

Súbory a systémy súborov

Portál: edu.ukf.sk - Vzdelávací portál - Univerzita
Konštantína Filozofa, Nitra

Kurz: Operačné systémy (KI/OS/15)

Kniha: Súbory a systémy súborov

Vytlačil(a): Zuzana Pavlendová

Dátum: Streda, 1 december 2021, 17:59

Opis

5 SÚBORY A SYSTÉMY SÚBOROV

5.1 Základné pojmy

5.2 Organizácia štruktúry adresárov

5.3 Generačné súbory

5.4 Prístupové práva v OS UNIX

5.4.1 Práva k súborom

5.4.2 Práva k adresárom

5.4.3 Práva setuid, setgid

5.5 Prístupové práva v OS Novell Netware

5.6 Uzamykanie súborov a ich častí

5.7 Zdieľanie súborov

5.8 Kvóty

Obsah

Úvod

Požiadavky

5.1 Základné pojmy

5.2 Organizácia štruktúry adresárov

5.3 Generačné súbory

5.4 Prístupové práva v OS UNIX

5.5 Prístupové práva v OS Novell Netware

5.6 Uzamykanie súborov a ich častí

5.7 Zdieľanie súborov

5.8 Kvóty

Precvičte sa

Úvod

V tejto kapitole sa dozviete:

- Čo je to súborový systém operačných systémov.
- A ko sú ukladané a organizované súbory na diskoch.
- Ako sú realizované prístupové práva pre rôzne operačné systémy.
- Prečo je systém ukladania súborov na veľkokapacitných záznamových zariadeniach dôležitou súčasťou dobre fungujúceho operačného systému.

Kľúčové slová tejto kapitoly:

Systémy súborov, adresár, operácie so súbormi, prístupové metódy, graf.

Požiadavky

V tejto kapitole sa dozviete:

- Čo je to súborový systém operačných systémov.
- A ko sú ukladané a organizované súbory na diskoch.
- Ako sú realizované prístupové práva pre rôzne operačné systémy.
- Prečo je systém ukladania súborov na veľkokapacitných záznamových zariadeniach dôležitou súčasťou dobre fungujúceho operačného systému.

Po jeho preštudovaní by ste mali byť schopní:

- Charakterizovať rôzne štruktúry adresárov.
- Poznať spôsoby vytvárania štruktúr adresárov u rôznych operačných systémov.
- Porozumieť spôsobu popísania vlastností súborov.
- Popísať prístupové práva k súborom pre rôzne operačné systémy.
- Definovať princípy uzamykania súborov.
- Porozumieť typom operácií, ktoré musia obsahovať operačný systém pri práci so súbormi.

Kľúčové slová tejto kapitoly:

Systémy súborov, adresár, operácie so súbormi, prístupové metódy, graf.

Doba potrebná k štúdiu: 4 hodiny

Sprievodca štúdiom

Štúdium tejto kapitoly nie je náročné a popisným spôsobom tu naštudujete základné princípy práce so súbormi. Na štúdium tejto časti si vyhradte 4 hodiny. Po celkovom preštudovaní a vyriešení všetkých príkladov odporúčame vypracovať korešpondenčný úkol. Pod systémom súborov chápeme dátové objekty, ktoré sa uchovávajú na vonkajších pamätiach počítača. V tejto kapitole sa budeme zaoberať len základnými princípmi.

5.1 Základné pojmy

Systémy súborov:

- Súbory uchovávajú dáta a programy. Operačný systém (OS) implementuje abstraktnú koncepciu súborov správou veľkokapacitných záznamových zariadení.

Organizácia systému súborov:

- OS oddeľuje používateľa od fyzických vlastností záznamových zariadení a vytvára logickú jednotku uloženia - **súbor** . Súbory sú mapované OS na fyzické zariadenie.

- Systém súborov sa skladá z dvoch častí: - sada súborov, obsahujúcich vlastné uložené informácie,

- štruktúra adresárov, obsahujúcich informácie o súboroch.

Koncepcia súborov:

- Súbor je sekvencia bitov, bytov, riadkov alebo záznamov, ich význam je definovaný zakladateľom a používateľom súboru.

- Súbor je obvykle pomenovaný a má nejaké atribúty (napr. typ, meno zakladateľa, dĺžka, čas poslednej zmeny, ...).

Podpora typov súborov OS:

- Výhodou je možnosť kontroly správneho použitia súboru používateľom, ďalšie prídavné funkcie (automatický make pri spustení programu, ...).

- Nevýhodou je veľkosť OS (pre každý typ súboru nejaký kód) a možnosť neobsiahnutia všetkých typov a následného zlého použitia.

- Opačným extrémom je nepodporovať OS žiadne typy súborov (napr. UNIX) a chápať súbor ako sekvenciu osembitových bytov.

Uloženie súborov na disku:

- Disky majú typicky definovanú veľkosť bloku určenú veľkosťou sektoru disku. Všetky I/O operácie sú vykonávané po blokoch a všetky bloky majú zhodné veľkosti. Logické bloky súboru sú potom zabalené do fyzických blokov disku.

- Obvykle dochádza ku strate určitej časti fyzických blokov, z dôvodu dĺžky súboru či rozdielu medzi dĺžkou fyzického a logického bloku.

Štruktúra adresárov:

- Adresár zariadenia je na každom fyzickom zariadení a popisuje všetky súbory uložené na tomto zariadení. Obsahuje fyzické atribúty súboru (dĺžka, ako je umiestnený na disku, ...).

- Štruktúra adresára súborov je logickou organizáciou súborov na všetkých zariadeniach. Obsahuje logické atribúty súboru (meno, typ, vlastník, prístupové práva, ...).

- Systém adresárov mapuje mená súborov na položky adresárov.

Operácie so súbormi :

- CREATE - vytvorí súbor. Nájdenie priestoru pre súbor a záznam novej položky do adresárovej štruktúry.
- OPEN - otvorí existujúci súbor. Prehľadanie adresárovej štruktúry a získanie atribútov súboru. Vracia handle.
- CLOSE - zavrie otvorený súbor.
- WRITE - zapíše do otvoreného súboru na pozícii kurzora súboru.
- READ - číta z otvoreného súboru z pozície kurzoru súboru.
- SEEK - nastavenie pozície kurzoru v súbore.
- DELETE - zmaže súbor.

Prístupové metódy:

- Sekvenčný prístup - čítanie a zápis na pozíciu kurzora v súbore s automatickým posunom kurzora.
- Priamy prístup - súbor je chápaný ako očíslovaná sekvencia záznamov a pomocou operácie SEEK sa dá nastaviť kurzor súboru na ľubovoľný záznam súboru .
- Pamäťové mapovanie - súbor je pomocou stránkovania namapovaný (je obrazom pamäte na disku) do virtuálneho adresového priestoru a s jednotlivými záznamami v súbore sa dá manipulovať ako s pamäťou.

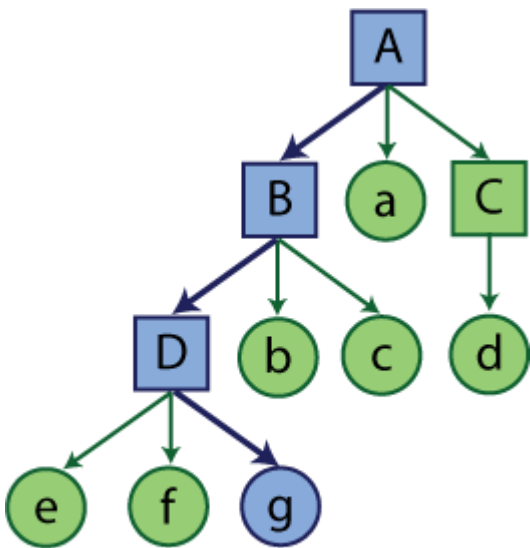
5.2 Organizácia štruktúry adresárov

Súborom sa priradujú mená. Systém ovládania súborov hľadá, na základe udania mena súboru, prístupovú cestu k súboru. Implementačne to obvykle znamená najst' riadiaci blok súboru. Dátová štruktúra používaná k zobrazeniu z množiny mien súborov do množiny riadiacich blokov súborov, je typická tabuľka, ktorú nazývame adresár alebo katalóg. Adresár sa môže vytvoriť ako nezatriedená alebo zatriedená tabuľka. Adresáre sa potom môžu uložiť ako súbor a usporiadať ich do štruktúr. Jedna položka adresára môže niesť údaj o umiestnení iného adresára. Jednoduchou formou takej štruktúry sú hierarchické adresáre. Prakticky najvýznamnejšie sú jednoúrovňové, dvojúrovňové a stromové adresáre. Rovnako sa často používajú adresáre s cyklickou štruktúrou.

Jednoúrovňový adresár sa obvykle implementuje „rozptýlene“ tak, že sa na jednotlivých zväzoch vytvorí tabuľky obsahujúce informácie o súboroch na týchto zväzoch.

Pri dvojúrovňovom adresári sú v hlavnom adresári uložené informácie o používateľských adresároch. Používateľ nemá prístup k hlavnému adresáru a tu už nemá používateľ **a** prístup k adresáru užívateľa **b**.

Prirodzeným rozšírením dvojúrovňových adresárov sú adresáre so stromovou štruktúrou. V koreni stromu je hlavný adresár. Používateľské adresáre môžu mať vytvorenú vlastnú hierarchiu adresárov a v nich logicky zoskupovať svoje súbory. Organizáciu takýchto adresárov si môžeme predstaviť v nejakej podobe grafu s koreňom.

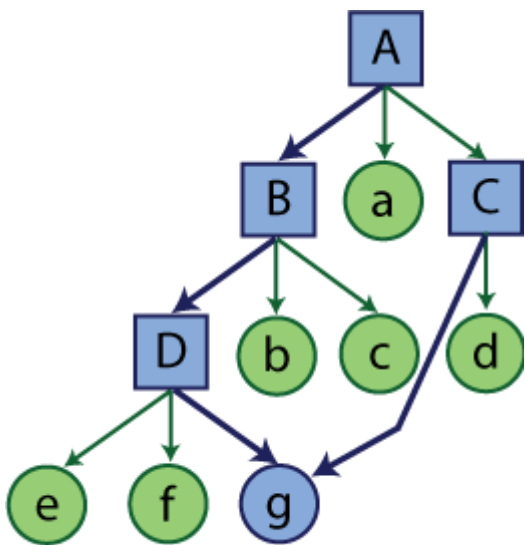


Stromová štruktúra /A/B/D/g

Pozícia používateľa v adresárovej štruktúre sa nazýva **aktuálny adresár**. Identifikácia používateľa, resp. meno jeho adresára a meno súboru tvorí úplnú prístupovú cestu k súboru. Hovoríme tiež o absolútnej prístupovej ceste.

Absolútna cesta je daná menom odboru spojeného s menami všetkých podadresárov ležiacich na ceste z koreňového adresára k danému súboru. Obdobne existuje relatívna cesta, ktorá je daná menom spojeným s menami všetkých podadresárov ležiacich na ceste z aktuálneho adresára k danému súboru. Obvykle vznikne pridaním špeciálnych mien adresárov pre aktuálny adresár a predchodcu adresára.

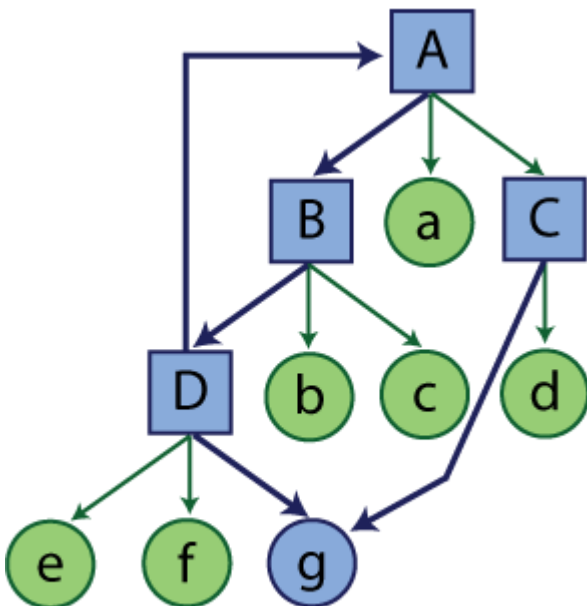
Ak zovšeobecníme stromovú štruktúru adresárov na acyklickú, dosiahneme to, že k jednému súboru môže viesť niekoľko prístupových ciest.



Acyklický graf

Například súbor určený prístupovou cestou **/A/B/D/g** je rovnako určený prístupovou cestou **/C/g**.

Ak môže byť štruktúra adresárov obecným grafom, môžu sa v prístupových cestách opakovať tie isté postupnosti.



Obecný graf

Například prístupové cesty **/A/B/D/g** a **/A/B/D/A/C/g** ukazujú na ten istý súbor.

Existujú rôzne prístupy k viacerým diskom:

- Oddelené systémy súborov na jednotlivých diskoch (A:, B:, C: - MS-DOS, Windows, podobne MacIntosh, VMS).
- Montovanie file systémov (UNIX). Na jednom disku je koreňový adresár. V hierarchickej sústave adresárov sa na niektorý adresár zavesí celý disk (Keď sa prepne do tohto adresára, ocitneme sa na inom fyzickom disku).

V starších OS boli značné obmedzenia na mená súborov: krátke mená (CP/M maximálne 6+3 znaky, MS-DOS 8+3, Unix 14), nerozlišujú sa malé a veľké písmená, množina znakov, ktoré môžu tvoriť meno súboru je značne obmedzená. Novšie OS často rozlišujú malé a veľké písmena, povoľujú špeciálne znaky (medzera, +, !, :) a národné znaky (ě, š, č, ř, ž, ý, á, í, é, ...). V pomenovávaní súborov vývoj smeruje k dlhým menám (desiatky znakov).

5.3 Generačné súbory

Udržujú sa i staršie verzie súborov. Pre odkaz na súbor sa môže používať buď iba meno (zodpovedá najnovšej verzii súboru) alebo meno generácie. V OS VMS sú generačné súbory súčasťou systému. V Unixe sa rieši programovo (RCS - Revision Control System).

OS spravidla o súbore udržuje ďalšie informácie:

- Atribúty (len pre čítanie, apod.). MS-DOS atribúty:
 - R - read-only, len pre čítanie,
 - H - hidden, skrytý,
 - S - systém, systémový,
 - A - archive, nebol archivovaný.
- Dátumy a časy; u niektorých systémov jeden u iných viac:
 - o vytvorení,
 - o poslednej zmene,
 - o poslednom prístupe.

U operačného systému Macintosh sú ešte ďalšie informácie o súboroch:

- 4 znaky typ súboru + 4 znaky program, ktorý ho vytvoril,
- tieto znaky nie sú súčasťou mena súboru,
- určuje, akou aplikáciou sa má súbor spracovávať.

V niektorých systémoch sú súbory členené na dátovú časť a resource časť (texty, obrázky, ikony, dialógy). Tieto OS spravidla umožňujú zmenu resource časti programu bez nutnosti nového prekladu programu. U počítačov Macintosh skutočne existujú dve časti súboru; programy v MS-Windows - všetko v jednom súbore.

Odkazy v systéme súborov (v Unixe link, na počítačoch Macintosh zástupcovi):

- dva alebo viac odkazov na ten istý súbor alebo adresár,
- odkazy môžu byť rovnakého alebo i rôzneho mena.

5.4 Prístupové práva v OS UNIX

Unix umožňuje pre každý súbor alebo adresár definovať práva pre vlastníka, skupinu a pre ostatných používateľov. Vlastník nepreberá práva skupiny a ostatných. Vlastník súboru môže byť len jeden. V niektorých Unixoch môže používateľ byť v danom okamihu členom viacerých skupín, v iných môže meniť členstvo v skupine, ale v každom okamihu môže byť členom iba jednej skupiny. V nasledujúcej tabuľke vlastník nemôže nič, skupina môže čítať a ostatný môžu všetko. Spravidla však nemá vlastník menšie práva než skupina a skupina nemá menšie práva než ostatný.

Vlastník	Skupina	Ostatní
---	r--	rwX

5.4.1 Práva k súborom

r - čítanie obsahu súboru

w - zmena obsahu súboru

x - spúšťanie súboru (všetky programy musia mať 'x') pre mazanie a premenovávanie súboru stačí mať práva 'wx' v adresári

5.4.2 Práva k adresárom

r - môže iba prečítať mená súborov a podadresárov daného adresára (nič iného, nedá sa ani zistiť, či sa jedná o súbory alebo podadresáre)

w - samo o sebe nie je k ničomu; spolu s právom x umožňuje vytvorenie, premenovania a zrušenia súborov a prázdnych podadresárov v adresári

x - umožňuje sa prepnúť do adresára, čítanie a zápis do súborov, zmena vlastníka a práv súborov a podadresárov, či nie je právo r, je potrebné poznať meno súborov v adresári

5.4.3 Práva setuid, setgid

Názov pochádza zo slov uid a gid (User Identifier = číslo používateľa, Group Identifier = číslo skupiny) nastaviť. Ak má program nastavené právo setuid (miesto práva x pre vlastníka sa vypisuje s), má používateľ pri jeho spustení rovnaké práva ako vlastník programu. Podobne funguje setgid pre skupinu.

Príklad použitia: v Unixe sú informácie o používateľoch (včítane zakódovaných hesiel) uložené v súbore /etc/passwd. Jeho vlastníkom je root (správca systému) a práva sú rw-r--r--. Súbor teda môže ktokoľvek čítať, meniť ho však môže iba root. Aby si používatelia mohli sami meniť heslá, má program, ktorý sa používa pre zmenu hesiel, nastavené právo setuid (a jeho vlastníkom je root). Vyššie zmienený program musí byť ale spoľahlivý a robiť len to čo má, inak narušíme bezpečnosť systému (napr. akýkoľvek používateľ bude môcť zmeniť iným používateľom heslá a potom zneužiť ich kontá).

U niektorých Unixov sa používa tzv. ACL (Acces Control List). Pre každý súbor a adresár existuje zoznam používateľov a skupín, u každého sú uvedené práva (podobne ako u Novellu).

5.5 Prístupové práva v OS Novell Netware

Právo	Adresár	Súbor
S (supervisory)	všetky práva pre adresár	všetky práva k súboru
	a celý podstrom neobmedzené	
	maskou zdedených práv	
	môže pridelovať práva ostatným	
R (read)	otvoriť a čítať súbory	otvoriť a čítať súbor
	(Ak nie je F, musí poznať meno	
	súboru)	
W (write)	otvoriť a zapisovať do exist.	zapisovať do súboru
	súborov	
C (create)	vytvárať súbory a adresáre,	obnoviť súbor (SALVAGE)
	zapisovať do novo vytvorených	po zmazaní
	súborov	
E (erase)	zmazať súbory alebo prázdne	zmazať súbor
	adresáre	
M (modify)	meniť mená a atribúty súborov	-> tiež
	a podadresárov (nie mazať a	
	meniť ich obsah)	
F (file scan)	je vidieť obsah adresára po DIR	vidieť súbor pri DIR
A (access control)	umožňuje meniť práva ostatných	-> tiež
	používateľov (s výnimkou S)	
	i tie čo sami nemáme	

U každého adresára a súboru sa dá definovať zoznam tvorený dvojicami: používateľ alebo skupina - práva.

Používateľ môže byť členom viacerých skupín. Výsledné práva sú zjednotením práv používateľa a všetkých skupín, ktorých je členom. Navyše sa práva preberajú z nadriadeného adresára (dedenia). Toto preberanie práv je možné pre jednotlivé adresáre obmedziť tzv. maskou dedičných práv.

Atribúty pre súbory a adresáre:

H hidden	skrytý súbor, nevypíše sa pri DIR
S system	systémový súbor, nevypíše sa pri DIR

D	delete inhibit	sa nedá zmazať
R	rename inhibit	sa nedá premenovať
P	purge	sa nedá obnoviť po zmazaní

Atribúty iba pre súbory:

A	archive	nebol archivovaný
Ro	read only	iba pre čítanie
Rw	read write	pre čítanie i pre zápis
X	execute only	sa nedá kopírovať, iba spustiť
I	indexed	pre urýchlenie prístupu k veľkým súborom
T	transactional	sa dá definovať transakcie = sady akcií, ktoré sa
		buď prevedú celé, alebo sa súbor automaticky
		obnoví do pôvodného stavu
S	shareable	zdieľateľný súbor -> môže ho otvoriť viac
		používateľov súčasne

5.6 Uzamykanie súborov a ich častí

Pri zápise do súboru (databázy) sa celý súbor alebo určitá časť súboru uzamkne a ostatné procesy alebo používatelia nemôžu po dobu uzamknutia obsah príslušnej oblasti meniť (alebo ani čítať). Používa sa napríklad pre prístup k databázam (príklad: databáza miesteniek prístupná z niekoľkých predpredajov v rôznych mestách).

5.7 Zdieľanie súborov

Pri otvorení súboru sa definuje:

- OPEN FOR (otvoriť pre ...)

- read

- write

- read & write

- DENY (ostatným po dobu otvorenia znemožniť ...)

- none -> môžu robiť všetko

- read -> znemožnenie čítania

- write -> znemožnenie zápisu

- all -> znemožnenie čítania i zápisu

5.8 Kvóty

Obmedzenie priestoru na disku (diskoch) obsadeného jedným používateľom. Používa sa u NOVELL, Unix, VMS.

Často sa používa tvrdá a mäkká kvóta. Tvrdú nemožno v žiadnom prípade prekročiť, mäkkú je možné prekročiť na obmedzenú dobu.

Precvičte sa

Kontrolné otázky:

- 5.1. Ktoré informácie operačný systém udržiava o súboroch?
- 5.2. Ktoré typy organizácie adresárov poznáte?
- 5.3. Aké prístupové metódy k súborom existujú?

Úlohy k zamysleniu:

- 5.1. O súboroch ukladaných na diskoch sa udržiava informácia o skutočnej dĺžke súboru a o dĺžke, ktorú zaberie súbor na disku. Prečo sú tieto údaje väčšinou rozdielne a ktorý tento údaj je väčší?
- 5.2. Zamyslite sa nad úlohou vyrovnávacej pamäte pri ukladaní a čítaní súborov na disk.

Korešpondenčná úloha:

- 5.1. Popíšte všetky typy operácií so súbormi, ktoré umožňujú realizovať operačný systém WINDOWS XP.

Zhrnutie obsahu kapitoly

V tejto kapitole ste sa zoznámili so spôsobmi ukladania súborov na vonkajších pamätiach, prevažne na diskoch. Dôraz v tejto kapitole bol kladený na pochopenie vlastností súborov, ktoré je nutné uchovávať pre jednoznačnú identifikáciu súborov.