

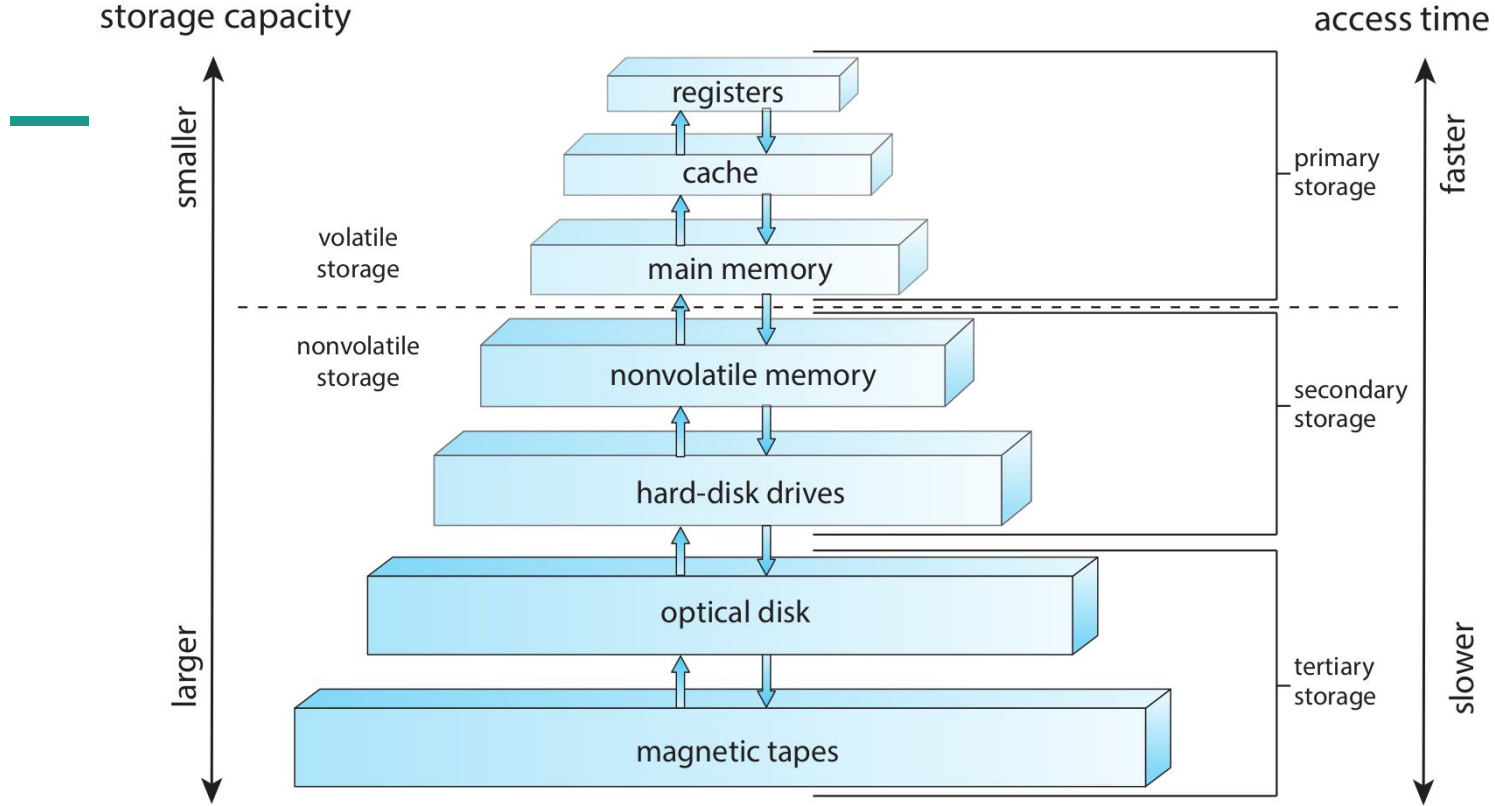


Systemy Operacyjne

Pamięć masowa

Dr hab. inż. Krzysztof Rzecki, prof. AGH

Na podstawie: Abraham Silberschatz, *Koncepcje systemów operacyjnych*



Źródło: A. Silberschatz, *Operating Systems Concepts Essentials*



Pamięć masowa

- Pamięć masowa to trwały nośnik pamięci, który w odróżnieniu do pamięci operacyjnej, może przechowywać (bez podtrzymywania elektrycznego) informacje przez długi czas.
- Adresowanie pamięci masowej zależy od jej typu, ale nie jest to adresowanie bezpośrednie, jak w przypadku pamięci operacyjnej.
- Rodzaje:
 - Pamięć magnetyczna: dysk twardy (HDD, ang. *hard disk drive*), ale też dyskietka (ang. *floppy disk drive*), taśma magnetyczna (ang. *magnetic tape*).
 - Pamięć optyczna: dysk CD (ang. *compact disk*), dysk DVD (ang. *Digital Video Disc* lub *Digital Versatile Disc*), dysk BD (ang. *Blu-ray Disc*).
 - Pamięć półprzewodnikowa: dysk SSD (ang. *Solid State Drive*), NVM (ang. *Non-volatile memory*): pamięć typu *flash*, pamięć USB, karty pamięci.

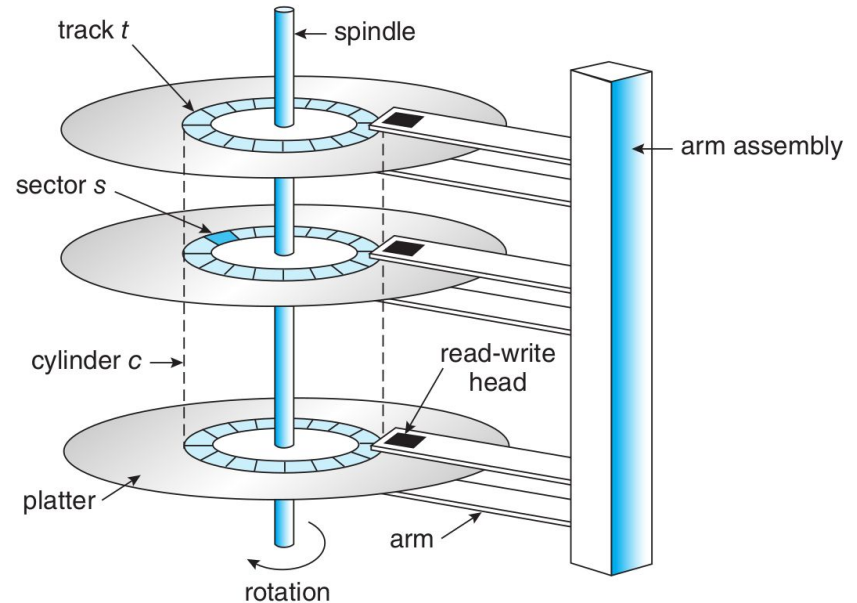


Pamięć masowa

- Brak możliwości bezpośredniego wykonywania programów.
- Cechy charakterystyczne (SSD/HDD):
 - Wielkość: 250 GB - 6+TB
 - Niska cena: ~ 0,05 - 0,10 zł / 1 GB
- Dane trwale zapisane bez podłączonego zasilania.
- Czas dostępu:
 - HDD: 20 ms
 - SSD: ~20 us
- Trwałość MTBF:
 - HDD: 1-2,5 mln godzin
 - SSD: 1,5-3 mln godzin

Budowa dysku twardego

- Dysk twardy składa się z:
 - Talerz (ang. *platter*): 1.8-3.5" - dwustronna powierzchnia magnetyczna.
 - Głowica czytająco-pisząca (ang. *head*).
 - Ramię z głowicami (ang. *arm*).
- Powierzchnia talerza:
 - Ścieżki (ang. *tracks*), które podzielone są na:
 - Sektory (ang. *sectors*), których zbiór przy danej pozycji ramienia to:
 - Cylinder (ang. *cylinder*).
- Każdy sektor posiada stały rozmiar i jest to najmniejsza jednostka transferu (obecnie najczęściej 4KB).
- Prędkość obrotowa: 5'400, 7'200, 10'000, 15'000.



Źródło: A. Silberschatz, *Operating Systems Concepts Essentials*



LBA - logical block addressing

$LBA = (\text{numer cylindra } C * \text{liczba głowic na cylinder } HPC + \text{numer głowicy } H)$

* liczba sektorów na ścieżkę SPT + numer sektora $S - 1$

$$C = LBA / (HPC * SPT)$$

$$H = (LBA / SPT) \bmod HPC$$

$$S = (LBA \bmod SPT) + 1$$



Interfejsy pamięci masowej

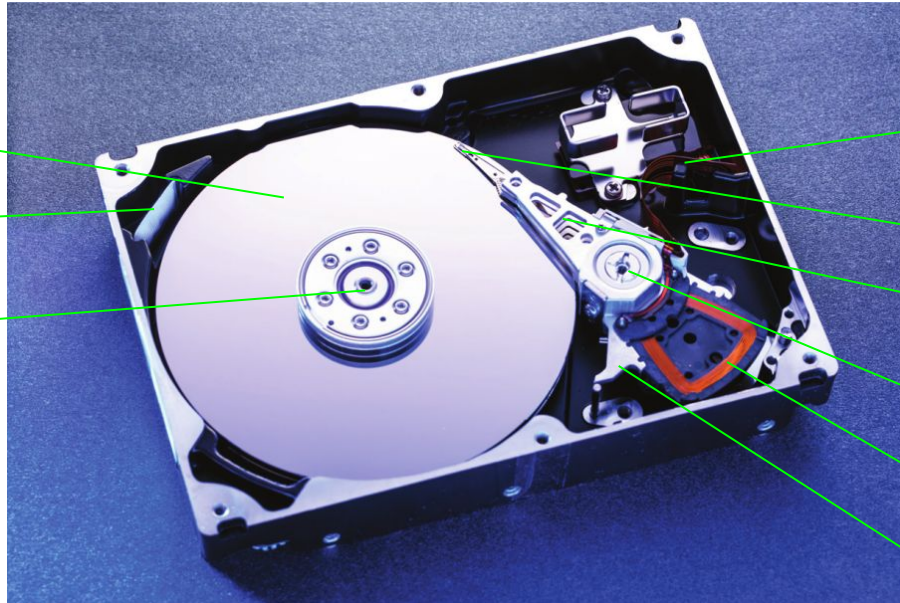
- ATA, ang. *Advanced Technology Attachment*
== IDE, ang. *Integrated Device Electronics*
- ATAPI, ang. *Advanced Technology Attachment Packet Interface*
- PATA, ang. *Parallel ATA* (ATA + kontroler)
- SATA, ang. *Serial ATA*
- eSATA, ang. *External SATA*
- SCSI, ang. *Small Computer Systems Interface*
- SAS, ang. *Serial Attached SCSI* - częściowo kompatybilny z SATA
- USB, ang. *Universal Serial Bus*
- M.2 (NGFF, ang. *Next Generation Form Factor*) i PCI Express dla urządzeń NVMe, ang. *Non-volatile Memory Express*

Dysk twardy HDD 3.5" (cale) - podobnie 2.5"

Talerz dysku
twardego

Filtr

Piasta silnika
(łożysko kulkowe)



Podłączenie
elektroniki

Głowica

Ramię

Pozycjoner

Cewka

Magnesy

Formaty dysków SSD (ang. *Solid-state drive*)



2,5"



mSATA



M.2



HHHL - Half Height, Half Length
(zwana także kartą rozszerzeń)



M.2
(wspiera wersję NVMe)



U.2
(dostępne tylko w NVMe)



Rozmiar sektora - dyski HDD i SSD

- Początkowo 512 bajtów, obecnie najczęściej 4KiB (4 kibibity).

- W systemie plików:

```
$ cat /sys/class/block/sdX/queue/physical_block_size => 4096
```

```
$ cat /sys/class/block/sdX/queue/logical_block_size => 512
```

- Programy:

```
$ LC_ALL=C fdisk -l /dev/sdX | grep 'Sector size'
```

```
$ sudo smartctl -a /dev/sdX | grep 'Sector Size'
```

```
$ sudo hdparm -l /dev/sda | grep 'Sector size:'
```



Rozmiar sektora - dyski NVMe

- W systemie plików:

```
$ sudo smartctl -a /dev/nvme0n1
```

```
Supported LBA Sizes (NSID 0x1)
```

```
Id Fmt Data Metadt Rel_Perf
```

```
0+ 512 0 0
```



Odczyt własności dysku twardego: HDD

```
# gdisk -l /dev/sdb
```

```
Disk /dev/sdb: 3907029168 sectors, 1.8 TiB
Model: ST2000DM008-2FR1
Sector size (logical/physical): 512/4096 bytes
Disk identifier (GUID): AB46240F-CCFF-4768-9379-8C1FAC47CD7B
Partition table holds up to 128 entries
Main partition table begins at sector 2 and ends at sector 33
First usable sector is 34, last usable sector is 3907029134
Partitions will be aligned on 2048-sector boundaries
Total free space is 3907029101 sectors (1.8 TiB)
```

```
# hdparm -g /dev/sdb
```

```
geometry      = 243201/255/63, sectors = 3907029168, start = 0
```



Odczyt własności dysku twardego: SSD

```
# gdisk -l /dev/sda
```

```
Disk /dev/sda: 937703088 sectors, 447.1 GiB  
Model: KINGSTON SUV500M  
Sector size (logical/physical): 512/4096 bytes  
Disk identifier (GUID): 2025DC72-437D-425C-A561-4D4EEB1FDB61  
Partition table holds up to 128 entries  
Main partition table begins at sector 2 and ends at sector 33  
First usable sector is 34, last usable sector is 937703054  
Partitions will be aligned on 2048-sector boundaries  
Total free space is 9139 sectors (4.5 MiB)
```

```
# hdparm -g /dev/sda
```

```
geometry      = 58369/255/63, sectors = 937703088, start = 0
```




Odczyt własności dysku twardego: NVMe

```
# gdisk -l /dev/nvme0n1
```

```
Disk /dev/nvme0n1: 1953525168 sectors, 931.5 GiB
Model: KINGSTON SA2000M81000G
Sector size (logical/physical): 512/512 bytes
Disk identifier (GUID): 15F70755-2100-47E8-B0DF-F5D42D92599B
Partition table holds up to 128 entries
Main partition table begins at sector 2 and ends at sector 33
First usable sector is 34, last usable sector is 1953525134
Partitions will be aligned on 2048-sector boundaries
Total free space is 28013 sectors (13.7 MiB)
```

```
# hdparm -g /dev/nvme0n1
```

```
geometry      = 953869/64/32, sectors = 1953525168, start = 0
```



Informacje diagnostyczne o dysku twardym

```
# smartctl -a /dev/sdb
```

- `Reallocated_Sector_Ct`- liczba realokowanych sektorów w miejsce zapasowe.
- `Reallocated_Event_Count`- liczba operacji realokowania (po kilka sektorów).
- `Current_Pending_Sector`- liczba sektorów oczekujących na remapowanie.
- `Offline_Uncorrectable`- liczba błędów, których nie można naprawić.



Mapowanie adresów

- Pamięć masowa adresowana jest jako jednowymiarowa macierz **bloków logicznych**.
- Blok logiczny to najmniejsza jednostka transferu.
- Każdy blok logiczny mapowany jest do sektora fizycznego lub strony pamięci półprzewodnikowej.
- W przypadku HDD logiczny sektor 0 to np. fizycznie pierwszy sektor najbardziej zewnętrznego cylindra. Adresowanie jest wzdłuż tego cylindra, a potem schodzi do cylindrów wewnętrznych.
- W przypadku NVM adresacja zaczyna się od pierwszych: chip, blok, strona.
- Na ogół używa się adresacji przez LBA, ang. *Logical Block Address*.

MBR, ang. *Master Boot Record*

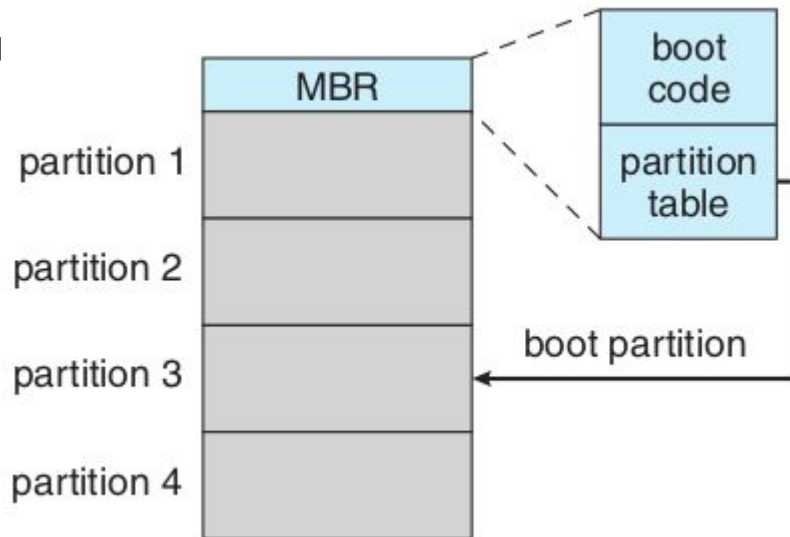
- Ładowanie systemu zaczyna się od programu **bootstrap**
- 446 bytes – Bootstrap.
- 64 bytes – Partition table.
- 2 bytes – Signature.

Kopia:

```
$ dd if=/dev/sdX of=file.img bs=512 count=1  
+ odzyskiwanie
```

Dla partycji rozszerzonych:

```
$ sfdisk -d /dev/sdX > /tmp/sdX.sfdisk  
$ sfdisk /dev/sdX < /tmp/sdX.sfdisk
```





Monitorowanie urządzeń we/wy

Narzędzia do monitorowania:

- iotop + watch
- iostat + watch

Przepływ danych:

- `dd if=/dev/{zero,random}`
- `cp`
- `sync`
- `pidstat -p `pidof ...``



Madahora