

Operációs rendszerek

A Windows NT file-rendszer (NTFS)

NTFS: Windows NT File System

Elvárások az NTFS-sel szemben

- Megbízható file-rendszer, visszaállíthatóság (*recoverability*).
- Állományok biztonságának garantálása, illetéktelen hozzáférés megakadályozása (*security*).
- Hibatűrés, redundáns tárolás lehetősége (*fault tolerance*).
- Nagy diszkek és nagy file-ok tárolásának lehetősége.

Megbízható file-rendszer

Megvalósítása:

- Redundáns tárolása a fontos adatoknak.
- Transaction processing
(műveletenkénti feldolgozás).

Redundáns tárolás

- A rendszer duplikálja a fontos adatállományokat a lemezen.
- A lemez egymástól távoli régióiban tárolja a másolatokat.

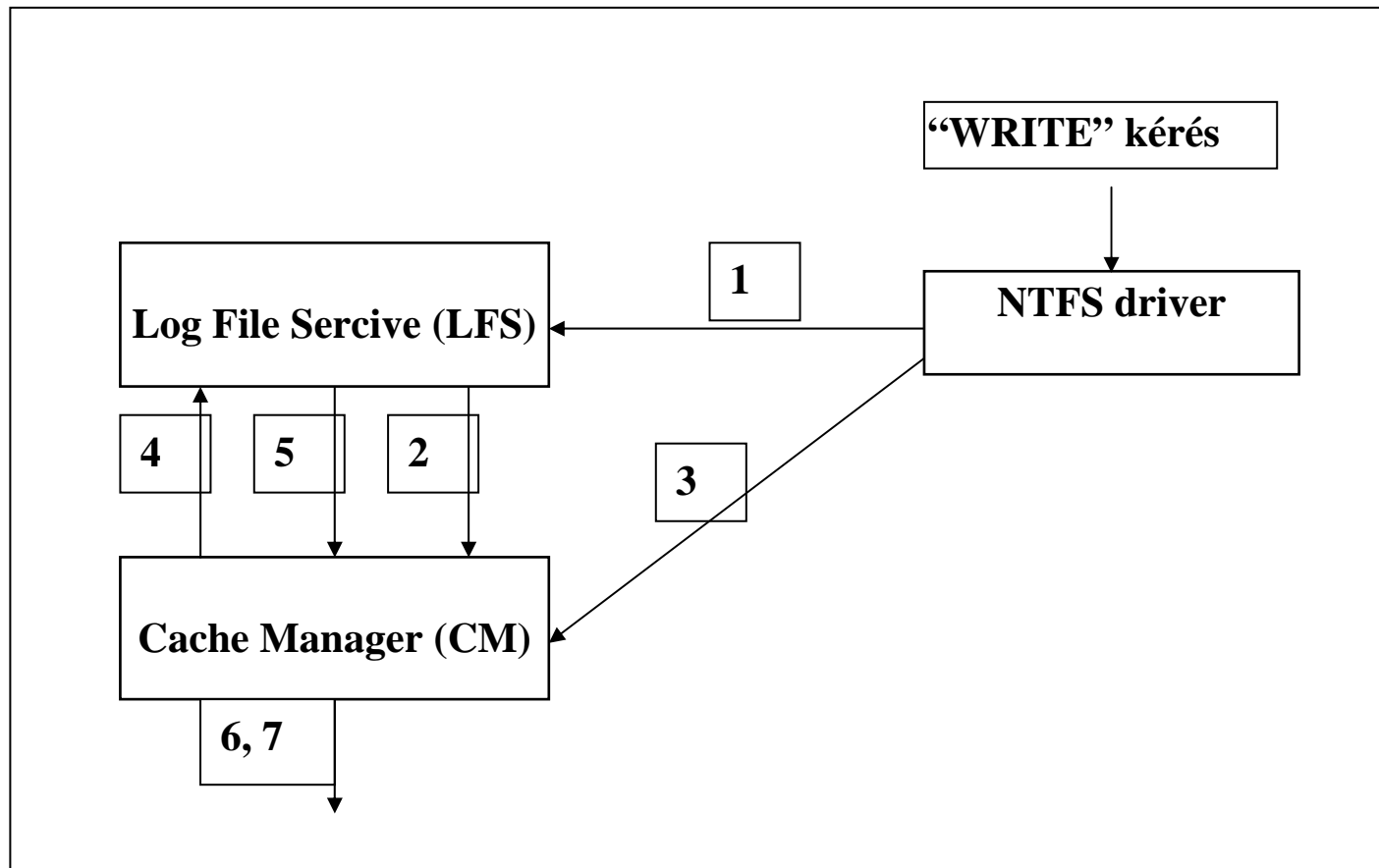
Transaction processing (műveletenkénti feldolgozás)

- "Mindent vagy semmit" elvre épül:
 - egy lemezművelet teljesen végrehajtódik vagy,
 - ha megszakadt, a rendszer visszaállítja a file-rendszer eredeti állapotát.
- Így mindig konzisztens a file-rendszer a lemezen.

Működés

- A file-rendszeren végzett műveletek lépésekre bontása.
- Redundáns adminisztrációs lépések beiktatása.
- Lépések olyan sorozata:
 - hogy csak konzisztens adatok vannak a lemezen,
 - vagy visszaállítható a kezdeti állapot.

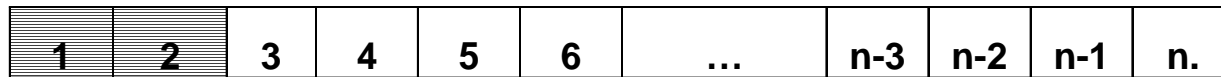
Egy lemezírási művelet kiszolgálása I.



Egy lemezírási művelet kiszolgálása II.

1. NTFS \rightarrow LFS : készítsen log rekordot.
2. Az LFS írja a cache-ben lévő log file-t.
3. Az NTFS végrehajtja a kért utasítást, írja a (cache-ben lévő) file-t.
4. A CM üzen, hogy az írás befejeződött, minden adat megvan.
5. Az LFS \rightarrow CM: milyen adatokat kell a cache-ből üríteni. (A megváltoztatott file-t + log file-t.)
6. A CM kiírja a lemezre a log file-t.
7. A CM kiírja a lemezre az adatokat, vagyis a megváltoztatott file-t.

A log file szerkezete



Újraindítási
információ
két
példányban.

Körkörösén íródó tömb.
Tranzakciónkénti log rekord.
(Újbóli elvégzéshez ill. a visszavonáshoz szükséges
információ.)

Réteg szerkezetű device driver struktúra használata

- Nagy file-rendszerek tárolása:
 - több lemezen tárolt file-rendszer (multivolume file system).
- Hibatűrő tárolás:
 - az NT által támogatott technikák:
 - RAID level 1: lemezek tükrözése,
 - RAID level 5: lemez-szeletek paritásos védelme.
- A RAID a Redundant Array of Inexpensive Disks (olcsó lemezegységek redundáns tára) rövidítése.
- Hibatűrő tárolás kommersz elemekkel.

Az NTFS további előnyös tulajdonságai I.

- Stream-ek használata:
 - egy file-hoz több függetlenül elérhető „adatfolyam” (stream) rendelhető,
 - ezek külön-külön írhatók és olvashatók.
- File nevek:
 - a UNICODE használata,
 - hossz: maximum 255 karakter,
 - szóközők, illetve pontok használata.

Az NTFS további előnyös tulajdonságai II.

- Az indexelés lehetősége:
 - adott tulajdonság alapján közvetlen hivatkozás (index buffer) az egyes file-okra.
- Dinamikus hibás (bad) szektor kezelés:
 - futás közbeni bad cluster észlelés => tiltás.
- A POSIX szabvány támogatása:
 - hard link,
 - case sensitive file-nevek,
 - időbélyegek (time stamp-ek) használata.

Adattípusok, adatszerkezetek a tároláshoz

Az NTFS által használt adattípusok, adatszerkezetek I.

- Volume (kötet):
 - a lemez egy logikai partíciója.
- Cluster (szektorok egymásutániséga):
 - az adattárolás alapegysége.
(Az NT csak cluster-eket tart nyilván.)
- Logical Cluster Number (LCN):
 - egy adatszerkezethez, pl. egy file-hoz tartozó cluster-ek sorszáma (pl.: 0-tól n-ig folyamatosan).

Az NTFS által használt adattípusok, adatszerkezetek II.

- Virtual Cluster Number (VCN):
 - a lemezen elhelyezkedő cluster-ek azonosítója,
 - a lemezen tárolt adatszerkezetek elérése:
 - az LCN-ek VCN-ekhez történő hozzárendelése alapján.

LCN:	0	1	2	...	n.
VCN:	6324	857243	9454	...	542

Az NTFS által használt adattípusok, adatszerkezetek III.

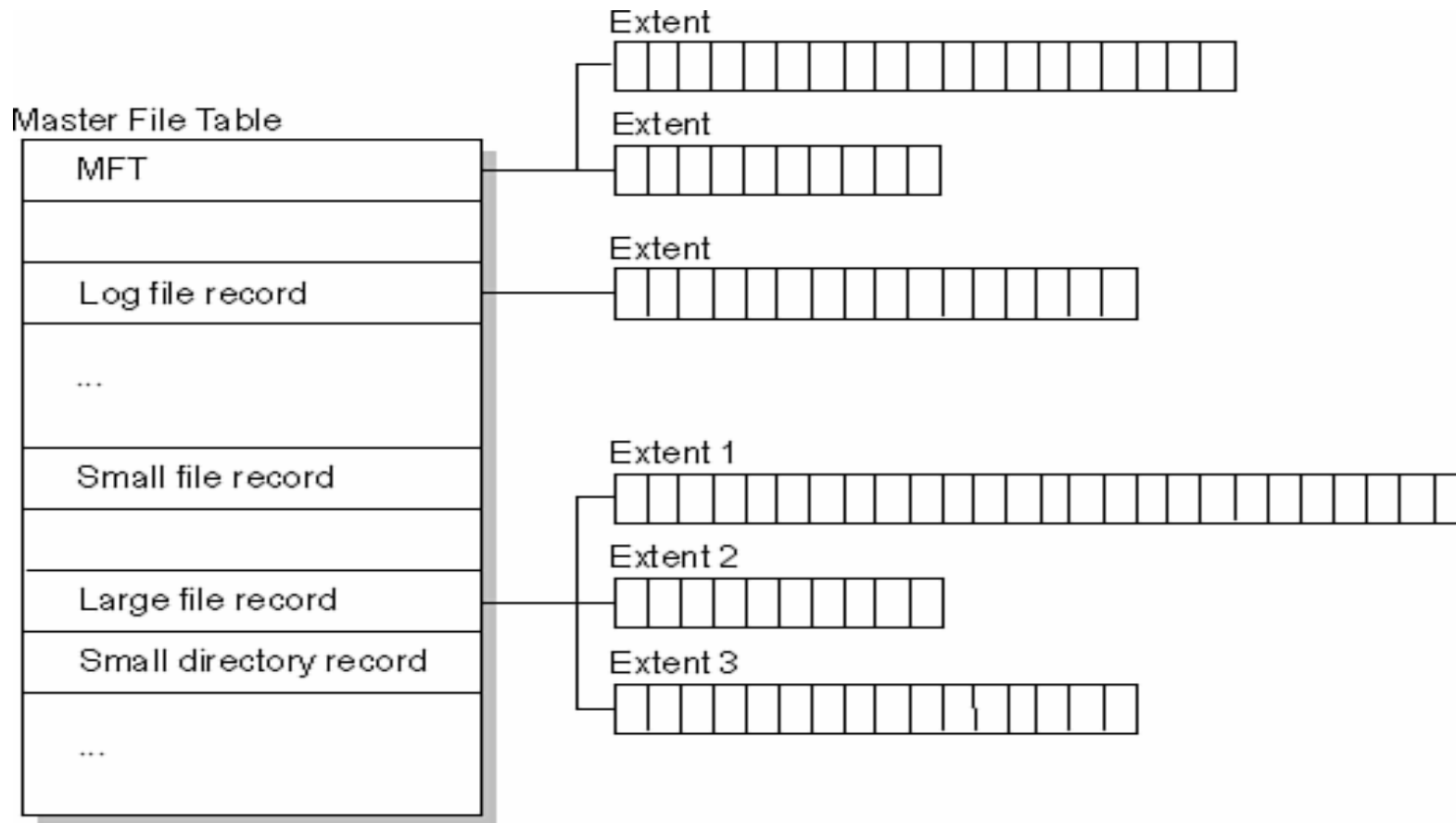
- NTFS metadata:
 - a file-rendszer kezeléséhez, ill. a file-ok eléréshez szükséges adatok gyűjtőneve,
 - elv:
minden adat a lemezen file-ként tárolódik:
 - pl.: a volume-leíró, a boot információ, a hibás szektorok leírása.

NTFS metadata információk

\$Boot file	A rendszer indulásakor használt információ.
\$Bad sector	A hibás cluster-ek sorszáma (LCN).
\$Bitmap	A cluster-ek foglaltsági térképe.
\$Logfile	A Log File Service által a lemeztranzakciók adminisztrálására használt file.
\$MFT	Master File Table.
\$MFT mirror	A Master File Table (részleges) biztonsági másolata.
\$Volume	A kötet leírását (pl.: fájlrendszer típusa) tartalmazó file.
\$Attribute	A kötet tulajdonságait, attribútumait tartalmazó file.

A tárolás logikája

A file-ok tárolása a Master File Table segítségével



MFT (Master File Table) I.

A Master File Table (MFT) a file-rendszerben tárolt file-ok leírása, elérésükhöz szükséges információ.

- Felépítése:
 - file record-ok sorozata,
 - az első 16 bejegyzést az OPR használja:
 - ezek az ún. rendszer-file-ok,
 - nevük \$ jellel kezdődik,
 - utána jönnek a felhasználói file-ok.

A Master File Table szerkezete

\$ MFT	0.
\$ MFT mirror	1.
\$ Logfile	2.
\$ Volume	3.
\$ Attrib	4.
...	...
\$...	15.
<hr/>	
USER FILE1	16.
USER FILE2	17.
...	...
USER FILEn	n.

MFT (Master File Table) II.

A könyvtárak leírása is egyszerű felhasználói file-okban tárolódik:

- táblázatszerűen: a könyvtárban levő file-ok nevének és a file-okhoz tartozó MFT-bejegyzés sorszámának összerendelése,
- a gyökérkönyvtár MFT- bejegyzés számát a fix helyen elérhető rendszer file-ok tartalmazzák.

A file-okhoz tartozó adatok tárolása az NTFS-ben

- A file: egymással összerendelt adatok (jellemzők, attribútumok) halmaza.
- Attribútumok pl.:
 - a file neve,
 - a keletkezési időpontja,
 - az elérhetősége,
 - a file tartalma.
- Az attribútumoknak minden file esetén meghatározott értéke van.

Megvalósítás I.

- A file-ok az ún. file record-okban tárolódnak.
- A file record:
 - a file attribútumok (file attribute) azonosítója (neve),
 - és az attribútumhoz tartozó adatmező.
- Az MFT = file record-ok sorozata.
- Az MFT csak a file record első 1K-s darabját tartalmazza.

Megvalósítás II.

Az MFT egy táblázat:

minden bejegyzés 1K-s:

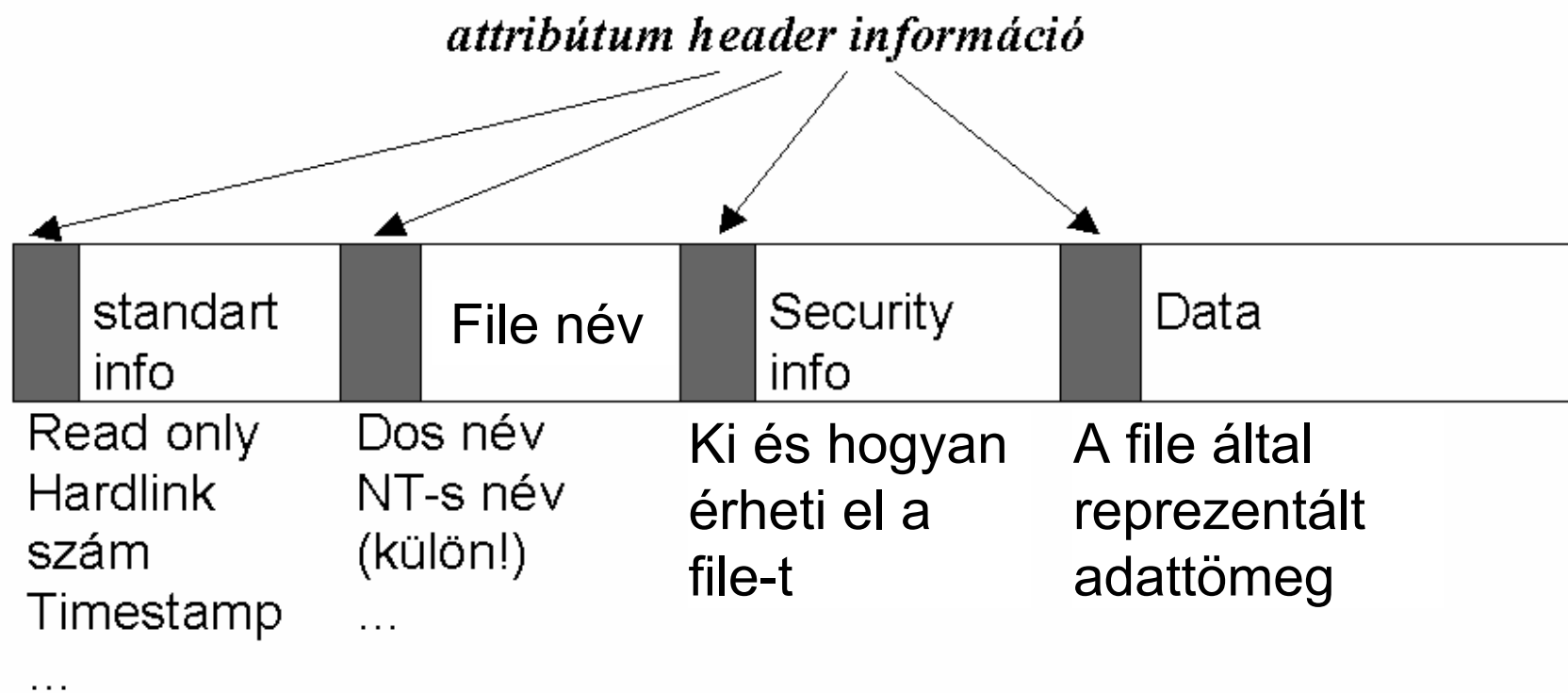
- minden MFT bejegyzés egy file-t azonosít,
- a file egyedi azonosítója:
a hozzá tartozó MFT bejegyzés sorszáma.
- a bejegyzések tartalma:
minden file-hoz tartozó információ,
- szerkezetük nem kötött.

A file record-ok tárolása

File record

- Az attribútumok értékének tárolása:
 - rezidens és
 - nem rezidens módon.
- Tárolás típusának meghatározása:
 - az attribútum értékét leíró adatok fizikai elhelyezkedése alapján.

Egy tipikus file record



A file attribútumainak tárolási módjai

- Rezidens tárolás:
 - az attribútum értéke közvetlenül az attribútum header (RESIDENT) után.
- Nem rezidens tárolás:
 - attribútum header után egy hivatkozás, hogy a lemez mely clustere tartalmazza az attribútum értékét,
 - ha az összes header nem fér bele az MFT rekordba, akkor egy új MFT bejegyzés jön létre, ezt egy speciális attribútum rögzíti, az ún. attribútum lista.

Rezidens tárolás

Általános felépítés:

attribútum (adat)	file név header	file név (adat)	security header	security (adat)
----------------------	--------------------	-----------------	--------------------	--------------------

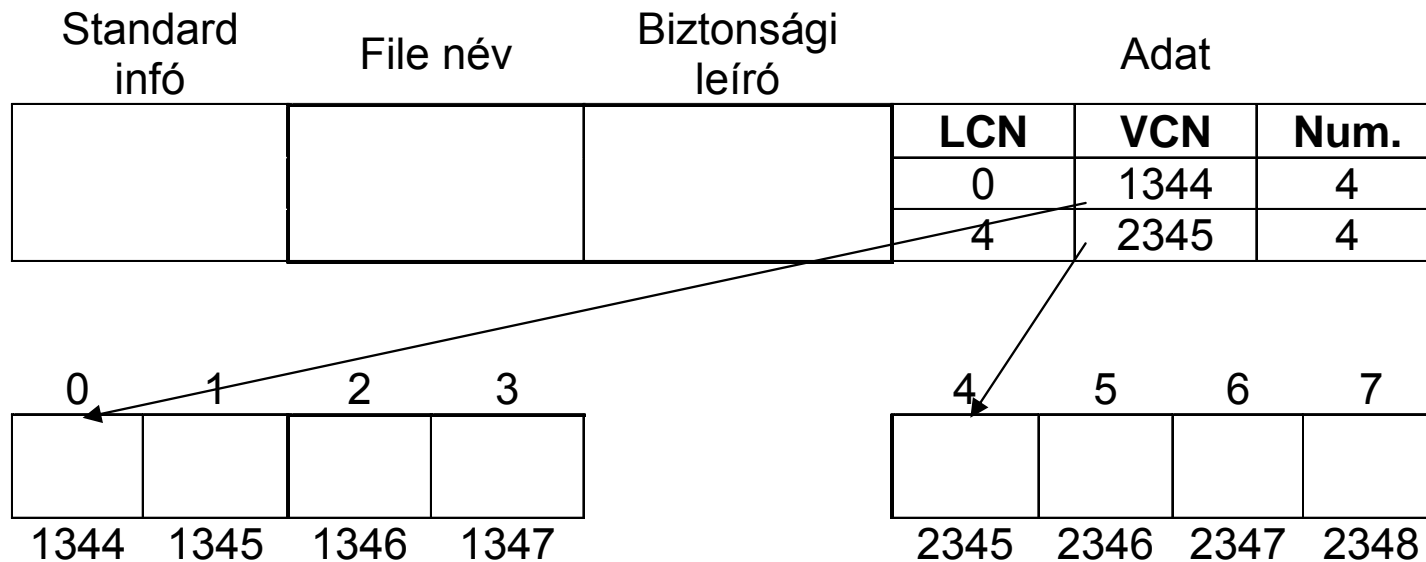
Egy konkrét példa: (az adatelemek fölött a lefoglalt byte-ok száma látható)

	0.....7	8.....21	22.....29	30 ...
	“RESIDENT” 8h (offset) 14h (length)	“MYFILE.DAT”	“RESIDENT” 8h (offset) 25h (length)	

Nem rezidens tárolás

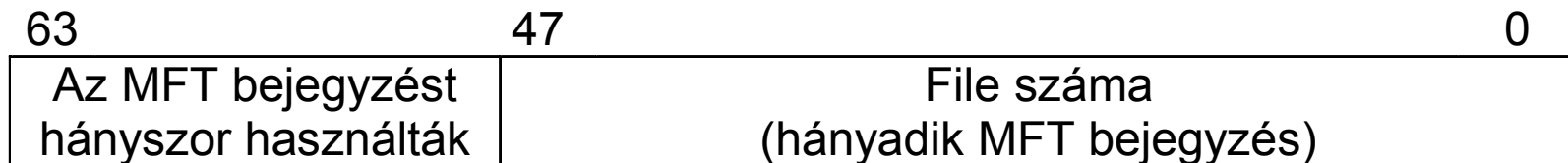
- A header jelzi a nem rezidens tárolás tényét (NONRESIDENT).
- A header után csak az attribútum értékét reprezentáló adatok helye van tárolva.
- Pl. az adatokat tartalmazó buffer címét tárolja a rendszer a rekordban.
- Adattárolás 2K-s (4K-s) buffer-ekben (run vagy extent).

Adat attribútum nem rezidens tárolása



Állományok azonosítása

- A file-ok azonosítása:
 - az ún. file referencia számmal (64 bit):
 - hányadik MFT rekord (48 bit),
 - az adott rekord helyet hányszor használták (16 bit). Így a maximális megkülönböztethető használati lehetőség 65.536 alkalom lehet.

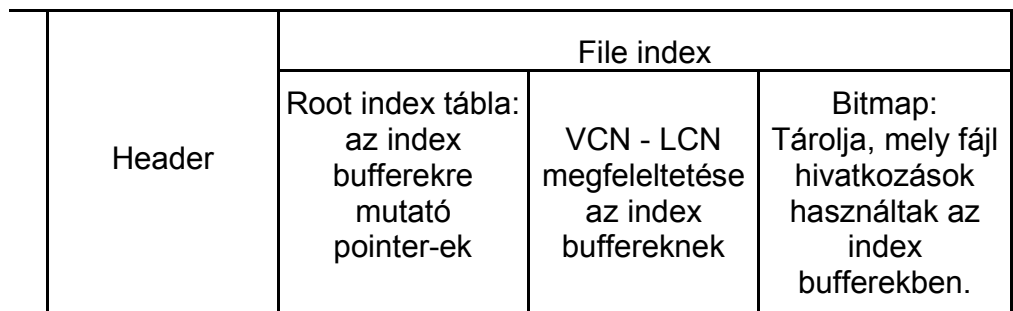


Könyvtár leírás rezidens tárolása

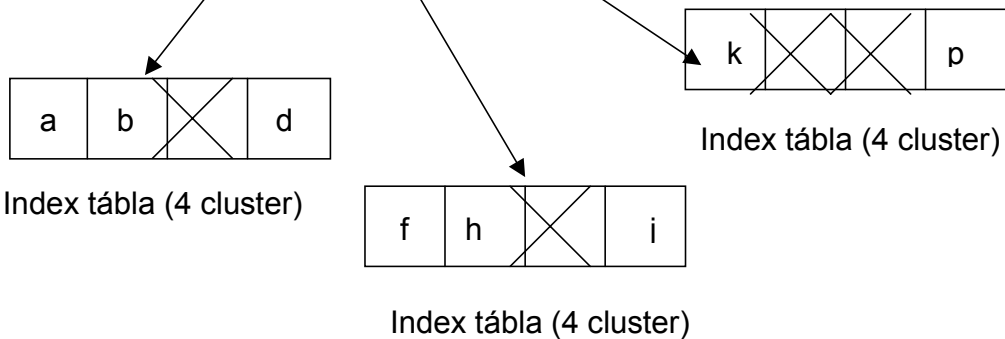
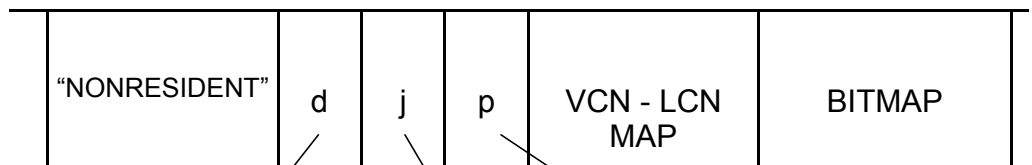
Standard infó	File név	Biztonsági leíró	Index root		
			file1	subdir1	file3
			referencia	referencia	referencia
			időbélyeg	időbélyeg	időbélyeg

Nem rezidens tárolás

Általános felépítés:



Konkrét példa:



Könyvtárak indexelt tárolása I.

- A header jelzi a nem-rezidens tárolás tényét.
- A file-ok tárolása: név szerint rendezve
- Root indextábla:
 - mely bufferek-ben milyen nevű file-ok találhatóak.
- VCN – LCN megfeleltetési táblázat:
 - tényleges helye a buffer-eknek.

Könyvtárak indexelt tárolása II.

- Bitmap:
 - táblázat: mely cluster-ek használtak és mely cluster-ek nem használtak a buffer-ekben.
- Struktúra: $b + fa$.
- Dinamikusan növekvő indextáblát használó egyszintű indexelés.

File-ok elérése

A file-ok elérése NTFS alatt I.

A lépések többségét csak a kötet (volume) rendszerindulás (boot-olás) utáni első elérésekor kell végrehajtani.

1. A \$Boot file elérése (a helye rögzített).
2. \$MFT elérése:
 - a \$Boot tartalmazza az MFT kezdetének a helyét,
 - a \$MFT mindig az első MFT-rekordhoz tartozik, a \$MFTMirr a másodikhoz, a \$LogFile a harmadikhoz, egészen a 15. rendszer file-ig.

A file-ok elérése NTFS alatt II.

3. \$MFT (\$MFTMirr) elérése, memóriába tárolása.
4. \$LogFile és más metafile-ok beolvasása.
5. A kötet rendszerindulás utáni első elérésekor végrehajtja az ún. recovery műveletet:
 - recovery: a file-rendszer konzisztens állapotának ellenőrzése, szükség esetén helyreállítása,
 - a logfile alapján ellenőrzi a rendszer, hogy kell-e tranzakciót "visszagörgetni" vagy újra végrehajtani,
 - recovery után a lemez konzisztens állapotba kerül.

A file-ok elérése NTFS alatt III.

6. A root “\” könyvtárhoz tartozó MFT-bejegyzés megkeresése, ill. a memóriában tárolása (későbbi gyorsítás).
7. A tényleges file-tranzakció végrehajtása.
Minden tranzakció bekerül a logfile-ba.

File attribúumok

File attribútumok I.

- Standard információ:
 - a file jellemzői, pl.: read only, archive,
 - időbélyegek: pl.: létrehozás, utolsó változtatás,
 - hard linkek száma,
 - mindig rezidens.

File attribútumok II.

- File név (\$FILENAME):
 - több lehet belőle:
 - POSIX hard link,
 - NT-s név (UNICODE!),
 - DOS-os név (automatikusan származtatott),
 - maximum 255 karakter,
 - névterek viszonya: POSIX - NT- DOS,
 - mindig rezidens.

File attribútumok III.

- Biztonsági (security) információ:
 - tulajdonos,
 - felhasználók elérési jogosultságai.
- Adat (\$DATA):
 - adat file: "default unnamed data attribute" (név nélküli adat),
 - több stream esetén néhány elnevezett adat attribútum,
 - könyvtárnak nincs adat attribútuma.

File attribútumok IV.

- Index root:
 - csak könyvtár leíró file esetén,
 - index attribútum a könyvtárbeli file-ok indexelésére,
 - mindig rezidens.
- Attribútum lista:
 - csak az egynél több MFT bejegyzést használó file-ok esetén,
 - a file attribútumainak leírása.

File attribútumok elérése

- Felhasználó:
 - attribútumokat ír és olvas,
 - szimbolikusan (\$NÉV) hivatkozik.
- Rendszer:
 - sorba rendezi és sorban tárolja az attribútumokat.

Biztonság

- Minden állomány-hozzáféréskor elérési jogosultság ellenőrzés.
- Az NT általános objektumelérést ellenőrző biztonsági rendszere.

Nagy lemezek kezelése

- 64 bit-es cluster-leíró használata.
- Elméleti korlátok:
 - egy file-rendszerben összesen 2^{48}
(281.474.976.710.656 = $\sim 281,5 \cdot 10^{12}$, azaz
 $\sim 281,5$ billió) db file lehet "egyidőben",
 - egy file mérete maximum 2^{64}
(18.446.744.073.709.551.616 = $16 \cdot 10^{18}$,
azaz 16 EB, exa= 10^{18}) byte lehet,
 - azaz file-rendszer maximális mérete:
 $2^{48} \cdot 2^{64} = 2^{112}$ byte, azaz $5,2 \cdot 10^{33}$ byte,
(azaz 5,2 kvintilliárd byte, azaz
4.294.967.296 Yotta byte (Tera-Peta-Exa-Zetta-Yotta)