

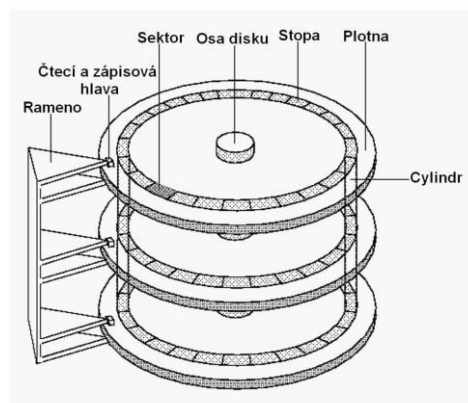
6. Pevné disky – fyzická a logická struktura pevného disku, princip magnetického zápisu a čtení, parametry pevných disků, RAID, SSD disky, princip zápisu a čtení optických disků.

HARDWARE A APLIKAČNÍ SOFTWARE

Struktura pevného disku

Fyzická struktura

- Skládá se z tuhých kovových disků (plotny), které jsou potaženy povlakem oxidu železa a uloženy na sebe ve vzduchotěsně uzavřeném prostoru.
- Ukládání dat probíhá tzv. magnetizováním a odmagnetizováním oblastí na desce.
- Fyzicky disk rozdělen na sektory, ty jsou spojeny do soustředných stop (kružnic), na povrchu každé plotny s tím, že na jednu plotnu připadá 305 a více stop.
- Pro obě strany každé desky má HD elektromagnetickou čtecí a zápisovou hlavu. Hlav je 2x více než ploten.
- Původní čtení dat bylo založeno na magnetické indukci, dnes se už data čtou jako sled odporů vyvolaných různou orientací magnetického pole (magnetických dipólů).
- Mechanika hlav je založena na:
 - Krokovém motoru – nespolehlivé
 - Vystavovací cívce – cívka vychylována změnami proudu
- Přístupová doba se skládá z doby vystavení (=doba vyhledávání, Track-to-Track Seek) a doby čekání (10 ms).
- Čtení a zápis probíhá po cylindrech.
- Disk se dělí na stopy a sektory, které se číslovají.
- V hlavičce sektoru se ukládají informace o obsahu daného sektoru.



Logická struktura

- Logická struktura je vytvořena vysokoúrovňovým formátováním, které umožňuje každý operační systém, je popisován soustavou tabulek.

Parametry disků

Pevný disk

= HDD / Hard Disk Drive

= Elektromechanické zařízení pro záznam a čtení

adresovatelných dat, paměť o velké kapacitě a s pomalejším přístupem než paměť RAM.

- **Kapacita:** 500 GB – 16 TB
- **Přístupová doba:**
 - od 4 ms
 - 8 ms (obvykle)
 - přes 10 ms (lacinější)
- **Rychlost otáčení:**



- 5400 rpm (lacinější)
- 7200 rpm (obvykle)
- 10000 rpm – 15000 rpm (serverové)
- **Přenosová rychlost:**
 - Desítky MB/s (náhodný přístup)
 - Přes 150 MB/s
- **RAM cache:** 8 – 64 MB RAM

SSD disk

= Solid-state drive/polovodičový disk

- Čistě elektronické ukládání dat.
- Odolnost vůči otřesům, změnám tlaku, nehlučnost, nižší spotřeba elektrické energie, a hlavně vyšší přenosová rychlost.
- **Části:** Datový a napájecí konektor, řadič disku, paměti NAND
- **Rozhraní:** SATA, M.2, PCI-Express
- **Typy buněk:**
 - **SLC (Single-level cell):** Vyšší cena i životnost, 1 bit na každou buňku
 - **MLC (Multi-level cell):** Více úrovní signálu na každou buňku (více bitů) – větší chybovost
- **Implementace příkazu TRIM:** OS informuje SSD, které datové bloky obsahují dále již nepoužívaná data.
- **Mezivrstva FTL (Flash Translation Layer):** Rovnoměrné opotřebování datových buněk.

Princip magnetického zápisu a čtení

- Zápis a čtení dat mají na starosti hlavičky, což jsou cívky navinuté na jádrech a na ramenech se vznášejí těsně nad povrchem ploten.

Zápisu

- Pokud cívkou prochází elektrický proud, dojde k vytvoření magnetického toku, který se uzavírá ve štěrbině mezi hlavičkou a plotnou a tím ovlivňuje i záznamovou vrstvu pevného disku.
- V závislosti na směru toku proudu při této operaci dochází k zmagnetizování daného místa určitým směrem.
- Mezi dvěma zmagnetizovanými místy vznikají tzv. magnetické reverzace.
- Jedná se o místa, v nichž se mění směr magnetizace, a právě ona jsou zodpovědná za vlastní funkčnost této metody zápisu.

Čtení

- Čtení dat probíhá opačným způsobem nežli zápis.
- Během pohybu hlaviček nad povrchem plotny reagují cívky právě na magnetické reverzace a ty následně vyvolávají magnetický tok, který je zpracováván na elektrický impuls a dále řídící elektroniky disku.
- Dnes se používá technika kolmého zápisu dat, kdy orientace magnetického pole je kolmá vůči povrchu plotny.
- Tato technologie zápisu využívá pro záznam magneticky tvrdší materiály společně s magneticky měkkou spodní vrstvou, která pomáhá hlavičce v zápisu dat. Výhodou je vysoká hustota zápisu.

Princip zápisu a čtení u optických disků

- Zápis a čtení následuje nepřerušovanou spirálu zaplňující celý povrch disku, směrem z vnitřního obvodu disku k vnějšímu obvodu disku.

Zápis

- Data se ukládají na disk pomocí laseru nebo lisovacího stroje.

Čtení

- Data lze přečíst, pokud je stopa dat osvětlena laserovou diodou v optické mechanice, kde je disk roztočen na rychlost kolem 4000 rpm a více (dle mechaniky, formátu disku a vzdálenosti čtecí hlavy od centra disku – vnitřní stopy jsou čteny vyšší rychlostí).
- Prohlubně nebo výstupky deformují odražené světlo laseru, proto má většina optických disků charakteristický duhový vzhled, tvořený odrazy drážek v odrazové vrstvě.

RAID

= Redundant Array of Independent Disks

- Zabezpečení dat proti selhání pevného disku.
- Uložená data jsou zachována i při selhání některého z disků (kromě RAID 0).
- RAID nenahrazuje zálohování dat.
- Jednotlivé způsoby ukládání vyjadřují čísla RAID.
- Způsoby ukládání dat jsou realizovány softwarově (snížení rychlosti) nebo hardwarově (řadič RAID).

Princip

- Při výpadku některého disku (resp. členu pole), dostane se pole do takzvaného degradovaného stavu – nižší výkon, uložená data jsou k dispozici.
- Nový disk namísto havarovaného se začlení zpět do pole - hot add.
- Začleněním nového disku začne rekonstrukce pole, při které jsou dopočítány chybějící údaje a zapsány na nový disk.
- Data jsou během rekonstrukce stále přístupná.
- Po dokončení rekonstrukce je RAID pole synchronizováno.
- Může být v poli trvale k dispozici rezervní disk – spare a rekonstrukce pole může být zahájena zcela automaticky.
- RAID vytváří logický (virtuální) úložný prostor, se kterým se dá pracovat stejným způsobem, jako by to byl jediný pevný disk.

RAID 0

- Nejedná se o skutečný RAID
- Neobsahuje žádné redundantní informace
- Neposkytuje uloženým datům žádnou ochranu (porucha členu znamená ztrátu dat)
- **Způsoby spojení:**
 - Lineární zřetězení – Just a Bunch Of Disks (JBOD)
 - Prokládání – data ukládána na disky cyklicky, zrychlení čtení i zápisu větších bloků dat – možnost současně číst blok z jednoho disku a zapisovat na jiný.

RAID 1

- Nejjednodušší, ale efektivní ochrana dat.
- Zrcadlení (mirroring) obsahu disku na druhý.

- V případě výpadku jednoho disku se pracuje s kopií, která je ihned k dispozici.
- **Nevýhody:**
 - Dvojnásobná disková kapacita na jedno množství dat.

RAID 2

- Méně používané.
- Složitější verze RAID 3, kde jsou data po bitech stripována mezi jednotlivé disky.

RAID 3

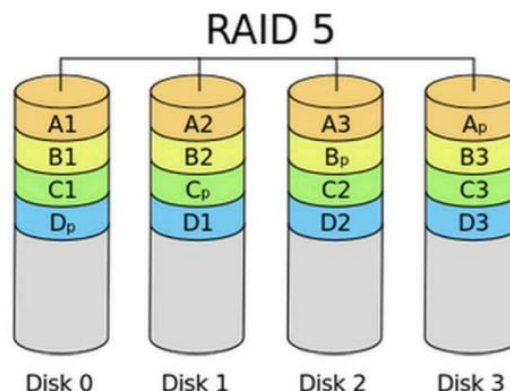
- Méně používané.
- N+1 stejných disků, na poslední disk je uložena XOR - parita. Při výpadku paritního disku jsou data zachována, při výpadku libovolného jiného disku je možno z ostatních disků spolu s paritním diskem ztracená data zrekonstruovat.
- **Nevýhody:**
 - Paritní disk je vytížen při zápisu na jakýkoliv jiný disk. (vyšší opotřebení, nižší spolehlivost).

RAID 4

- Méně používané.
- Disky jsou stripovány po blocích (ne po bitech) a parita je na paritním disku ukládána po blocích.
- Výhody a nevýhody jsou stejné jako u RAID 3.

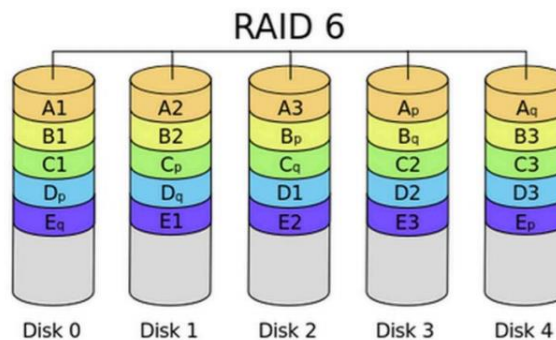
RAID 5

- Vyžaduje alespoň 3 členy diskového pole.
- Kapacitu jednoho členu zabírají samo opravné kódy, které jsou uloženy na členech střídavě (a ne pouze na jednom, čímž je odstraněna nevýhoda RAID 4).
- Lze využít paralelního přístupu k datům - delší úsek dat je rozptřčen mezi více disků - rychlejší čtení.
- Pomalejší zápis (nutnost výpočtu samo opravného kódu).
- Odolný vůči výpadku jednoho disku.



RAID 6

- Na rozdíl od RAID 5 používá dva paritní disky.
- Na každém z nich je samo opravný kód vypočten jiným způsobem.
- Paritní data uložena střídavě na všech discích.
- Odolnost proti výpadku dvou disků.
- Rychlost čtení srovnatelná s RAID 5.
- Zápis pomalejší než u RAID 5 - výpočet dvou sad paritních informací.
- Možno sestavit minimálně ze čtyř disků (zrcadlení lepší).

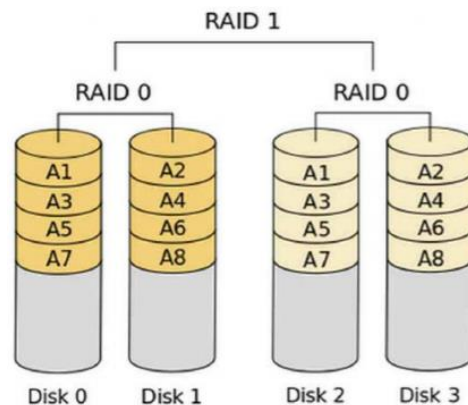


RAID 7

- Méně používané.
- Odvozené od RAID 3 a RAID 4, liší se zejména tím, že k nim přidává vyrovnávací paměť.

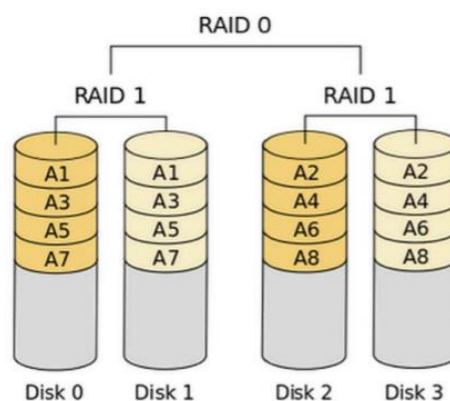
Víceúrovňový RAID 01

- RAID 01 = RAID 0+1
- Data uložená prokládaně (stripováním) na dvě dvojice disků a tyto dvojice jsou zrcadleny.
- Rozložení zátěže mezi více disků při čtení a zápisu.
- Data jsou uložena zrcadleně a dají se po chybě obnovit.
- **Nevýhody:**
 - Využití pouze 50 % celkové diskové kapacity pro data.
 - Při výpadku jednoho ze čtyř disků ztráta redundance dat.



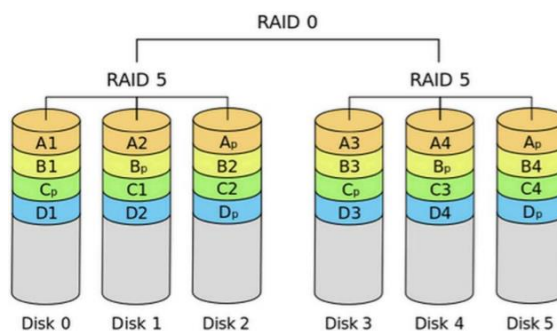
Víceúrovňový RAID 10

- RAID 10 = RAID 1+0
- Data uložená zrcadleně na dvě dvojice disků a tyto dvojice jsou stripovány.
- Výhody podobné jako u RAID 01.
- Rychlejší obnova dat než u RAID 01.
- Využití pro hodně vytížené databázové aplikace.
- **Nevýhody:**
 - Využití pouze 50 % celkové diskové kapacity.



Víceúrovňový RAID 50

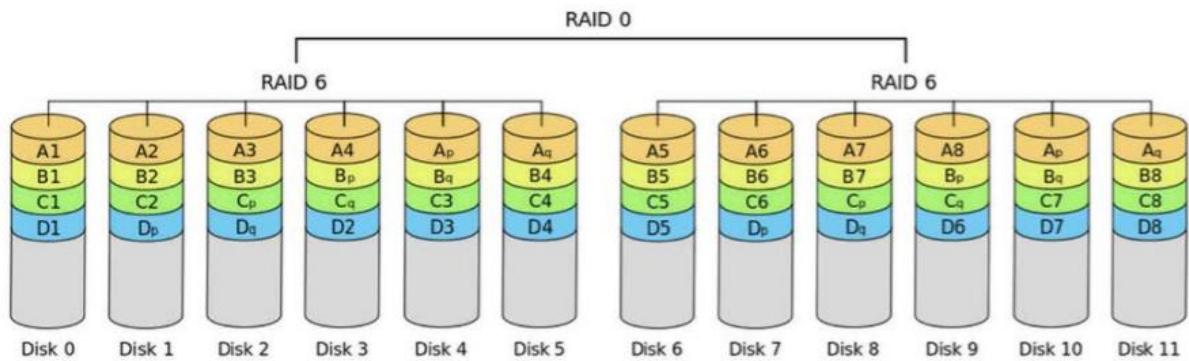
- Dvouúrovňové pole, vytvořené prokládáním několika RAID 5 polí.
- Prokládání zvyšuje rychlost oproti jednoúrovňovému RAID 5.
- V každém pod-poli je potřeba jeden disk navíc na paritní data.
- Pole je odolné proti selhání jednoho disku každého pod-pole.
- kapacita = $(n - 1) \cdot c \cdot P$; (n – počet disků, c – kapacita disku; P – počet podřazených polí)



Víceúrovňový RAID 60

- Dvouúrovňové pole, vytvořené prokládáním několika RAID 6 polí.
- Prokládání zvyšuje rychlost, oproti RAID 50 poskytuje vyšší zabezpečení.
- Minimální počet disků je 8 (tehdy je rychlejší RAID100).
- Jedno z nejbezpečnějších řešení pro vysoké výkony.

- velikost = $(n - 2) \cdot c \cdot P$, (n – počet disků, c – kapacita disku; P – počet podřazených polí)



Víceúrovňový RAID 100

- RAID 100 = RAID 10+0
- Tříúrovňové pole, vytvořené dvouúrovňovým prokládáním dat na zrcadlené pod-pole (RAID 1).
- Minimální počet disků je 8.
- Větší přenosové rychlosti díky prokládání.
- Odolnost proti výpadku jednoho disku v každém pod-poli.
- **Nevýhody:**
 - Využití pouze 50 % celkové diskové kapacity.
- velikost = $(n \cdot c) / 2$ (n – počet disků, c - kapacita disku)

