

Pliki i systemy plików

Tomasz Lewicki

WWSIS, Wrocław

kwiecień 2007

Plik – co to takiego?

- Logiczny zbiór informacji powiązanych ze sobą w sposób określony przez twórcę pliku
- Zbiór ten jest niezależny od fizycznych właściwości nośnika
- Pliki przechowują programy oraz dane różnego rodzaju
- Wewnętrzna budowa pliku może być dowolna lub ściśle określona
- Plik możemy traktować jako ciąg bitów, bajtów, wierszy lub rekordów

Rodzaje plików

Zwykłe

- ASCII oraz binarne
- Wykonywalne (programy, skrypty, pliki wsadowe)

Specjalne

- Dowiązania symboliczne (skrótów, aliasy)
- Urządzenia blokowe (transmisja równoległa: twarde dyski, płyty CD/DVD)
- Urządzenia znakowe (transmisja szeregową: terminale, drukarki, interfejsy sieciowe, ...)
- Potoki, kolejki, gniazda

Katalogi

Rozszerzenia

- .exe, .com, .bat – pliki wykonywalne i wsadowe
- .txt, .doc, .odt – pliki tekstowe
- .jpg, .png, .tif – pliki graficzne
- .avi, .mov, .mp3 – pliki multimedialne
- .dll – biblioteki systemowe
- MNÓSTWO innych rozszerzeń

Atrybuty

- Odczyt
- Zapis
- Wykonalność

„Magiczna liczba” w nagłówku pliku

Pierwsza linia w plikach tekstowych

Prawa do plików – Unix

| | | Plik | Katalog |
|---|---|--------------------|----------------------------|
| 4 | r | prawo do odczytu | prawo do listowania |
| 2 | w | prawo do zapisu | prawo do tworzenia plików |
| 1 | x | prawo do wykonania | prawo do otwarcia katalogu |

Przykłady:

```
drwxr-xr-x 2 stalker users 4096 sty 16 23:51 WLAN
-rw-r--r-- 1 stalker users 1201170 sty 9 21:35 firestarter-1.0.3.tar.gz
-rw----- 1 stalker users 454649 sty 19 13:52 image011.gif
```

Prawa do plików – Unix

Sesja Edycja Widok Zakładki Ustawienia Pomoc

The screenshot shows a terminal window with a file manager interface. The main window displays a list of files and directories in a table format. A modal dialog titled "Polecenie chmod" is open in the center, showing options for setting permissions. The terminal output at the bottom shows the command "stalker@fomalhaut: ~/Temp" and a list of keyboard shortcuts.

| Nazwa | Rozmiar | Czas | Nazwa | Rozmiar | Czas |
|--------------------|---------|--------------|----------------------------|---------|--------------|
| .. | | | .. | | |
| ./Skype | 4096 | 9 sty 21:51 | ./Dzwiek | 4096 | 17 sty 19:51 |
| ./beagle | 4096 | 19 sty 18:18 | ./WLAN | 4096 | 16 sty 23:51 |
| ./beryl | 4096 | 11 sty 23:19 | ./XGL | 4096 | 17 sty 19:53 |
| ./config | 4096 | 12 sty 00:01 | ./firestarter-1.0.3 | 4096 | 14 sty 12:44 |
| ./cnerald | 4096 | 11 sty 23:19 | ./firestarter-1.0.3.tar.gz | 1201170 | 9 sty 21:35 |
| ./fontconfig | 4096 | 17 sty 22:28 | ./image011.gif | 454649 | 19 sty 13:52 |
| ./fonts | 4096 | 17 sty 22:28 | | | |
| ./gconf | | | | | |
| ./gconfd | | | | | |
| ./gkrellm-reminder | | | | | |
| ./gkrellm2 | | | | | |
| ./gnome2 | | | | | |
| ./gnome2_private | | | | | |
| ./gnupg | | | | | |
| ./gphoto | | | | | |
| ./kadu | | | | | |
| ./kde | | | | | |
| ./local | | | | | |
| ./lyx | | | | | |
| ./macromedia | | | | | |
| ./mc | | | | | |
| ./mcpop | | | | | |
| ./mozilla | | | | | |
| ./mplayer | | | | | |
| ./ooo-2.0 | | | | | |
| ./opera | | | | | |
| ./qt | | | | | |
| ./skel | 4096 | 7 sty 01:39 | | | |
| ./ssh | 4096 | 7 sty 22:13 | | | |
| ./stellarium | 4096 | 8 sty 19:28 | | | |
| ./superkaramba | 4096 | 17 sty 18:51 | | | |
| ./texlive2005 | 4096 | 11 sty 19:14 | | | |
| ./.. | | | ./WLAN | | |

Polecenie chmod

| | |
|---|----------------------------|
| Uprawnienia | Plik |
| <input type="checkbox"/> ustaw UID przy wykonaniu | Nazwa |
| <input type="checkbox"/> ustaw GID przy wykonaniu | WLAN |
| <input type="checkbox"/> bit "sticky" | Uprawnienia (0senkowo) |
| <input checked="" type="checkbox"/> odczytywanie przez właśc. | 40755 |
| <input checked="" type="checkbox"/> zapisywanie przez właśc. | Nazwa właściciela |
| <input checked="" type="checkbox"/> wykonywanie przez właśc. | stalker |
| <input checked="" type="checkbox"/> odczytywanie przez grupę | Nazwa grupy |
| <input type="checkbox"/> zapisywanie przez grupę | users |
| <input checked="" type="checkbox"/> wykonywanie przez grupę | |
| <input checked="" type="checkbox"/> odczytywanie przez innych | Spacja przełącza |
| <input type="checkbox"/> zapisywanie przez innych | opcję, klawisze kursora |
| <input checked="" type="checkbox"/> wykonywanie przez innych | przenoszą pomiędzy opcjami |
| | zaś T lub Ins zaznacza |

[< Ustaw >] [Amuluj]

Porada: Użycie kopiowania i wklejania za pomocą myszy może wymagać Shifta.
stalker@fomalhaut: ~/Temp
1Pomoc 2Menu 3Podgl 4Zycia 5kopia 6Przen 7Twórz 8Usu 9Rozwi 10Kończ

Prawa do plików – c.d.

Listy kontroli dostępu

Listy kontroli dostępu (ACLs – Access Control Lists) pozwalają na bardziej szczegółowe określenie uprawnień do plików i katalogów dla pojedynczych użytkowników lub ich grup. Zaletą jest lepsza kontrola nad przydziałem uprawnień do operacji na obiektach systemu plików, znacznie szersza niż trzy podstawowe prawa dostępu.

Wadą natomiast jest stopień skomplikowania list kontroli dostępu, szczególnie w dużych systemach z tysiącami użytkowników i setkami grup. W tym sensie ACLs są „kosztowne” w zarządzaniu. Występują również problemy z archiwizacją ACLs i ich późniejszym odzyskaniem.

Prawa do plików – c.d.

Podójście mieszane

W nowoczesnych systemach operacyjnych stosuje się podejście dualne (mieszane): najpierw definiuje się ogólniejsze prawa dostępu (odczyt, zapis, wykonanie dla użytkowników, grup i „świata”), dodając w razie potrzeby listy kontroli dostępu dla wymagających tego plików i katalogów.

Atrybuty plików

- Nazwa
- Typ (plik, katalog, dowiązanie, ...)
- Rozmiar
- Właściciel i grupa
- Prawa dla różnych użytkowników
- Data utworzenia, ostatniej modyfikacji, ostatniego dostępu
- Położenie na nośniku danych (ogólniej: urządzeniu)

Operacje na plikach

- Utworzenie
- Otwarcie
- Zapisywanie, dopisywanie, skracanie
- Odczytywanie
- Usuwanie
- Zamykanie
- Przesunięcie wskaźnika pozycji wewnątrz pliku
- Nadawanie i zmiana praw oraz atrybutów

Dyskowe systemy plików

Dyskowy system plików jest warstwą pośrednią między „surowym” nośnikiem danych a plikami, katalogami i dowiązaniem. Ułatwia organizację plików i katalogów na urządzeniu przechowywania danych.

Możemy go sobie wyobrazić jako strukturę złożoną z warstw:

- nośnik (urządzenie)
- partycje
- katalogi i pliki

Przykłady:

- FAT16, FAT32, NTFS – rodzina systemów Microsoft Windows
- Ext2, Ext3, ReiserFS – różne dystrybucje Linuksa
- HFS – Mac OS
- ISO9660 – płyty CD i DVD

Montowanie systemu plików

Zanim system plików zostanie użyty, trzeba go **zarejestrować** i **zamontować**.

Rejestracja polega na załadowaniu do pamięci operacyjnej modułu obsługującego dany system plików. Tę operację można wykonać podczas startu systemu bądź tuż przed użyciem systemu plików.

Montowanie polega na jawnym wskazaniu urządzenia fizycznego oraz miejsca w strukturze katalogów, w którym ma być przyłączony system plików. W trakcie montowania system operacyjny sprawdza, czy wywoływane urządzenie zawiera obsługiwany system plików.

Montowanie systemu plików – c.d.

Microsoft Windows

Dwupoziomowa struktura katalogów; urządzeniom pamięci masowej przypisywane są kolejne litery. Systemy plików są montowane podczas startu systemu oraz po dołączeniu urządzeń w trakcie działania SO.

Unix/Linux

Informacje o urządzeniach i systemach plików są zawarte w pliku *fstab*, który jest odczytywany podczas startu systemu. Nowe urządzenia można montować w sposób jawny również w trakcie działania systemu, przypisując je do dowolnego miejsca w strukturze plików.

MacOS X

Urządzenia są analizowane w chwili dołączania do systemu operacyjnego i montowane automatycznie.

Sieciowe systemy plików

Pojęcie „system plików” jest używane nie tylko w odniesieniu do fizycznych nośników danych. Mamy z nimi do czynienia również w sieciach. Sieciowe i „zwykłe” systemy plików różnią się tym, że w sieciach nie określa się fizycznych formatów zapisu danych. W tym znaczeniu sieciowy system plików jest niezależny od „zwykłego”. Między innymi dlatego różne systemy mogą współdzielić zasoby charakterystyczne dla siebie.

Przykłady:

- SMB (Server Message Block) – rodzina systemów Microsoft Windows
- NFS (Network File System) – różne odmiany Uniksa
- NCP (Network Core Protocol) – Novell Netware
- Andrew FS i Coda – implementacje w różnych systemach operacyjnych

Katalog – co to takiego?

Katalog (folder, kartoteka) to logiczny sposób organizacji plików na nośniku danych.

Powody tworzenia katalogów:

- **Wygoda** – użytkownik grupuje pliki według własnego uznania
- **Wydajność** – szybkie i łatwe wyszukiwanie konkretnych plików, zarówno przez użytkownika, jak i system operacyjny

Organizacja katalogów

Organizacja katalogów i plików w nich zawartych może mieć różny stopień złożoności:

- **jeden poziom** – wszystkie pliki w jednym (tzw. głównym) katalogu
Wady: bałagan, konieczność nadawania plikom unikalnych nazw
- **dwa poziomy** – pliki systemowe i pliki użytkowników w osobnych katalogach
Wada: trudność współdzielenia plików między użytkownikami
- **drzewo** – katalog główny (*root*) zawierający podkatalogi z plikami systemu i katalogami oraz podkatalogami użytkowników
Zaleta: możliwość współdzielenia plików i katalogów przez użytkowników dzięki dowiązaniom symbolicznym do jednej kopii pliku (katalogu) fizycznie istniejącej na nośniku (uwaga: Windows i Unix odmiennie traktują dowiązania!)

Operacje na katalogach

- Utworzenie
- Odczytywanie zawartości (tzw. listowanie)
- Usuwanie
- Zamykanie
- Wyszukiwanie pliku
- Dowiązanie pliku
- Nadawanie praw dostępu

Metody dostępu do plików

- **Sekwencyjny** – przetwarzanie rekordów „jeden za drugim”; pojęcie *wskaźnika pozycji*.
Przykład: taśma magnetyczna
- **Swobodny (bezpośredni)** – przetwarzanie rekordów w dowolnej kolejności.
Można łatwo zasymulować dostęp sekwencyjny.
Przykład: dysk, płyta, pamięć USB
- **Indeksowany** – oparty na dostępie swobodnym. Istnieje tablica rekordów, w której zapisane są informacje o pozycji szukanych danych.
Przykład: bazy danych. Analogia: indeks w książce

Backup i archiwizacja

System plików może ulec uszkodzeniu, najczęściej wskutek awarii nośnika fizycznego. Z tego względu należy zadbać o właściwe **zabezpieczenie danych** przed awarią. Można to realizować różnymi sposobami i na różnych poziomach.

„Bierną” formą ochrony przed utratą danych (spowodowaną zarówno czynnikami sprzętowymi, jak i w mniejszym lub większym stopniu niezależnymi od sprzętu) jest **wykonywanie kopii zapasowych** oraz **archiwizacja** danych.

Backup i archiwizacja – c.d.

Kopie zapasowe dzielimy na **całościowe** (pełne), **różnicowe** i **przyrostowe**.

Przyrostowe wykonuje się częściej, kopiując tylko te pliki, które uległy zmianie bądź powstały od czasu wykonania poprzedniej kopii pełnej lub przyrostowej.

Całościowe wykonuje się rzadziej, kopiując wszystkie pliki, niezależnie od czasu ich powstania lub ostatniej modyfikacji. *Różnicowe* zawierają pliki zmienione od czasu wykonania poprzedniej kopii całościowej. Kopie zapisuje się na taśmach magnetycznych, dodatkowych dyskach twardych, nośnikach optycznych, a obecnie również w sieciach SAN (Storage Area Network).

Zarówno częstotliwość wykonywania obu rodzajów kopii, jak i ich zawartość jest determinowana przez charakter zabezpieczanych danych. Wiele programów użytkowych na bieżąco tworzy kopie bezpieczeństwa plików otwartych przez użytkownika.

Redundant Array of Independent/Inexpensive Disks czyli **nadmiarowa macierz niezależnych dysków** to zespół dwóch lub więcej dysków twardej, współpracujących ze sobą w celu zapewnienia większego bezpieczeństwa danych.

Inne, niejako uboczne korzyści z łączenia dysków w macierz RAID to podniesienie prędkości transferu danych w operacjach odczytu i zapisu oraz zwiększenie łącznej pojemności dysków pozostających do dyspozycji systemu operacyjnego (te dwie cechy zależą od poziomu RAID!).

Macierz RAID można zbudować wykorzystując kontroler sprzętowy lub mechanizmy zawarte w systemie operacyjnym. Kontrolery sprzętowe są na ogół bardziej wydajne, ale drogie.

RAID 0

Połączenie co najmniej dwóch dysków fizycznych w jeden dysk logiczny (inaczej *striping*, paskowanie). Dyski nie muszą mieć takiej samej pojemności, nie muszą być nawet tego samego typu, ale jest to wskazane ze względu na wydajność (macierz będzie tak szybka jak najwolniejszy z dysków).

Zalety:

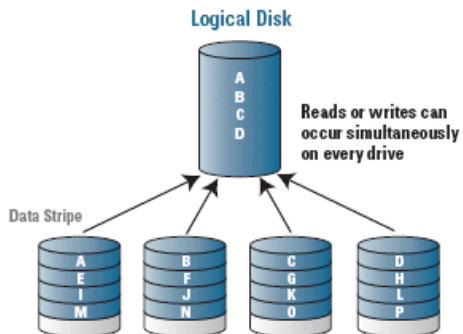
- większa sumaryczna przestrzeń dyskowa
- szybsze operacje odczyt/zapis w porównaniu do pojedynczego dysku

Wady:

- brak nadmiarowości, czyli brak zabezpieczenia danych

Typowe zastosowanie: przechowywanie dużych plików (np. multimedia), logi systemowe, szybko zmieniające się dane

RAID 0



RAID 1

Połączenie dwóch dysków fizycznych tak, by każdy z nich był **lustrzanym odbiciem** drugiego (stąd również nazwa *mirroring*).

Zalety:

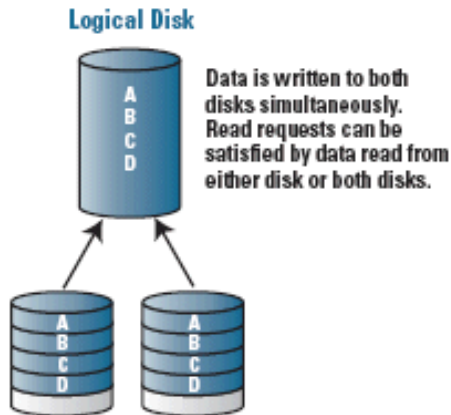
- odporność na awarię $N - 1$ dysków przy macierzy N dysków
- przy odpowiednim dobraniu dysków i sposobie odczytu można zwiększyć prędkość czytania danych

Wady:

- mniejsza prędkość zapisu
- „utrata” pojemności $N - 1$ dysków

Typowe zastosowanie: systemy operacyjne, transakcyjne bazy danych

RAID 1



RAID 0+1

Połączenie czterech dysków fizycznych w układzie: dwa dyski w RAID 0 połączone w *mirroringu*.

Zalety:

- szybkość odczytu/zapisu jak w RAID 0
- bezpieczeństwo jak w RAID 1

Wady:

- awaria jednego dysku pociąga za sobą konieczność odbudowy całej macierzy RAID 1 (czyli naprawienie kopii lustrzanej)
- awaria jednego dysku powoduje utratę bezpieczeństwa całej macierzy
- stosunkowo duży koszt

Typowe zastosowanie: systemy operacyjne, transakcyjne bazy danych

RAID 1+0 (RAID 10)

Połączenie czterech dysków fizycznych w układzie: dwie pary *mirrorowanych* dysków połączone w RAID 0.

Zalety:

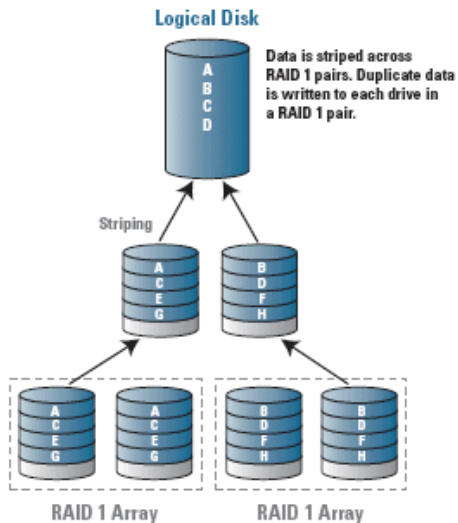
- takie jak w RAID 0+1
- lepsze zabezpieczenie danych w porównaniu z RAID 0+1 – w przypadku awarii jednego dysku należy odbudować tylko fragment macierzy

Wady:

- stosunkowo duży koszt (taki jak w RAID 0+1)

Typowe zastosowanie: bazy danych, serwery aplikacji

RAID 1+0



RAID 5

Połączenie co najmniej trzech dysków z zapisem kodów parzystości.

Zalety:

- odporność na awarię jednego dysku
- szybsze operacje odczytu

Wady:

- mniejsza prędkość operacji zapisu z uwagi na konieczność obliczania sum kontrolnych
- uszkodzenie dysku powoduje spadek prędkości odczytu danych – trzeba obliczać sumy kontrolne
- odbudowa macierzy może długo trwać, w tym czasie wydajność systemu jest niższa

Typowe zastosowanie: hurtownie danych, serwery WWW, archiwizacja

RAID 5

