Technical note:
How create a SWAP file on a Linux server

Gilles Celli & Nicolas d'Oreye

European Centre for Geodynamics and Seismology (ECGS), 19 rue Josy Welter, L-7256 Walferdange, Luxembourg
National Museum of Natural History (NMNH), 19 rue Josy Welter, L-7256 Walferdange, Luxembourg

RATIONALE:

The present document explains how to create a swap file on a Linux computer e.g. when there is not enough RAM to process a MSBAS inversion.

The procedure was set up by Gilles Celli for the specific case of a server (named “dellrack”) that had only 512 Gb of RAM at ECGS, thought it had 2 separate SSD (Samsung PM981 NVME of 1024GB).

When Kubuntu was installed on the server, a 128GB SWAP partition was allocated on the first SSD (/dev/nvmen0). Resizing this partition could be a risky task when a OS is already installed.

A solution is to add a 512GB swapfile. A drawback is that a SWAP file on a SSD could reduce the SSD'slifespan. A swap file on a mechanical Harddisc would be better, however the RAID partition is XFS formatted, which doesn't allow Swap file creation (hole error!).

So it was decided to create a 512GB swap file on /dev/nvme1 (second SSD)

Note: on Mac computer, swap should be automatic…

TABLE OF CONTENT

CHECK SWAP FIRST

First we check the swap already used: this displays the 127GB swap partition on the first SS /dev/nvme0

```
#sudo swapon --show
```

```
NAME                                TYPE        SIZE USED PRIO
/dev/nvme0n1p4                      partition 127,7G   0B   -2
```

```
#df -h
```

```
Filesystem Size Used Avail Use% Mounted on
tmpfs 51G 4,4M 51G 1% /run
/dev/nvme0n1p3 808G 164G 603G 22% /
tmpfs 251G 0 251G 0% /dev/shm
tmpfs 5,0M 4,0K 5,0M 1% /run/lock
/dev/nvme0n1p1 796M 23M 774M 3% /boot/efi
dellrack-data-pool 49T 640K 49T 1% /mnt/dellrack_data
tmpfs 51G 64K 51G 1% /run/user/501
/dev/nvme0n1p2 5,0G 3,2G 1,9G 63% /media/nicolas/OS
/dev/nvme1n1 938G 515G 376G 58% /media/nicolas/Data
//hp-storeeasy.ecgs.welter/hp-D3602-Data_RAID5 81T 71T 9,5T 89% /mnt/3602
//hp-storeeasy.ecgs.welter/hp-D3600-Data_Share1 55T 47T 8,2T 86%
/mnt/3600
//syno-data.ecgs.welter/Syno_Data_1 26T 6,4T 19T 26% /mnt/syno_data_1
//syno-sar.ecgs.welter/DataSAR 44T 35T 9,3T 79% /mnt/syno_sar
//hp-storeeasy.ecgs.welter/hp-D3601-Data_RAID6 73T 63T 11T 87% /mnt/3601
//hp-storeeasy.ecgs.welter/hp-1650-Data_Share1 55T 41T 14T 75% /mnt/1650
//hpe1660.ecgs.welter/hp1660 73T 6,8T 67T 10% /mnt/1660
//hpe1660.ecgs.welter/D3610 91T 501G 91T 1% /mnt/3610
//syno-congo.ecgs.welter/DATARDC 11T 4,0T 6,5T 39% /mnt/syno_congo
```

ADD ADDITIONAL SWAP FILE (NOT A PARTITION!) TO THE SECOND SSD ON DELLRACK

```
#sudo mkdir /media/nicolas/Data/SWAP512GB

#sudo fallocate -l 512GB /media/nicolas/Data/SWAP512GB/swapfile

#sudo chmod 600 /media/nicolas/Data/SWAP512GB/swapfile

#sudo mkswap /media/nicolas/Data/SWAP512GB/swapfile
```

```
Setting up swapspace version 1, size = 512 GiB (549755809792 bytes)
no label, UUID=b5f15fd2-38f0-4ba6-8c86-b2cb58a957bc
```

```
#sudo swapon /media/nicolas/Data/SWAP512GB/swapfile
```

After enabling the swapfile it will show up with command `swapon --show`

```
#sudo swapon --show
```

```
NAME                                      TYPE        SIZE USED PRIO
/dev/nvme0n1p4                            partition 127,7G   0B   -2
/media/nicolas/Data/SWAP512GB/swapfile file        512G   0B   -3
```

or check with: `# free -h`

```
total used free shared buff/cache available
Mem: 501Gi 4,4Gi 495Gi 80Mi 1,2Gi 494Gi
Swap: 639Gi 0B 639Gi
```

MOUNT SWAP FILE AT BOOT

Edit /etc/fstab and check if the SSD disc is first mounted
(/dev/nvme1n1 on mount /media/nicolas/Data)

Note: I had some issues mounting the SWAP file because the mount point of /dev/nvme1n1 was not present in `/etc/fstab`. /dev/nvme1n1 was only mounted when user "nicolas" logged in from KDE Plasma.

By adding the line `UUID=6e48816a-67f2-4e07-bfd8-e6836f255deb /media/nicolas/Data ext4 errors=remount-ro 0 1` to `/etc/fstab` the partition was correctly mounted to `/media/nicolas/Data` and after that we could mount the 512GB swap file.

```
# /dev/nvme1n1 (Second SSD of two SSDs) which only contains Data & OS
UUID=6e48816a-67f2-4e07-bfd8-e6836f255deb /media/nicolas/Data    ext4
errors=remount-ro 0       1
```

The add the swap file by editing again `/etc/fstab` by adding the lines:

```
# Swap file of 512GB on second SSDs (dev/nvme1n1)
/media/nicolas/Data/SWAP512GB/swapfile none swap    sw  0   0
```

ENABLE ZSWAP - 2023.06.08

- **ZSWAP:** The frontswap system hooks attempts to swap out pages and uses zswap as write-back-cache for a HDD/SSD swap device: An attempt is made to compress the page and if it contains poorly compressible data it is directly written to the disk. If the data is compressed, it is stored in the pool of zswap memory. If pages are swapped out of memory when the total compressed pages in RAM exceeds a certain size, the Least Recently Used (LRU) ***compressed*** page is written to the disk as it is unlikely to be required soon.

- **Benefits:** Very efficient use RAM and disk based swap. Minimizes Disk I/O by both reducing the number of writes and reads required (data is compressed and held in RAM) and by reducing the bandwidth of these I/O operations as the data is in a compressed form.

- **Limitations:** It is an enhancement of disk based swap systems and hence depends on a swap partition on the hard disk.

- **Status:** Merged into the 3.11 mainline linux kernel.

Edit /etc/default/grub

Add `zswap.enabled=1` to line: GRUB_CMDLINE_LINUX_DEFAULT

```
GRUB_CMDLINE_LINUX_DEFAULT="quiet splashi zswap.enabled=1"
```

Update grub

Finally, update Grub using the **update-grub** command.

```
sudo update-grub
```

Tune Swap space settings (Swappiness, cache pressure)

Firstly, check the value of the “swappiness” parameter, which indicates how often the system is going to swap data out of the RAM to the added swap space. The parameter’s value will lie between “0” and “100”.

By default it's 100:

```
cat /proc/sys/vm/swappiness
```

```
100
```

Another parameter’s value that we are going to tune is “vfs_cache_pressure”. This parameter controls the cache dentry and inode information.

```
#cat /proc/sys/vm/vfs_cache_pressure
```

```
100
```

Change values of swappinness vfs_cache_pressure in /etc/sysctl.conf

Edit /etc/sysctl.conf and add following lines:

```
# Gilles - added vm.max_map_count and swappiness and cache_pressure - 2023.06.01
# See: https://www.digitalocean.com/community/tutorials/how-to-add-swap-space-on-ubuntu-22-04
vm.max_map_count=524288
vm.swappiness=20
vm.vfs_cache_pressure=40
```

Activate with command (or reboot):

```
sudo sysctl -p
```

RESOURCES

https://ubunlog.com/en/mejorar-rendimiento-ubuntu-gracias-zswap/

https://askubuntu.com/questions/471912/zram-vs-zswap-vs-zcache-ultimate-guide-when-to-use-which-one

https://linuxhint.com/add-swap-space-ubuntu-22-04/

https://www.digitalocean.com/community/tutorials/how-to-add-swap-space-on-ubuntu-22-04