

Operativni sistemi

Milena Vujošević-Janičić

www.matf.bg.ac.rs/~milena

Arhitektura i operativni sistemi
Beograd, 24. decembar, 2013.

Fajl sistem

- Fajl sistem je sistem koji omogućava skladištenje, hijerarhijsku organizaciju, manipulaciju, navigaciju i pristup podacima.
- Linux podržava veliki broj fajlsistema.
- Moguć je istovremen rad sa više različitih fajlsistema koji su predstavljeni u okviru jedinstvenog drveta koje služi da korisniku predstavi sve dostupne fajlove na uniforman način bez obzira na karakteristike fajlsistema na kome se oni nalaze.

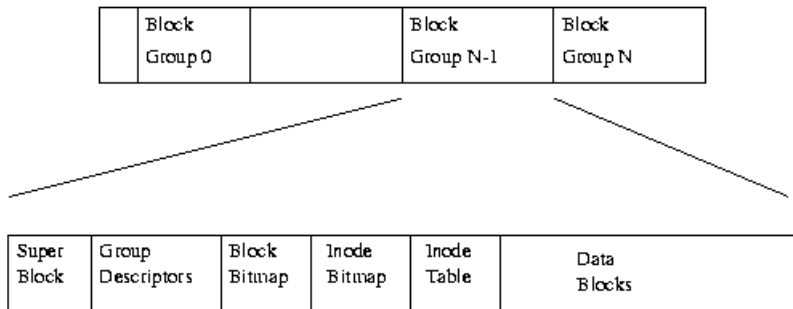
ext2

- Jedan od najpopularnijih fajlsistema je **ext2** nastao 1993. na osnovu **ext** fajlsistema.
- Slično drugim fajlsistemima pretpostavlja da se podaci u fajlovima čuvaju u okviru blokova podataka iste dužine.
- Blok je najmanja jedinica podataka i veličina fajla se zaokružuje naviše na ceo broj blokova.
- U proseku se zbog ovoga gubi pola bloka po fajlu.

ext2

- Sa tačke gledišta fajlsistema, svaki uređaj je niz blokova koji se mogu čitati i pisati.
- O fizičkom rasporedu ovih blokova njihovom pristupanju stara se drajver uređaja od koga fajlsistem po potrebi traži određeni broj blokova.
- ext2 deli particiju na kojoj se nalazi na grupe blokova.

ext2



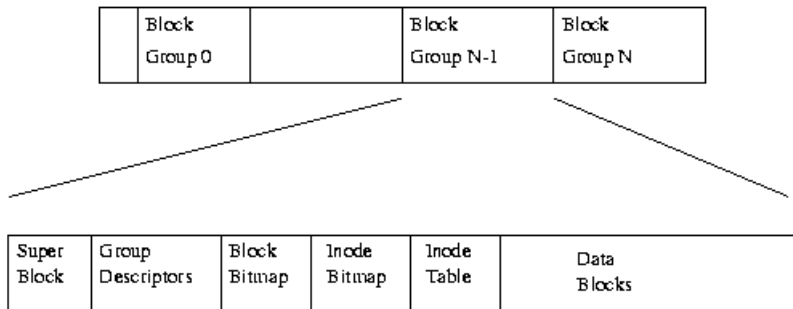
Superblok

- Superblok sadrži osnovne informacije o fajlsistemu koje su potrebne za njegovo održavanje i upravljanje.
- Obično se čita superblok grupe 0, ali je on umnožen i u ostalim grupama radi povećanja pouzdanosti.

Superblok

- Neki od podataka koji se čuvaju u superbloku su:
 - Veličina bloka,
 - Broj blokova po grupi,
 - Broj slobodnih blokova u fajl sistemu,
 - Broj slobodnih inodova u fajl sistemu,
 - Broj prvog inoda na fajlsistemu (koreni direktorijum).
 - Trenutni i maksimalni broj vezivanja (mount) koji se pamte zbog automatskih provera fajlsistema.
 - ...

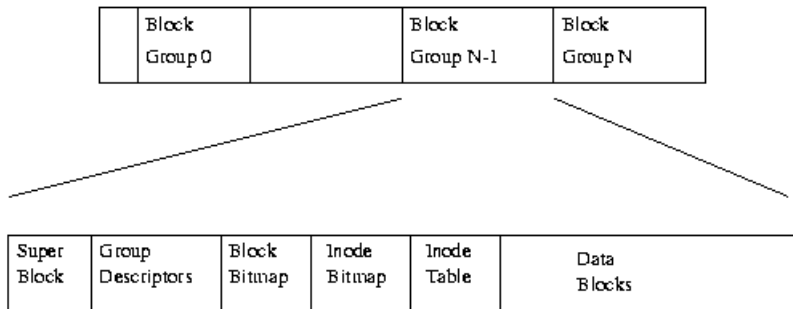
ext2



Deskriptori grupa

- Svaka grupa blokova ima strukturu koja je opisuje — **deskriptor**.
- Svi deskriptori grupa čine tabelu deskriptora grupa koja je duplirana u svim grupama (kao i superblok), a koristi se tabela deskriptora grupe 0.
- Deskriptor čuva podatke o rasporedu slobodnih i zauzetih blokova i inodova (tj adrese odgovarajućih bitmap-a), adresu tabele inodova, broj slobodnih blokova, broj slobodnih inodova, broj korišćenih direktorijuma.

ext2



Block bitmap, Inode bitmap, Inode table

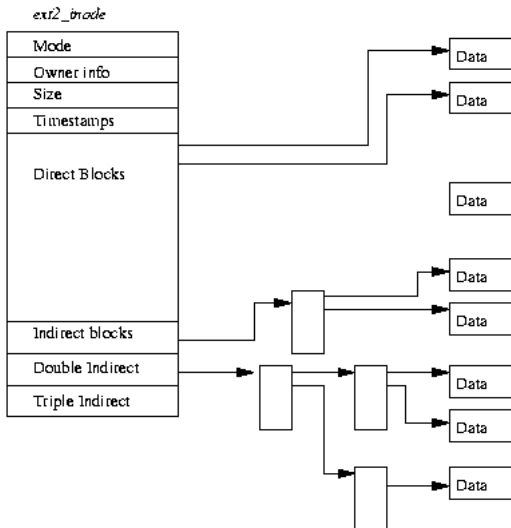
- ext2 prati slobodan prostor koristeći bitmape
- **Block bitmap** prati slobodne blokove
- **Inode bitmap** prati slobodne inodove
- **Inode table** sadrži listu svih inodova u grupi bloka

Inode

- Inode je osnovni gradivni blok ext2 sistema.
- Svaki fajl i direktorijum sistema je opisan jednim inodom.
- Inode se identifikuje svoji brojem — `ls -li`

Inode

- Inode sadrži podatke o
 - Tipu fajla: običan fajl, direktorijum, simbolička veza, blokovski/karkaterski specijalni fajl...
 - Grupi i vlasiniku fajla, pravima pristupa
 - Veličini fajla i vremenima (timestamps)
 - Broj inoda
 - Pokazivači (15) na blokove podataka (prvih 12 su direktni, 13 je indirektan, 14 je duplo a 15 trostruko indirektan).



Inode

Pretpostavimo da je realna veličina sistemskog bloka (1KB) i da su pokazivači 32-bitni. Jedan indirektni blok može sadržati najviše $1\text{KB}/4\text{B}=256$ 32-bitnih pokazivača. Maksimalne veličine datoteka koje se mogu dobiti alokacijom prostora pomoću direktnih i 32-bitnih indirektnih pokazivača:

- samo sa direktnim pokazivačima: $12 * 1 \text{ KB} = 12 \text{ KB}$
- + indirektni pokazivač: $12 \text{ KB} + 256 * 1 \text{ KB} = 268 \text{ KB}$
- + dvostruki ind. pokazivač: $268 \text{ KB} + 256^2 * 1 \text{ KB} \approx 64 \text{ MB}$
- + trostruki ind. pokazivač: $64 \text{ MB} + 256^3 * 1 \text{ KB} \approx 16 \text{ GB}$

Inode

Pretpostavimo da je realna veličina sistemskog bloka (8KB) i da su pokazivači 32-bitni. Jedan indirektni blok može sadržati najviše $8\text{KB}/4\text{B}=2048$ 32-bitnih pokazivača. Maksimalne veličine datoteka koje se mogu dobiti alokacijom prostora pomoću direktnih i 32-bitnih indirektnih pokazivača:

- samo sa direktnim pokazivačima: $12 * 8 \text{ KB} = 96 \text{ KB}$
- + indirektni pokazivač: $96\text{KB} + 2048 * 8 \text{ KB} \approx 16 \text{ MB}$
- + dvostruki ind. pokazivač: $16 \text{ MB} + 2048^2 * 8 \text{ KB} \approx 32 \text{ GB}$
- + trostruki ind. pokazivač: $32 \text{ GB} + 2048^3 * 8\text{KB} \approx 64 \text{ TB}$

Inode

- `stat ime_fajla` — prikazuje neke podatke koji se nalaze u inodu
- I specijalni fajlovi imaju inode, ali nemaju podatke na dsiku, već inodove koriste za pristup uređajima.

Direktorijumi

- Direktorijumi su specijalni slučajevi fajlova. Sadržaj ovog fajla je lista fajlova koji se nalaze u direktorjumu, tj u ext2 unosi sadrže sledeće informacije:
 - Broj inoda fajla
 - Dužina unosa u bajtovima (inode — 4 bajta, ovo polje — 2 bajta, dužina imena fajla — 1 bajt, tip fajla — 1 bajt, ime fajla — promenljiv broj bajtova)
 - Dužina imena fajla
 - Tip fajla
 - Ime fajla

Direktorijumi

Prva dva unosa su uvek "." (ovaj direktorijum) i ".." (prethodni direktorijum).

inode	rec_len	name_len	file-type	name							
13	12	1	2	.	\0	\0	\0				
2	12	2	2	.	.	\0	\0				
18	16	5	2	M	u	s	i	c	\0	\0	\0
15	16	8	1	t	e	s	t	.	t	x	t
19	12	3	2	s	r	c	\0				

Fajlsistemi sa evidencijom transakcija (journaled)

- Ovi fajlsistemi evidentiraju sve operacije koje se vrše na disku (tzv. transakcije) u poseban log fajl.
- Prvo se upisuje transakcija, pa se vrši stvarna operacija.
- U slučaju pada sistema, operacije koje su polovično izvršene se mogu ponovo izvršiti na osnovu evidencije iz log fajla.
- [ext3](#), [ext4](#)

Virtuelni fajlsistem (VFS)

- Virtualni fajlsistem je mehanizam predstavljanja različitih fajlsistema korisniku na uniforman način.
- Korisnik ne mora da zna ništa o karakteristikama i lokaciji fajlsistema na kome se nalazi željeni fajl ukoliko je on vezan na virtualni fajlsistem.

Virtuelni fajlsistem (VFS)

- VFS ima sličnu organizaciju kao i ext2. Takođe koristi superblokove i inodove.
- Prilikom inicijalizacije operativnog sistema fizički fajlsistemi se registruju kod VFS koji čita njihove superblokove i konstruiše adekvatan VFS superblok koji obuhvata strukturu fajlsistema koji su na njega vezani.
- Takođe i inodovi realnih fajlsistema se moraju preslikati u VFS inode.

- proc fajlsistem demonstrira snagu VFS-a.
- Iako se on registruje kod VFS prilikom podizanja operativnog sistema, fajlovi i direktorijumi prikazani u proc fajlsistemu ne postoje fizički već se njihov prikaz po potrebi generiše na osnovu podataka iz kernela.
- On u stvari predstavlja jednostavan način da se pristupi informacijama iz kernela.

- Slično fajlovima proc fajlsistema i specijalni fajlovi se ne nalaze stvarno na disku, već samo predstavljaju način pristupa drajverima uređaja prisutnih na sistemu.

Particije

- Tradicionalno sistemi nalik Unix-u dele diskove na particije.
- Na svakoj se može nalaziti fajlsistem, a oni se mogu kombinovati u jedinstven VFS.
- Particije predstavljaju podskupove cilindra diska.

Veze

- Veze omogućavaju pristup istom sadržaju pomoću više imena. Mogu biti
 - Tvrde (hard link) i predstavljaju obične unose u direktorijumima koji imaju broj odgovarajućeg inoda. Broj tvrdih veza se može videti pomoću `ls -l`.
 - Meke (soft link) i predstavljaju fajlove koji sadrže putanju do nekog fajla.
- Tvrde veze moraju da pokazuju na fajlove na istom fizičkom fajlsistemu, dok meke ne moraju.
- `cp -s file symLink` kreira simbolički link
`ln file hardLink` kreira tvrdi link za dati fajl

Veze

- Kreirati praznu datoteku knjiga
- Napraviti simbolički link za tu datoteku
- Proveriti sa `ls -l` (obratiti pažnju na tip datoteke)
- Napraviti tvrd link za datu datoteku pod imenom roman
- Proveriti sa `ls -l`, obratiti pažnju na broj tvrdih linkova u ovom slučaju
- Proveriti sa `ls -il` brojeve inodova za roman i knjigu
- Izmeniti sadržaj romana
- Proveriti sadržaj knjige
- Obrisati napravljene datoteke

Vezivanje fajlsistema

- Vezivanje je proces u kome se fizičke particije pripremaju za pristupanje i dodeljuju im se lokacije u okviru drveta fajl sistema.
- Komanda: `mount`
- Najčešća upotreba: `mount -t tip uredjaj direktorijum`
- Uređaj se može vezati i na više lokacija u fajlsistemu (`-bind`).
- Uklanjanje: `umount direktorijum`

Vezivanje fajlsistema

- Argument `umount` komande može biti i uređaj ako je vezan na samo jednom mestu u sistemu.
- Direktorijum ne mora biti prazan da bi se na njega vezao fajlsistem, ali njegov sadržaj će biti sakriven dok je fajlsistem vezan.
- Informacije o povezanim sistemima se pišu u `/etc/mstab` (mounted file systems table), format fajla sličan `fstab`-u.

fstab

- **fstab** sadrži statičke informacije o fajlsistemima, informacije potrebne za automatizaciju procesa povezivanja (mount) particija. Unosi ove datoteke su oblika:

```
device mount_point file_system_type options dump pass
```

- **device** — uređaj koji sadrži fajlsistem, identifikuje se sa /dev/sdXY, labelom ili sa UUID (universal unique identifier)
- **mount point** — direktorijum u okviru korenog direktorijuma preko kojeg je moguće pristupiti sadržaju uređaja
- **file system type** — tip fajlsistema (auto, vfat, ntfs-3g, jfs, reiserfs, ext2, ext3, ext4, uidf, iso9660, swap...)
- **dump** određuje da li je potrebno praviti backup fajl sistema (0 ne, 1 da), ovo se retko koristi, tj najčešće je 0.

fstab

- **fstab** sadrži statičke informacije o fajlsistemima, informacije potrebne za automatizaciju procesa povezivanja (mount) particija. Unosi ove datoteke su oblika:

```
device mount_point file_system_type options dump pass
```

- **pass** određuje redosled proveravanja uređaja prilikom podizanja sistema (0 se ne proverava, 1 proverava se prvo, 2 proverava se sledeće).
- Neke od mogućih opcija su: `auto`, `default`, `users`, `owner`, `group`, `noauto`, `nouser`, `ro`, `rw`
- **mount -a** vezuje sve fajlsisteme iz `fstab`-a osim ako imaju `noauto` opciju. Ova komanda se izvršava prilikom inicijalizacije sistema da bi povezala ove fajlsisteme.

Upravljanje particijama i fajlsistemom

- Particije se mogu kreirati korišćenjem programa kao što su `fdisk` i `cdisk`.
- Fajlsistem na particiji se kreira komandom `mkfs`. Za `ext2/ext3` možemo koristiti `mke2fs (-bcjLS)`.
- Sistem se proverava pomoću komande `fsck (-Ay)`, `e2fsck (-bcfpy)` za `ext2/ext3`, za `ext4` proverava se sa `fsck.ext4`
- Podešavanja za `ext2/ext3` pomoću `tune2fs (-cL)`.
- `df` (disk free) daje podatke o zauzetom prostoru na particiji ili o inodovima (opcija `-i`).

Swap particija

- Swap particija služi za proširenje RAM-a, onda kada nema dovoljno slobodne memorije.
- Swap particija se nalazi na disku, pa je zato značajno sporija od RAM memorije.
- Swap particija se priprema komandom `mkswap`.

Pitanja

- Šta je fajl sistem?
- Da li je pod Linux-om moguć istovremen rad sa više različitih fajl sistema? Ukoliko jeste, kako se to odražava na korisnika, ukoliko nije objasniti zašto nije.
- Navesti imena 4 različita fajl sistema?
- Šta je inode i koje podatke sadrži?
- Šta je direktorijum i koji je njegov sadržaj?
- Šta je superblok, gde se nalazi i koji superblok se koristi?

Pitanja

- Šta su deskriptori grupa i gde se nalaze?
- Šta je virtualni fajl sistem?
- Navesti primer fajlsistema sa evidencijom transakcija.
- Šta je particija?
- Kojom komandom se vrši vezivanje a kojom uklanjanje fajlsistema?
- Šta sadrži fajl /etc/fstab, a šta /etc/mtab?
- Šta je swap particija?