Partitioner och filsystem 1

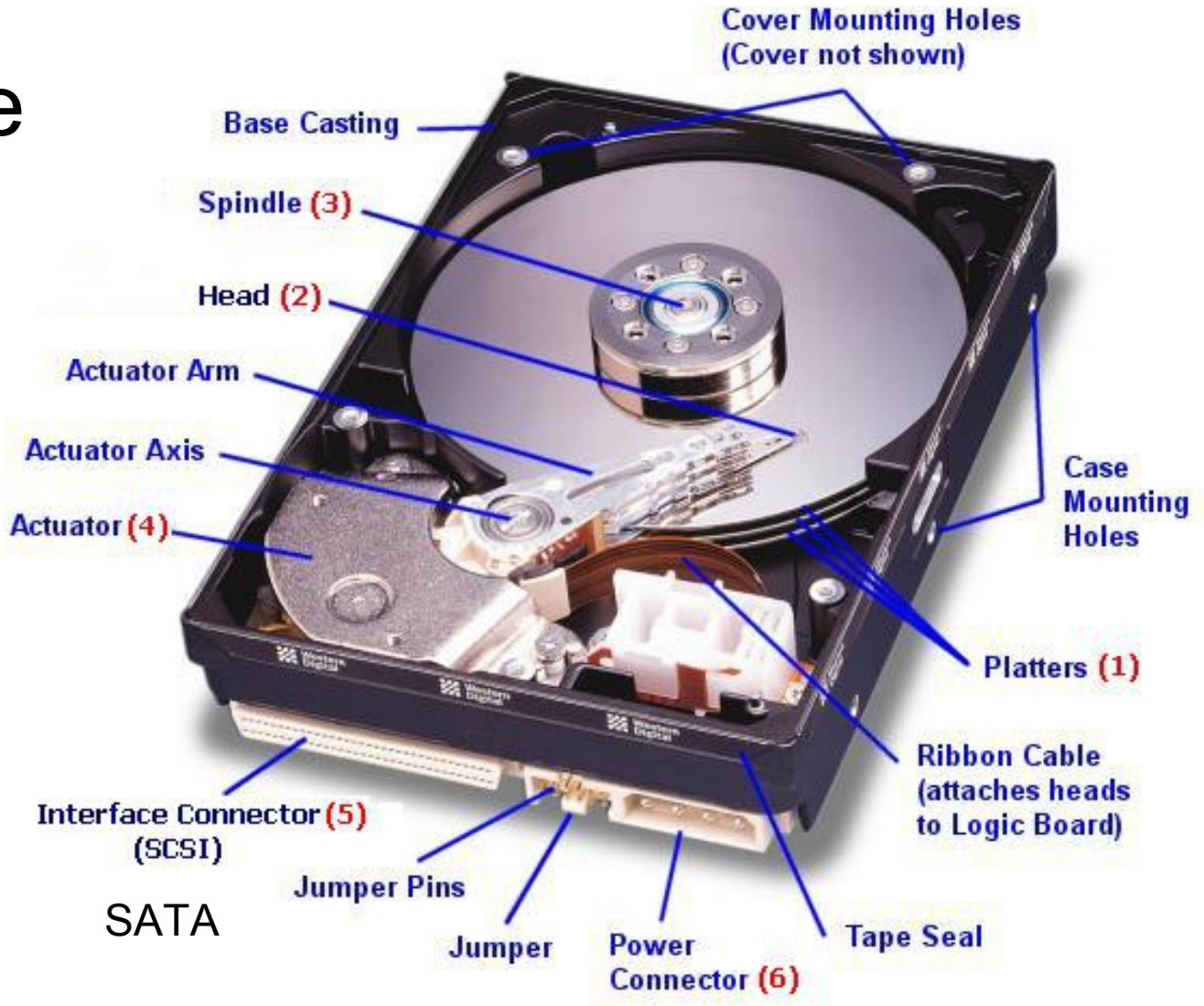
Hard drives

Partitions

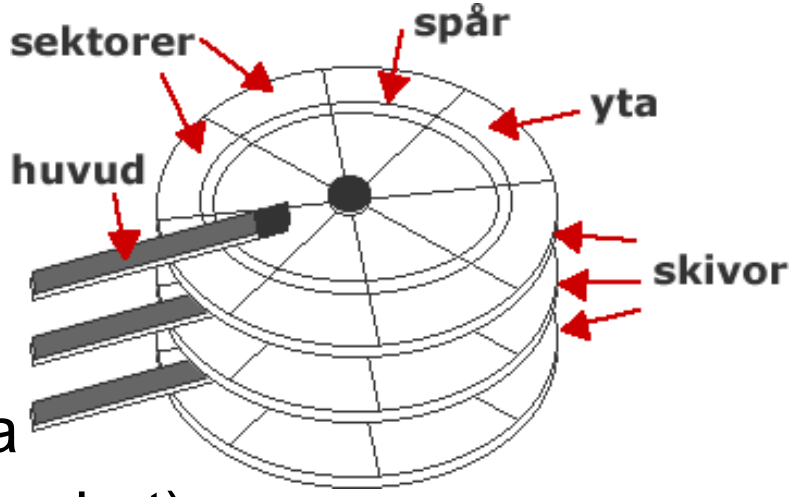
NAS and SAN

Hårddiskens uppbyggnad I

- 6 tillverkare
  - Seagate
  - Western Digital
  - Hitachi
  - Samsung
  - Fujitsu
  - Toshiba
  - Fler?



Hårddisken uppbyggnad II

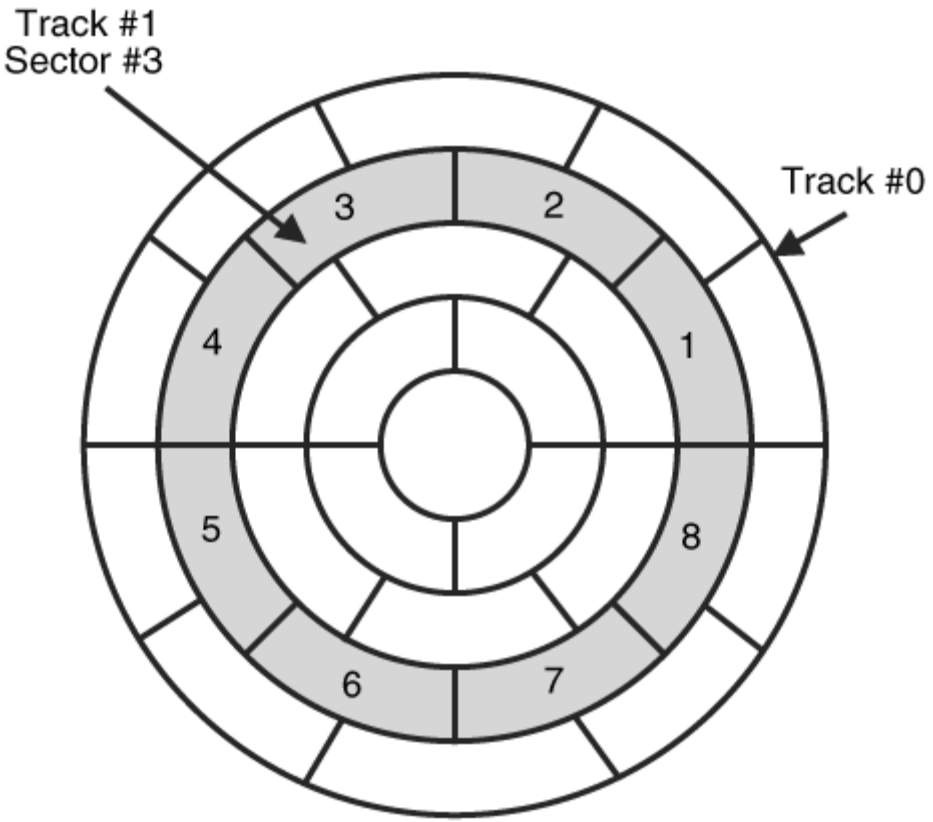


- Hårddiskens geometri är informationen om
  - Cylindrar (spår, tracks)
  - Huvud (heads), ett för varje yta
  - Sektorer (minsta skriv/läsbara enhet)
- Filsystem organiserar data logiskt i kluster, kluster skapas med en viss storlek beroende på
  - Storleken på disken
  - Filsystemets typ
  - Användarens kommando vid formatering
  - Kolla med cmd shell: format /?

Track, sektorer och kluster

- Formaterade hårddiskar består av sektorer
  - Block size
  - 512 byte för det mesta (1024 byte kan förekomma)
- EN ISO 9660 CD använder 2048 byte blocks

- Kluster byggs upp av ***n*** antal sektorer
  - T.ex sector 1-4
- En fil består av ***m*** antal kluster som ofta är spridda över hela hårddisken
- En fil kan inte dela ett kluster med någon annan fil



Standard klusterstorlekar i Windows

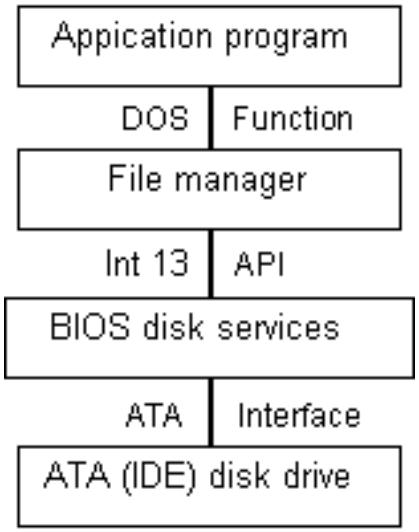
http://support.microsoft.com/kb/140365

- Kan även bero på OS-version
  - Fat 12 har 4 KB klusterstorlek

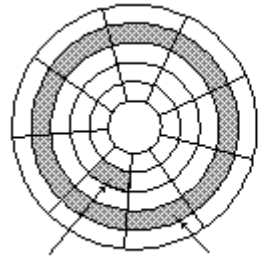
Volume size	FAT16 cluster size	FAT32 cluster size	NTFS cluster size
7 MB–16 MB	2 KB	Not supported	512 bytes
17 MB–32 MB	512 bytes	Not supported	512 bytes
33 MB–64 MB	1 KB	512 bytes	512 bytes
65 MB–128 MB	2 KB	1 KB	512 bytes
129 MB–256 MB	4 KB	2 KB	512 bytes
257 MB–512 MB	8 KB	4 KB	512 bytes
513 MB–1,024 MB	16 KB	4 KB	1 KB
1,025 MB–2 GB	32 KB	4 KB	2 KB
2 GB–4 GB	64 KB	4 KB	4 KB
4 GB–8 GB	Not supported	4 KB	4 KB
8 GB–16 GB	Not supported	8 KB	4 KB
16 GB–32 GB	Not supported	16 KB	4 KB
32 GB–2 TB	Not supported	Not supported	4 KB

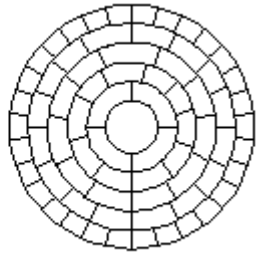
Hårddisk begränsning

- DOS-BIOS modellen
- Int 13 registren är 8 bitars och används när en läsning etc. beordras
- Max storlek är därför
  - 1024 **C**ylinders (2^{10}) kan adresseras
  - 256 **H**eads (2^{8}) kan adresseras
  - 63 **S**ectors (2^{6}-1) kan adresseras
  - Vid 512 bytes per sektor ger det en max teoretisk kapacitet av ca: 8.4 GB
- Logical geometry
  - Zoned Bit Recording (ZBR) eller Zone Density Recording
  - Allt översätts internt av hårddiskens styrprogramvara
  - Alla CHS värden är i princip falska på moderna diskar



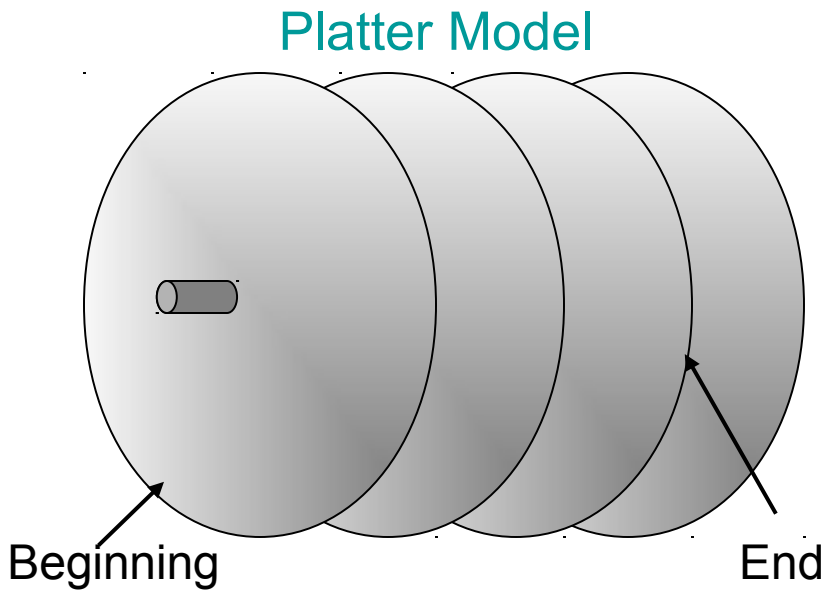
Bitar	Register
8 (low)	Cylinder Low
2 (high) +6 bit sector	Cylinder High/Sector Number
8 bit	Head Number

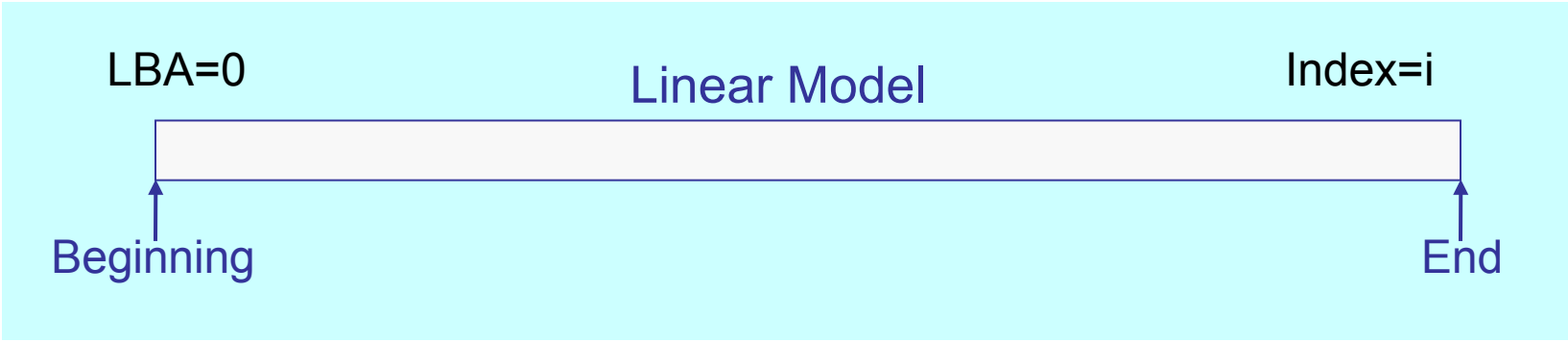




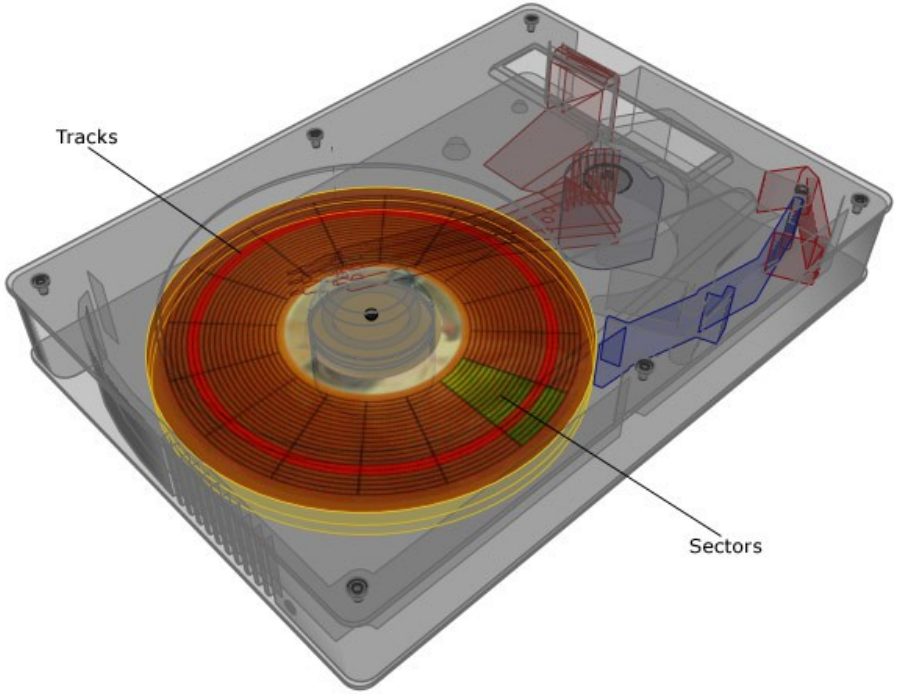
LBA (Logical Block Addressing)

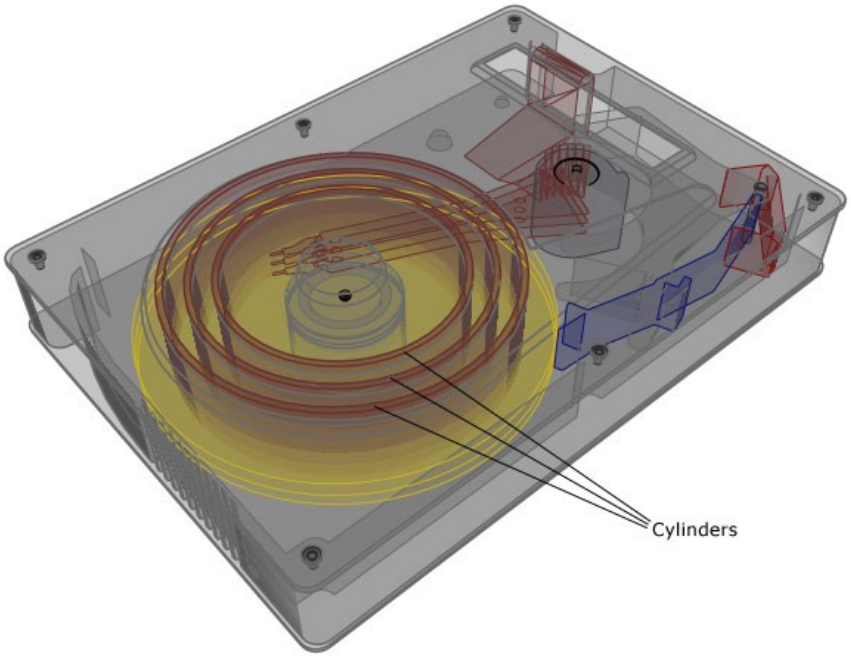
- LBA, kallas även **Linear Base Address**
  - Är den vanligaste adresserings metoden
- Cylinder-Head-Sector (CHS) schema
- LBA har 48 bits adressering
  - Geometri translation i BIOS
  - 128 pebibyte (PiB, 2^50) om 512 byte sektorer används





Hårddiskens uppbyggnad III





- Interface
  - IDE
    - PATA
    - SATA
  - SCSI/SAS
    - Ett gäng
- Exempel 120 GB
  - Cyl: 238216 *
  - Heads: 16 *
  - SPT: 63

  = LBA: 240121728

LBA * 512 / 1024^3

= 114,49

disk storleken i GB

LBA <-> CHS översättning

- **LBA = linjära basadressen för blocket**
- **CHS = Cylinders – Heads – Sectors**
- SPT = Sectors Per Track
- HPC = Heads Per Cylinder
- SECT = Värdet av sector för CHS koordinaten
- HEAD = Värdet av head för CHS koordinaten
- CHS beskrivning av data på media behöver ibland göras av OS för vissa filsystem, t.ex. disketter
- Från LBA till CHS

  CYL = LBA / (HPC * SPT)
  TEMP = LBA % (HPC * SPT)
  HEAD = TEMP / SPT
  SECT = TEMP % SPT + 1

Konverterings formler

- Från CHS till LBA (används inte ofta)

  LBA = (((CYL * HPC) + HEAD) * SPT) + SECT – 1

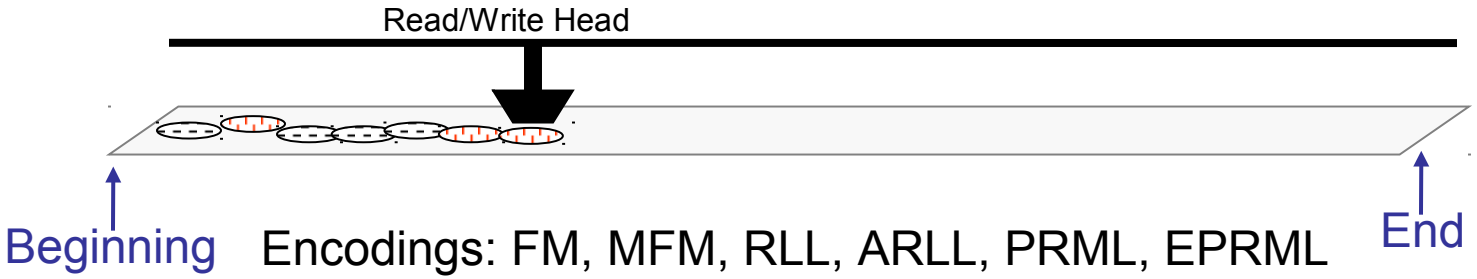
Fixed/Removable Media

- Skriv/läs process (**mycket** förenklad)
  - Write Process

    Digital signals are encoded (for timing recovery) and transformed into analog signals that drive the magnetic field on the write head
  - Read Process

    Analog magnetic field is sensed, timing is recovered, and sampled signals are converted into digital data
- Performance
  - Disk performance (time to retrieve data) can be measured in terms of several important characteristics
  - **Disk access time** is the sum of **(spin-up time** of the disk, **seek time** of the arm to reach the track, **rotational latency/speed** of the disk**)**

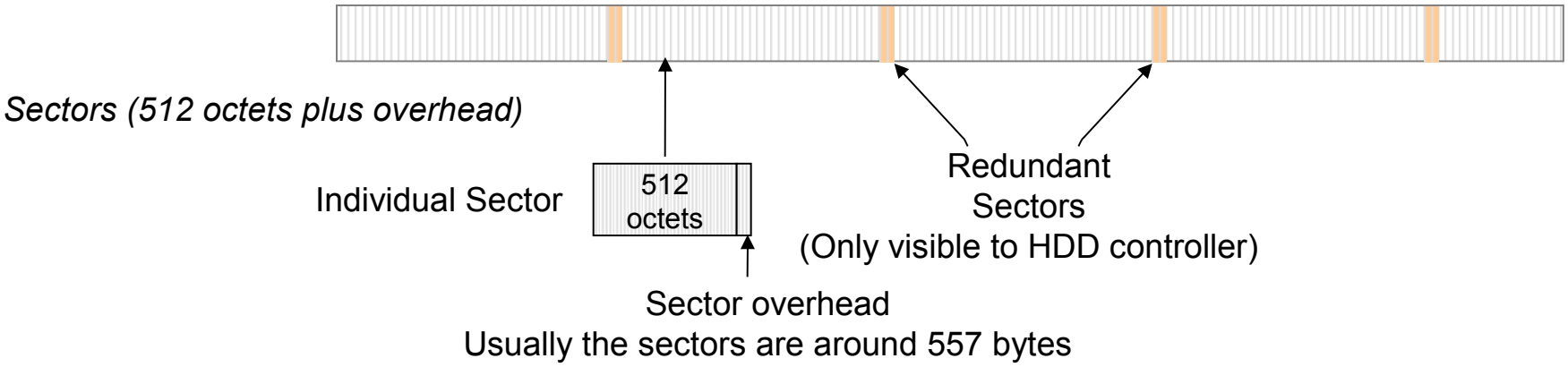
Linear Model



Encodings: FM, MFM, RLL, ARLL, PRML, EPRML

Fixed/Removable Media

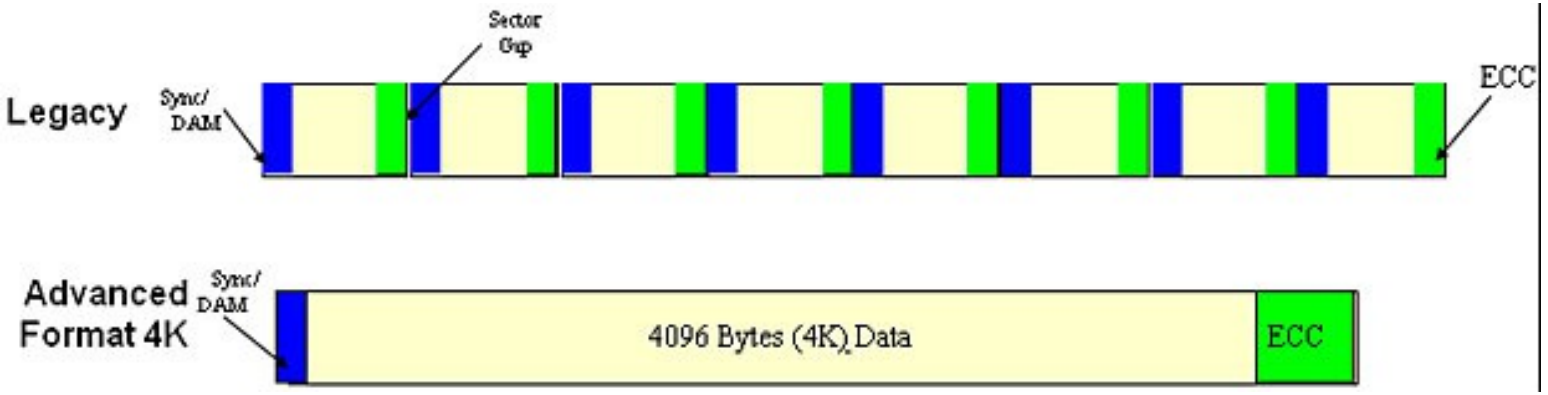
- Low Level Format
  - Performed at factory
  - Low-level formatting adds indivisible units of storage called sectors
  - Most modern HDDs use 512+ octet (byte) sectors
    - The + accounts for sector overhead bytes (differs by manufacturer)
  - Overhead bytes provide **error correction** and **timing recovery** functions
  - Bad sectors are automatically remapped to redundant sectors by the HDD controller



Western Digital brings Advanced Format to Caviar Green

http://anandtech.com/storage/showdoc.aspx?i=3691

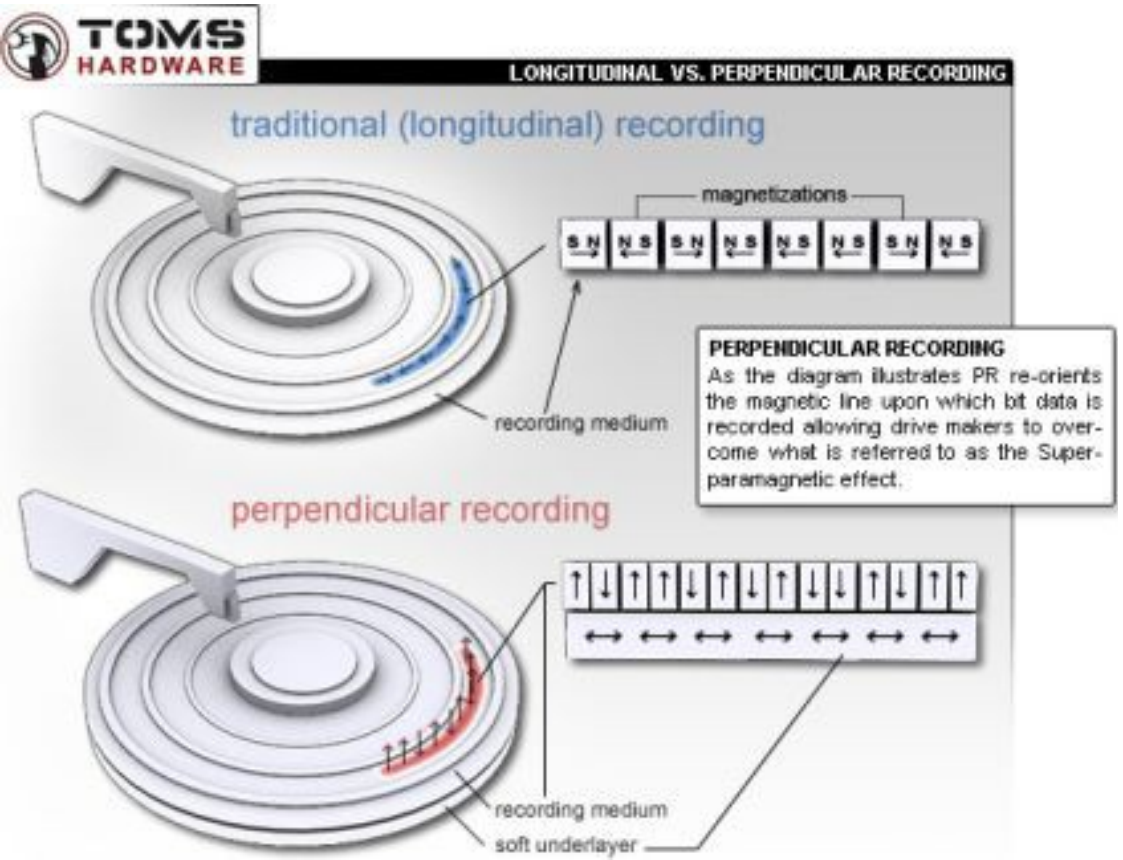
- Advanced format skips using the sector gap
- Stores 4096 bytes in each sector
- Increase capacity with about 10%
- Not all OS can handle advanced format
  - http://www.sweclockers.com/nyhet/10733-storre_sektorer_ger_samre_prestanda_i_windows_xp
- HD firmware is capable of translating to old format for old OS
- New HD needed – factory prepared

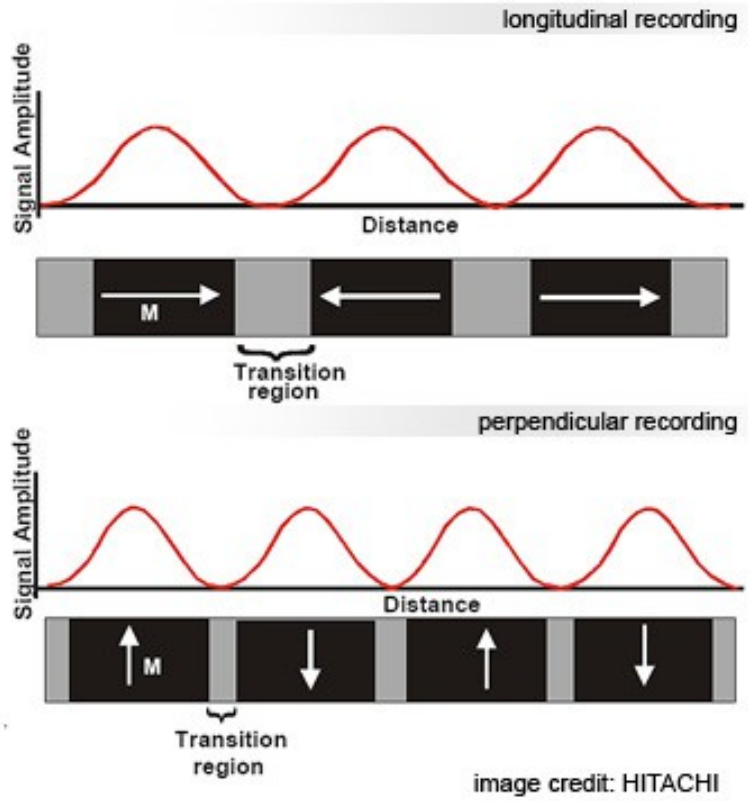


Perpendicular Magnetic Recording (PMR)

http://www.hitachigst.com/hdd/research/recording_head/pr/PerpendicularAnimation.html
http://en.wikipedia.org/wiki/Perpendicular_recording

- Traditionella hårddiskar lagrar data linjärt, i längdriktningen på skivans yta
- Med perpendikulär lagring sker processen vertikalt på skivan istället
- Perpendikulär lagring ökar datadensiteten och förväntas öka mängden tillgänglig lagringsyta med upp till 5-10 ggr. (2005 ->)





DOS/MBR Partitioner

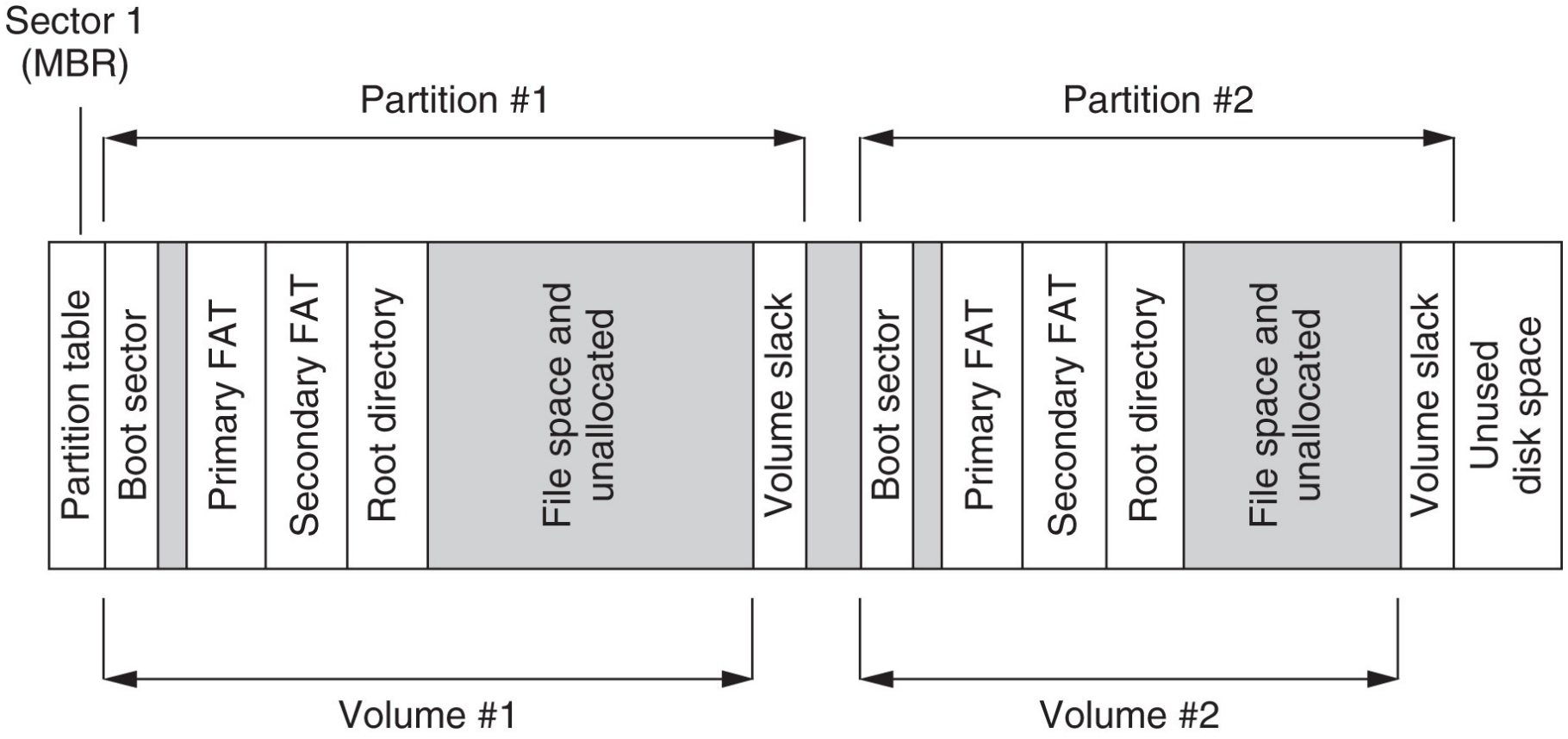


FIGURE 15.6 Simplified depiction of disk structure with two partitions, each containing a FAT formatted volume.

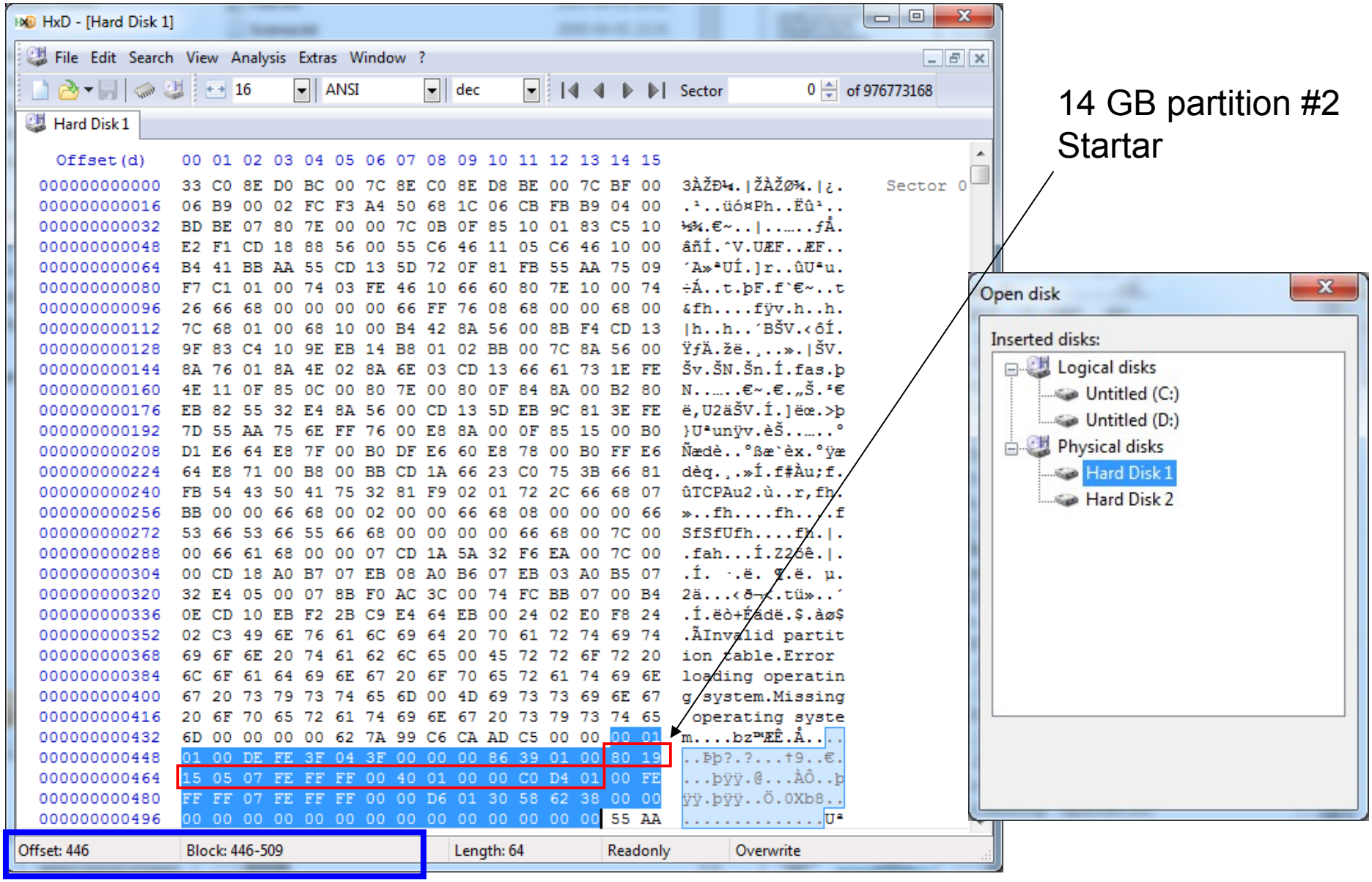
DOS/MBR Partitioner

- Partitionsinformationen är alltid lagrad på cylinder 0, head 0, sector 1 (sector 0 i vissa program) dvs. den första sektorn
- Master Boot Record (MBR) – de första 512 byten

Table 5.1. Data structures for the DOS partition table.

Byte Range	Description	Essential
0–445	Boot Code	No
446–461	Partition Table Entry #1 (see Table 5.2)	Yes
462–477	Partition Table Entry #2 (see Table 5.2)	Yes
478–493	Partition Table Entry #3 (see Table 5.2)	Yes
494–509	Partition Table Entry #4 (see Table 5.2)	Yes
510–511	Signature value (0xAA55)	No

MBR 446-509

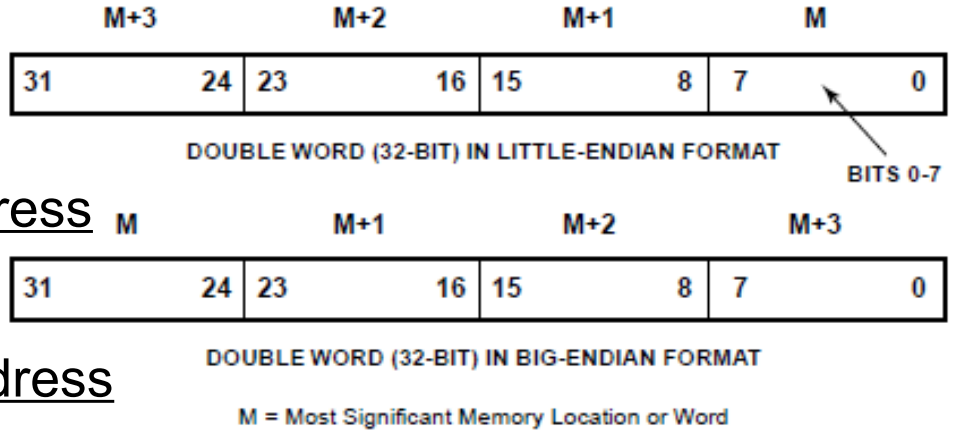


14 GB partition #2
Startar

Endianness

http://en.wikipedia.org/wiki/Endianness

- Byte och bit ordningen för att representera data
- Little endian
  - Intel x86, LSB at lowest address
- Big endian
  - Motorola, MSB at lowest address
- Bi endian
  - I princip alla andra processorarkitekturer - inställningsbar
  - ARM, MIPS, SPARC, IA64 etc.
- Word exempel
  - 0x39 LE
  - 0x93 BE



Word

Data	6	9	3	5
Address	0x1023	0x1024	0x1025	0x1026

DOS/MBR Partitioner

http://en.wikipedia.org/wiki/Master_boot_record

- Disk/partition parameters 446 - 509
- 4 primära partitioner kan skapas i PC datorer
  - VBR (Volume Boot Record) or Boot sector

Table 5.2. Data structure for DOS partition entries.

Byte Range	Description	Essential	
0–0	Bootable Flag	No	= 0x80 = boot, 00 no boot
1–3	Starting CHS Address	Yes	= 0x051519
4–4	Partition Type (see Table 5.3)	No	= 0x07
5–7	Ending CHS Address	Yes	= 0xFFFFFE
8–11	Starting LBA Address	Yes	= 0x00014000 = 81920
12–15	Size in Sectors	Yes	= 0x01D4C000 = 30720000

14 GB partition #2
i föregående slide

Tydning av ett 16 byte partition entry – läs little endian på rätt offset

14 GB =
30720000 * 512
/ 1024^3

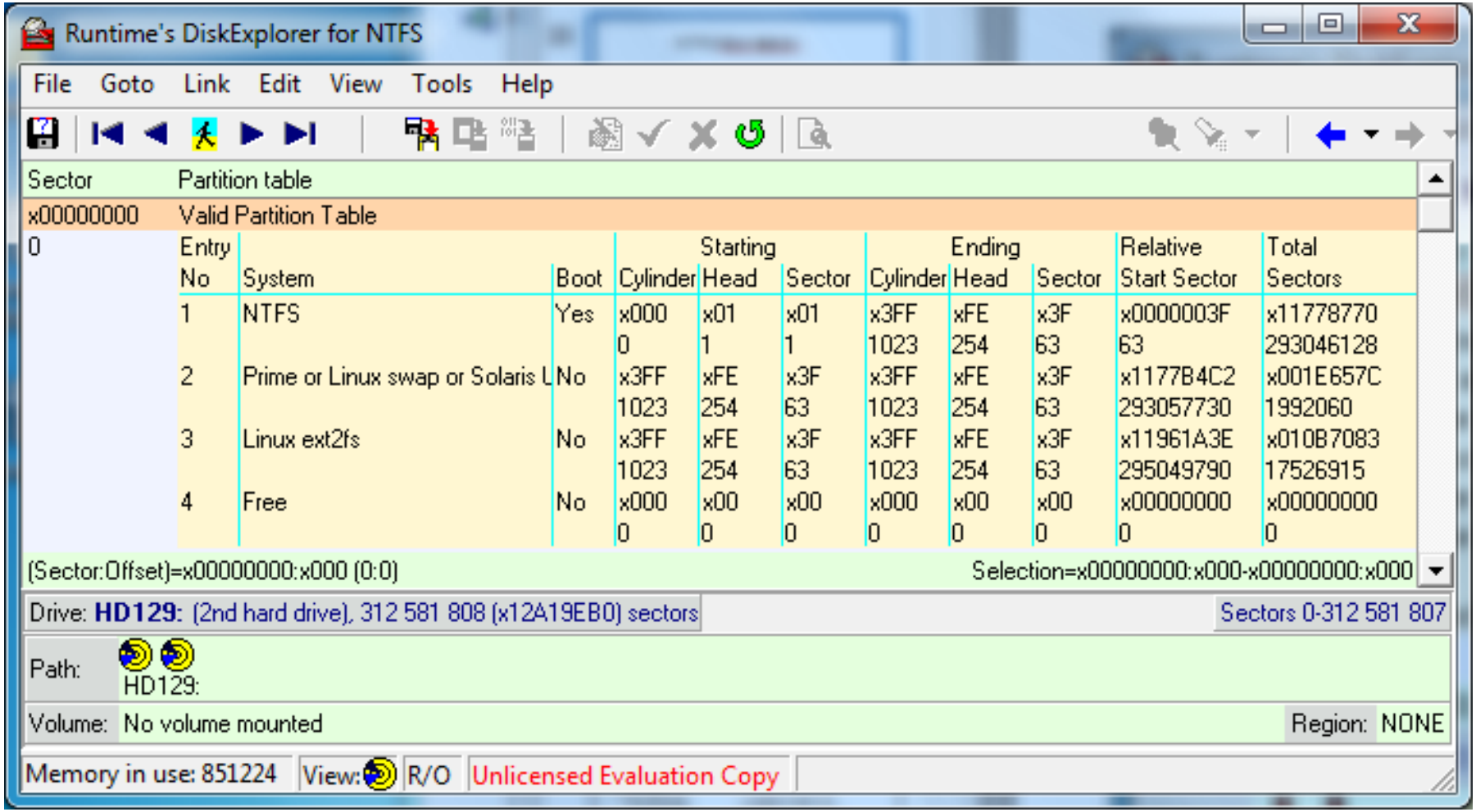
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
80	19	15	05	07	FE	FF	FF	00	40	01	00	00	C0	D4	01

Några DOS Partition typer

Table 5.3. Some of the type values for DOS partitions.

Type	Description
0x00	Empty
0x01	FAT12, CHS
0x04	FAT16, 16–32 MB, CHS
0x05	Microsoft Extended, CHS
0x06	FAT16, 32 MB–2GB, CHS
0x07	NTFS
0x0b	FAT32, CHS
0x0c	FAT32, LBA
0x0e	FAT16, 32 MB–2GB, LBA
0x0f	Microsoft Extended, LBA
0x11	Hidden FAT12, CHS
0x14	Hidden FAT16, 16–32 MB, CHS
0x16	Hidden FAT16, 32 MB–2GB, CHS
0x1b	Hidden FAT32, CHS
0x1c	Hidden FAT32, LBA

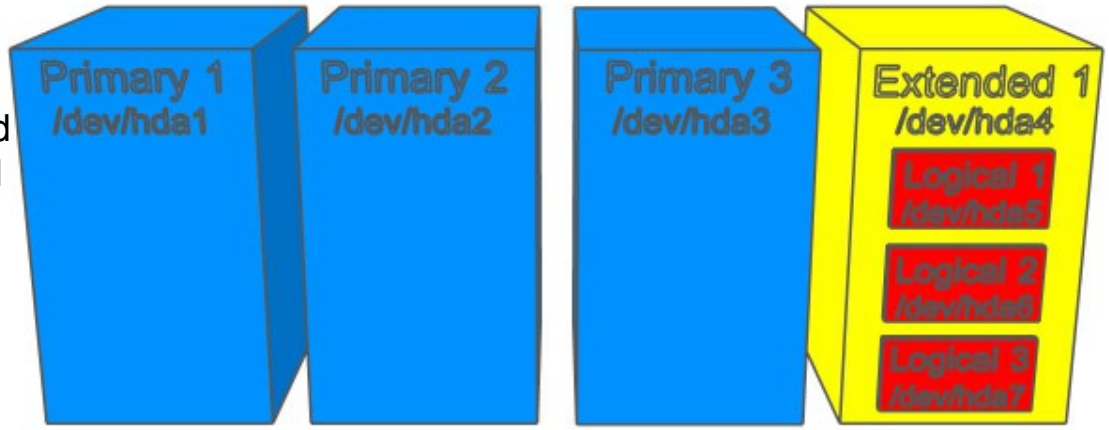
DiskExplorer MBR



MBR/DOS Partitionstyper primary, extended, logical

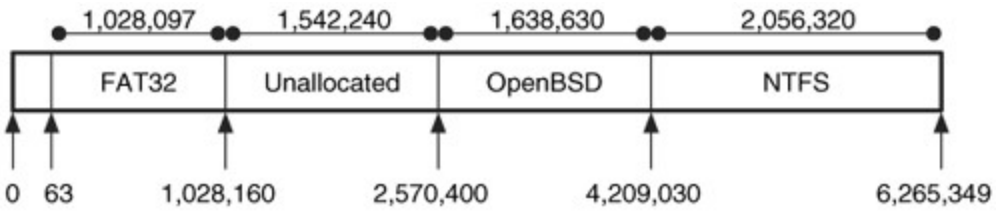
- Linux använder t.ex. 8 bitar för att adressera diskar (major och minor nummer för drivrutinen)
  - För IDE används 2 bitar till primära adresseringen och 6 bitar för logiska adresseringen = 63 logiska partitioner per disk
  - För SCSI används 4 bitar till primära adresseringen (SCSI kan ha 16 enheter) och 4 bitar för logiska = 15 logiska partitioner per disk

Utdrag av partitionstyper, siffran till vänster är Hex
0x01 DOS 12-bit fat
…
0x07 Windows NT NTFS
0x0b WIN95 OSR2 32-bit FAT
0x0c WIN95 OSR2 32-bit FAT, LBA-mapped
0x0e WIN95: DOS 16-bit FAT, LBA-mapped
…
0x82 Solaris x86
0x82 Linux swap
0x83 Linux native (usually ext2fs)
0x84 OS/2 hidden C: drive
0x84 Hibernation partition
0x85 Linux extended partition
…
Osv -> 0xFF



Partitioner och verktyg

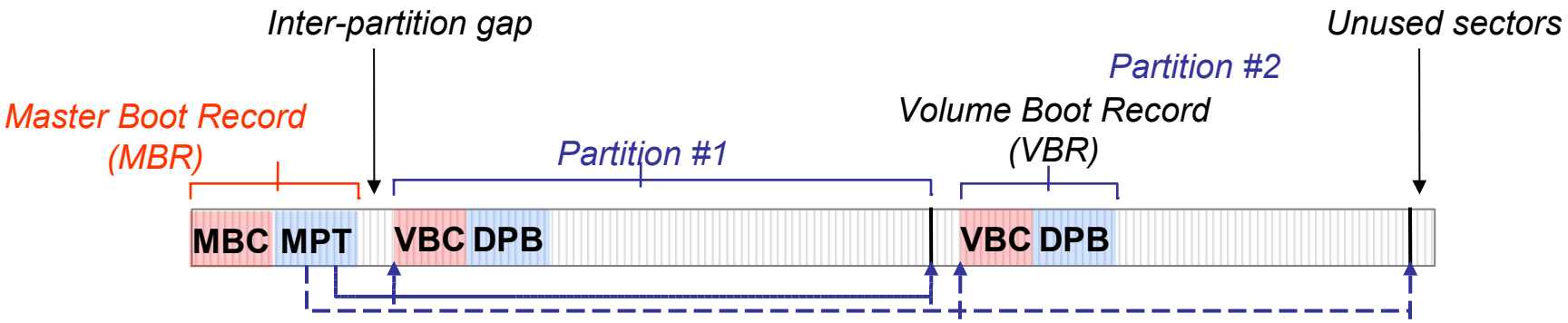
- Logiska volymer liknar en länkad lista
- 63 sektorer allokeras typiskt för MBR iom. att en partition måste starta på en cylinder gräns
  - I verkligheten är det *många/variabelt* fler sektorer per track hårddisken gör egna beräkningar...



- Fdisk, Cfdisk
- MS ResKit – sector inspect
- Partition Magic – numera Symantec/Norton
- Acronis Disk Director – www.acronis.com
- GParted LiveCD - http://gparted.sourceforge.net
- Lista på fria partitions editorer
  - http://www.thefreecountry.com/utilities/partitioneditors.shtml

Partitioner – VBR (Boot sector)

- The Master Boot Record (MBR) is created and includes the Master Boot Code (MBC) and the Master Partition Table (MPT) – always at sector 1 on any bootable media
- The MBC is executed at boot if the HDD is designated as the boot device
- The MPT contains information about logical volumes (partitions), including the active partition, the partition whose Volume Boot Code (VBC) will be executed
- Each partition has a Disk Parameter Block (DPB) that stores information about extended partitions, file system type, date and time last mounted, etc.
- Inter-partition gaps are a collection of unused sectors
- Some sectors are unused due to addressing issues



On each partition a VBR contains Volume Boot Code and a Disk Parameter Block

Removable media

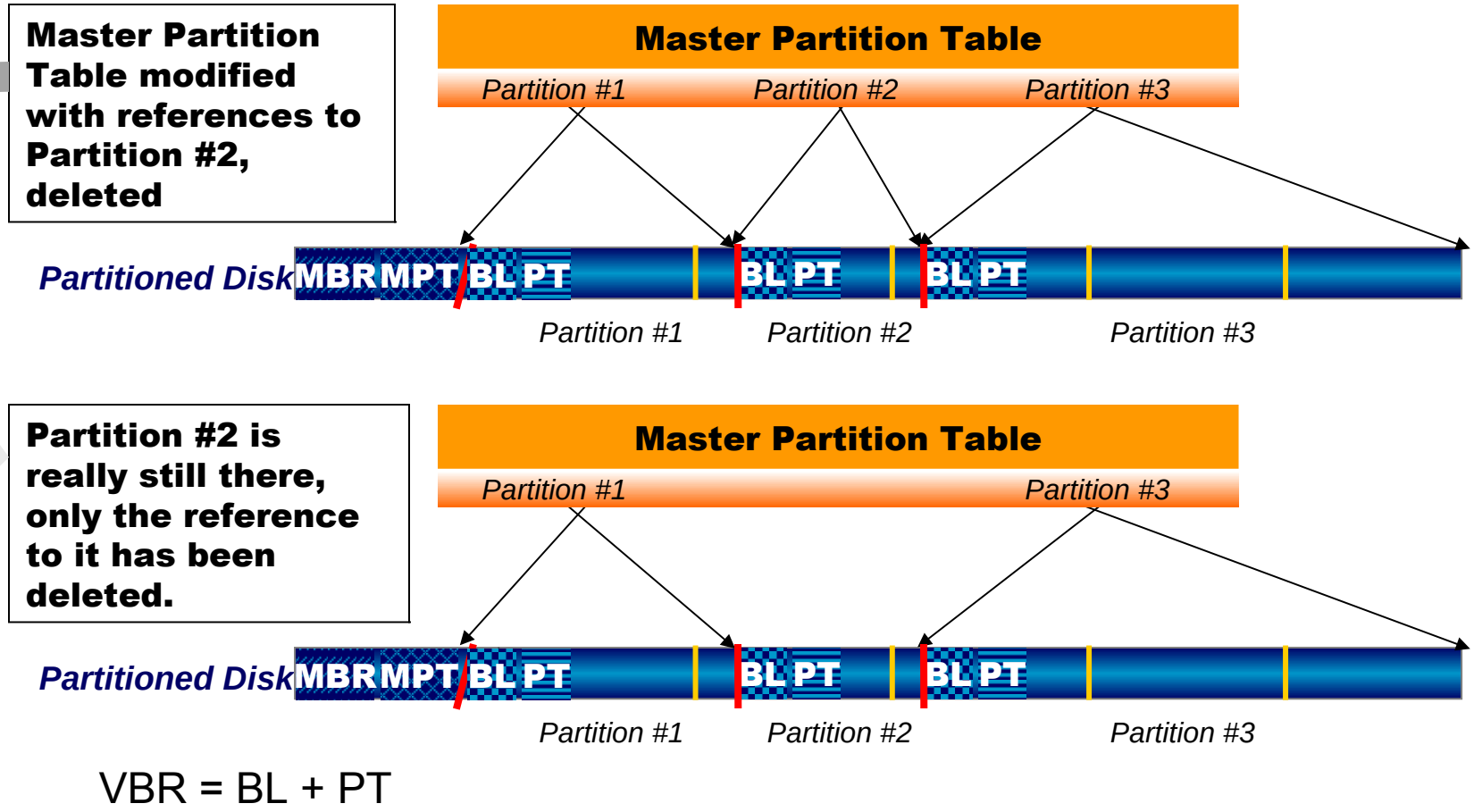
- De flesta lösa media har partitioner (som hårddiskar)
  - USB minnen, kamera minnen, ZIP diskar etc.
  - Undantaget är floppy som bara kan ha en partition, har en VBR (om man skall vara korrekt)
- CD/DVD är komplext
  - Många variationer
  - Hybrider (OS)
  - Sessioner (data)

Formatering av diskar

http://support.microsoft.com/?kbid=302686

- Quick format (Windows)
  - Tar bara bort FAT/MFT etc.
  - Allt annat är kvar på disken
- Regular format (Windows)
  - Som quick format men scannar även efter bad sectors
- ”Low level format” utility från disktillverkare
  - Gör write-read verify och andra kontroller av disken
- ”Low level format” från fabrik
  - Servo, sector layout, defect management etc.
  - Håller för diskens livstid, kan inte göras om
- Återställa MBR – odokumenterat kommando
  - FDISK /MBR (DOS/Windows)
- VISTA/7 från Windows RE (Recovery Environment)
  - bootrec /FixMbr och bootrec /FixBoot

Raderade Partitioner

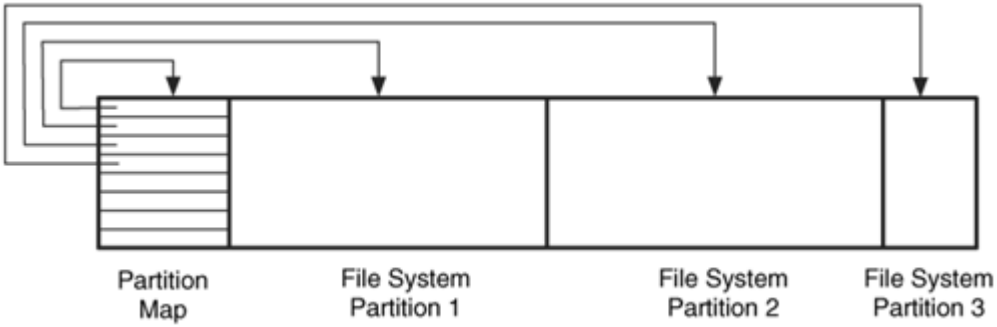


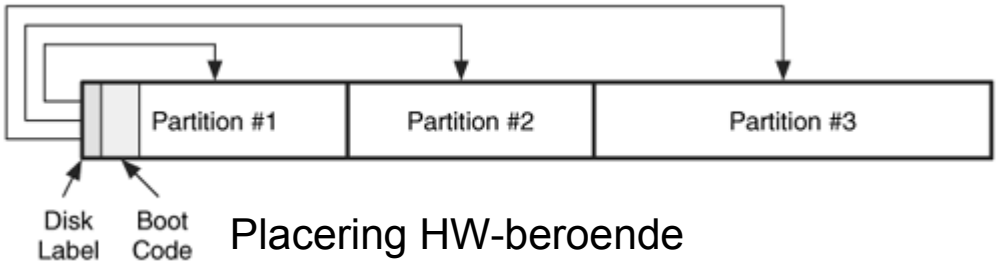
VBR = BL + PT

Forensic Analysis: Look for inconsistencies between the end of one partition, and the start of another

Andra partitioner

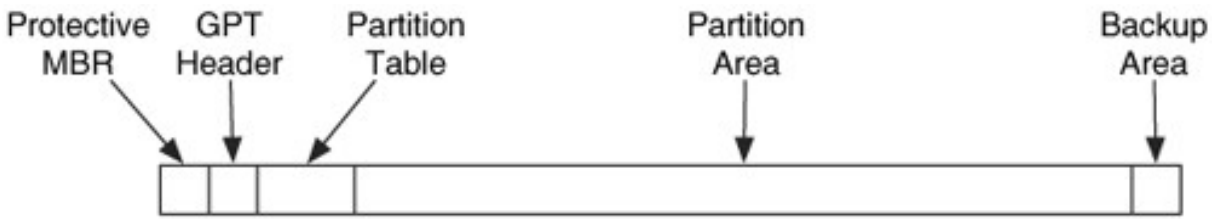
- PC baserade
  - Apple partitioner
  - Startar på andra sektorn
  - Obegränsat antal
- Server baserade
  - BSD (som DOS)
  - SUN Solaris (slice)





Placering HW-beroende

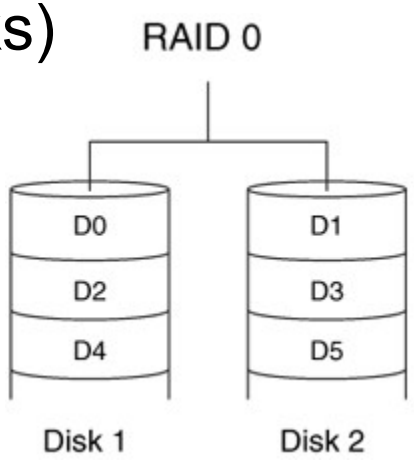
  - IA64 GUID Partition Table (GPT)
  - DOS MBR för legacy – riktiga infon i areorna efter



Avancerade Disksystem

http://en.wikipedia.org/wiki/RAID

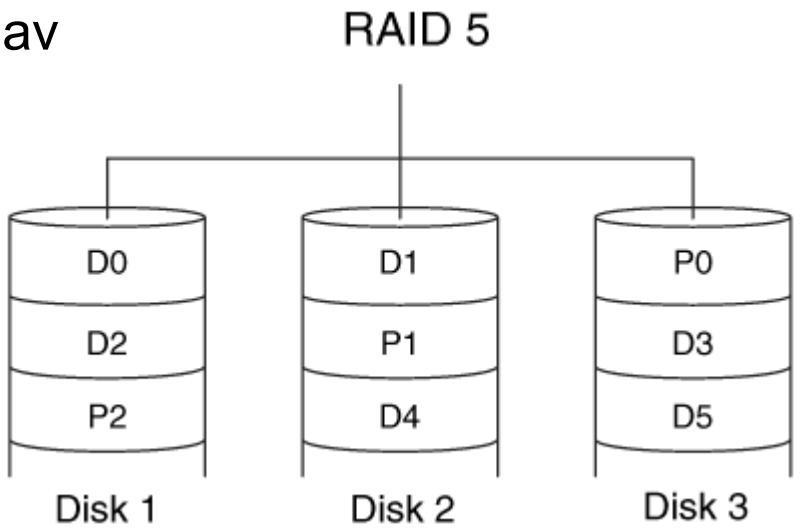
- RAID (Redundant Array of Independent Disks)
  - I hårdvara eller mjukvara
  - Stripe width/block >= sektor size
  - 0 - striping i block med viss storlek
  - 1 - mirror i block med viss storlek
  - 2 - är mycket ovanligt
  - 3 - kräver minst 3 diskar, som 0 men med dedikerad paritetsdisk och stripe storlek i byte
  - 4 - som 3 men med stripe storlek i block
  - 5 - som 4 men utan dedikerad paritetsdisk
  - 6 - som 5 men kräver 4 diskar, klarar 2 felaktiga diskar
  - 10 - kombinerar RAID 1 och 0



RAID 5 och 6

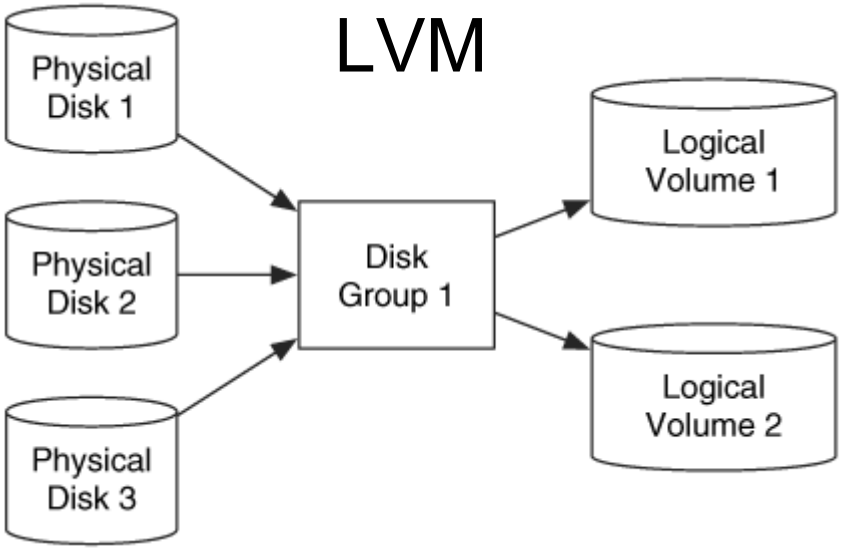
- Ingen dedikerad paritetsdisk (distribuerad och dual distribuerad)
- Alla diskar innehåller data och paritet alternerat
  - Ökar prestanda?
- RAID 5/6 är den mest vanliga formen på servrar
- Pariteten tillför redundans
  - Oftast används XOR (exclusive or), producerar SANT om endast en av operanderna är sann

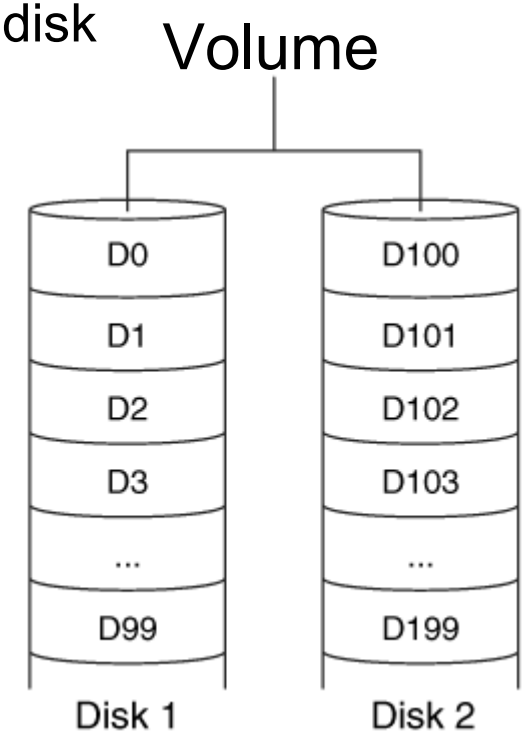
Input 1	Input 2	Output
0	0	0
0	1	1
1	0	1
1	1	0



Disk spanning

- De flesta RAID lösningar kan fixa detta
  - Ger en stor volym utan redundans eller prestandavinster
  - Vissa lösningar medger dynamisk hantering
  - Linux MD (linear RAID) och LVM (Logical Volume Manager)
  - LDM (Logical Disk Manager) är ansvarigt för att hantera logiska volymer i Windows, använder dynamic disk

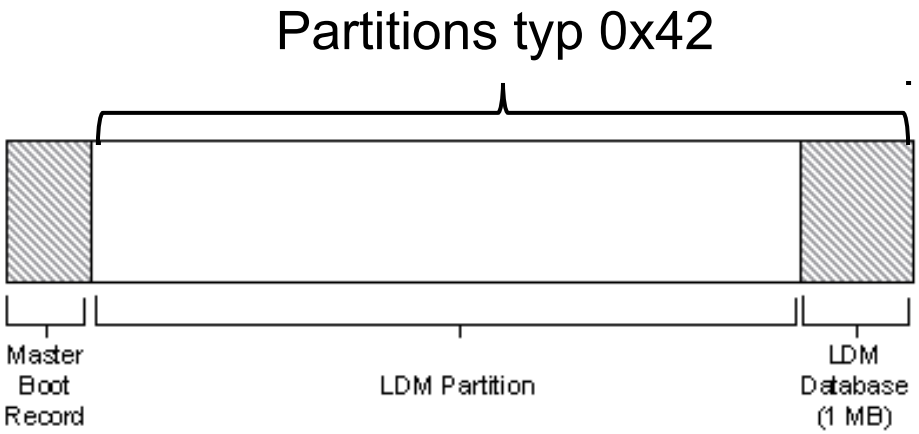




Basic vs. dynamisk disks (Windows)

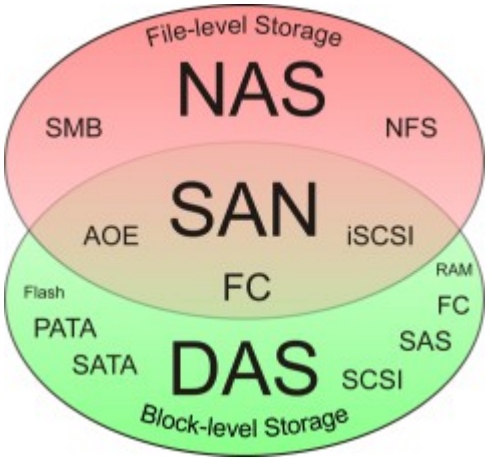
http://msdn.microsoft.com/en-us/library/aa363785%28VS.85%29.aspx

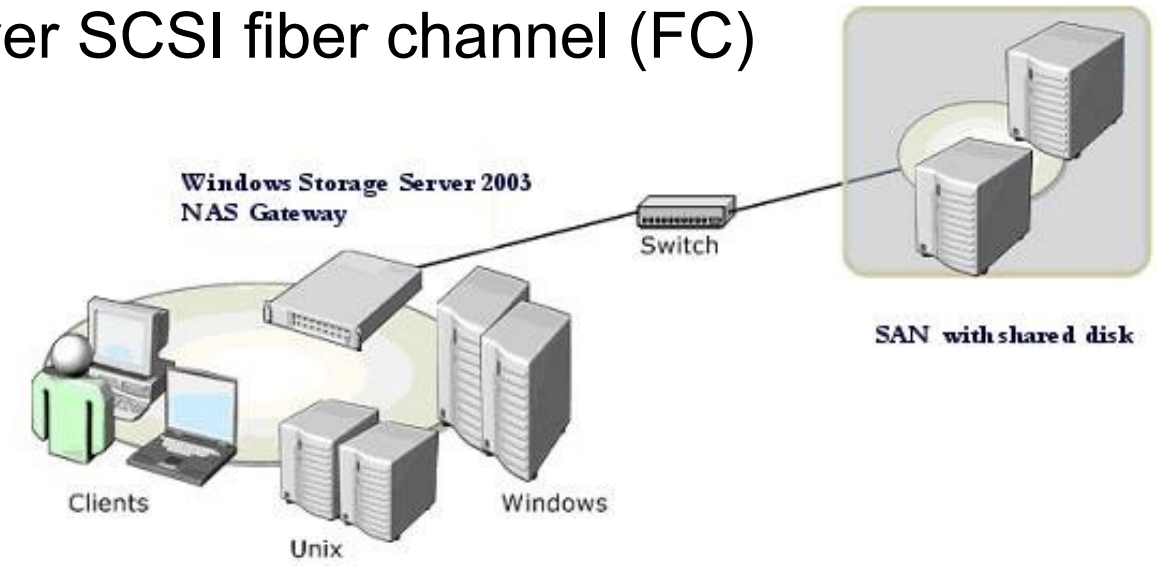
- Basic disk (upp till 4 partitioner – DOS/GPT)
  - En partition, markerad aktive som innehåller OS boot kod
- Diskar kan konverteras till dynamiska diskar
  - Envägs process
  - Ger tillgång till RAID-0, 1 och 5 på servrar
  - Har inte en normal partitionstabell (finns bara ett entry för kompabilitet) – därmed ej bootbara
  - Använder en 1 MB LDM (Logical Disk Manager) databas vid slutet på varje volym för att lagra konfigurationsinformation (redundant distribuerat)
- Hanteras via Disk Management lövet i Computer Management programvaran



NAS (Network Attached Storage) vs. SAN (Storage Area Network)

- NAS är en enhet (RAID Disk Array) som fungerar som lagringsenhet när den kopplas till ett nätverk
  - En typisk NAS kan vara en enkel dator med ett antal hårddiskar ofta i RAID och ett eller flera nätverkskort. För kommunikation med enheten eller enheterna används typiskt TCP/IP och nätverksprotokollen AppleTalk, SMB/CIFS, NFS eller FTP/HTTP
- SAN är i princip en NAS men adresserar och kommunicerar istället disk block över SCSI fiber channel (FC)





Partitions sammanfattning

- Partition Magic och andra disk verktyg kan dölja information för OS:et
- Att förstå partitioneringen av diskar är viktigt om man skall hitta gömd information
- Partitionsinformationen är lagrad på Cylinder ?, Huvud ?, Sektor ?
  - I MBR arean som är de ??? första bytes på disken
  - Information om upp till ? primära partitioner lagras här
  - Kunna räkna ut partitionens lagringsstorlek och LBA sektoradressen var den börjar
- Med t.ex. Acronis Disk Director eller annat liknande program kan man enkelt lära sig mer om partitioner