File Systems Used:

1. ext4
2. ZFS

Feature Used :

Deduplication

ext4 doesn’t support deduplication whereas ZFS does.

By use od deduplication, ZFS doesn’t store duplicate data. It stores a hash table of chunks, and if the chunk came with the data that it has stored earlier, it only stores the pointer from previous chunk, rather than storing the same data again.

ZFS implement block level dedupliaction and is synchronus.

Implementation details:

1. Mounted a pendrive with ZFS file system , pendrive size = 32 GB

command to mount as zfs:

1. cd /
2. mkdir <dir-name>
3. zpool create -f <dir-name> <mount-point>

2. Created a 2MB file in the system.

3. Copy 1000 copies of same file with different name in ZFS file system

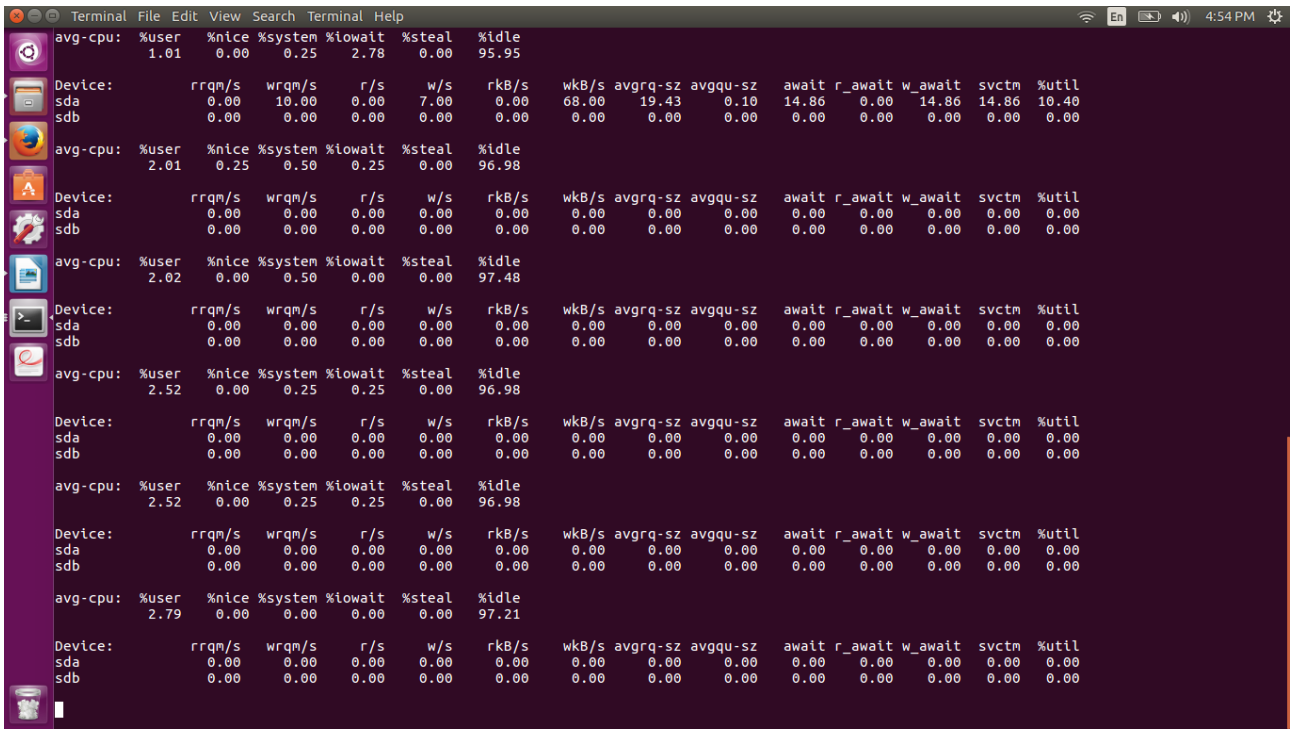
command:

for i in `seq 1 1000`; do cp <orig.txt> <dir-name/foo$i.txt>;done

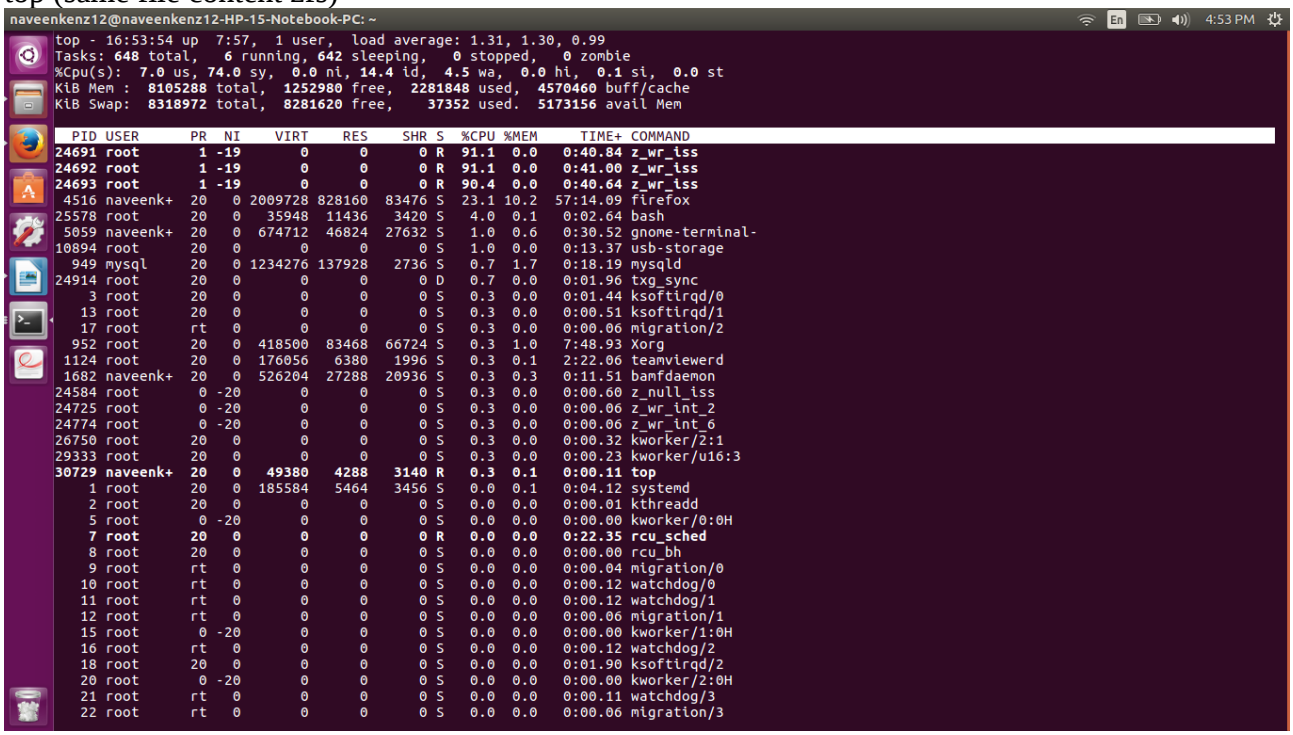
4. captured screenshot of iostat -x 1 10, and top command

Done same for ext4 file system.

Iostate -x 1 10 (for same file content zfs)

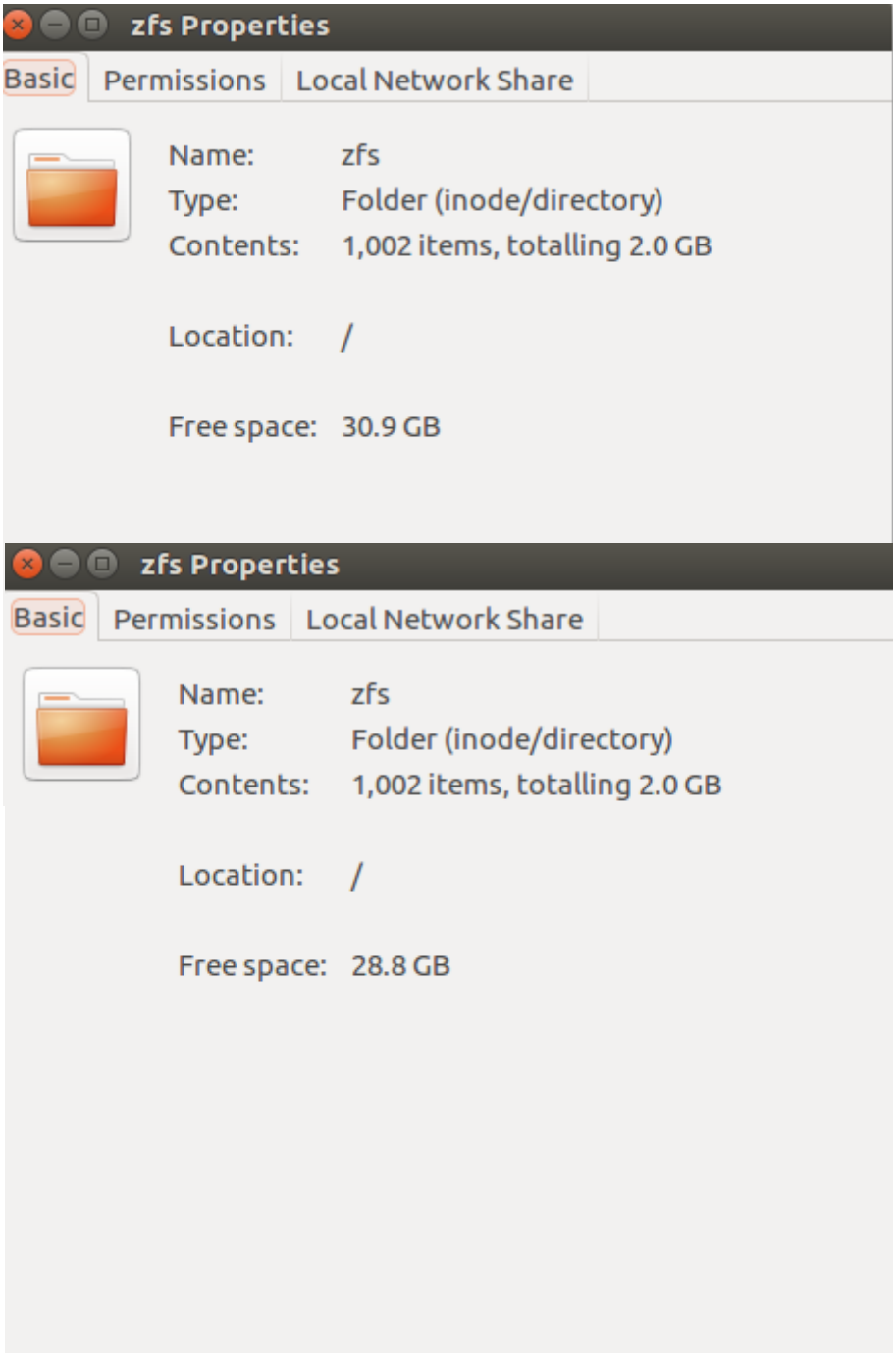


top (same file content zfs)



CPU extensive

size of 1 file = 2MB
total = 1000 files with same content
total = 2GB

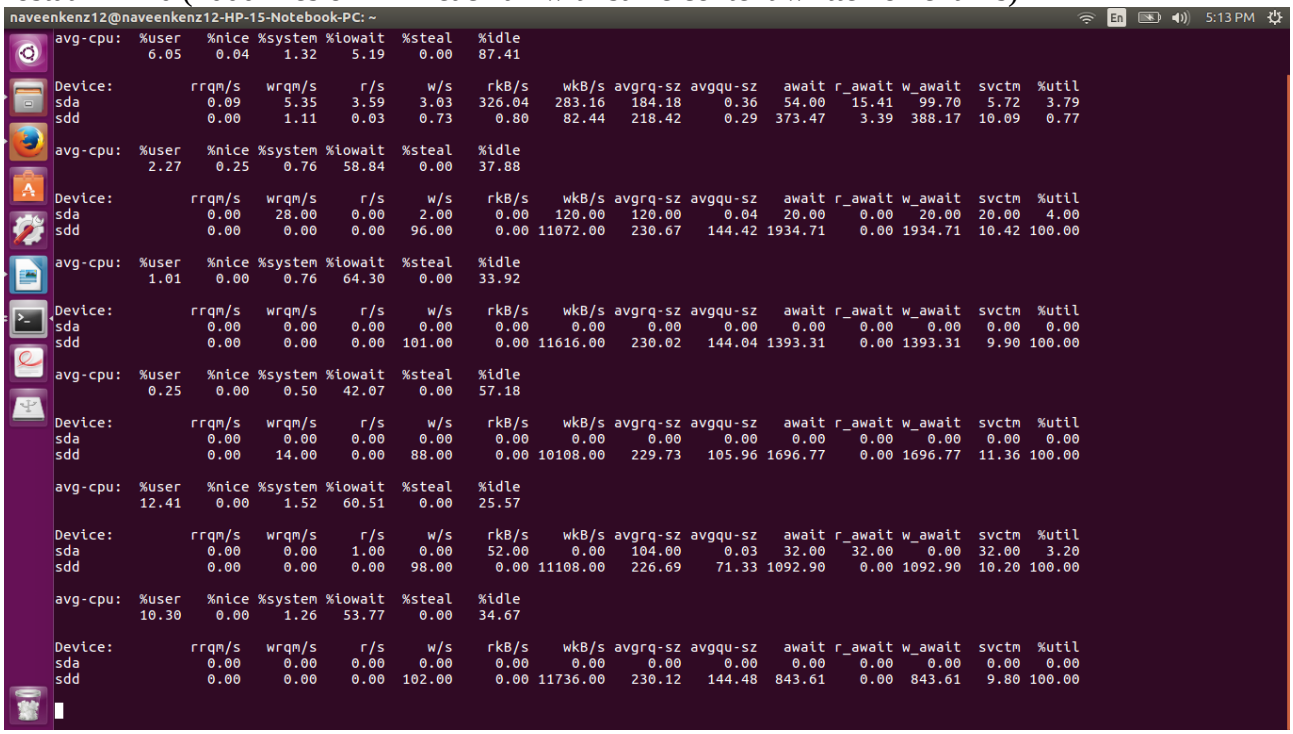


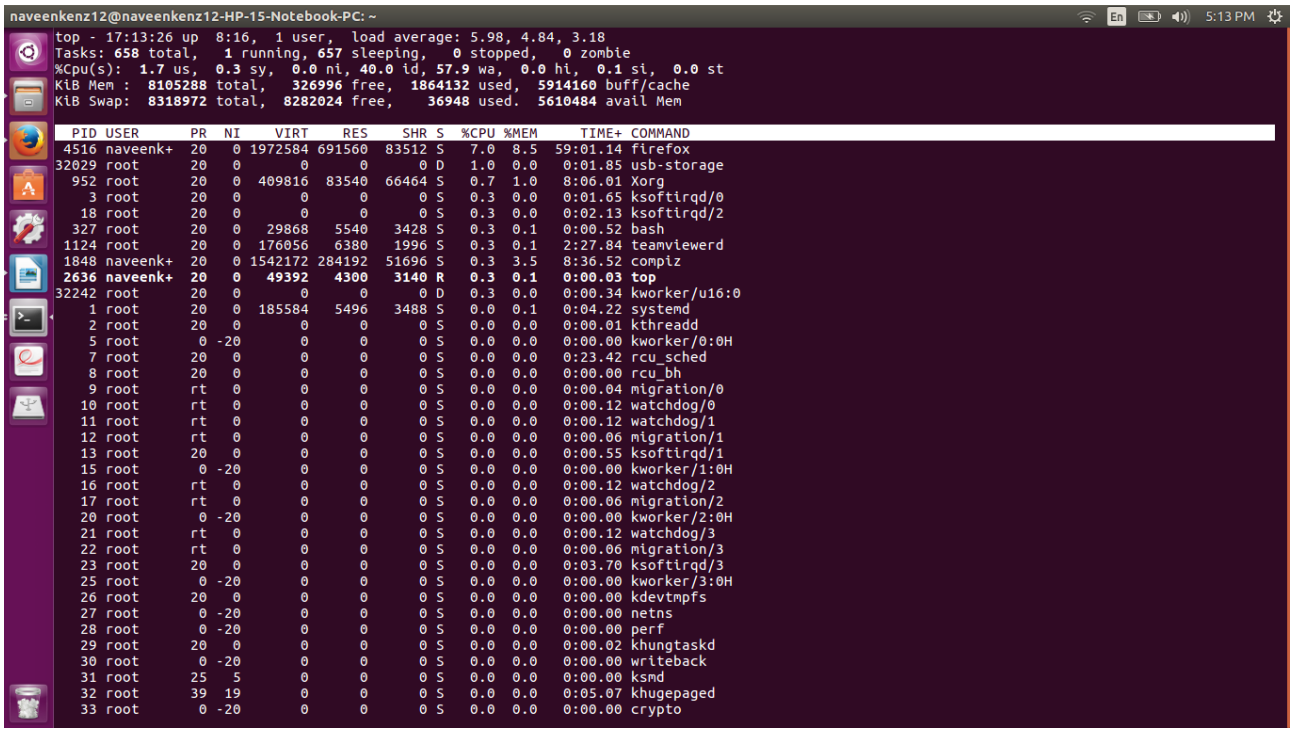
When 2GB(1000 files of 2MB) having same content was written, free size was = 30.9GB
but when 2GB different content was written , free size = 28.8 GB

so, there was deduplication, in case of zfs file system.

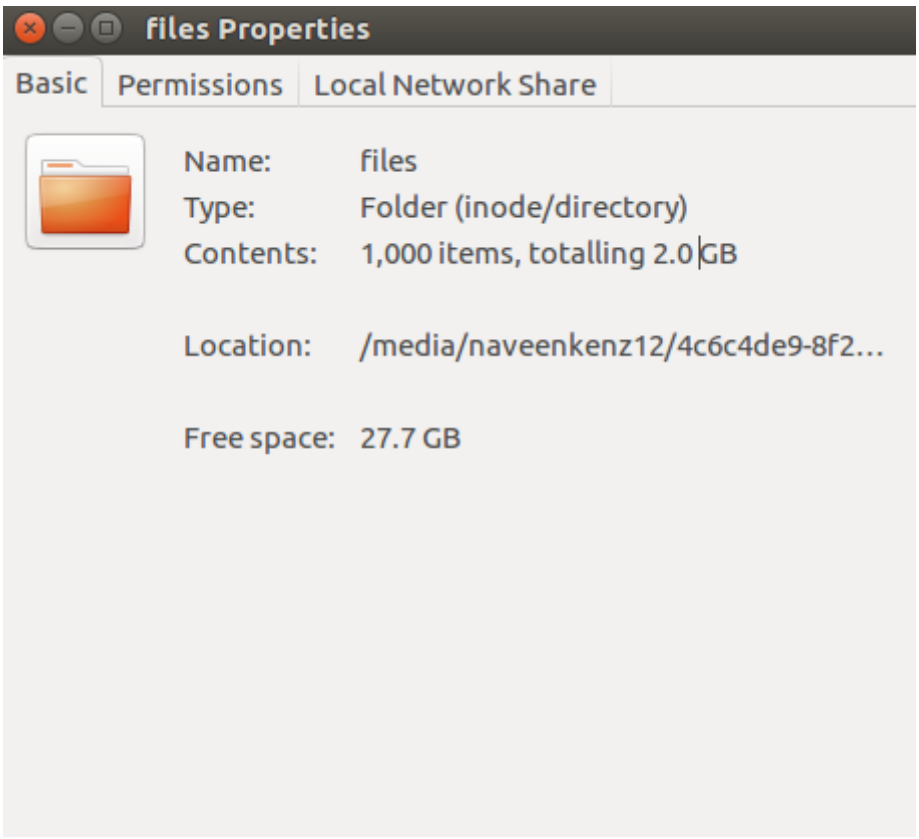
Zpool List:
Alloc = 4.47 MB (after copying 1000 2MB files)
Free = 29.7 GB(same before and after copying) (as only 4.47MB was subtracted)
Dedup = 1000x

Iostat -x 1 10 (1000 files of 2MB each all with same content written on ext4 fs)





no cpu intensive process (Top command)



No deduplication in case of ext4 file system