

Pevné disky, fyzická a logická struktura, princip magnetického zápisu a čtení, SSD disky, parametry pevných disků, RAID

Pevné disky

- Pevné disky jsou jedním z medií pro ukládání dat
- Kapacita je obvykle od 128 GB až do 18 TB
- Přístupová doba se pohybuje od 4ms do 10 ms
- Rychlost otáčení – laciné desktopové 5400, obvykle ale 7200 a serverové 10 až 15 tisíc otáček za minutu
- Přenosová rychlost desítky MB/s u náhodného přístupu, přes 150 MB/s u sekvenčního čtení
- RAM cache (vyrovnávací paměť) – 8 až 512 MB

Fyzická struktura disku

Hlavy

- Používají se pro čtení a zápis z magnetické vrstvy
- Dříve se používaly magnetodynamické hlavy nyní se používá krystal
- Proti mechanickému poškození při doteku hlavy bývá povrch magnetické záznamové vrstvy chráněn vrstvou teflonu popřípadě vrstvou napařeného uhlíku
- Operace nutné pro čtení nebo zápis dat:
 - Vystavení čtecí hlavy na průměr žádané stopy
 - Vyčkání na utlumení rozkmitu způsobeného setrvačností hlav (řádové milisekundy)
 - Vyčkání než se žádaný sektor či blok otáčením disku dostane pod čtecí hlavu
- Průměrný čas, za který je disk připraven číst nebo zapisovat se označuje jako přístupová doba. V současnosti u běžných disků okolo 8,5 ms, u serverových pod 4 ms
- **Magnetický záznam a čtení – princip:**
 - Zápis:
 - Pokud cívkou nebo krystalem prochází el. proud dojde k vytvoření magnetického toku
 - Tok se uzavírá ve štěrbině mezi hlavičkou a plotnou a tím ovlivňuje i záznamovou vrstvu pevného disku
 - Zmagnetizováním daného místa, vznikají tzv. magnetické rezervace – mění se směr magnetizace a dochází k zápisu
 - Čtení:
 - Probíhá opačným způsobem než zápis
 - Cívky nebo krystaly reagují na magnetické rezervace a magnetický tok je zpracován na el. impuls

Diskové plotny

- Data jsou na pevném disku uložena v podobě magnetického záznamu
- Disk tvoří kovové nebo keramické desky (plotny), které jsou pokryté tenkou magneticky tvrdší vrstvou
- Plotny jsou neohebné na rozdíl od ohebných ploch v disketách
- V současných discích bývá většinou jedna o průměru 3,5“, 2,5“ nebo 1,8“ a otáčí se na vřetenu poháněném elektromotorem
- Maximální počet plotýnek u disků:
 - 2 u 1,8“ disků
 - 5 u 2,5“ disků
 - 10 u 3,5“ disků

Pohon disku, hřídel

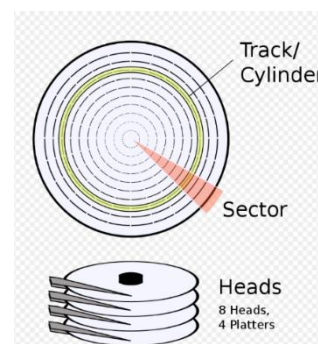
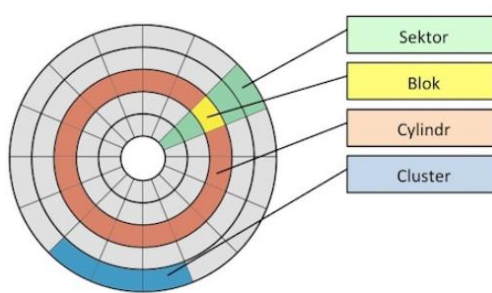
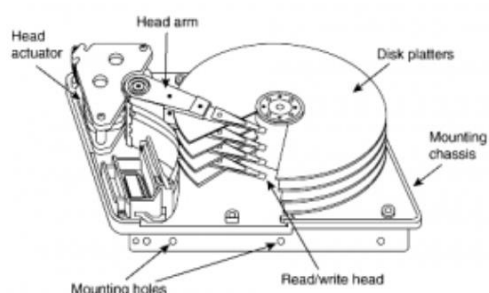
- Na 1 společné hřídeli může být umístěno několik diskových kotoučů
- Na pohonu disku, otáčkách hřídele je závislá průměrná čekací doba (čím větší otáčky tím menší čekací doba)
- Avšak platí: Čím více otáček, tím horší provozní podmínky (vyšší tvorba tepla uvnitř pouzdra disku)

Další části disku

- Mechanika pro pohyb raménka s hlavou
- Elektronika řídící práci disku
- Rozhraní disku pro připojení k základní desce
- Proces magnetického dělení disku = fyzické formátování

CHS (Cylinder-head-sector)

- Jedná se o 3D souřadnicový systém pro určování adresních pozic bloků na disku
- Head = disk má počet hlav, který se rovná dvojnásobku počtu diskových ploten, v CHS systému určuje vertikální souřadnici
- Cylinder = jedná se o „horizontální“ souřadnici, která se propisuje na všechny plotýnky nad sebou, stejný průměr
- Sector = vybírá konkrétní sektor ze kterého mají být data z cylindru vybrána, počet sektorů je roven podle toho na kolik sektorů je rozdělen cylindr, 512 Byte



- Cluster – několik sektorů, běžně 4 kB, využití například u adresování celého videa

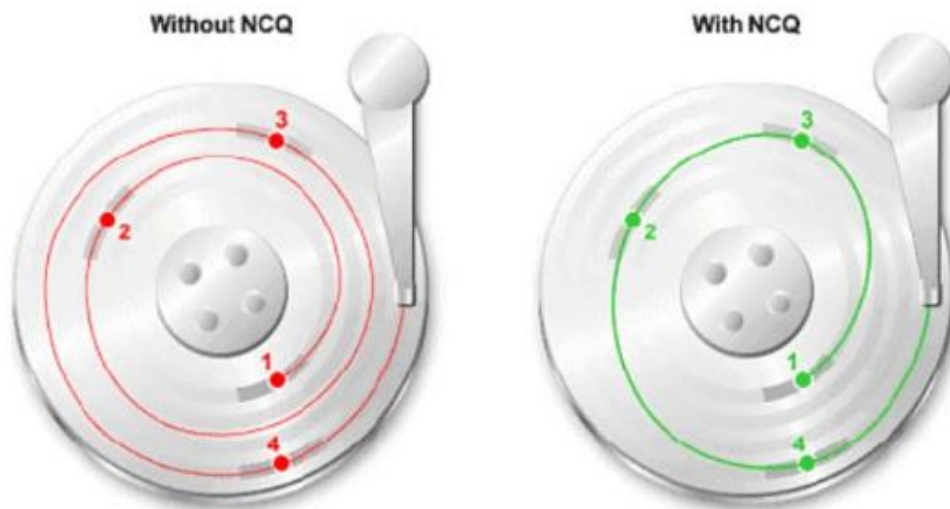
- FAT (File Allocation Table) – Tabulka obsahující informace o umístění dat na disku, jednoduchý souborový systém
 - Několik verzí podle podpory velikosti disků a velikosti clusterů
 - FAT 32 – podpora 8TB diskových oddílů a souborů až 4GB, nevodný tedy pro velké soubory, využití hlavně u přenosných uložišť (USB flash, SD karta)

LBA (Linear Block Addressing)

- Nahrazuje adresaci CHS
- Bloky se číslují lineárně od 0 – 48b adresa

NCQ

- Neoptimálnější cesta přečtení dat na disku



PATA (starší)

- Paralelní ATA
- Sběrnice pro připojování disků

IDE

- Rozhraní ATA
- Dovoluje připojení 2 disků, které musely být rozděleny na master a slave

EIDE

- Lepší IDE, dovoluje 4 disky

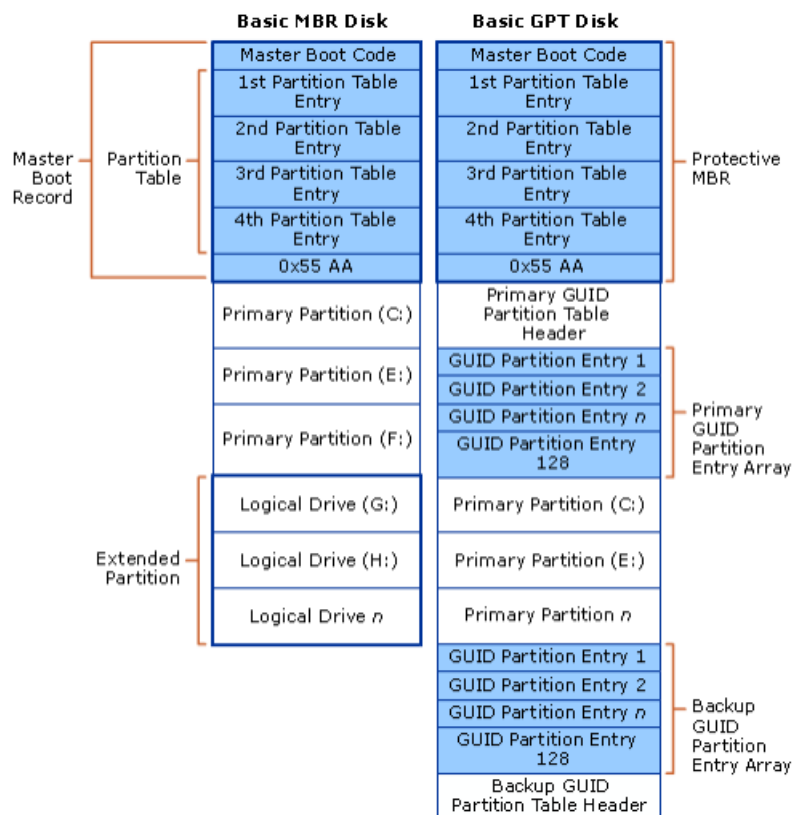
ATAPI

- = ATA Packet Interface
- Rozšíření o packet interface
- Umožňuje přenášet příkazy

Logická struktura pevného disku

- Zorganizování uložených údajů pro rychlé nalezení dat
- Informace uloženy podle Firmware PC:
 - BIOS – Master Boot Record (MBR) – **hlavní zaváděcí záznam** obsahuje tabulku 4 oblastí, pouze jedna může být aktivní, tam hledá **OS zavaděč**. Tabulka rozdělení disku s maximálně 232 sektory po 512 bajtech, Max. kapacita disku tak může být jen něco přes 2 TB

- UEFI – pracuje s tabulkou nazvanou GUID (Globally Unique Identifier) Partition Table (GPT) s adresami dlouhými 64bitů a zvládne až 264 sektorů – adresovat tak může až 9 zettabajtů



SSD disky

- Solid-state drive (polovodičový disk)
- Zařízení pro čistě elektronické ukládání dat
- Nástupce pevných disků
- Neobsahuje pevné části = odolnost vůči otřesům a změnám tlak
- Nehlučnost
- Nižší spotřeba
- Vyšší přenosové rychlosti – spotřebitelské 200 až 2500 MB/s, podnikové až několik gigabytů za sekundu
- Téměř okamžitý start
- Nízká latence čtení
- Náhodný přístup běžně pod 100 μs
- Rozhraní: SATA, PCI-Express (PCIe; NVMe)
- Běžně nevyžadují žádné zvláštní chlazení
- Elektromagnetický impuls může poškodit data na disku
- Jsou velmi lehké ale pro lepší náhradu hdd používají stejné palcové formáty
- Díky absenci pohyblivých částí je spolehlivost a životnost ovlivněna jen podle počtu cyklů pro vymazání a přepsání

- Kapacity od 120 GB do 60 TB

SATA (Serial Advanced Technology Attachment)

- Označuje sběrnici, která využívá datové rozhraní pro připojení paměťových zařízení
- 3 generace sata:
 - První generace: Standard SATA 1,5 Gb/s, datová propustnost 150 MB/s, frekvence 1,5 GHz
 - Druhá generace: Standard SATA 3 Gb/s, datová propustnost 300 MB/s, frekvence 3 GHz
 - Třetí generace: Standard SATA 6 Gb/s, datová propustnost 600 MB/s, frekvence 6 GHz
- SATA disky požadují 2 kabely: SATA napájecí kabel a SATA datový kabel (pokud se avšak nejedná o M.2 SATA SSD)
- Většina SATA disků postupuje v klasickém HDD rozměrovém formátu = 3,5" a 2,5"
- SATA kabel:
 - 7 vodičů – 4 datové, 3 zem
 - Full-duplex
 - Malé napětí 250mV
 - Délka do 1 m

PCIe (PCI-Express, NVMe)

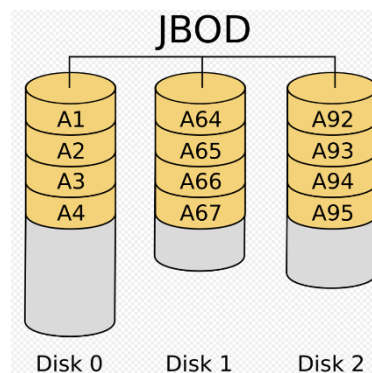
- Peripheral component interconnect express
- Označuje se jako standard systémové sběrnice což není úplně pravda jelikož se jedná pouze o dvou bodové spoje
- Verze PCIe 1.0 – PCIe 7.0
- Pro SSD je tato sběrnice velmi významná díky využití u M.2 SSD NVMe disků
 - NVMe (Non-Volatile Memory express) = rozhraní pro komunikaci mezi flash pamětí a jejím řadičem, maximální využití nízké latence a vysoké propustnosti dat přes PCIe
 - M.2:
 - Standart definující rozměry a rozhraní pro připojení rozšiřujících karet k počítači
 - Nejčastěji využito pro připojení právě SSD disků
 - NVMe SSD disky:
 - Rychlosti až 3 500 MB/s u PCIe generace 3.0
 - Rychlosti až 7 500 MB/s u PCIe generace 4.0
 - Vyšší cena ale nejrychlejší zápisy a čtení na trhu

RAID (Redundant Array of Inexpensive/Independent Disks)

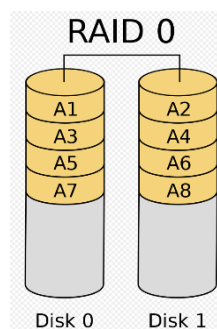
- Česky: Vícenásobné pole levných/nezávislých disků
- Jedná se o metodu ochrany/zabezpečení dat proti selhání pevného disku
- Realizace pomocí ukládání dat na více nezávislých disků
- Úroveň zabezpečení se liší podle zvoleného typu RAID
- Nejčastější typy RAID = RAID 0, RAID 1, RAID 5 nebo RAID 6
- Využíván také pro navýšení výkonu diskového subsystému

RAID 0

- Nejedná se vlastně o skutečný RAID jelikož neobsahuje žádné redundantní informace = neposkytuje datům žádnou ochranu
- Jednotlivá zařízení jsou jen spojena do logického celku a vytváří tak kapacitu součtu všech členů
- Spojení může být realizováno dvěma způsoby – Zřetězení a Prokládání
- Zřetězení (JBOD)
 - Just a Bunch Of Disks („jen hromada disků“)
 - Data jsou ukládána postupně na několik disků
 - Po zaplnění prvního jdou data do druhého a tak dále
 - Při výpadku jednoho členu nějaká data nemusí být zasažena



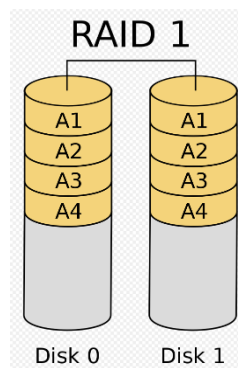
- Prokládání
 - Data jsou ukládána cyklicky/střídavě
 - Při poruše disku není pravděpodobné, že by nějaký soubor zůstal nepoškozen
 - Prokládání může zrychlit čtení i zápis větších bloků dat a to i oproti RAID 1
 - Max. kapacita úložiště je rovna velikosti nejmenšího disku krát počet zapojených disků



RAID 1 (zrcadlení)

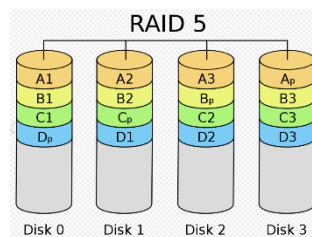
- Nejjednodušší, ale poměrně efektivní ochrana dat
- Provádí se zrcadlením = obsah se zaznamenává na dva disky
- V případě výpadku jednoho disku se pracuje s kopií, která je ihned k dispozici
- Může se zvýšit rychlost čtení a o něco snížit odezva
- Naopak se může zpomalit zápis samotných dat kvůli dvojitému zápisu

- Potřeba dvojnásobné diskové kapacity



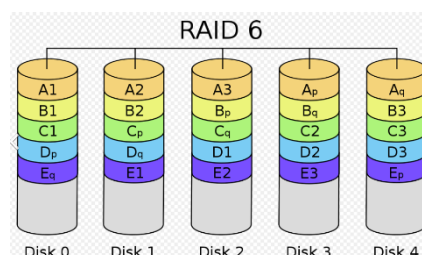
RAID 5

- Vyžaduje alespoň 3 členy
- Kapacita jednoho členu je zabrána samoopravným kódem neboli paritou, které jsou ukládány na členech střídavě
- Znovu je čtení rychlejší ale zápis pomalejší



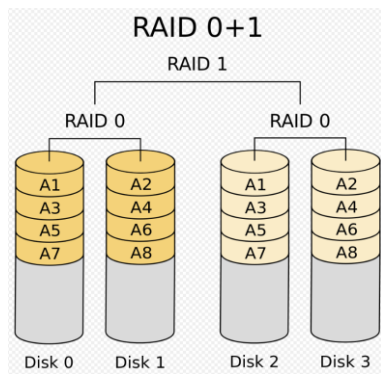
RAID 6

- Používá 2 paritní bloky na každém z použitých disků, přičemž na každém je samoopravný kód vypočten jiným způsobem
- Znovu jsou paritní bloky ukládány střídavě
- Rychlost čtení je srovnatelná s RAID 5 ale je pomalejší zápis
- Minimálně 4 disky

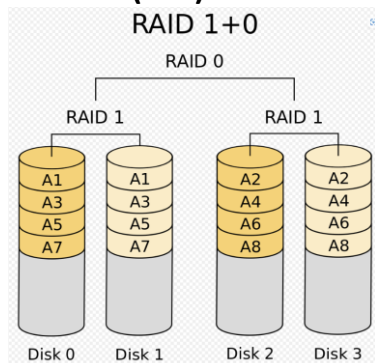


Víceúrovňové typy RAID

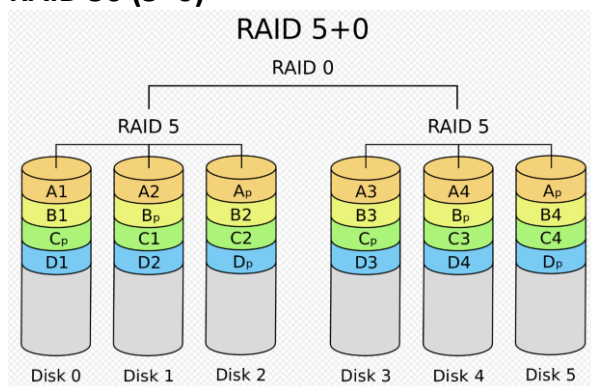
RAID 01 (RAID 0+1)



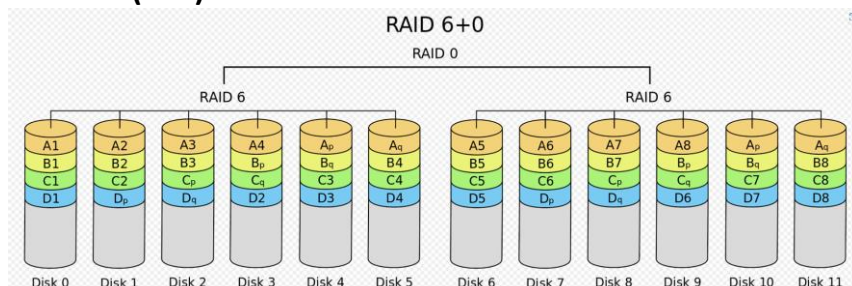
RAID 10 (1+0)



RAID 50 (5+0)



RAID 60 (6+0)



RAID 100 (10+0)

