



15. Storage and File

▼ Storage

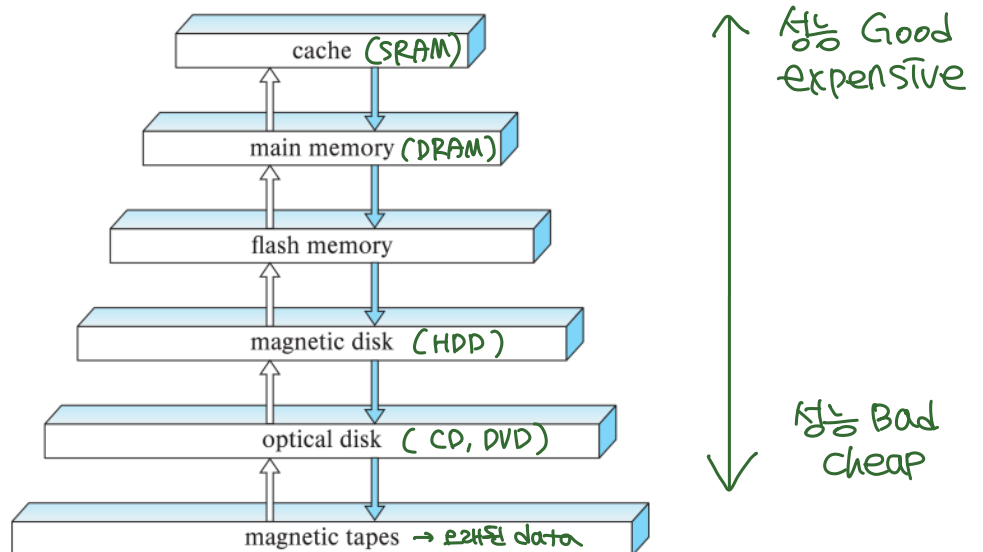
▼ differentiate

1. volatile storage : power가 꺼졌을 때 content 저장 x
2. non-volatile storage : power가 꺼져도 content 유지 → 비휘발성
 - main memory의 back-up으로 secondary, tertiary storage로 사용

▼ Storage 선택은 어떤 기준으로?

1. 접근 속도(성능)
2. 단위 용량 당 가격(유지 비용)
3. reliability(유지 비용) → 서버장은 고장... 얼마나 잘 나눌지...

▼ Storage hierarchy



- High
1. primary storage : fast media, volatile But, expensive
 2. secondary storage : non-volatile, fast access time
 - on-line storage라고도 불림 (ex. flash memory, magnetic disks)

Low 3. tertiary storage : non-volatile, slow access time

- off-line storage라고도 불림
- archival storage : 빈번하게 접근되지 않는 놈을 저장

ex) magnetic tape

→ sequential access, 1 to 12TB capacity

→ few drives with many tapes(외장하드) → 막 petabyte 단위임

▼ Storage interfaces

- Disk와 연결하는 interface
 - SATA(serial ATA) : SATA3는 6GB/sec까지 속도 가능 → serial port 필요
 - SAS(serial Attached SCSI) : SAS version3은 23GB/sec까지 속도 가능
 - NVMe(Non-Volatile Memory Express) : 낮은 latency와 높은 transfer rate로 PCIe connector로 연결 가능
⇒ 더 높은 성능 ← ⇒ 병용적으로 사용
 - 24GB/sec까지 속도 가능

∴ DISK → computer system과 직접적으로 연결 됨

- Storage Area Networks(SAN) : 많은 수의 disk가 많은 server에 높은 속도로 연결 (disk 기반) → 같은 network에 있음
- Network Attached Storage(NAS)
 - networked storage가 disk system interface를 제공하지 x
 - networked file system protocol을 이용하여 file system interface 을 제공

▼ SSD → Flash Storage flash memory 기반

- Solid state disks
 - block 단위의 disk inference를 쓰지만 내부적으로는 multiple flash storage에 data 저장
 - SATA → 500MB/sec , NVMe와 PCIe → 3GB
- flash storage
 - NAND flash (↔ NOR flash) ⇒ page 존재 (↔ DRAM)
 - NOR flash보다 싸기 때문에 더 wide하게 쓰인다
 - 한 번 읽을 때 하나의 page를 읽음(20ms ~ 100ms) → page : 512B ~ 4KB
 - page는 한 번만 쓰일 수 있다 → rewrite를 하기 위해서는 반드시 지워져야 함

- erase block ~~☆~~

→ 128 ~ 256개의 page

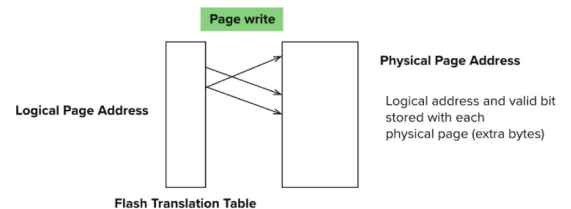
- erase는 erase block 단위로 일어남 (2 ~ 5ms, 256KB ~ 1MB)
- 10만개에서 1백만번의 erase → erase block 더이상 사용 불가
⇒ wear leveling 필요 : data가 골고루 작성되어 다 비슷한 level이 되도록

- Remapping

- logical page address → physical page address로 다시 mapping
⇒ erase를 기다리는 것을 막아줌

- Flash translation table

- mapping 정보 저장
- flash page의 label field 저장
- flash translation layer에 의해 remapping



- SSD performance metrics → 보통 Random한 R/W?

- no SSD (typical)

- 4KB read → 10,000 IOPS (초당 1만번 read)
- 4KB write → 40,000 IOPS

- SSD → 병렬 작업 가능

- 4KB read
→ 100,000 IOPS on SATA (QD-32)
→ 350,000 IOPS on NVMe PCIe
- 4KB write → 100,000 IOPS on QD-32

- data 전송 속도도 NVMe, PCIe가 더빠름 ↔ SATA-3

⇒ Hybrid disks : 큰 magnetic disk에 적은 양의 flash cache 합쳐해서 사용