

15. Souborové systémy FAT a FAT32, NTFS a exFAT

- Struktura logického disku, číslování sektorů, cluster
- Zavaděcí záznam, hlavní adresář, typy FAT, struktura podadresářů
- Rozdíly FAT, VFAT 16 a VFAT 32, řešení dlouhých názvů
- Chyby FAT systémů (programy k tomu určené a způsoby jejich odstraňování)
- Bezpečné odstraňování dat, fragmentace a defragmentace disku
- Charakteristika a vnitřní struktura NTFS, porovnání s FAT
- Metasoubory a jejich funkce, struktura logického disku NTFS

1. Struktura logického disku, číslování sektorů, cluster

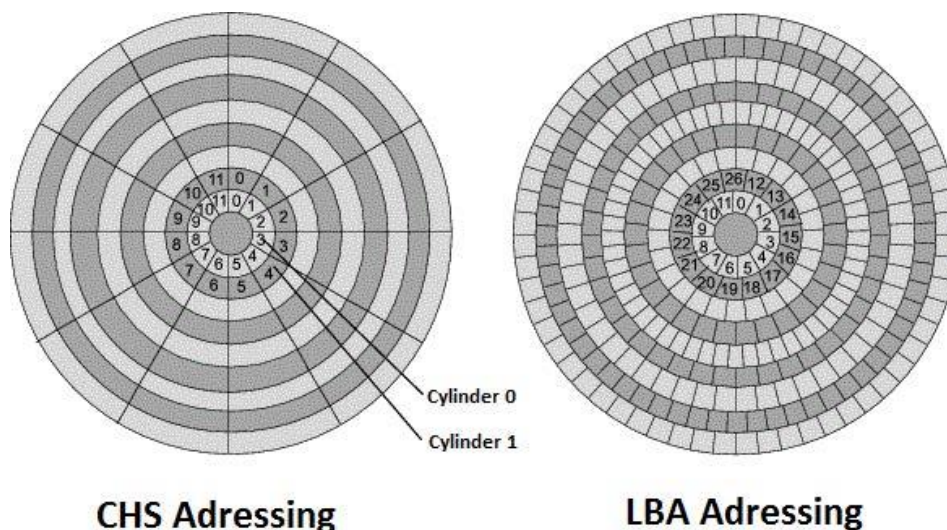
- Souborový systém je sada pravidel, podle kterých OS organizuje data na logickém disku

- **Struktura logického disku**

- Skládá se ze 2 hlavních oblastí
 - Organizační oblast
 - Boot Record
 - 1. FAT
 - 2. FAT
 - Root Directory – hlavní adresář, kořenový adresář
 - vytvářejí se vysokoúrovňovým formátováním logického disku (příkazem format)
 - Datová oblast
 - Cluster – prostor pro soubory

- **Číslování sektorů**

- CHS – Cylinder Hlava Sektor
 - Číslování podle geometrie disku
 - Hlava a cylindr se číslují od 0
 - Sektor se čísluje od 1
- LBA – Logical Block Addressing
 - Sektory na disku se číslují lineárně
 - Sektory se číslují postupně od 0
 - Velikost bloku je 512B
 - Délka adresy je 28b → možné za adresovat 2^{28} sektorů



- Cluster

- Více bloků za sebou
- Nejmenší adresovatelná jednotka na logickém disku, kterou může OS vyčlenit pro soubor (adresář)
- Velikost clusteru se odvíjí od velikosti disku a použitého FAT systému (12, 16, 32)
- Celkový počet bloků v clusteru lze vypočítat jako Kapacita logického disku / velikost jednoho bloku
- Počet bloků v clusteru lze vypočítat jako Celkový počet bloků / $2^{\text{šířka záznamu ve FAT}}$ (v případě FAT32 to je 2^{32})
- Cluster se skládá z určitého počtu sektorů a obsahuje data souborů nebo adresářů
- Každý cluster má právě jeden jedinečný záznam ve FAT tabulce v datové oblasti
- Čísluje se od 2, jelikož 0 značí prázdný cluster a 1 byl historicky vyhrazen pro systém

2. Zaváděcí záznam, hlavní adresář, typy FAT, struktura podadresářů

- Zaváděcí záznam

- Skládá se ze 2 částí
 - Spouštěcí kód svazku
 - Blok parametrů disku
- Obsahuje krátký program spuštěný při startu PC BIOSem
- Jeho úkolem je načíst tabulku oblastí a najít aktivní oblast, ze které se načte OS
- Vytváří se rozdělením disku na oddíly
- Tvoří základ logické struktury, je součástí MBR
- Záznam je umístěný na začátku disku (CHS 001)

- Hlavní adresář

- Nachází se za kopií FAT2 (nachází se hned za FAT1 a je její kopií)
- Obsahuje informace o uložených souborech
 - Jméno souboru/adresáře
 - Přípona
 - Velikost, aby se nemusela pokaždé počítat
 - Datové známky (vytvoření, editace, poslední přístup)
 - Atributy (vlastnosti pro čtení/zápis, skrytý soubor)
 - Číslo 1. clusteru, ve kterém soubor/adresář začíná
 - Kontrolní součet

3. Typy FAT, rozdíly FAT, VFAT 16 a VFAT 32, řešení dlouhých názvů (6. Rozdíly FAT, VFAT 16 a VFAT 32)

- FAT 12

- Používá konvenci 8.3 (max 8 znaků název a max 3 znaky přípona)
- 12bit
- Používán na disketách
- Bez podpory podadresářů
- Jeden záznam tabulky má šířku 12b
- Celkový počet záznamů je tedy 2^{12}

- **FAT 16**

- Používá konvenci 8.3
- 16bit
- Podpora podadresářů

- **FAT 32**

- Používá konvenci 8.3
- Rozšířená verze FAT a VFAT
- Má omezení velikosti souboru na 4GB a oddílu 8TB

- **VFAT (Virtual File Allocation Table)**

- Podpora dlouhých jmen (delší než původních 8 znaků a 3 znaky přípon) a 4 znaků přípony
 - VFAT 32 se liší umístěním kořenového adresáře do datové oblasti do 2 clusterů
- Toho bylo dosaženo tak, že každý soubor s dlouhým názvem obsahoval v seznam souborů v adresáři více záznamů (každý o velikosti 13 znaků) umístěných bezprostředně za sebou, přičemž až poslední záznam je standardním záznamem typu FAT. Předchozí záznamy mají příznaky souboru nastavené tak, aby byly staršími operačními systémy ignorovány (Těm se většinou zobrazoval zkrácený název uložený v posledním záznamu)
 - Takže jeden název souboru mohl zabrat až 21 záznamů
 - Kořenový adresář má limit 512 souborů (v případě použití jednoho záznamu na soubor)
 - Problémem je ale, že souborů s maximální délkou názvu byste do kořenového adresáře dostali jen 24. Proto by se měly dlouhé názvy v kořenovém adresáři používat zřídka. (ostatní adresáře takové omezení nemají)
- možnost až 255 znaků včetně mezer a teček

- **exFAT**

- vytvořený Microsoftem speciálně pro flash disky, SSD disky
- Počet bloků se svazky až $2^{64}-1$ (16EiB – Exbibyte)
- Zvýšený výkon přidělování volného prostoru díky bitmapy
- Podpora pro transakce
- Clustery je možné označit jako špatné a tím zabránit zápisu na ně
- možnost až 255 znaků

- **NTFS**

- New Technology File System
- Umí s ním pracovat pouze OS Windows a zřídka ho podporují jiné OS, takže je vhodný pouze na interní disky
- Oproti exFAT a FAT32 umí podporu oprávnění k souborům a vede si zápis o změnách do souboru (žurnalování).
 - V případě chyby se OS vrátí v žurnálu na poslední stabilní verzi a pokusí se znovu provést operaci
 - V případě, že se mu nepodaří převést uživatelské soubory, tak je smaže
- Maximální velikost souboru je 16TB
- Používá MFT – Master File Table
 - V případě, že byl soubor dostatečně malý, jsme schopni jej uložit přímo do ní

- Relační databáze obsahující záznamy o všech souborech, adresářích a metadatech a jelikož je MFT soubor, tak i o sobě samotné.
- Nachází se hned za boot sektorem
- Aby se předešlo fragmentaci tabulky, systém kolem něj udržuje zónu volného místa

• Struktura podadresářů

- Kořenový adresář je vytvořen uživatelem nebo OS
 - je umístěn v organizační oblasti
 - Má tzv. 0 velikost a atribut DIR
- Podadresář je vytvořen uživatelem nebo OS v prvním volném clusteru
- Záznam pro cluster ve FAT se změní z 0 na EOF a podadresář vždy po založení zabere 1 cluster i když má v root dir 0 velikost
- V tomto clusteru se ihned po založení vytvoří první 2 záznamy
 - .a..
 - Tečka → číslo 1 clusteru ukazuje na daný adresář (znovunačtení)
 - Dvě tečky → číslo 1 clusteru ukazuje na nadřazený adresář (přechod o úroveň výš)

Příklad struktury: (čísla v závorkách jsou čísla clusterů ve kterých jsou adresáře uloženy)

```

— C:\SPSE (100)      \sut (101)
                     \priz (102)      \sekret (103)
                                      \knih (104)
                                      \jaz (105) \soubor.txt 1 KB (106)
  
```

```

C:\ SPSE\priz
.          DIR      102 (č. cl. daného adresáře – znovunačtení)
..         DIR      100 (č. cl. nadřazeného adresáře )
sekret     DIR      103
knih       DIR      104
jaz        DIR      105
  
```

```

C:\ SPSE\priz\jaz\
.          DIR      105 (č. cl. daného adresáře – znovunačtení)
..         DIR      102 (č. cl. nadřazeného adresáře )
soubor.txt 1 KB    106
  
```

4. Chyby FAT systémů (programy k tomu určené a způsoby jejich odstraňování)

- Ztracený cluster = na cluster neodkazuje žádný jiný záznam ve FAT tabulce, i když v něm jsou data
- Překřížený cluster = Na jeden cluster odkazují dva záznamy ve FAT tabulce
- Poškozený FAT = Pokud je souboru přiřazen blok několika clusterů, avšak ukazatel v některém z těchto clusterů ukazuje na konec disku nebo oddílu
- Fragmentace

- Fragmentovaný soubor je takový, který není uložený do řetězce clusterů následujících za sebou – je rozházený po disku (leží na několika různých cylindrech).
- Takový soubor bude z disku načítán pomaleji.
- Program pro záchranu dat má menší šanci opravit případné chyby vznikající při zápisu správně
- Fragmentace vzniká častým mazáním a zápisem nových souborů, které jsou delší než uvolněné místo po těch vymazaných

5. Bezpečné odstraňování dat, fragmentace a defragmentace disku

- Je rozdíl mezi vymazáním souboru a bezpečným odstraněním souboru z disku
 - Vymazáním souboru se pouze přepíše první znak názvu na znak E5h a OS pak dále přepíše celý řetězec záznamu ve FAT tabulce číslem 0
- Bezpečné odstranění dat – skartování
 - Neprovádí jej OS, ale speciální programy
 - Dojde nejen k přepsání prvního znaku názvu souboru, ale taky k odstranění celého záznamu názvu v adresáři k přepsání obsahu souboru novým obsahem
 - Metody skartování dat
 - Rychlá – přepis 1x nějakým novým obsahem (nulou)
 - Bezpečné podle amerického MO (Ministerstvo Obrany) – Přepis 3x nulou, 3x Ffh a nakonec F6h a celý proces se opakuje několikrát.
 - Gurmanova metoda – extrémně bezpečná, ale dost pomalá
- Defragmentace
 - Defragmentace znamená, že program spojí jednotlivé fragmenty souboru do jednoho celku tím, že jej přesune na místo, kam se soubor vleze celý
 - Defragmentační programy
 - V OS je to defragmentace
 - O&O Defrag, Diskeeper

7. Metasoubory a jejich funkce, struktura logického disku NTFS

- Metasoubory a jejich funkce
 - Prvních 16 metasouborů jsou speciální systémové soubory zodpovědné za systémové operace. Metasoubory jsou v rootu NTFS, jejich název začíná znakem \$, takže je obtížné pomocí běžných prostředků o nich získat nějaké informace.
 - \$MFT – samotná tabulka
 - \$MFTmirr – kopie prvních 16 záznamů MFT umístěna do středu disku
 - \$LogFile – Soubor pro protokolování
 - \$Volume – interní informace – název oddílu, verze systému souborů, atd.
 - \$Boot – boot sektor (bootovatelný oddíl)
- Struktura logického disku NTFS
 - Podporuje rekonstrukci poškozených souborů
 - Podporuje automatickou kompresi dat a podporu diskových kvót a šifrování
 - Je symbolicky rozdělen do dvou částí. Kromě uživatelských dat přidává NTFS i metadata. Ta jsou ukládána do 16 speciálních skrytých souborů. Jedním z nich je i Master File Table.