

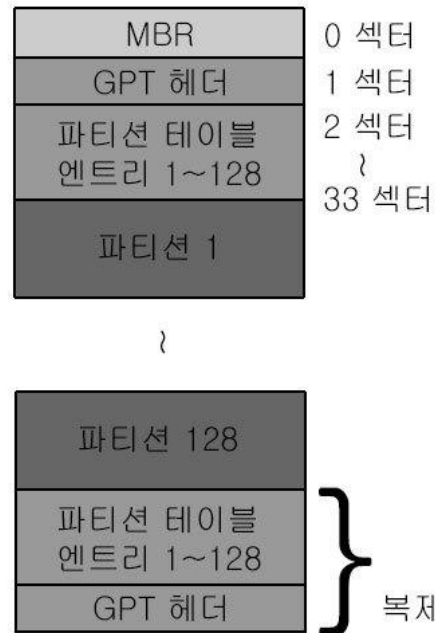
Volume and Filesystem

Volume



Partition

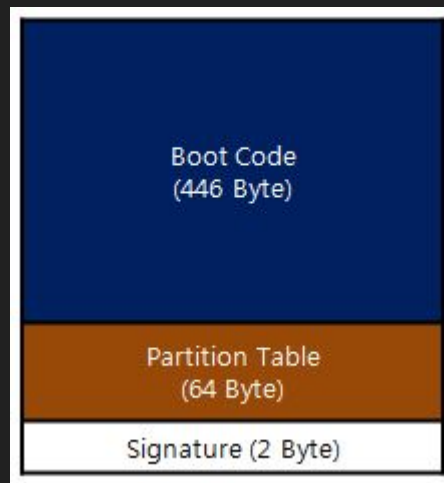
- 그냥 자른거
- Not Volume



MBR

- Master Boot Record

| 범위 (Byte Range) | | 설명 (Description) | 크기 (Size) |
|-----------------|-----------------|--------------------------|-----------|
| 10 진수 | 16 진수 | | |
| 0 – 445 | 0x0000 – 0x01BD | Boot code | 446 bytes |
| 446 – 461 | 0x01BE – 0x01CD | Partition table entry #1 | 16 bytes |
| 462 – 477 | 0x01CE – 0x01DD | Partition table entry #2 | 16 bytes |
| 478 – 493 | 0x01DE – 0x01ED | Partition table entry #3 | 16 bytes |
| 494 – 509 | 0x01EE – 0x01FD | Partition table entry #4 | 16 bytes |
| 510 – 511 | 0x01FE – 0x01FF | Signature (0x55AA) | 2 bytes |



Partition entry

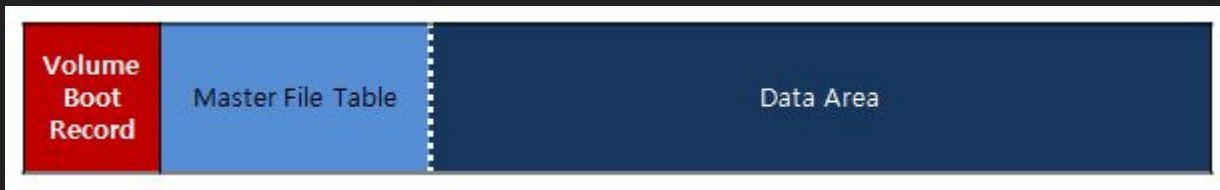
| 범위 (Byte Range) | | 설명 (Description) | 크기 (Size) |
|-----------------|-----------------|--|-----------|
| 10 진수 | 16 진수 | | |
| 0 – 0 | 0x0000 – 0x0000 | Boot Indicator 00 = do not use for booting 80 = system partition | 1 byte |
| 1 – 3 | 0x0001 – 0x0003 | Starting CHS address | 3 bytes |
| 4 – 4 | 0x0004 – 0x0004 | Partition type | 1 byte |
| 5 – 7 | 0x0005 – 0x0007 | Ending CHS address | 3 bytes |
| 8 – 11 | 0x0008 – 0x000B | Starting LBA address | 4 bytes |
| 12 – 15 | 0x000C – 0x000F | Total sectors | 4 bytes |

MBR Sector

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|-------|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----------|-----------|-----------|----|---|---|----|---|---|---|---|---|
| 00033 | C0 | 8E | D0 | BC | 00 | 7C | FB | 50 | 07 | 50 | 1F | FC | BE | 1B | 7C | 3ÅZDw | | üP | · | P | · | ü | · | 1 | | | |
| 016BF | 1B | 06 | 50 | 57 | B9 | E5 | 01 | F3 | A4 | CB | BD | BE | 07 | B1 | 04 | 2 | · | PW | · | Å | · | 0x | E | 4 | 1 | | |
| 03238 | 6E | 00 | 7C | 09 | 75 | 13 | 83 | C5 | 10 | E2 | F4 | CD | 18 | 8B | F5 | 8n | | u | · | f | Å | · | 0 | I | < | 0 | |
| 04883 | C6 | 10 | 49 | 74 | 19 | 38 | 2C | 74 | F6 | A0 | B5 | 07 | B4 | 07 | 8B | fE | · | It | · | 8 | t | 0 | u | 0 | < | 0 | |
| 064F0 | AC | 3C | 00 | 74 | FC | BB | 07 | 00 | B4 | 0E | CD | 10 | EB | F2 | 88 | a | 0 | · | t | ü | 0 | · | · | · | · | · | 0 |
| 0804E | 10 | E8 | 46 | 00 | 73 | 2A | FE | 46 | 10 | 80 | 7E | 04 | 0B | 74 | 0B | N | · | e | F | · | s | · | p | F | · | 0 | 0 |
| 09680 | 7E | 04 | 0C | 74 | 05 | A0 | B6 | 07 | 75 | D2 | 80 | 46 | 02 | 06 | 83 | 0 | · | · | · | · | · | · | · | · | · | · | 0 |
| 11246 | 08 | 06 | 83 | 56 | 0A | 00 | E8 | 21 | 00 | 73 | 05 | A0 | B6 | 07 | EB | F | · | · | f | V | · | 0 | · | · | · | · | 0 |
| 128BC | 81 | 3E | FE | 7D | 55 | AA | 74 | 0B | 80 | 7E | 10 | 00 | 74 | C8 | A0 | 4 | 0 | p | U | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 144B7 | 07 | EB | A9 | 8B | FC | 1E | 57 | 8B | F5 | CB | BF | 05 | 00 | 8A | 56 | 0 | 0 | 0 | 0 | W | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 16000 | B4 | 08 | CD | 13 | 72 | 23 | 8A | C1 | 24 | 3F | 98 | 8A | DE | 8A | FC | 0 | 0 | I | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17643 | F7 | E3 | 8B | D1 | 86 | D6 | B1 | 06 | D2 | EE | 42 | F7 | E2 | 39 | 56 | C | 0 | a | 0 | N | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1920A | 77 | 23 | 72 | 05 | 39 | 46 | 08 | 73 | 1C | B8 | 01 | 02 | BB | 00 | 7C | w | 0 | r | 0 | 9 | F | 0 | 0 | 0 | 0 | 0 | 0 |
| 2088B | 4E | 02 | 8B | 56 | 00 | CD | 13 | 73 | 51 | 4F | 74 | 4E | 32 | E4 | 8A | < | 0 | N | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 22456 | 00 | CD | 13 | EB | E4 | 8A | 56 | 00 | 60 | BB | AA | 55 | B4 | 41 | CD | V | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 24013 | 72 | 36 | 81 | FB | 55 | AA | 75 | 30 | F6 | C1 | 01 | 74 | 2B | 61 | 60 | 0 | 0 | r | 6 | 0 | U | 0 | 0 | 0 | 0 | 0 | 0 |
| 2566A | 00 | 6A | 00 | FF | 76 | 0A | FF | 76 | 08 | 6A | 00 | 68 | 00 | 7C | 6A | j | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 27201 | 6A | 10 | B4 | 42 | 8B | F4 | CD | 13 | 61 | 61 | 73 | 0E | 4F | 74 | 0B | 0 | 0 | j | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 28832 | E4 | 8A | 56 | 00 | CD | 13 | EB | D6 | 61 | F9 | C3 | 49 | 6E | 76 | 61 | 2 | 0 | a | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3046C | 69 | 64 | 20 | 70 | 61 | 72 | 74 | 69 | 74 | 69 | 6F | 6E | 20 | 74 | 61 | l | id | partition | ta | | | | | | | | |
| 32062 | 6C | 65 | 00 | 45 | 72 | 72 | 6F | 72 | 20 | 6C | 6F | 61 | 64 | 69 | 6E | ble | Error | loadin | | | | | | | | | |
| 33667 | 20 | 6F | 70 | 65 | 72 | 61 | 74 | 69 | 6E | 67 | 20 | 73 | 79 | 73 | 74 | g | operating | syst | | | | | | | | | |
| 35265 | 6D | 00 | 4D | 69 | 73 | 73 | 69 | 6E | 67 | 20 | 6F | 70 | 65 | 72 | 61 | em | Missing | opera | | | | | | | | | |
| 36874 | 69 | 6E | 67 | 20 | 73 | 79 | 73 | 74 | 65 | 6D | 00 | 00 | 00 | 00 | 00 | ting | system | 0000 | | | | | | | | | |
| 38400 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 40000 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 41600 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 43200 | 00 | 00 | 00 | 00 | 00 | 2C | 44 | 63 | 1C | 20 | 1C | 20 | 00 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 44801 | 00 | 07 | FE | FF | FF | 3F | 00 | 00 | 00 | 00 | 62 | 04 | 53 | 07 | 00 | FE | 00000000 | | | | | | | | | | |
| 464FF | FF | 05 | FE | FF | FF | A1 | 04 | 53 | 07 | E0 | 40 | C9 | 15 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 48000 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00000000 | | | | | | | | | | | |
| 49600 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 00 | 55AA | 00000000 | | | | | | | | | | |

VBR

- 클러스터 크기
- MFT 위치
- 전체 섹터 등 해당 볼륨의 추가적인 정보
- 부팅에 필요한 시스템 파일의 위치와 실행할 수 있는 코드 포함



| Cluster Size (Byte) | VBR Size (Sector) |
|---------------------|-------------------|
| 512 | 1 |
| 1K | 2 |
| 2K | 4 |
| 4K | 8 |

Jump command

OEM ID

BIOS Parameter Block

| Offset | Jump command | OEM ID | BIOS Parameter Block | Hex | ASCII |
|--------|--|----------------------------------|-------------------------|-------------------------|---------------------|
| 000 | EB 52 90 | 4E 54 46 53 20 20 20 20 | 00 02 08 00 00 | 00 00 00 00 00 00 00 00 | äRONTFS |
| 016 | 00 00 00 00 00 | F8 00 00 3F 00 FF 00 3F 00 00 00 | 00 00 00 00 00 00 00 00 | 00 00 00 00 00 00 00 00 |ø...?·ÿ?... |
| 032 | 00 00 00 00 80 00 80 00 5F A2 98 25 00 00 00 00 | | | |€·€·_·c~&.... |
| 048 | 00 0C 00 00 00 00 00 00 91 22 01 00 00 00 00 00 | | | |"..... |
| 064 | F6 00 00 00 01 00 00 00 8E D5 02 10 08 03 10 F0 | | | | ö.....ŽÕ.....ä |
| 080 | 00 00 00 00 FA 33 C0 8E D0 BC 00 7C FB B8 C0 07 | | | |ú3ÄŽĐ· ú,À· |
| 096 | 8E D8 E8 16 00 B8 00 0D 8E C0 33 DB C6 06 0E 00 | | | | Ž0è...· ŽÀ3ŮE... |
| 112 | 10 E8 53 00 68 00 0D 68 6A 02 CB 8A 16 24 00 B4 | | | | ·èS·h· hĵ·ĚŠ·\$·` |
| 128 | 08 CD 13 73 05 B9 FF FF 8A F1 66 0F B6 C6 40 66 | | | | ·Í·s··ÿÿŠñf·ŦE@f |
| 144 | 0F B6 D1 80 E2 3F F7 E2 86 CD C0 ED 06 41 66 0F | | | | ·ŦŇĈâ?·â·í·Á·Af· |
| 160 | B7 C9 66 F7 E1 66 A3 20 00 C3 B4 41 BB AA 55 8A | | | | ·Ěf·áff··Ä·A»·UŠ |
| 176 | 16 24 00 CD 13 72 0F 81 FB 55 AA 75 09 F6 C1 01 | | | | ·\$·Í·r·ŦŦU·u ōÁ· |
| 192 | 74 04 FE 06 14 00 C3 66 60 1E 06 66 A1 10 00 66 | | | | t·p·...Äf·...f;...f |
| 208 | 03 06 1C 00 66 3B 06 20 00 0F 82 3A 00 1E 66 6A | | | | ...f;· · · · · ·fj |
| 224 | 00 66 50 06 53 66 68 10 00 01 00 80 3E 14 00 00 | | | | ·fP·Sfh·...€>... |
| 240 | 0F 85 0C 00 E8 B3 FF 80 3E 14 00 00 0F 84 61 00 | | | | ...·è·ÿ€>...·„a· |
| 256 | B4 42 8A 16 24 00 16 1F 8B F4 CD 13 66 58 5B 07 | | | | ·BŠ·\$·...<ōÍ·fX[· |
| 272 | 66 58 66 58 1F EB 2D 66 33 D2 66 0F B7 0E 18 00 | | | | fXfX·ë-f3Ŧf·..... |
| 288 | 66 F7 F1 FE C2 8A CA 66 8B D0 66 C1 EA 10 F7 36 | | | | f·ñp·ÄŠĚf·ĐfÁê·÷6 |
| 304 | 1A 00 86 D6 8A 16 24 00 8A E8 C0 E4 06 0A CC B8 | | | | ··+ŦŠ·\$·ŠèÄ· Ì, |
| 320 | 01 02 CD 13 0F 82 19 00 8C C0 05 20 00 8E C0 66 | | | | ··Í·...·ŦÄ· ·ŽÀf |
| 336 | FF 06 10 00 FF 0E 0E 00 0F 85 6F FF 07 1F 66 61 | | | | ÿ·...ÿ·...·...oÿ·fa |
| 352 | C3 A0 F8 01 E8 09 00 A0 FB 01 E8 03 00 FB EB FE | | | | Ä ø·è · ů·è·ûëp |
| 368 | B4 01 8B F0 AC 3C 00 74 09 B4 0E BB 07 00 CD 10 | | | | ··<ä-<·t ··»·Í· |
| 384 | EB F2 C3 0D 0A 41 20 64 69 73 6B 20 72 65 61 64 | | | | èòÄ A disk read |
| 400 | 20 65 72 72 6F 72 20 6F 63 63 75 72 72 65 64 00 | | | | error occurred· |
| 416 | 0D 0A 4E 54 4C 44 52 20 69 73 20 6D 69 73 73 69 | | | | NTLDR is missi |
| 432 | 6E 67 00 0D 0A 4E 54 4C 44 52 20 69 73 20 63 6F | | | | ng· NTLDR is co |
| 448 | 6D 70 72 65 73 73 65 64 00 0D 0A 50 72 65 73 73 | | | | mpressed· Press |
| 464 | 420 43 74 72 6C 2B 41 6C 74 2B 44 65 6C 20 74 6F | | | | Ctrl+Alt+Del to |
| 480 | 20 72 65 73 74 61 72 74 0D 0A 00 00 00 00 00 00 | | | | restart |
| 496 | 00 00 00 00 00 00 00 00 83 A0 B3 C9 00 00 55 AA | | | |f ·É·U· |

윈도우즈 부팅 과정

1. CPU 레지스터인 Program Counter(PC)를 초기화
2. CPU 이상유무를 테스트
3. POST(Power On Self-Test) 작업을 수행하기 위한 기본적인 테스트
4. System Bus가 정상적으로 동작하는지 테스트
5. RTC(Real-Time Clock; or system clock)을 테스트
6. 비디오 구성요소들(비디오 메모리 등)을 테스트
7. RAM을 테스트
8. 키보드가 정상적으로 연결되어 있는지 혹은 눌려진 키가 없는지 테스트
9. 연결된 모든 드라이브(플로피, CD, 하드디스크 등) 테스트
10. POST의 결과가 CMOS(RTC/NVRAM)에 저장된 값과 일치하는지 검사
11. SCSI BIOS와 같은 추가적인 BIOS를 RAM에 로드
12. Plug and Play를 실행하여 운영체제 로드를 위한 기본적인 구성을 RAM에 로드
13. 드라이브 내에서 첫 번째 섹터 MBR(Master Boot Record)를 읽는다
14. MBR에서 부팅 가능한 파티션을 찾는 작업
15. VBR 참조
16. NTLDR, BOOTMGR
17. NTOSKRNL.EXE 이후 커널(Kernel), HAL(Hardware Abstraction Layer), 시스템 레지스트리 정보를 로드
18. TCP/IP와 관련된 네트워크 드라이버들을 로드
19. 로그인
20. 사용자에게 대한 레지스트리 정보를 가져와 사용자 환경을 구성

What is Volume?

- Logical drive
- A single accessible storage area
- A single filesystem
- Block device - 일정크기(block) 단위로 접근하는 장치
- Filesystem - 파일 단위 접근

1. 파일시스템으로부터 Block IO 요청을 받고
2. 스케줄링
3. block device로 commit

I/O

- Read (disk -> memory)
- Write (memory-> disk)
- Buffer VS Cache
- Filesystem과 Disk에서 r/w access 하는 단위가 다름 (bio or IRP, sector)
- 물리장치가 가지고 있는 특성에 맞는 request 조절이 필요

IO Scheduler (Linux)

- Filesystem으로부터 들어온 request들을 merge & sort
- 보통 HDD의 특성에 맞춰져 있음
 - Seek & rotation time 최소화
- SSD의 특성을 반영한 수정도 활발히 이루어짐

1. Noop
2. Deadline
3. CFQ(Complete Fair Queuing)



File system

(Local)

What is File system?

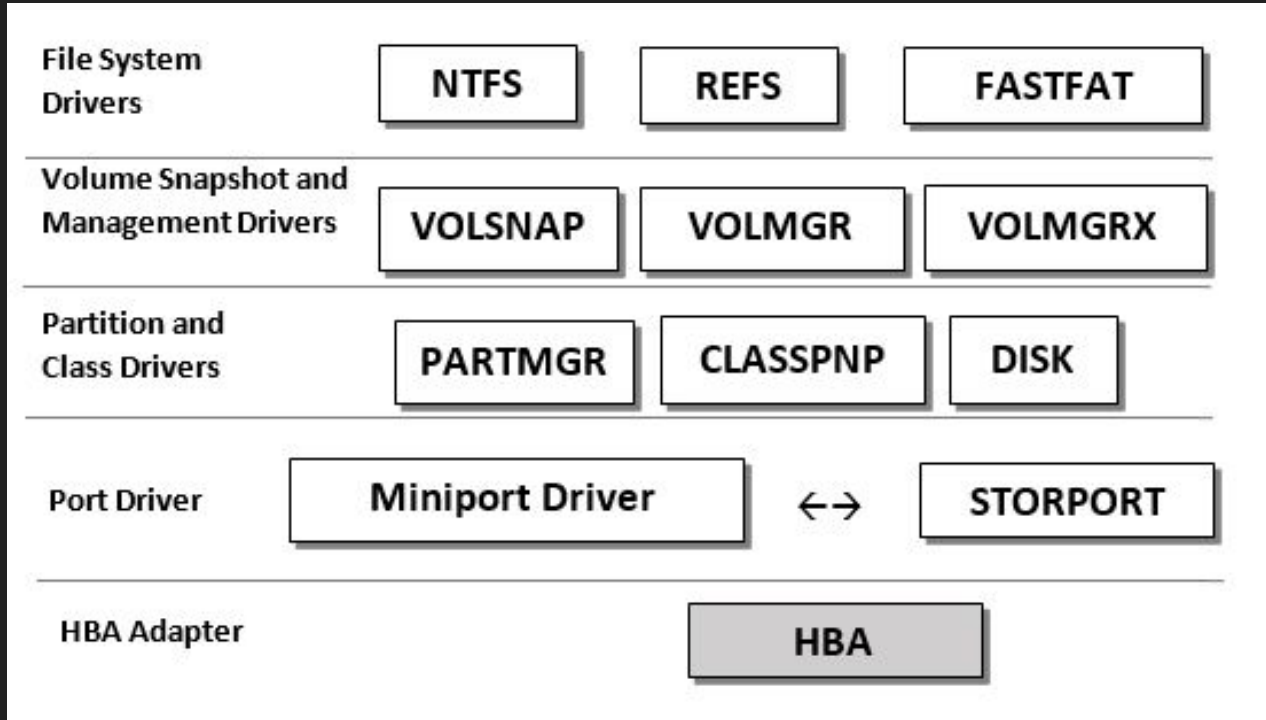
“파일 시스템은 컴퓨터에서 파일이나 자료를 쉽게 발견 및 접근할 수 있도록 보관 또는 조직하는 체제를 가리키는 말이다.” - Wiki

- 기본요소
 - 클러스터
 - 파일
 - 디렉토리
- 부가요소
 - 소유권
 - 동기화
 - 일관성(Consistency) 체크
 - 저널링
 - 보안
 - 압축

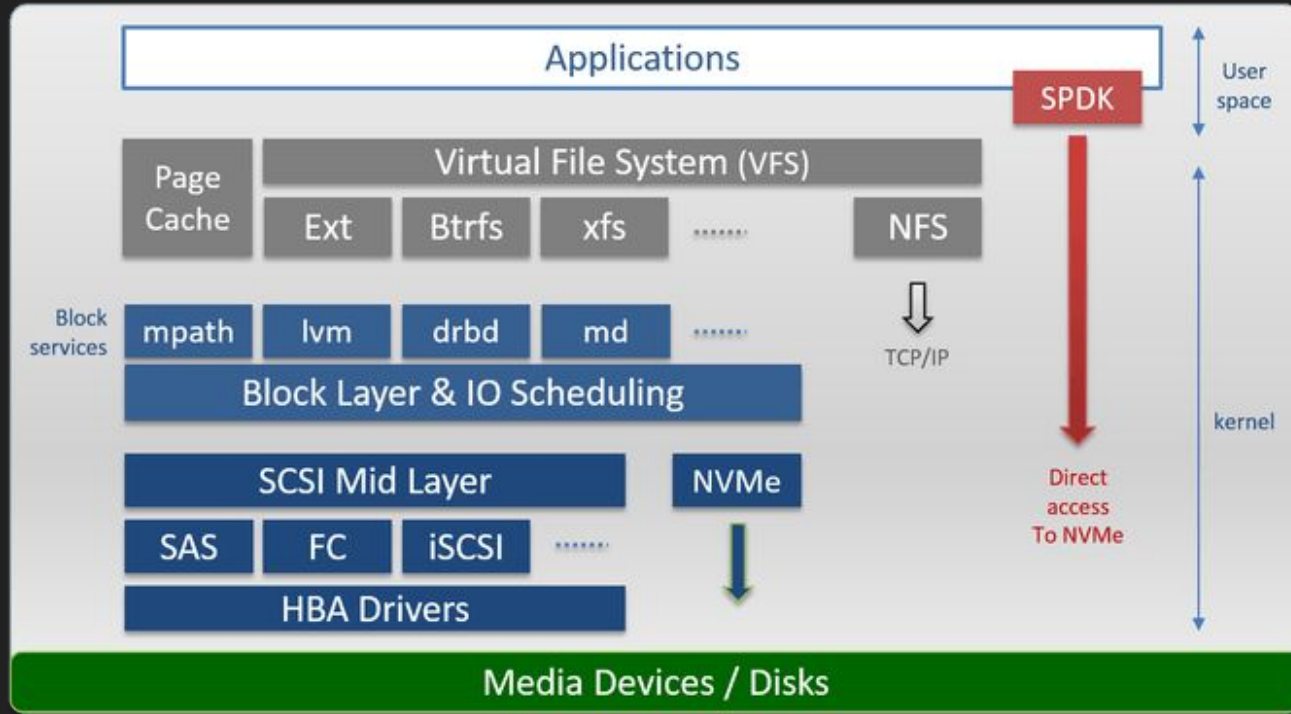
File Systems

- Disk file systems
 - Unix 계열
 - Linux
 - ext, ext2, ext3, ext4, XFS, ReiserFS, BTRFS, ZFS
 - Solaris
 - ZFS
 - MS Windows
 - FAT, NTFS, ReFS
- Database file systems
- Network file systems
 - NFS, SMB
- Special file systems
 - Proc, sysfs, udev

Windows Storage Stack

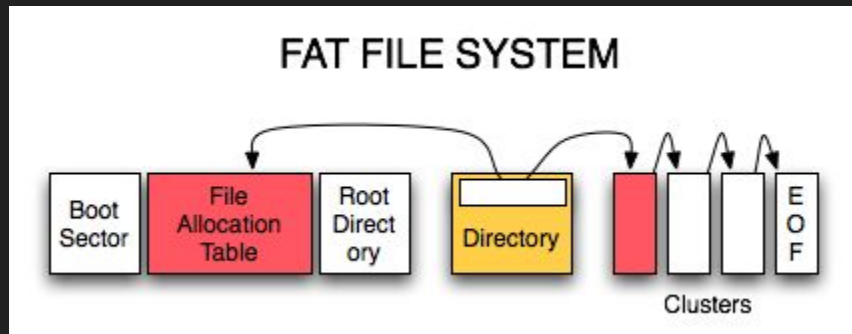


Linux Storage Stack



FAT (File Allocation Table)

- FAT12, FAT16, FAT32, exFAT
- 초기 MS의 OS에서 사용



NTFS (New Technology File System)

- 현재 MS 윈도우즈의 표준 파일 시스템
- 정확한 기술공개를 하지 않고 있음
- Versions
 - V1.0 (1993) - NT 3.1
 - V1.1 (1995) - NT 3.51. +압축
 - V1.2 (1996) - NT 4.0
 - V3.0 - Windows 2000. +encrypting, sparse, quotas
 - V3.1 (2001) - Windows XP. Expanded MFT
 - 이후는 기능 추가만
 - Symbolic links, Transactional, partition shrinking, self-healing and so on
- B+ tree

NTFS Features

- Scalability
- Journaling
- Alternate data streams
- File compression
- Sparse files
- Volume Shadow Copy
- Transactions
- Security, Encryption
- Reparse points
- Resizing

ReFS (Resilient File System)

복원력 있는 파일 시스템

- 체크섬을 포함한 메타데이터 **무결성**
- 선택적 사용자 데이터 무결성을 제공하는 **무결성** 스트림
- 강력한 디스크 업데이트를 위한 쓰기 시 할당 **트랜잭션** 모델
- **대용량** 볼륨, 파일 및 디렉터리 크기
- **저장소 풀** 구성과 **가상화**로 파일 시스템 생성과 관리 용이성 개선
- 성능 개선(대역폭을 관리할 수 있음)을 위한 데이터 **스트라이핑** 및 장애 복구에 대비한 **이중화**
- 숨겨진 디스크 **오류로부터 보호**해 주는 디스크 스크러빙
- ‘복원’을 통해 **손상에 대한 복원력**을 제공하여 모든 경우에 최대의 볼륨 가용성 유지
- 다수의 시스템상에서 공유 저장소 풀을 구성하여 **장애 복구** 능력과 **부하 균형** 조절 능력 개선
 - MSDN 문서중

Linux Filesystem

- Ext (Extended file system)
 - inode
 - Minix에서 사용되던 파일시스템에서 벗어나 Linux용으로 나온 파일시스템
 - ext4 까지 나오는 동안 리눅스 주력 파일시스템
 - VFS와 연동
- Ext3
 - Ext2에 journaling 기능 추가
 - Hashed btree
- Ext4
 - 대형 파일 시스템 (1 EB지원, 16 TB 파일)
 - Extent 단위 (ext2, 3은 block mapping)
- ReiserFS
- XFS

ZFS, BTRFS

- 차세대 파일시스템
 - 하지만 아직 갈길이 멀다
- 기존 파일시스템이 가져야할 요소는 다 가지고 있다.
- COW
- 파일 시스템이 RAID 지원
- 파일 시스템이 Dedup 지원
- 파일 시스템이 Snapshot 지원
 - Incremental backup 지원
- 파일 시스템이..... 하여간 많다.

Reference

- [\[Linux\] 블록 장치 I/O 동작 방식 1~4](#)
- [Linux 2.6 I/O Scheduler Configuration](#)
- [Operating System Inside - General](#)
- [Forensic-proof](#)
- [차세대 Windows 파일 시스템 개발: ReFS](#)
- [확장성, 복원력 및 효율성을 위한 저장소 가상화](#)
- 『[임베디드 개발자를 위한 파일시스템의 원리와 실습](#)』