID 223) in 11154 ms on ip-172-31-23-225.ec2.internal (75/75)
16/03/03 08:25:33 INFO scheduler.DAGScheduler: ResultStage 2 (saveAsTextFile at <console>:30) removed in 467.228 s
16/03/03 08:25:33 INFO scheduler.DAGScheduler: Job 1 finished: saveAsTextFile at <console>:30, took 709.013839 s
end_time: Long = 1459326333927
Spark time for sorting : 770164ms
```

Valsort result

Name Node

NameNode 'ip-172-31-19-106.ec2.internal:9000'

Started: Wed Mar 30 07:53:37 UTC 2016
Version: 1.0.4, r1393290
Compiled: Wed Oct 3 05:13:58 UTC 2012 by hortonfo
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#) [Namenode Logs](#)

Cluster Summary

82 files and directories, 185 blocks = 267 total. Heap Size is 101.5 MB / 889 MB (11%)

Configured Capacity	:	30.03 GB
DFS Used	:	18.79 GB
Non DFS Used	:	2.85 GB
DFS Remaining	:	8.39 GB
DFS Used%	:	62.58 %
DFS Remaining%	:	27.93 %
Live Nodes	:	1
Dead Nodes	:	0
Decommissioning Nodes	:	0
Number of Under-Replicated Blocks	:	110

NameNode Storage:

Storage Directory	Type	State
/mnt/ephemeral-hdfs/dfs/name	IMAGE_AND_EDITS	Active

This is [Apache Hadoop](#) release 1.0.4

B. 16 Node for 100 GB

Total time taken

```
abhijeet@abhijeet-Inspiron-5547:~/Downloads/C5533/Assignment2/Spark/spark-1.6.0-bin-hadoop2.6/c2
16/03/30 06:11:42 INFO scheduler.TaskSetManager: Finished task 728.0 in stage 2.0 (TID 2218) in 10366 ms on ip-172-31-39-137.ec2.internal (707/745)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Starting task 739.0 in stage 2.0 (TID 2229, ip-172-31-40-51.ec2.internal, partition 739.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Finished task 700.0 in stage 2.0 (TID 2230) in 32684 ms on ip-172-31-40-51.ec2.internal (708/745)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Starting task 711.0 in stage 2.0 (TID 2231, ip-172-31-39-149.ec2.internal, partition 711.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Finished task 714.0 in stage 2.0 (TID 2204) in 20315 ms on ip-172-31-39-254.ec2.internal (709/745)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Starting task 741.0 in stage 2.0 (TID 2231, ip-172-31-34-177.ec2.internal, partition 741.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Finished task 730.0 in stage 2.0 (TID 2220) in 11028 ms on ip-172-31-34-177.ec2.internal (710/745)
16/03/30 06:11:43 INFO scheduler.TaskSetManager: Starting task 742.0 in stage 2.0 (TID 2232, ip-172-31-36-196.ec2.internal, partition 742.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Finished task 705.0 in stage 2.0 (TID 2195) in 31882 ms on ip-172-31-36-196.ec2.internal (711/745)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Starting task 743.0 in stage 2.0 (TID 2233, ip-172-31-33-131.ec2.internal, partition 743.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Finished task 721.0 in stage 2.0 (TID 2211) in 18619 ms on ip-172-31-33-131.ec2.internal (712/745)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Starting task 744.0 in stage 2.0 (TID 2234, ip-172-31-42-219.ec2.internal, partition 744.PROCESS_LOCAL, 1894 bytes)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Finished task 768.0 in stage 2.0 (TID 2194) in 33387 ms on ip-172-31-42-219.ec2.internal (713/745)
16/03/30 06:11:44 INFO scheduler.TaskSetManager: Starting task 715.0 in stage 2.0 (TID 2215) in 23022 ms on ip-172-31-32-149.ec2.internal (714/745)
16/03/30 06:11:45 INFO scheduler.TaskSetManager: Finished task 720.0 in stage 2.0 (TID 2210) in 23961 ms on ip-172-31-32-149.ec2.internal (715/745)
16/03/30 06:11:45 INFO scheduler.TaskSetManager: Finished task 709.0 in stage 2.0 (TID 2199) in 29460 ms on ip-172-31-33-131.ec2.internal (716/745)
16/03/30 06:11:47 INFO scheduler.TaskSetManager: Finished task 712.0 in stage 2.0 (TID 2202) in 27767 ms on ip-172-31-40-54.ec2.internal (717/745)
16/03/30 06:11:47 INFO scheduler.TaskSetManager: Finished task 722.0 in stage 2.0 (TID 2212) in 20486 ms on ip-172-31-39-254.ec2.internal (718/745)
16/03/30 06:11:47 INFO scheduler.TaskSetManager: Finished task 707.0 in stage 2.0 (TID 2197) in 33823 ms on ip-172-31-37-139.ec2.internal (719/745)
16/03/30 06:11:47 INFO scheduler.TaskSetManager: Finished task 798.0 in stage 2.0 (TID 2198) in 32229 ms on ip-172-31-45-119.ec2.internal (720/745)
16/03/30 06:11:48 INFO scheduler.TaskSetManager: Finished task 723.0 in stage 2.0 (TID 2213) in 21631 ms on ip-172-31-32-4.ec2.internal (721/745)
16/03/30 06:11:49 INFO scheduler.TaskSetManager: Finished task 724.0 in stage 2.0 (TID 2214) in 21097 ms on ip-172-31-32-4.ec2.internal (722/745)
16/03/30 06:11:49 INFO scheduler.TaskSetManager: Finished task 725.0 in stage 2.0 (TID 2215) in 20408 ms on ip-172-31-37-110.ec2.internal (723/745)
16/03/30 06:11:50 INFO scheduler.TaskSetManager: Finished task 715.0 in stage 2.0 (TID 2087) in 26049 ms on ip-172-31-45-119.ec2.internal (724/745)
16/03/30 06:11:50 INFO scheduler.TaskSetManager: Starting task 726.0 in stage 2.0 (TID 2216) in 19846 ms on ip-172-31-32-149.ec2.internal (725/745)
16/03/30 06:11:52 INFO scheduler.TaskSetManager: Finished task 725.0 in stage 2.0 (TID 2217) in 23938 ms on ip-172-31-37-139.ec2.internal (726/745)
16/03/30 06:11:52 INFO scheduler.TaskSetManager: Finished task 703.0 in stage 2.0 (TID 2193) in 40978 ms on ip-172-31-34-177.ec2.internal (727/745)
16/03/30 06:11:56 INFO scheduler.TaskSetManager: Finished task 726.0 in stage 2.0 (TID 2216) in 26563 ms on ip-172-31-33-248.ec2.internal (728/745)
16/03/30 06:12:01 INFO scheduler.TaskSetManager: Finished task 732.0 in stage 2.0 (TID 2222) in 23276 ms on ip-172-31-32-141.ec2.internal (729/745)
16/03/30 06:12:02 INFO scheduler.TaskSetManager: Finished task 719.0 in stage 2.0 (TID 2209) in 37093 ms on ip-172-31-32-141.ec2.internal (730/745)
16/03/30 06:12:02 INFO scheduler.TaskSetManager: Finished task 731.0 in stage 2.0 (TID 2221) in 31511 ms on ip-172-31-43-172.ec2.internal (731/745)
16/03/30 06:12:04 INFO scheduler.TaskSetManager: Finished task 718.0 in stage 2.0 (TID 2208) in 41986 ms on ip-172-31-37-110.ec2.internal (732/745)
16/03/30 06:12:04 INFO scheduler.TaskSetManager: Finished task 716.0 in stage 2.0 (TID 2206) in 43310 ms on ip-172-31-40-54.ec2.internal (733/745)
16/03/30 06:12:11 INFO scheduler.TaskSetManager: Finished task 736.0 in stage 2.0 (TID 2226) in 30154 ms on ip-172-31-39-194.ec2.internal (734/745)
16/03/30 06:12:12 INFO scheduler.TaskSetManager: Finished task 737.0 in stage 2.0 (TID 2227) in 30078 ms on ip-172-31-33-248.ec2.internal (735/745)
16/03/30 06:12:12 INFO scheduler.TaskSetManager: Starting task 729.0 in stage 2.0 (TID 2218) in 28056 ms on ip-172-31-32-149.ec2.internal (736/745)
16/03/30 06:12:13 INFO scheduler.TaskSetManager: Finished task 721.0 in stage 2.0 (TID 2231) in 26692 ms on ip-172-31-34-177.ec2.internal (737/745)
16/03/30 06:12:14 INFO scheduler.TaskSetManager: Finished task 748.0 in stage 2.0 (TID 2230) in 30413 ms on ip-172-31-39-254.ec2.internal (738/745)
16/03/30 06:12:14 INFO scheduler.TaskSetManager: Finished task 735.0 in stage 2.0 (TID 2225) in 33707 ms on ip-172-31-40-51.ec2.internal (739/745)
16/03/30 06:12:14 INFO scheduler.TaskSetManager: Finished task 738.0 in stage 2.0 (TID 2228) in 32146 ms on ip-172-31-39-137.ec2.internal (740/745)
16/03/30 06:12:15 INFO scheduler.TaskSetManager: Finished task 734.0 in stage 2.0 (TID 2224) in 35038 ms on ip-172-31-42-219.ec2.internal (741/745)
16/03/30 06:12:17 INFO scheduler.TaskSetManager: Finished task 742.0 in stage 2.0 (TID 2232) in 33262 ms on ip-172-31-36-196.ec2.internal (742/745)
16/03/30 06:12:25 INFO scheduler.TaskSetManager: Finished task 744.0 in stage 2.0 (TID 2234) in 40666 ms on ip-172-31-42-219.ec2.internal (743/745)
16/03/30 06:12:29 INFO scheduler.TaskSetManager: Finished task 733.0 in stage 2.0 (TID 2223) in 50443 ms on ip-172-31-36-196.ec2.internal (744/745)
16/03/30 06:12:29 INFO scheduler.DAGScheduler: ResultStage 2 (saveAsTextFile at <console>:30) finished in 564.622 s
16/03/30 06:12:29 INFO scheduler.TaskSchedulerImpl: Removed TaskSet 2.0, whose tasks have all completed, from pool
16/03/30 06:12:29 INFO scheduler.DAGScheduler: Job 1 finished: saveAsTextFile at <console>:30, took 824.809698 s
end_time: Long = 1459318349472
Spark time for sorting : 874567ms
scala>
```

Valsort result and Head tail entries

Name Node

Hadoop NameNode ip-172-31-34-57.ec2.internal:9000 - Mozilla Firefox

EC2 Management... WhatsApp Web FINISHED Applications Namenode Information getMerge - abhije... Cloud Computing ... Inbox (1,539) - ak... Rihanna - Work (E... Hadoop NameNod...

ec2-52-91-119-132.compute-1.amazonaws.com:5070/dfshealth.jsp Search

NameNode 'ip-172-31-34-57.ec2.internal:9000'

Started: Wed Mar 30 04:05:46 UTC 2016
Version: 1.0.4, r1393290
Compiled: Wed Oct 3 05:13:58 UTC 2012 by hortonfo
Upgrades: There are no upgrades in progress.

[Browse the filesystem](#)
[Namenode Logs](#)

Cluster Summary

752 files and directories, 1844 blocks = 2596 total. Heap Size is 151 MB / 889 MB (16%)

Configured Capacity	:	480.5 GB
DFS Used	:	375.86 GB
Non DFS Used	:	45.63 GB
DFS Remaining	:	59.01 GB
DFS Used%	:	78.22 %
DFS Remaining%	:	12.28 %
Live Nodes	:	16
Dead Nodes	:	0
Decommissioning Nodes	:	0
Number of Under-Replicated Blocks	:	0

NameNode Storage:

Storage Directory	Type	State
/mnt/ephemeral-hdfs/dfs/name	IMAGE_AND_EDITS	Active

This is [Apache Hadoop](#) release 1.0.4