Infrastructure

Partition, Filesystem

2019/03/11 @TsungHan

Outline

- Big Picture
- Disk Management
  - Partition
  - RAID, LVM
- Filesystem
  - Filesystem Type
  - Mount
  - Filesystem Table
  - Linux VFS
- Exercise : VM Partitioning

Before We Start ...

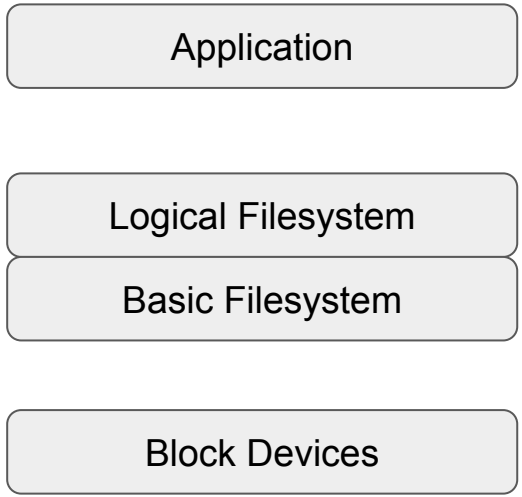
- Download CentOS 7 ISO in advanced

  http://ftp.yzu.edu.tw/Linux/CentOS/7.6.1810/isos/x86_64/CentOS-7-x86_64-Minimal-1810.iso

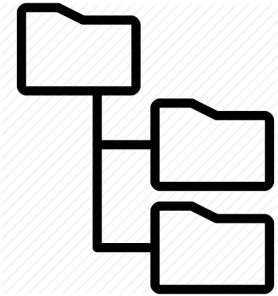
- The In-class Exercise is easy today
  - Takes within 20 minutes
  - You can pay more attention on the lecture =)

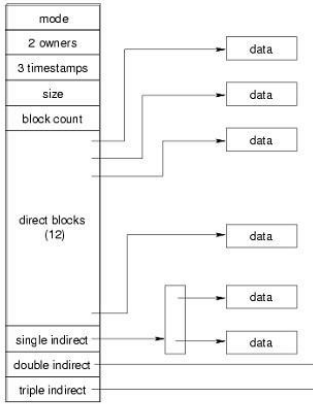
Big Picture

- Application
  - Read / Write on Filesystem
- Filesystem
  - Handle lock, permission, journal, index …
  - Ex : ext 2/3/4, FAT32, NTFS, XFS ...
- Block Devices
  - Disk
  - What really store things



Example





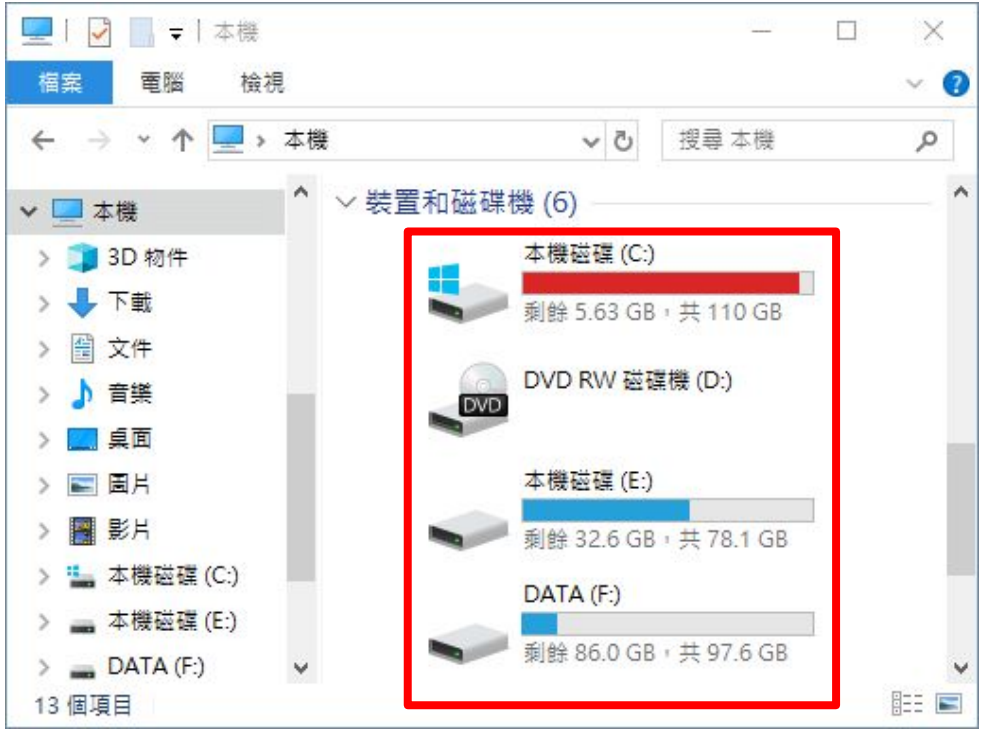
Application
Write a program

Logical Filesystem
Save it at ~/code/a.c

Basic Filesystem
ext4 inode implementation

Block Device
Where the data is really saved

Partition


本機
檔案
電腦
檢視
本機
搜尋 本機
本機
3D 物件
下載
文件
音樂
桌面
圖片
影片
本機磁碟 (C:)
本機磁碟 (E:)
DATA (F:)
13 個項目
裝置和磁碟機 (6)
本機磁碟 (C:)
剩餘 5.63 GB，共 110 GB
DVD RW 磁碟機 (D:)
DVD
本機磁碟 (E:)
剩餘 32.6 GB，共 78.1 GB
DATA (F:)
剩餘 86.0 GB，共 97.6 GB

Why Partitioning

- Separate OS and personal data
- Multiple operating systems
- Make backup more flexible
- Improve performance (smaller partition)
- ...

Linux Partition

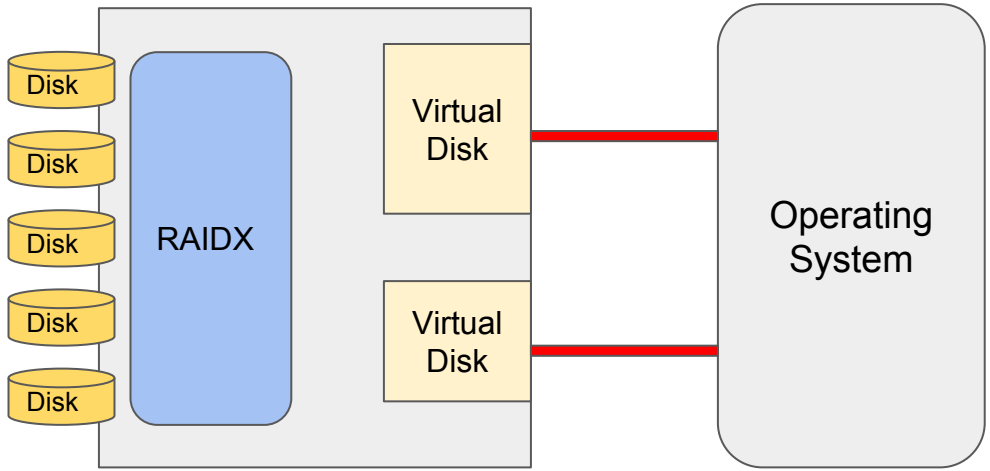
- All devices, executable programs … are regarded as file
  - /dev/sd[a-z][1-128]
  - /dev/mapper/centos-root
  - ...
- fdisk / gdisk / parted
  - Partition using CLI
- lsblk
  - see block devices as well as partitions even not mounted
  - Try it on CSIE workstation now !!!

Linux Partition Cont.

- MBR
  - Support up to 4 primary partition
  - Up to 1 extended partition (logical partition)
  - OS cannot read block devices over 2.2T
- GPT
  - More partition
  - Support 2 level boot

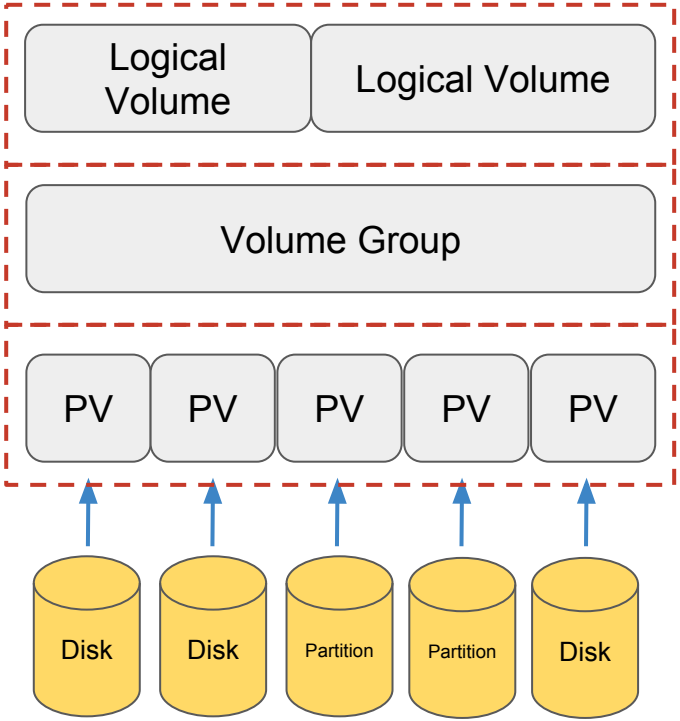
RAID

- Redundant Arrays of Independent Disks
- Why RAID ?
  - Performance
  - Reliability
  - Storage Pooling



LVM

- Logical Volume Manager
- Why LVM
  - Storage Pooling
  - Resizing
  - Snapshots



LVM Cont.

- Useful Commands

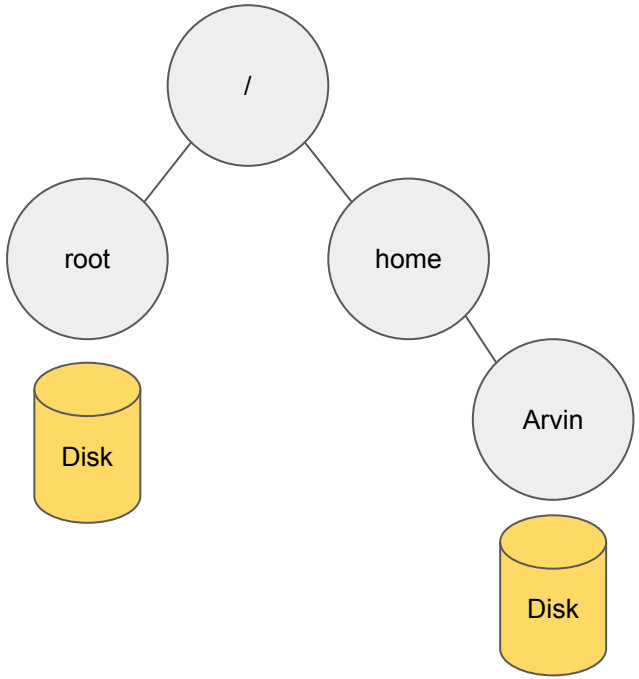
```
# create physical volume
pvcreate <partition>

# create volume group
vgcreate -s <size> <vgname> <pv_name>

# create logical volume
lvcreate -L <size> -n <lv_name> <vg_name>
lvresize -L +/-<size> <lv_path>
```

Filesystem Type

- FAT, NTFS
- EXT 2/3/4 (Linux)
- XFS (CentOS 7 default)
- NFS
- . . .

Mount

- Mounting
  - Makes files and directories on a storage device available for users to access via the computer’s file system
- Mount Point
  - The attached location within the file structure for a file system



Mount Cont.

```
# build a Linux filesystem on a device

mkfs -t <filesystem> <device_path>


# mounting process

mount <device_path> <mount_path>
```

Filesystem Table

```
$ cat /etc/fstab

# <file system>                           <dir> <type>  <options>  <dump> <pass>

/dev/mapper/centos-root                   /       xfs     defaults        0 0
UUID=a89516e2-5356-4c28-b871-24ab85fe6c3b /boot   xfs     defaults        0 0
/dev/mapper/centos-swap                   swap    swap    defaults        0 0

nas-master.sa.csie.ntu.edu.tw:/maillog-smtp1 /nas/log/maillog       nfs local_lock=all
nas-master.sa.csie.ntu.edu.tw:/fail2ban-smtp1 /nas/log/fail2ban.log nfs local_lock=all
```

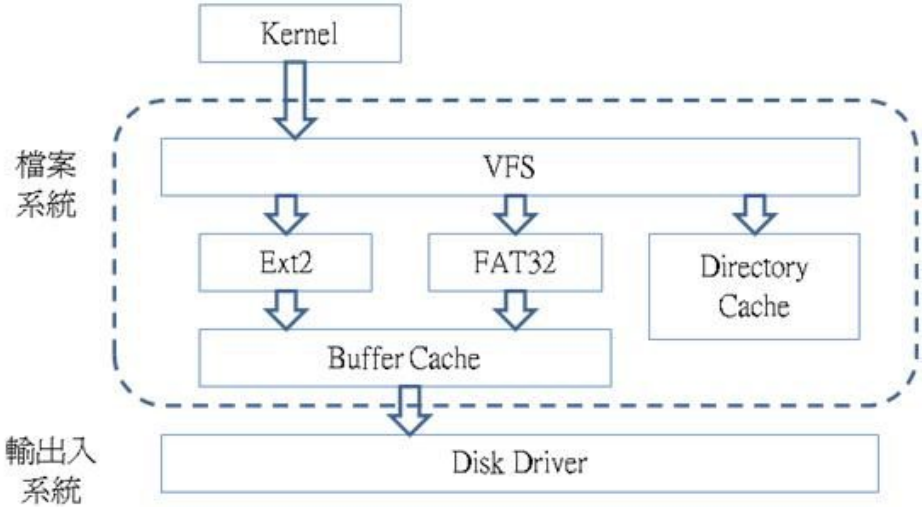
See usage of mounted partitions

```
$ df -h
```

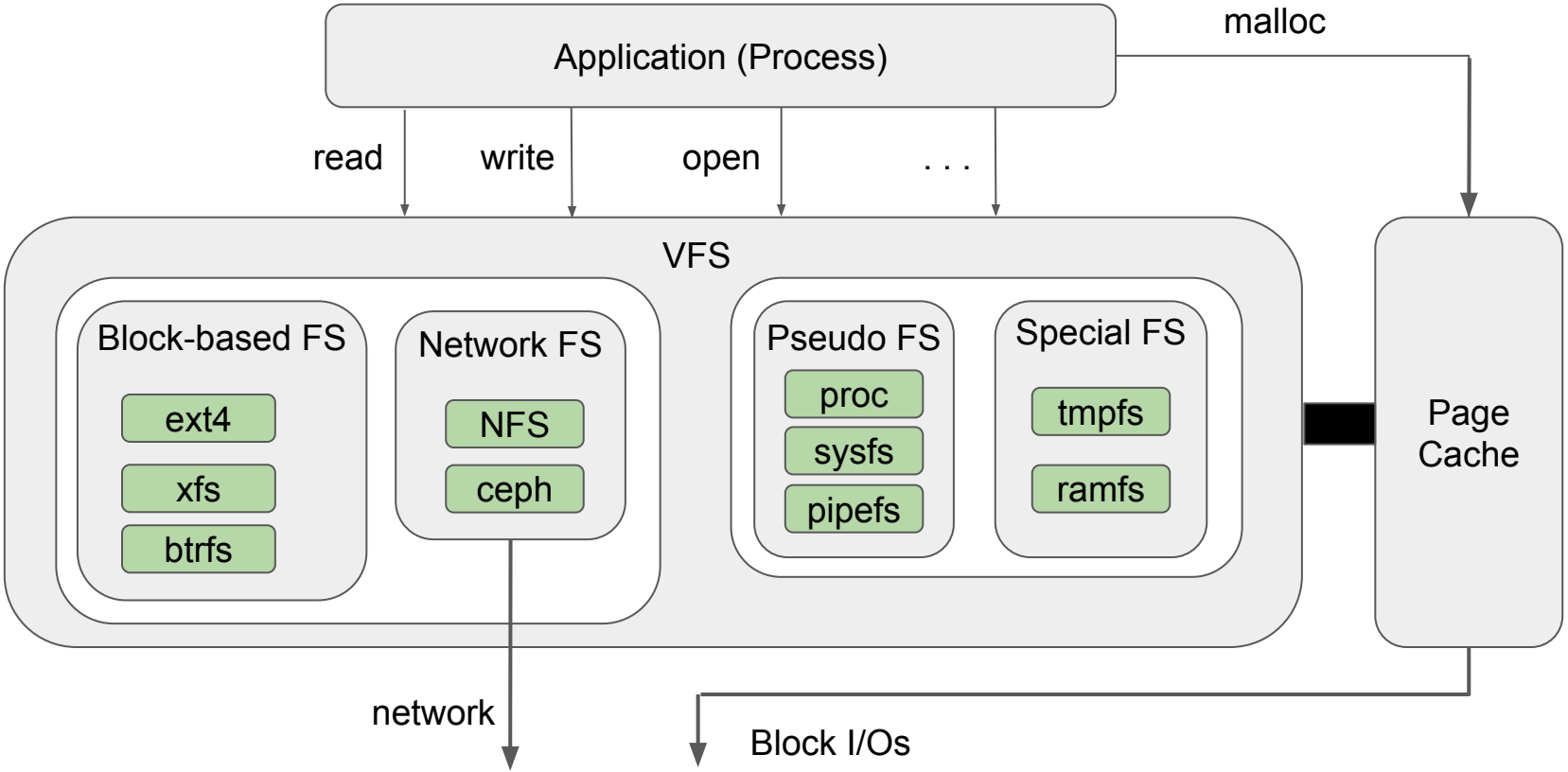
```
檔案系統                                          容量  已用  可用 已用% 掛載點
dev                                               32G     0   32G    0% /dev
run                                               32G  1.4M   32G    1% /run
/dev/mapper/ws-root                               54G   28G   24G   55% /
tmpfs                                             32G     0   32G    0% /dev/shm
tmpfs                                             32G     0   32G    0% /sys/fs/cgroup
tmpfs                                             32G  212K   32G    1% /tmp
/dev/mapper/ws-log                               4.9G   27M  4.6G    1% /var/log/wslab
/dev/mapper/ws-journal                            15G  1.7G   13G   12% /var/log/journal
/dev/mapper/ws-tmp2                              845G   28G  774G    4% /tmp2
home-new.sa.csie.ntu.edu.tw:/e/undergrad         837G  609G  186G   77% /nfs/undergrad
home-new.sa.csie.ntu.edu.tw:/e/ta                443G  189G  233G   45% /nfs/ta
home-new.sa.csie.ntu.edu.tw:/e/alumni_master     1.6T  350G  1.2T   23% /nfs/alumni_master
...
```

Linux VFS

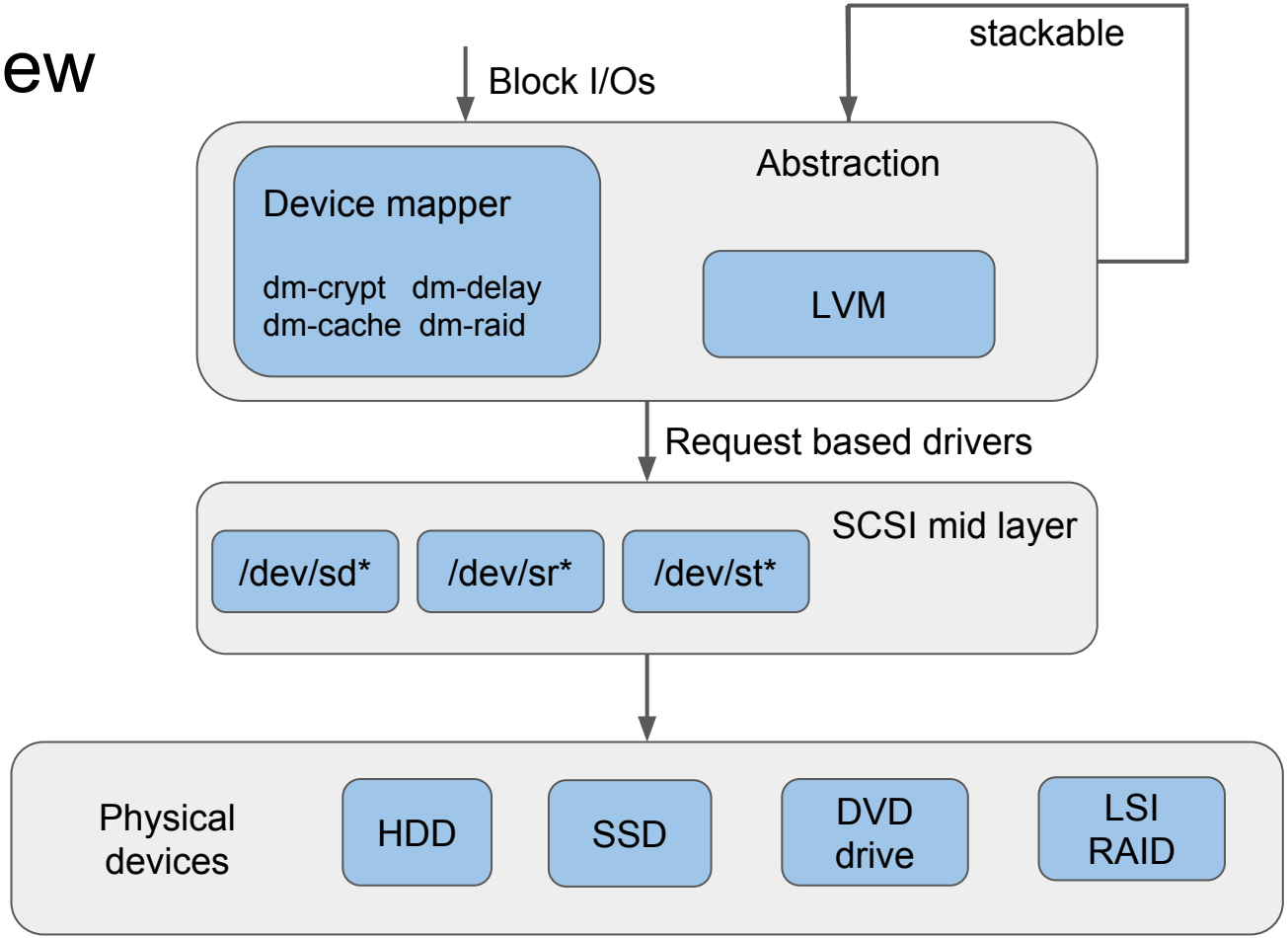
- Abstract layer on top of a more concrete filesystem
- Specify an interface between kernel and concrete filesystem
- Easy to add support for new file system types to the kernel simply by fulfilling the contract



Review


Application (Process)
malloc
read
write
open
. . .
VFS
Block-based FS
ext4
xfs
btrfs
Network FS
NFS
ceph
Pseudo FS
proc
sysfs
pipefs
Special FS
tmpfs
ramfs
Page Cache
network
Block I/Os

Review

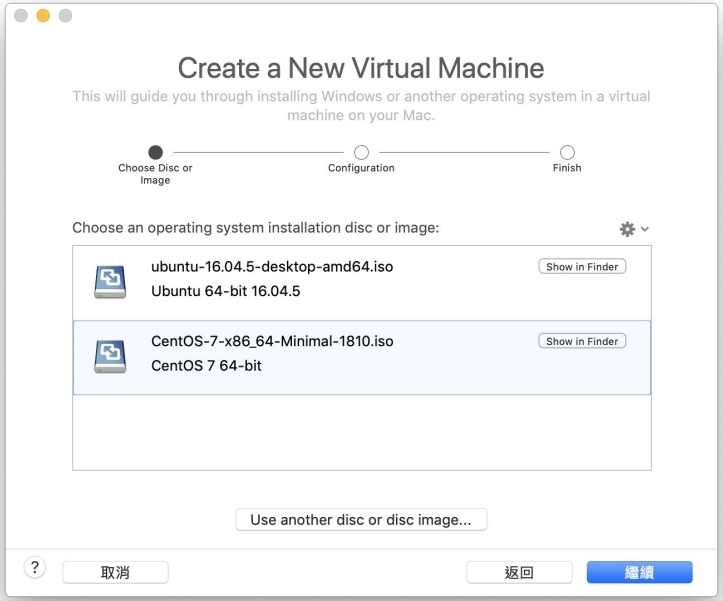

stackable
Block I/Os
Abstraction
Device mapper
dm-crypt dm-delay
dm-cache dm-raid
LVM
Request based drivers
SCSI mid layer
/dev/sd*
/dev/sr*
/dev/st*
Physical devices
HDD
SSD
DVD drive
LSI RAID

Exercise - VM Partitioning

1. Install CentOS 7 VM on VirtualBox or VMware using ISO image
2. Partition the VM using graphical interface (You can use CLI if you want XD)
    - sda : A block with 20G initially
        - sda1 : `boot partition, 1G, fs : xfs, mount point : /boot`
        - sda2 : LVM (Volume Group Name : your student ID)
            - `<student ID>-root,   5G, fs : ext4, mount point : /`
            - `<student ID>-swap,   2G, fs : swap, mount point : [SWAP]`
            - `<student ID>-home,  10G, fs : ext4, mount point : /home`
            - `<student ID>-backup, 2G, fs : xfs,  mount point : /backup`
3. Set the networking such that you can access your VM via ssh

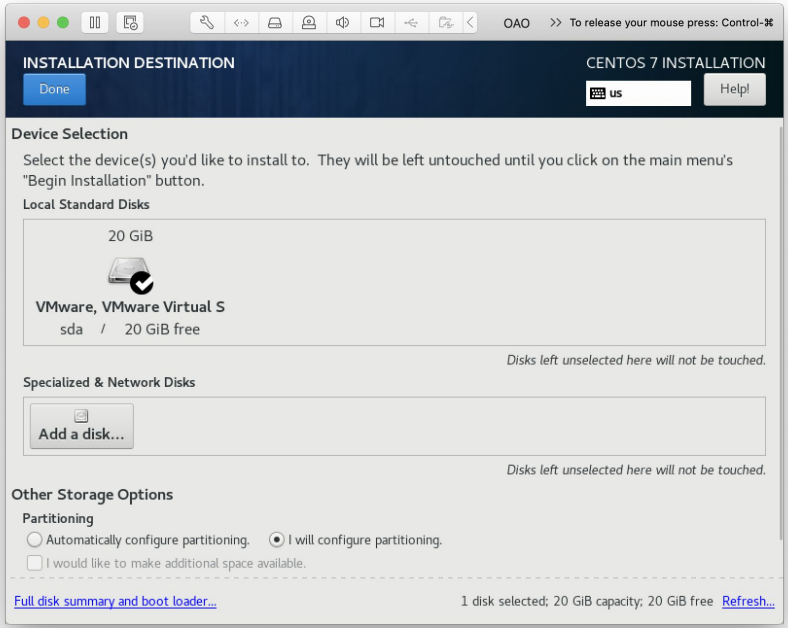
Exercise - VM Partitioning

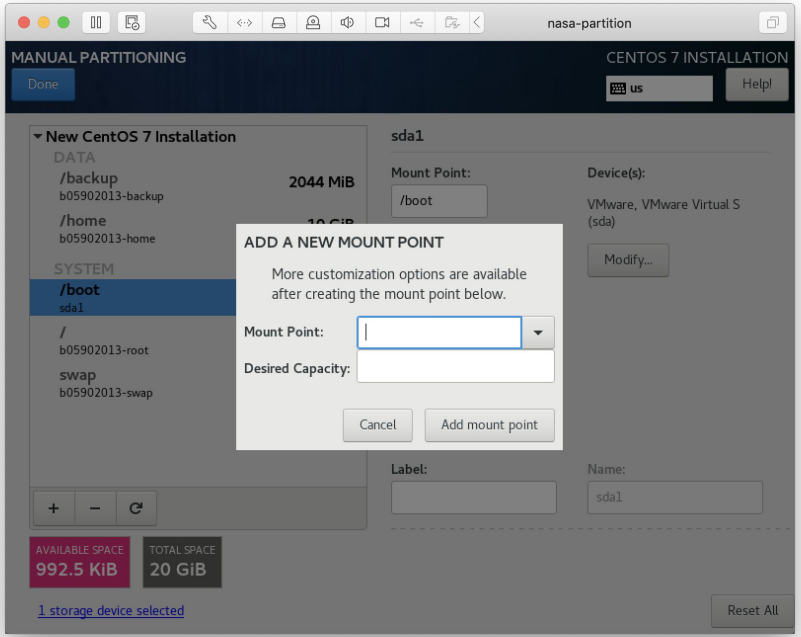
- Step 1 : Create a new Virtual Machine



Exercise - VM Partitioning

- Step 2 : Configure Partitioning





Exercise - VM Partitioning

- Step 3 : Set the networking
  - Hints : nmtui and set automatically connect
  - After that, you can access your virtual machine via ssh

Exercise - Sample Solution

```
[root@localhost ~]# lsblk
NAME                  MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
sda                     8:0    0   20G  0 disk
├─sda1                  8:1    0    1G  0 part /boot
└─sda2                  8:2    0   19G  0 part
  ├─b05902013-root    253:0    0    5G  0 lvm  /
  ├─b05902013-swap    253:1    0    2G  0 lvm  [SWAP]
  ├─b05902013-backup  253:2    0    2G  0 lvm  /backup
  └─b05902013-home    253:3    0   10G  0 lvm  /home
sr0                    11:0    1 1024M  0 rom
```

```
[root@localhost ~]# cat /etc/fstab
/dev/mapper/b05902013-root /                       ext4    defaults        1 1
/dev/mapper/b05902013-backup /backup               xfs     defaults        0 0
UUID=8e58d89b-4be0-4d42-a427-dc836cfe7c80 /boot    xfs     defaults        0 0
/dev/mapper/b05902013-home /home                   ext4    defaults        1 2
/dev/mapper/b05902013-swap swap                    swap    defaults        0 0
```