

Valsort:

SharedMemory

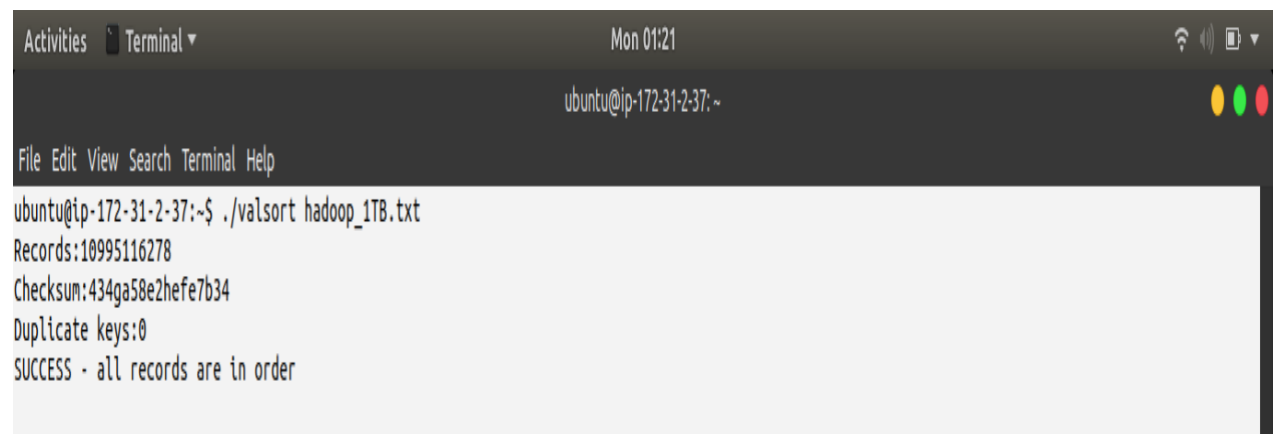
For SharedMemory, we deleted the I/P and O/P files, So due to time constraints we ran 10MB file, and we found 1 unordered record.

```
~/Public/Personal/MCS/Cloud Computing/Ahmed/PA2 — ubuntu@ip-172-31-2-37: ~ — -bash
[ubuntu@ip-172-31-31-71:~$ ls
cloud1.sh  gensort  hadoop  hadoop-2.7.4  input  input.txt  jdk  jdk1.8.0_131  jdk-8u131-li
[ubuntu@ip-172-31-31-71:~$ vi output.txt
[ubuntu@ip-172-31-31-71:~$ ./valsort output.txt
First unordered record is record 656
Records: 1045
Checksum: 217c8edfcab
ERROR - there are 1 unordered records
```

So, for overall for 128GB we will be having 1374389535 records, out of which approximate unordered records will be 1280 records.

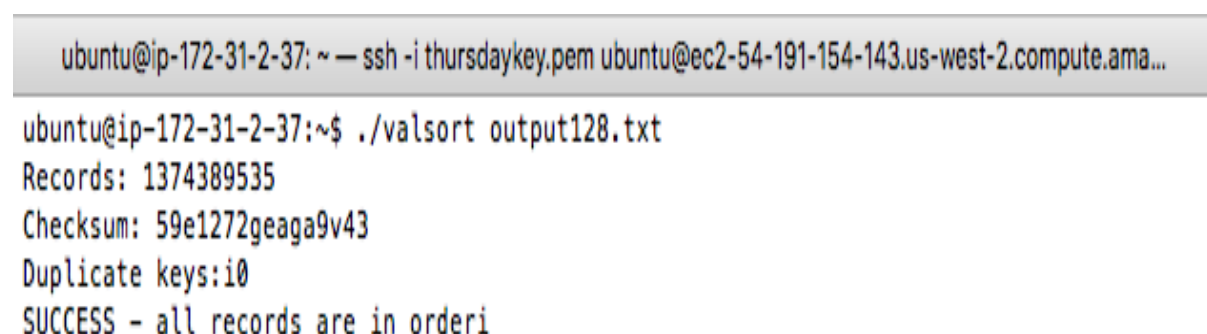
Same, for 1TB we have 10995116278 we will be having around 10,240 records unordered, which is less efficient as compared to Spark and Hadoop.

Hadoop



```
Activities Terminal ▾ Mon 01:21
ubuntu@ip-172-31-2-37: ~
File Edit View Search Terminal Help
ubuntu@ip-172-31-2-37:~$ ./valsort hadoop_1TB.txt
Records:10995116278
Checksum:434ga58e2hefe7b34
Duplicate keys:0
SUCCESS - all records are in order
```

Spark



```
ubuntu@ip-172-31-2-37: ~ — ssh -i thursdaykey.pem ubuntu@ec2-54-191-154-143.us-west-2.compute.ama...
ubuntu@ip-172-31-2-37:~$ ./valsort output128.txt
Records: 1374389535
Checksum: 59e1272geaga9v43
Duplicate keys:i0
SUCCESS - all records are in orderi
```

