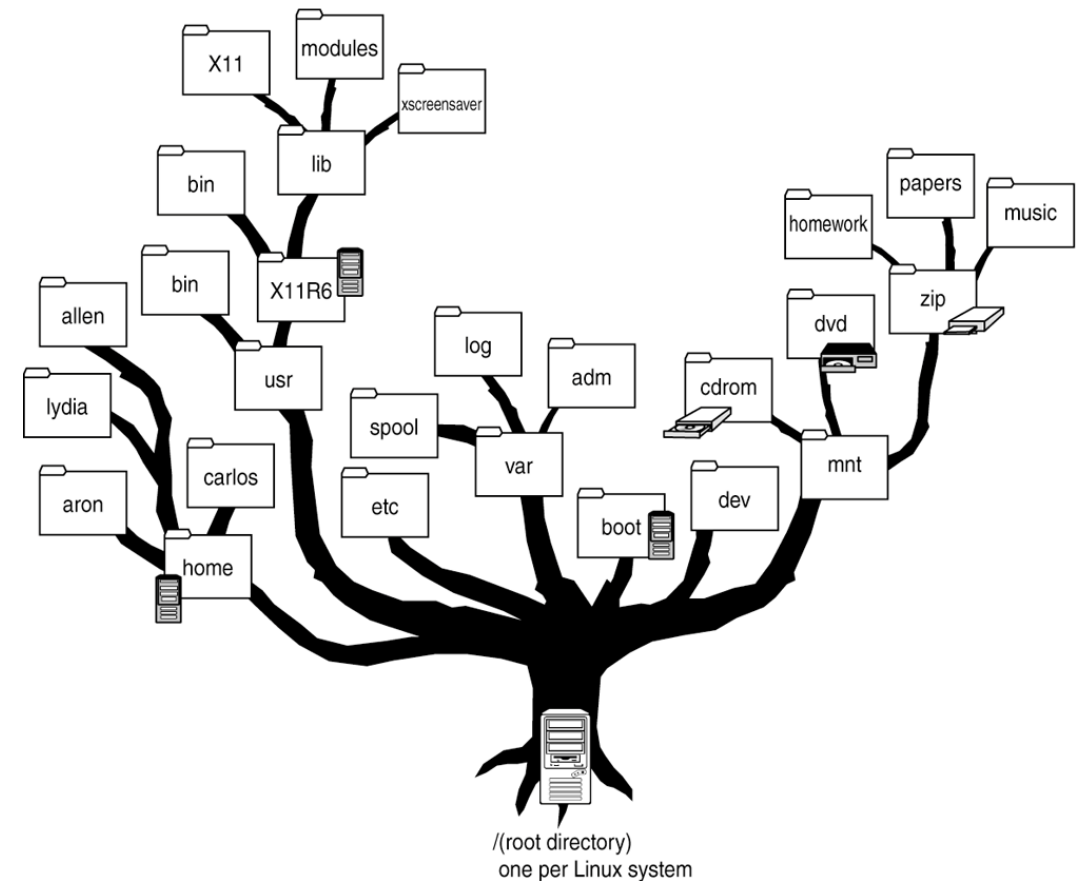# File Systems

File System

- A file system or filesystem (often abbreviated to fs) is a method and data structure that the operating system uses to control how data is stored and retrieved.

Data Storage forms

- **Primary storage** (also known as main memory, internal memory, or prime memory), often referred to simply as memory, is the only one directly accessible to the CPU.

- **Secondary storage** (also known as external memory or auxiliary storage) differs from primary storage in that it is not directly accessible by the CPU.

- **Tertiary storage** or tertiary memory is a level below secondary storage. Typically, it involves a robotic mechanism which will mount (insert) and dismount removable mass storage media into a storage device according to the system's demands; such data are often copied to secondary storage before use.
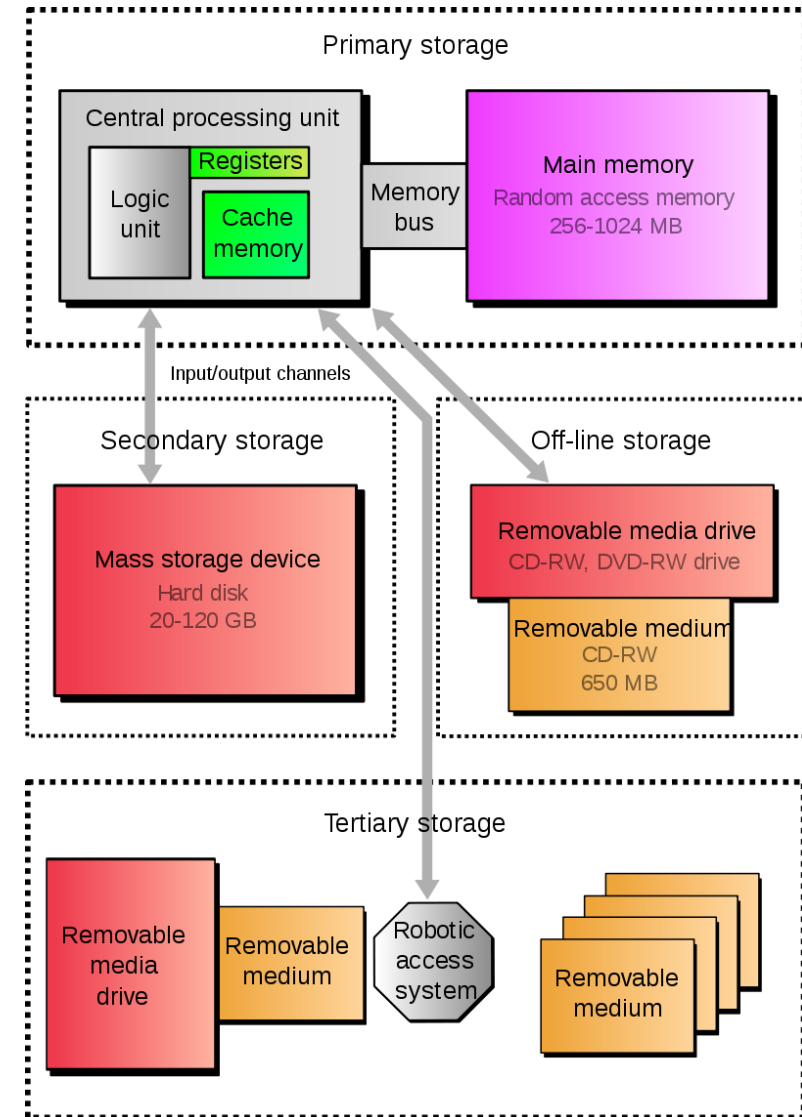
Hard disk drive

- A hard disk drive (HDD) is an electro-mechanical data storage device that stores and retrieves digital data using magnetic storage with one or more rigid rapidly rotating platters coated with magnetic material.

- Introduced by IBM in 1956.

HDD History Examples

- IBM 350 (First Hard disk)
  - Introduced in 1956.
  - Size: two large refrigerators.
  - Capacity: 3.75 Mbyte.
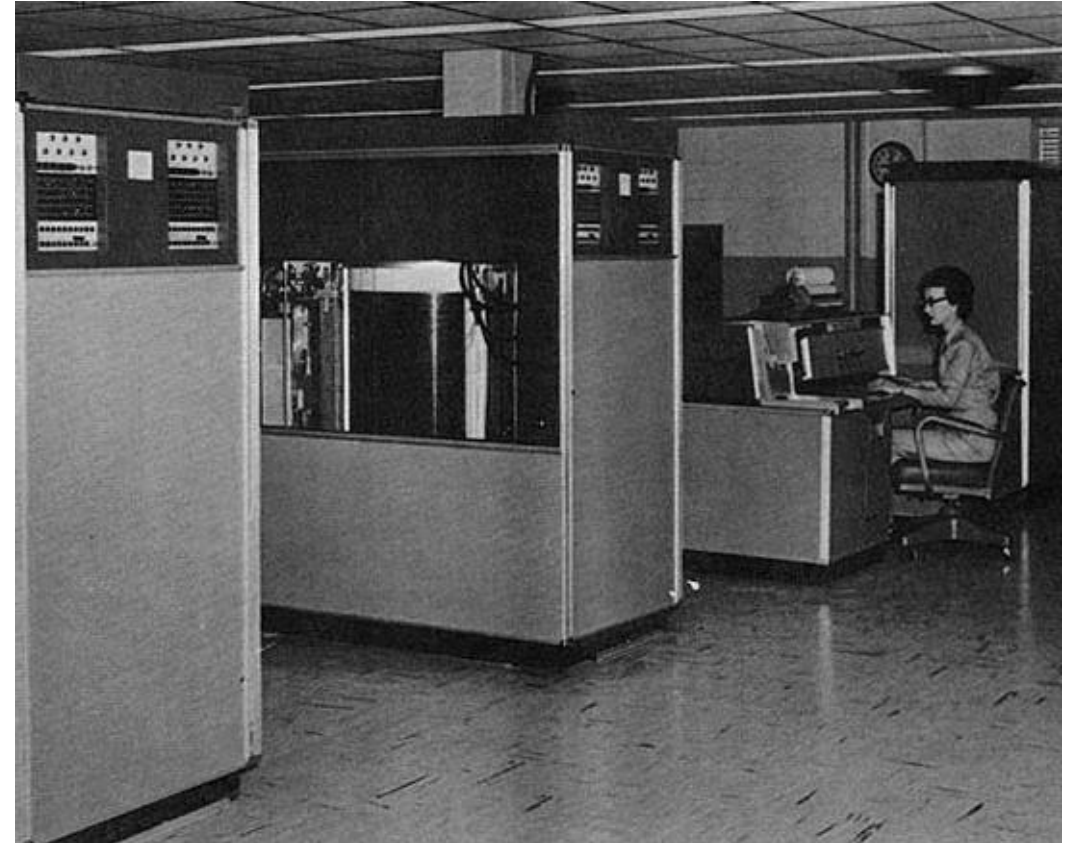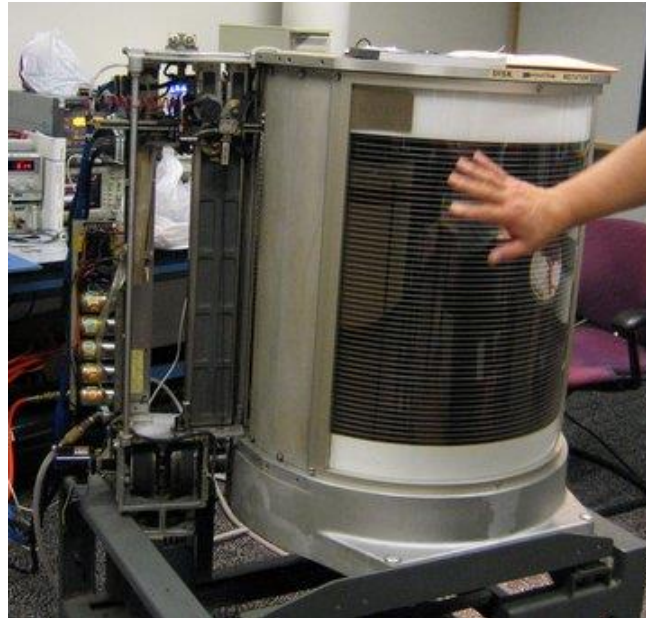
HDD History Examples

- IBM 1311
  - Introduced in 1962.
  - Size: washing machine.
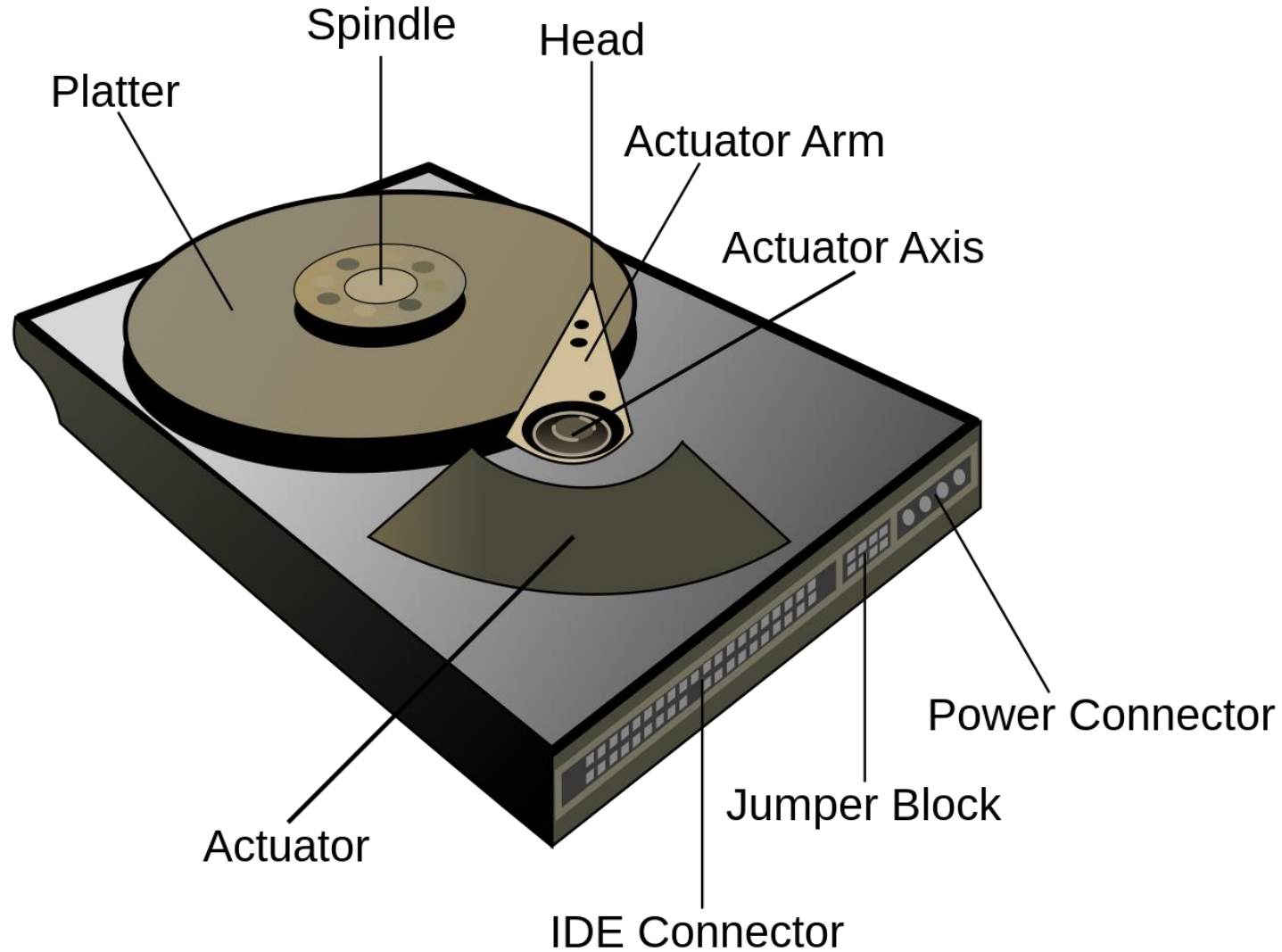 - Capacity: 1.43 Mbyte.
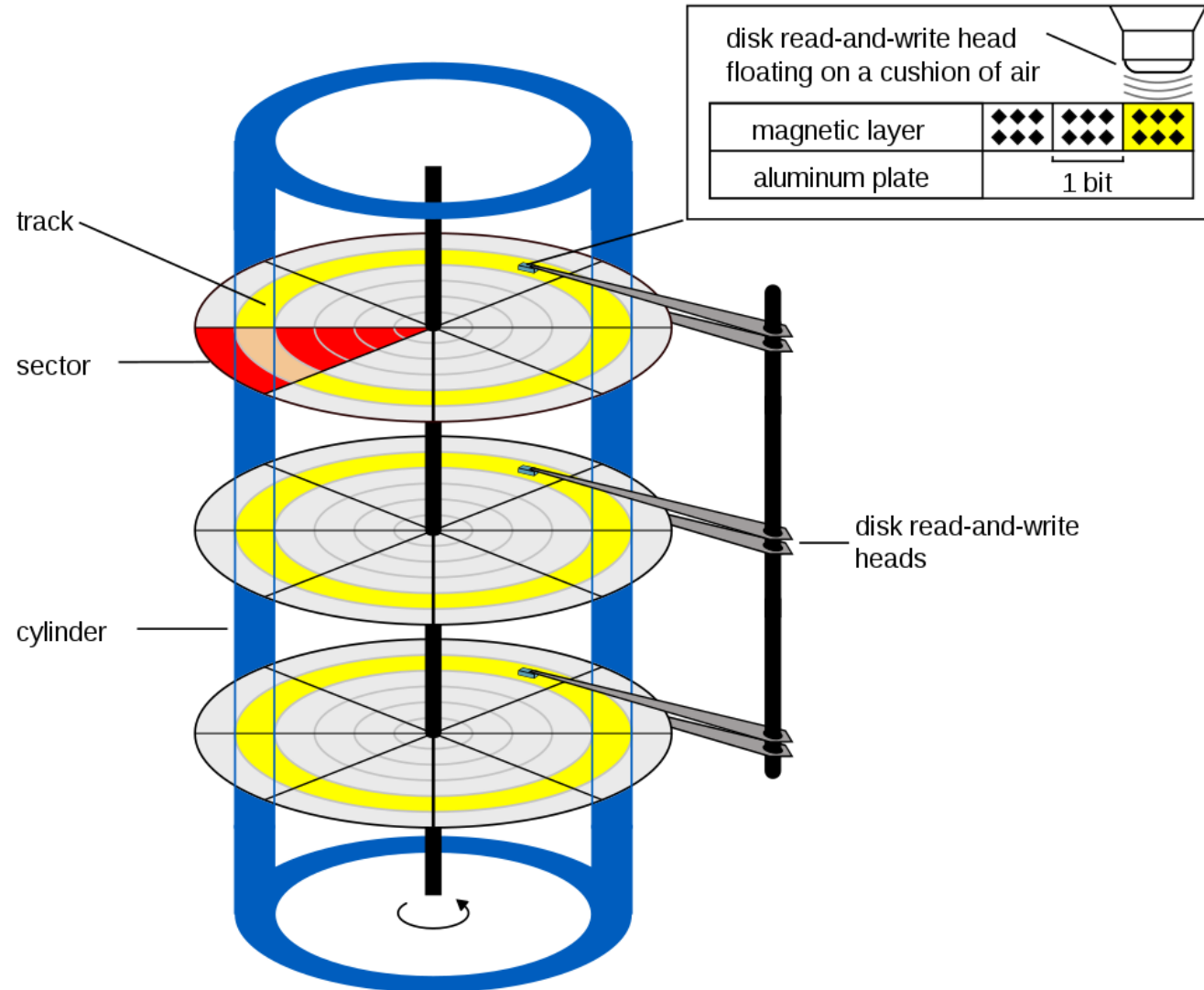
HDD Improvements over years

Improvement of HDD characteristics over time

Parameter	Started with (1957)	Improved to	Improvement
Capacity (formatted)	3.75 megabytes[18]	22 terabytes (as of 2023)[19]	5.86-million-to-one[c]
Physical volume	68 cubic feet (1.9 m³)[d][6]	2.1 cubic inches (34 cm³)[20][e]	56,000-to-one[f]
Weight	2,000 pounds (910 kg)[6]	2.2 ounces (62 g)[20]	15,000-to-one[g]
Average access time	approx. 600 milliseconds[6]	2.5 ms to 10 ms; RW RAM dependent	about 200-to-one[h]
Price	US$9,200 per megabyte (1961; US$83,107 in 2021)[21]	US$0.024 per gigabyte by 2020[22][i][23]	3.46-billion-to-one[j]
Data density	2,000 bits per square inch[24]	1.3 terabits per square inch in 2015[25]	650-million-to-one[k]
Average lifespan	c. 2000 hrs MTBF[citation needed]	c. 2,500,000 hrs (~285 years) MTBF[26]	1250-to-one[l]

Components of HDD

Spindle

Head

Platter

Actuator Arm

Actuator Axis

Power Connector

Actuator

Jumper Block

IDE Connector

Components of HDD

disk read-and-write head
floating on a cushion of air

magnetic layer			
aluminum plate		1 bit	

track

sector

cylinder

disk read-and-write
heads
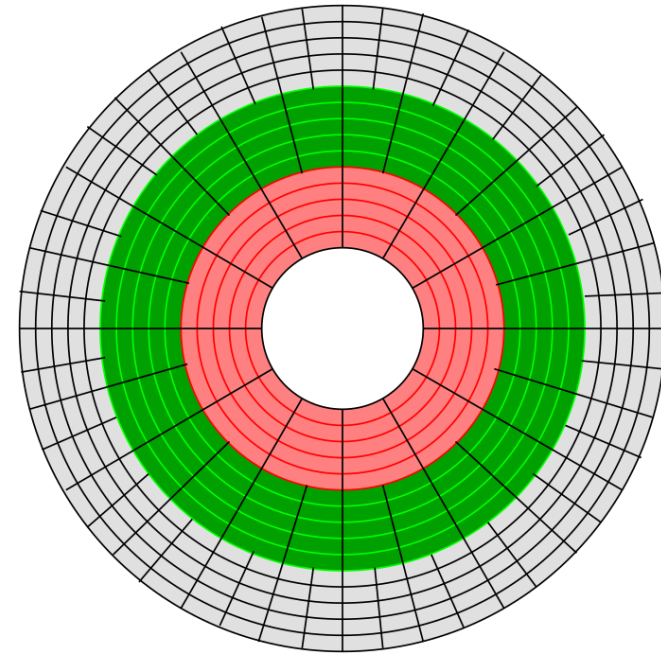
Components of HDD

- Zone bit recording (ZBR) is a method used by disk drives to optimize the tracks for increased data capacity.

- Zone recording was pioneered and patented by *Chuck Peddle* in 1961 while working for General Electric.

HDD Addressing Schemes

- CHS (Cylinder – Head - Sector)
  - Limited by:
    - 512 bytes / sector.
    - 63 sectors / track.
    - 255 heads (tracks/Cylinder).
    - 1024 Cylinders.
  - Maximum disk size: 8064 MB.

- LBA (Logical Block Addressing)
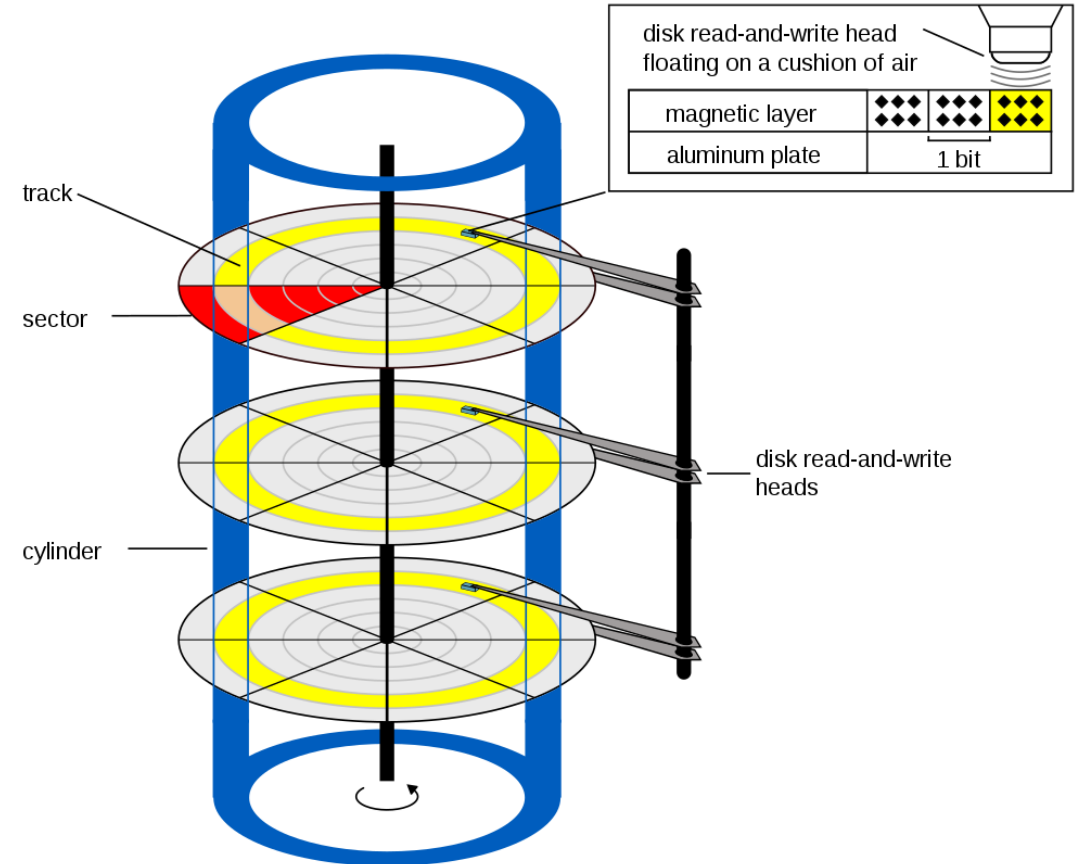
SSD

- A solid-state drive (SSD)
  - is a solid-state storage device that uses integrated circuit assemblies to store data persistently, typically using flash memory, and functioning as secondary storage in the hierarchy of computer storage.
- Outperforms the HDD in:
  - Speed.
  - Noise.
  - Capacity.
  - Durability.
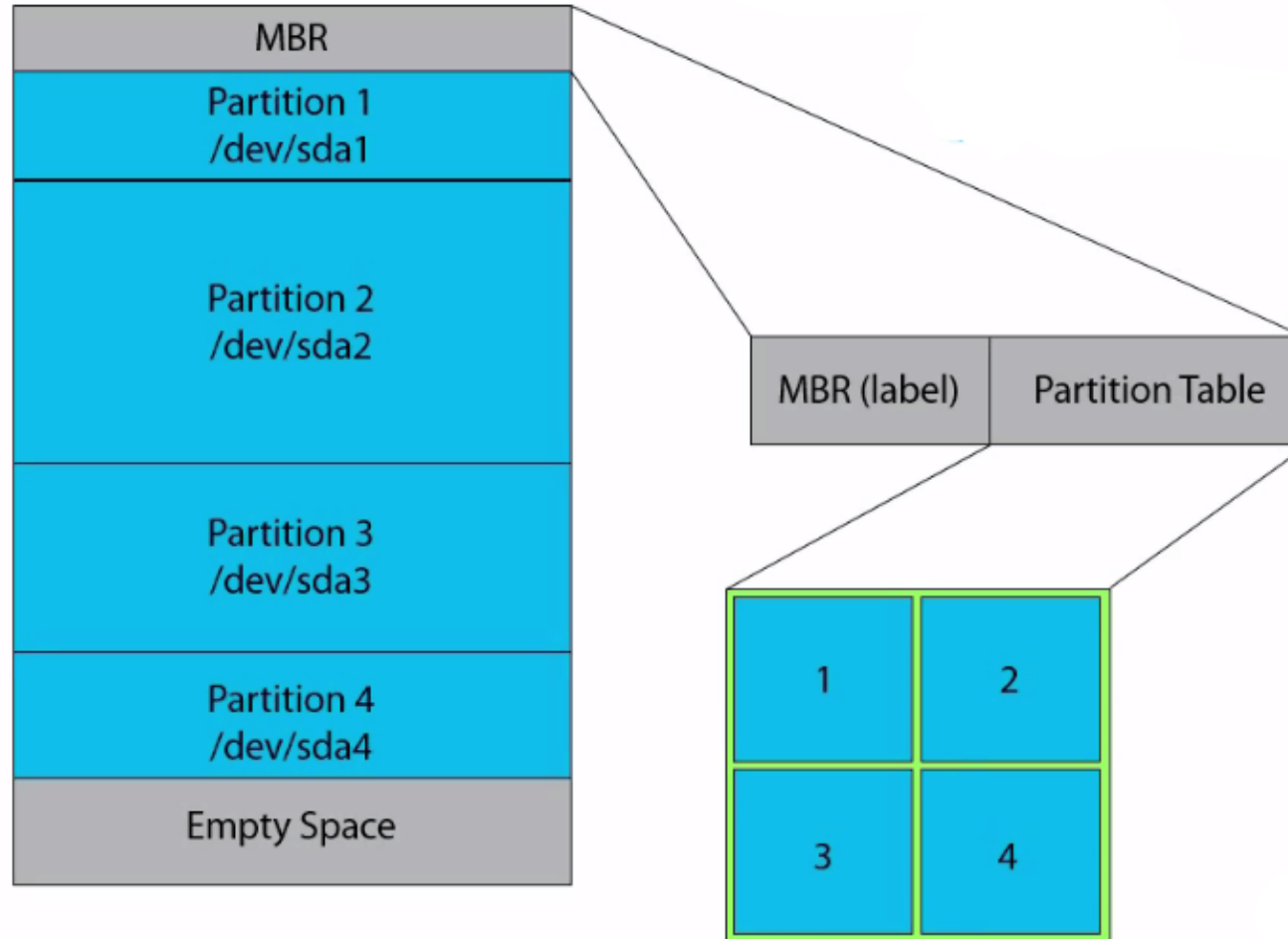 - Size.
  - Fragmentation.
  - Price.

Disk Partitioning

- **Disk partitioning** is the creation of one or more regions on secondary storage, so that each region can be managed separately. These regions are called partitions.

- Disk Partitioning reasons:
  - Organize data.
  - Security.
  - Avoid crossing size limits.
  - Implement quotes.
  - Supporting multiple filesystems.

- Partition Types:
  - Data Partitions.
  - Swap Partitions.

Disk Partitioning

- **Partition Table:** is a table maintained on a disk by the operating system that outlines and describes the partitions on that disk.

- Partition Table Types:
  - MBR (Master Boot Record).
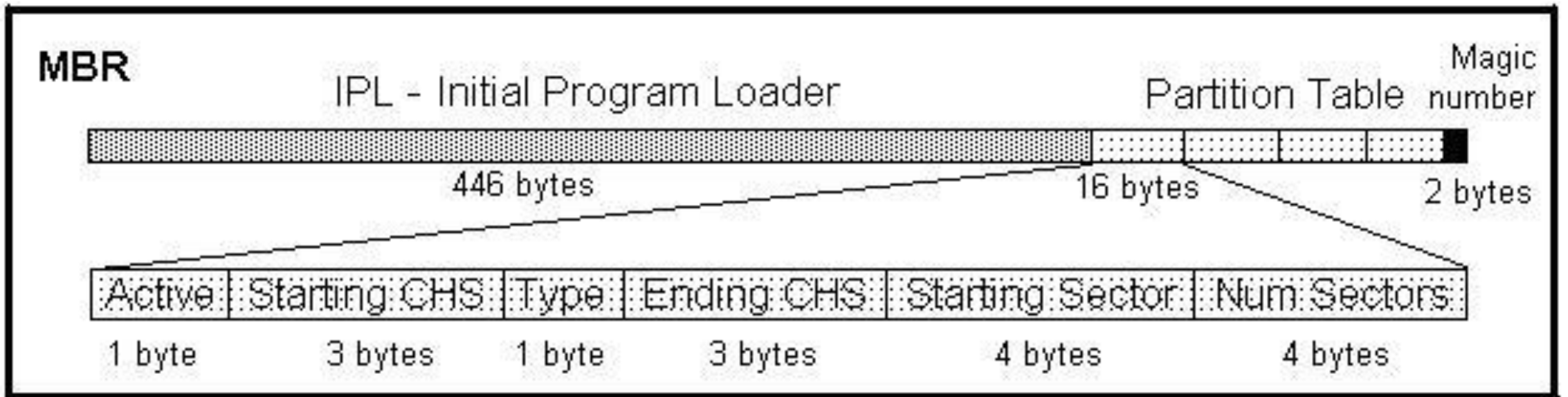 - GPT (GUID (Globally Unique IDentifier) Partition Table).

Master Boot Record

Master Boot Record

Master Boot Record

MBR with an Extended Partition

Boot code

1
2 } Partition table
3
4

Partitions within an extended partition

Boot partition

Partition 1   Partition 2   Partition 3 (Extended)   Partition 4
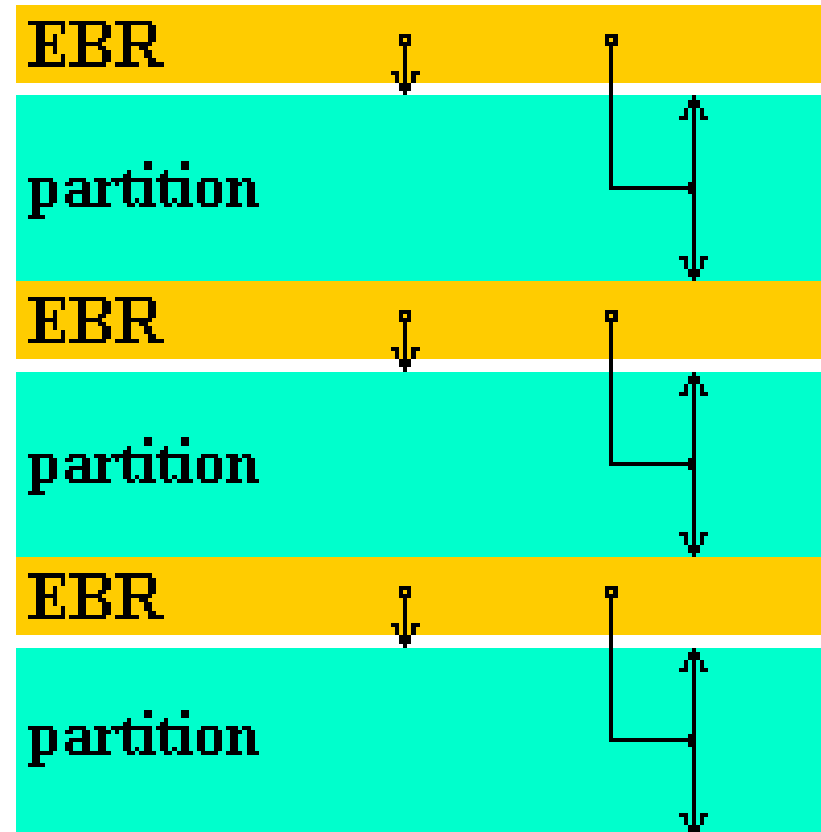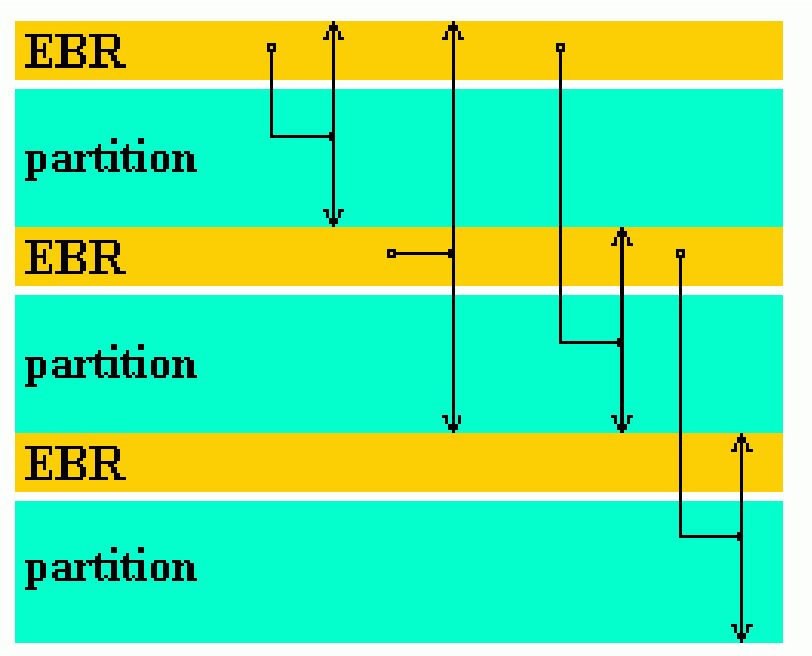
MBR   Boot sector   Extended partition boot record
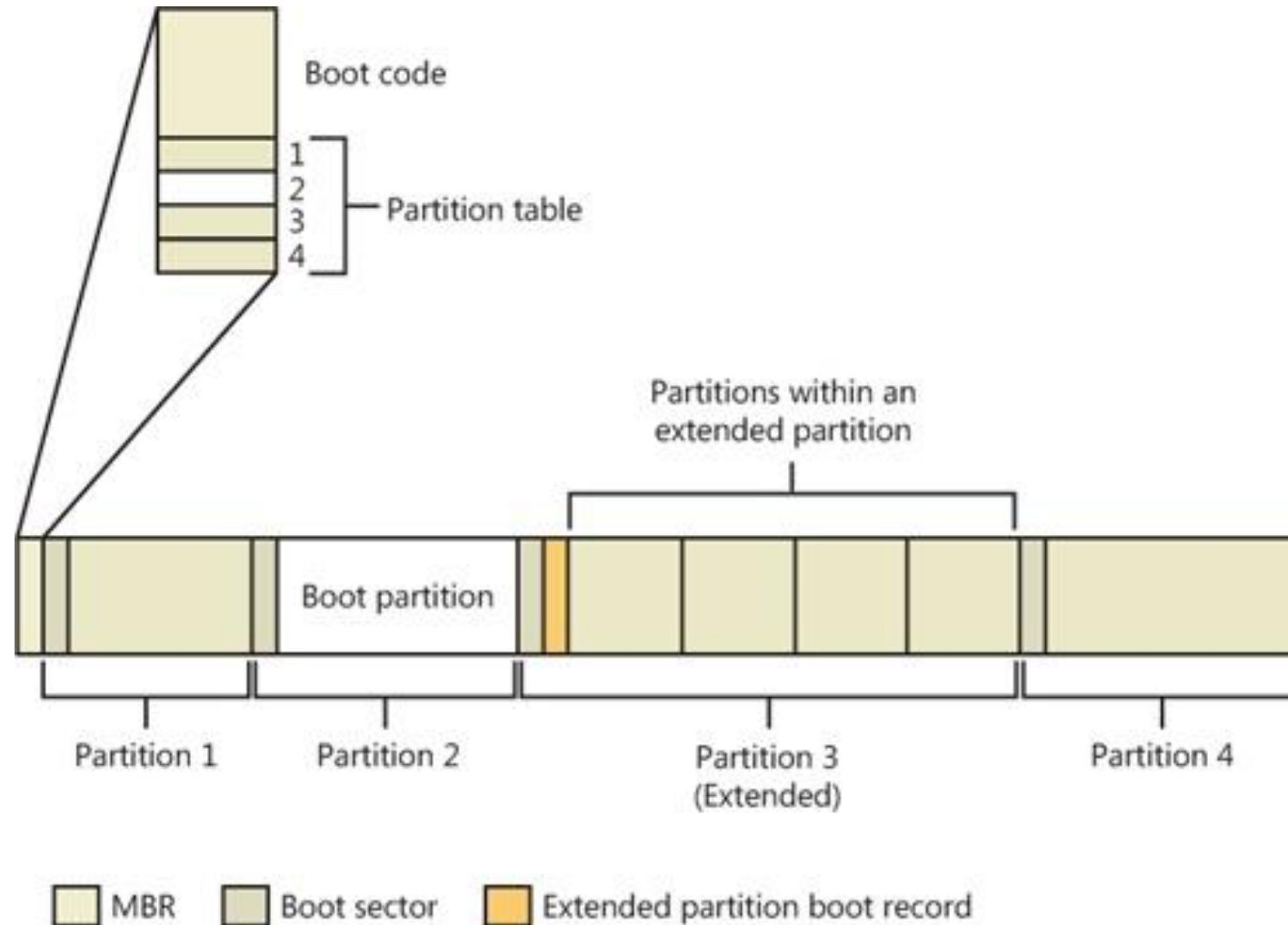
Extended Boot Record (First entry)

Extended Boot Record (Second entry)

GUID Partition Table (GPT)

- Up to 128 partition.

- Partition size: up to 9.44 ZiB (10^9 TiB).

- GPT table is repeated at the end of the hard disk.

GUID Partition Table Scheme

LBA 0	Protective MBR
LBA 1	Primary GPT Header
LBA 2	Entry 1 | Entry 2 | Entry 3 | Entry 4
LBA 33	Entries 5–128
LBA 34	Partition 1
	Partition 2
	Remaining Partitions
LBA −34	
LBA −33	Entry 1 | Entry 2 | Entry 3 | Entry 4
LBA −2	Entries 5–128
LBA −1	Secondary GPT Header

Primary GPT

Secondary GPT
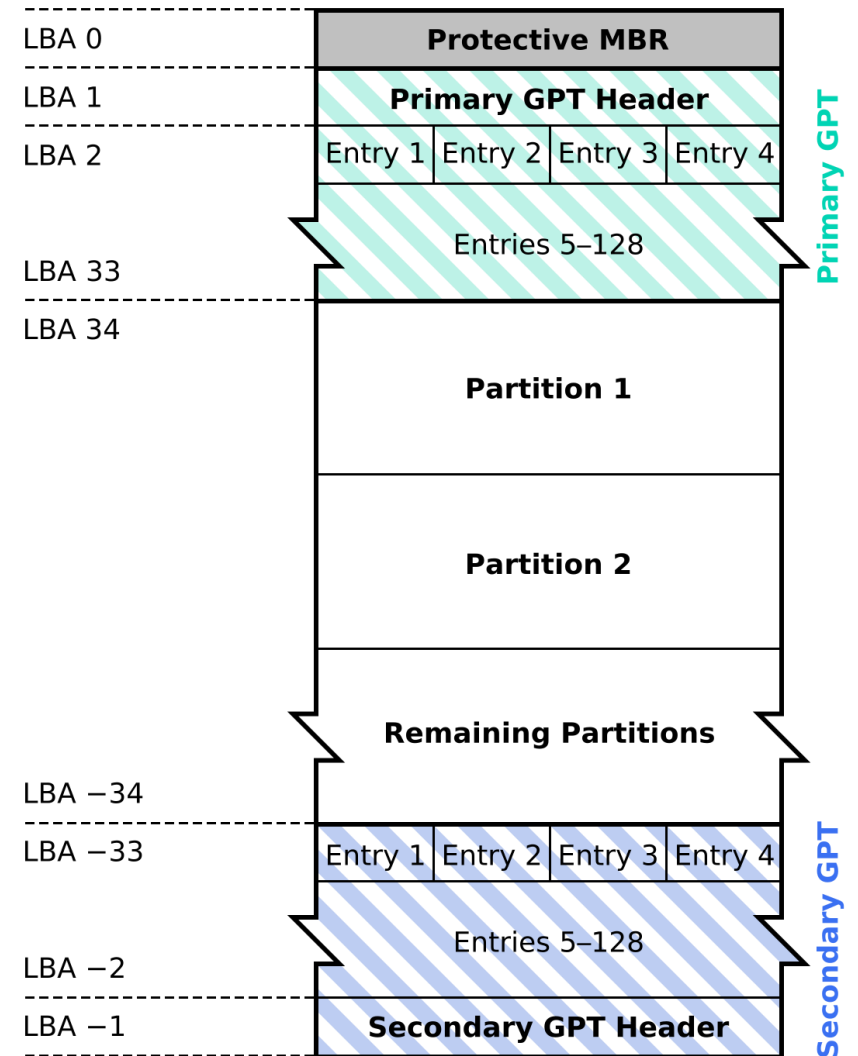
GPT Header

GPT header format

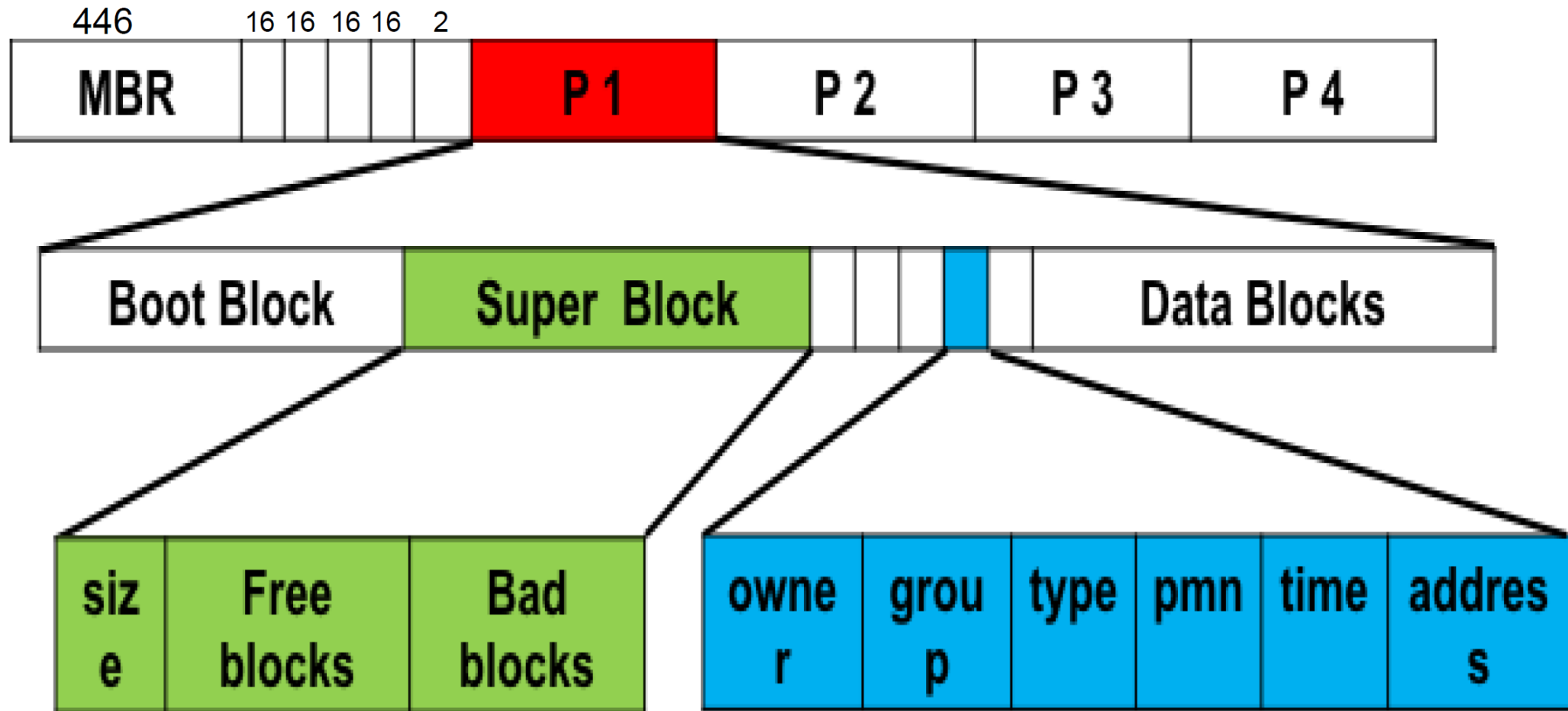| Offset	Length	Contents
0 (0x00)	8 bytes	Signature ("EFI PART", 45h 46h 49h 20h 50h 41h 52h 54h or 0x5452415020494645ULL[a] on little-endian machines)
8 (0x08)	4 bytes	Revision number of header - 1.0 (00h 00h 01h 00h) for UEFI 2.10
12 (0x0C)	4 bytes	Header size in little endian (in bytes, usually 5Ch 00h 00h 00h or 92 bytes)
16 (0x10)	4 bytes	CRC32 of header (offset +0 to +0x5b) in little endian, with this field zeroed during calculation
20 (0x14)	4 bytes	Reserved; must be zero
24 (0x18)	8 bytes	Current LBA (location of this header copy)
32 (0x20)	8 bytes	Backup LBA (location of the other header copy)
40 (0x28)	8 bytes	First usable LBA for partitions (primary partition table last LBA + 1)
48 (0x30)	8 bytes	Last usable LBA (secondary partition table first LBA − 1)
56 (0x38)	16 bytes	Disk GUID in mixed endian[12]
72 (0x48)	8 bytes	Starting LBA of array of partition entries (usually 2 for compatibility)
80 (0x50)	4 bytes	Number of partition entries in array
84 (0x54)	4 bytes	Size of a single partition entry (usually 80h or 128)
88 (0x58)	4 bytes	CRC32 of partition entries array in little endian
92 (0x5C)	*	Reserved; must be zeroes for the rest of the block (420 bytes for a sector size of 512 bytes; but can be more with larger sector sizes)

GPT Partition Entry

GUID partition entry format

Offset	Length	Contents
0 (0x00)	16 bytes	Partition type GUID (mixed endian[12])
16 (0x10)	16 bytes	Unique partition GUID (mixed endian)
32 (0x20)	8 bytes	First LBA (little endian)
40 (0x28)	8 bytes	Last LBA (inclusive, usually odd)
48 (0x30)	8 bytes	Attribute flags (e.g. bit 60 denotes read-only)
56 (0x38)	72 bytes	Partition name (36 UTF-16LE code units)

Unix Filesystem

Inode structure