TRƯỜNG ĐẠI HỌC THỦY LỢI KHOA CÔNG NGHỆ THÔNG TIN

NGUYÊN LÝ HỆ ĐIỀU HÀNH

Giảng viên: TS. Đoàn Thị Quế Bộ môn Mạng và an toàn thông tin

Chương 6: Hệ thống file

- File
- Thu muc
- Thực hiện hệ thống file
- Ví dụ về các hệ thống file

5.1. File

- Khái niệm file
- Tên file
- Cấu trúc file
- Kiểu file
- Truy cập file
- Các thuộc tính của file
- Các thao tác với file

Khái niệm file

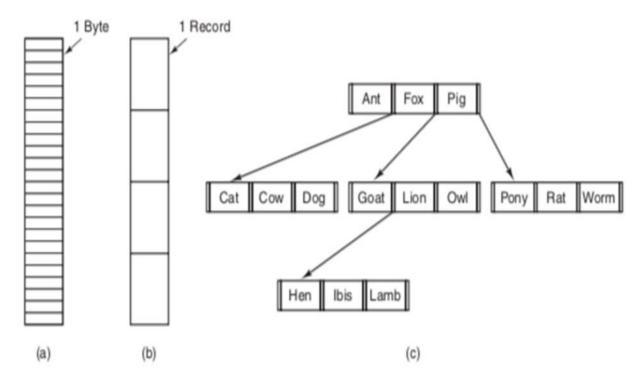
- File là đơn vị logic của thông tin được tạo bởi tiến trình (process). Mỗi File có thể được coi như một loại không gian địa chỉ. Thông tin lưu trữ trong file được lưu lại và không bị ảnh hưởng bởi quá trình tạo và kết thúc tiến trình.
- File được quản lý bởi hệ thống file (File System) của hệ điều hành.

Tên file

- Tên file là thuộc tính quan trọng nhất của file giúp người dùng lưu trữ và tìm lại được nơi thông tin được lưu trữ trong ổ đĩa.
- Tên file trong các hệ thống hiện đại thường gồm 2 phần ngăn cách bởi dấu chấm gồm tên file và phần mở rộng. Phần mở rộng là mô tả về file và giúp hệ điều hành nhận ra phần mềm để mở file.
- Bài tập: Tìm hiểu qui tắc đặt tên file của một số HĐH

Cấu trúc file

File có thể cấu trúc theo một số cách sau: chuỗi byte không có cấu trúc, một chuỗi các bản ghi độ dài cố định, cây các bản ghi kích thước khác nhau.



Ba loại cấu trúc file: (a) Chuỗi byte, (b) Chuỗi bản ghi, (c) Cây các bản ghi.

Kiểu file

- File thông thường là các file chứa thông tin người dùng.
- File chứa thông tin về thư mục, được dùng cho mục đích quản lý file
- File đặc biệt ký tự: liên quan tới quá trình vào/ra và sử dụng để mô hình hóa các thiết bị vào ra nối tiếp như thiết bị cuối, máy in và và thiết bị mạng.
- File đặc biệt khối: sử dụng cho mô hình hóa các ổ đĩa.

Truy cập file

- Những hệ điều hành thời kỳ đầu chỉ cung cấp một chế độ truy cập file là truy cập tuần tự (sequential access).
 - Thông tin được đọc/ghi lần lượt từ đầu file, ví dụ theo từng byte hay từng bản ghi.
 - Thích hợp với phương tiện lưu trữ là băng từ.
- Khi sử dụng đĩa để lưu trữ file, việc đọc byte hay bản ghi không cần theo thứ tự. File và các bản ghi có thể được truy cập theo chế độ truy cập trực tiếp (direct access).
 - file được xem như gồm các khối byte hay các bản ghi được đánh số.
 - Việc đọc ghi các khối dược tiến hành theo thứ tự tùy ý.

Các thuộc tính của file

 Mỗi file bao gồm tên, dữ liệu và một số thông tin khác như kích thước, thời gian thay đổi lần cuối, ... được gọi là các thuộc tính file (attributes hay metadata)

Thuộc tính	Ý nghĩa		
Protection	Quy định người có quyền truy cập và cách thức truy cập file		
Password	Mật khẩu cần thiết để truy cập file		
Creator	ID của người tạo file		
Owner	Người sở hữu hiện tại		
Read-only flag	0-cho phép đọc/ghi; 1-chỉ cho phép đọc		
Hidden flag	0-là file bình thường; 1-sẽ không hiến thị file		
System flag	0-là file bình thường; 1-là file hệ thống		
Archive flag	0-là file đã sao lưu; 1-là file cần được sao lưu		
ASCII/binary flag	0-là file ASCII; 1-là file nhị phân		
Access flag	0-chỉ cho phép truy nhập tuần tự; 1-cho phép truy nhập trực tiếp		
Temporary flag	0-là file bình thường; 1-file sẽ bị xóa khi tiến trình kết thúc		
Lock flags	0-không khóa; khác 0-đã bị khóa		
Record length	Số byte trong một bản ghi		
Key position	Vị trí của khóa trong mỗi bản ghi		
Creation time	Ngày và giờ tạo file		
Time of last access	Ngày và giờ truy nhập file lần cuối		
Time of last change	Ngày và giờ thay đổi file lần cuối		
Current size	Số byte trong file		
Maximum	Số byte cực đại mà file có thể đạt được		

Các thao tác với file

- Create tao file
- Delete xóa file
- Open mở file
- Close đóng file
- Read đọc file
- Write ghi vào file
- Append ghi thêm vào cuối file

- Seek định vị
- Get attributes lấy thuộc tính file
- Set attributes thiết
 lập thuộc tính file
- Rename đổi tên file

5.2. Thư mục

- Khái niệm thư mục
- Cấu trúc hệ thống thư mục
- Đường dẫn
- Các thao tác với thư mục

Khái niệm thư mục

- Để quản lý các file trên mỗi đĩa lôgic, thông tin về file được lưu trong hệ thống thư mục.
- Thư mục (directory hay folder) được tạo thành từ các khoản mục (entry), mỗi khoản mục ứng với một file. Khoản mục chứa các thông tin về file như tên file, kích thước, vị trí, kiểu file và các thuộc tính khác hoặc chứa con trỏ tới nơi lưu trữ những thông tin này.

Các cách lưu thông tin về file trong thư mục:

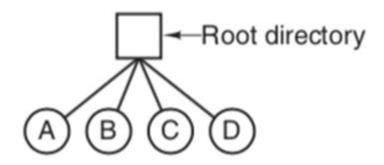
- C1: Toàn bộ thuộc tính của file được lưu trong thư mục, bản thân file chỉ chứa dữ liệu. Ví dụ: MS-DOS
- C2: Thư mục chỉ lưu thông tin tối thiểu cần thiết cho việc tìm vị trí của file trên đĩa. Các thuộc tính file còn lại được lưu trữ luôn cùng với dữ liệu của file. Ví dụ: thư mục sử dụng trong hệ thống file EXT2 và EXT3 của Linux

Cấu trúc hệ thống thư mục

- Hệ thống thư mục đơn cấp
- Hệ thống thư mục đa cấp

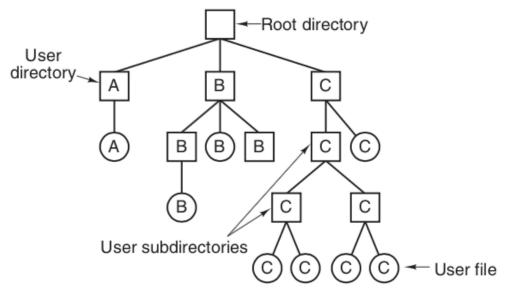
Hệ thống thư mục đơn cấp

- Là hình mẫu đơn giản nhất của hệ thống thư mục.
- Còn được gọi là thư mục gốc "root directory".



Hệ thống thư mục đa cấp

- Trong các hệ thống phức tạp hàng nghìn file, các file cần được nhóm lại vào các thư mục.
- Người dùng có thể tạo một số lượng tuỳ ý các thư mục con cung cấp công cụ mạnh mẽ để tổ chức công việc.



Đường dẫn

- Khi hệ thống file được tố chức theo cây thư mục, có 2 cách để xác định file.
 - Cách thứ nhất dùng đường dẫn tuyệt đối là đường dẫn từ gốc cây thư mục tới file.

```
Windows \usr\ast\mailbox \UNIX /usr/ast/mailbox
```

- Cách thứ hai là sử dụng đường dẫn tương đối. Là đường dẫn từ thư mục làm việc hiện hành tới file.
 - Giả sử đang ở trong thư mục hiện hành là /usr, chỉ cần sử dụng đường dẫn tương đối ast/mailbox

Các thao tác với thư mục

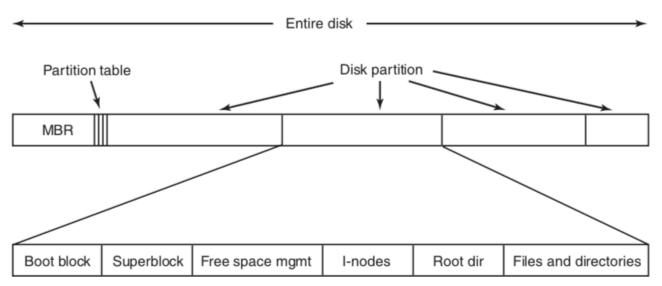
- Create
- Delete
- Opendir
- Closedir
- Readdir trả về giá trị của khoản mục tiếp theo trong một thư mục đã mở
- Rename
- Link kết nối thư mục với file
- Unlink ngắt bỏ kết nối thư mục với file

5.3. Thực hiện hệ thống file

- Cấu trúc hệ thống file
- Tổ chức thực hiện file
- Tổ chức thực hiện thư mục

Cấu trúc hệ thống file

Hệ thống File được lưu trữ trong ổ đĩa. Ô đĩa được chia thành nhiều phân khu (partitions) khác nhau với hệ thống file độc lập trên từng phân khu.



Mỗi hệ thống file trên từng phân khu chứa một số thành phần như: khối khởi động (Boot block), khối đặc biệt (superblock) chứa các tham số chính, ...

Sector 0 chứa:

- MBR (Master Boot Record Bản ghi khởi động chủ) dùng để khởi động máy tính
- Bảng phân khu (partition table): chứa thông tin về địa chỉ bắt đầu và kết thúc của từng phân khu.
- Khối khởi động (Boot block): chương trình trong khối khởi động sẽ nạp hệ điều hành từ phân khu đó vào đĩa khởi động

Tổ chức thực hiện file

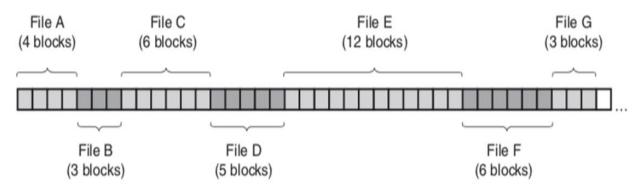
- Số lượng các sector trên một phân khu đĩa cứng là rất lớn, khó quản lý.
- Người ta thường nhóm nhiều sector trên phân khu đĩa lại thành một block (còn gọi là khối đĩa, nhóm sector, cluster, liên cung...)
- Block là đơn vị nhỏ nhất để cấp phát đĩa

Tổ chức thực hiện file

- File được chia vào một số các block nhớ.
- Thực hiện file là quá trình kiểm soát việc block nào được sử dụng để lưu trữ dữ liệu của file.
- Một số phương thức cấp phát:
 - Cấp phát liên tục
 - Cấp phát bằng danh sách liên kết
 - Cấp phát bằng danh sách liên kết sử dụng bảng nhớ
 - Cấp phát sử dụng nút chỉ số (I-node)

Cấp phát liên tục

- Mỗi file được chứa vào một dãy liên tiếp các khối đĩa.
- Vị trí file trên đĩa được xác định bởi vị trí khối đầu tiên và độ dài hoặc số khối mà file đó chiếm.



Tên file	Thư mục Bắt đầu	Độ dài	
File A	0	4	
File B	4	3	
File C	7	6	
• • •			

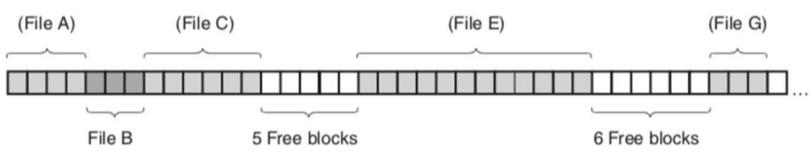
Cấp phát liên tục

Uu điểm:

- Đơn giản và dễ thực hiện
- Tốc độ đọc dữ liệu cao

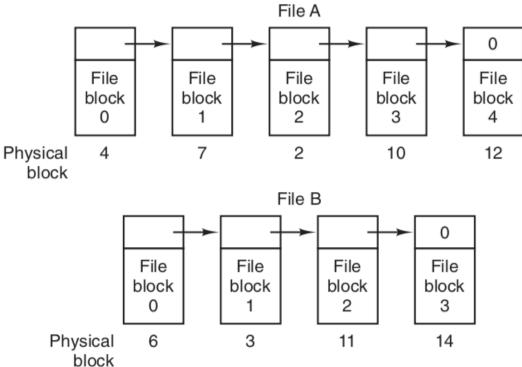
Nhược điểm:

- Khi tạo file phải biết được kích thước file để lựa chọn vùng trống.
- Lãng phí không gian đĩa do hiện tượng phân mảnh ngoài



Cấp phát bằng danh sách liên kết

File được lưu bằng một danh sách liên kết các khối đĩa. Từ nhớ đầu tiên của mỗi khối đĩa được sử dụng để trỏ tới khối tiếp theo, phần còn lại của khối là dữ liệu.



Cấp phát bằng danh sách liên kết

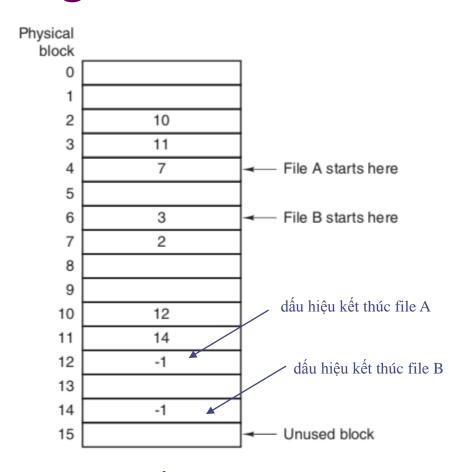
• Ưu điểm: không xảy ra lãng phí không gian nhớ do phân mảnh ổ đĩa. Dễ dàng quản lý thư mục chỉ cần bằng lưu địa chỉ của block đầu tiên.

Nhược điểm:

- Đọc chậm do cần nhiều thao tác tìm kiếm.
- File càng lớn thì phần dữ liệu thêm để định vị block tiếp theo càng lớn – lãng phí không gian lưu trữ.

Cấp phát bằng danh sách liên kết sử dụng bảng nhớ

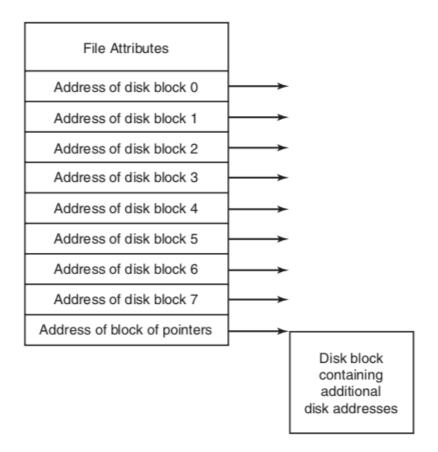
 Hai nhược điểm của giải pháp phân phối theo danh sách liên kết có thể được loại bỏ bằng cách tách con trỏ khỏi các khối đĩa, rồi lưu chúng vào một bảng trong bộ nhớ - Bảng phân phối file (FAT - File Allocation Table)



File A: các khối đĩa theo trật tự 4, 7, 2, 10 và 12 File B: các khối đĩa theo trật tự 6, 3, 11 và 14

Cấp phát sử dụng nút chỉ số (I-node)

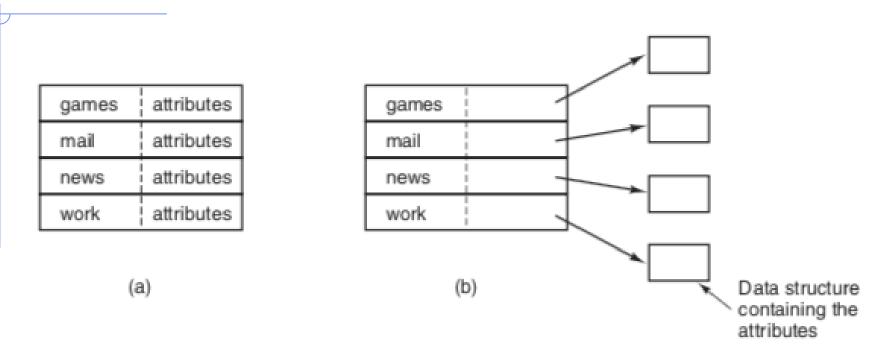
- Thực hiện liên kết mỗi file với một cấu trúc dữ liệu gọi là nút chỉ số (Inode). Cấu trúc này là một danh sách các thuộc tính và địa chỉ các khối đĩa của file.
- Thay vì lưu toàn bộ danh sách liên kết bằng FAT trong bộ nhớ chính, i-node chỉ cần đưa vào bộ nhớ khi file mở.



Tổ chức thực hiện thư mục

- Trong bất cứ hệ thống nào, chức năng chính của hệ thống thư mục là ánh xạ tên file sang các thông tin cần thiết để định vị dữ liệu của file trên đĩa.
- Hệ điều hành sẽ sử dụng đường dẫn mà người dùng cung cấp để xác định entry (khoản mục) của bảng thư mục. Entry thư mục sẽ cung cấp thông tin (với cấp phát liên tục là địa chỉ của toàn bộ file, với hệ thống danh sách liên kết là số của khối đầu tiên hoặc số hiệu của nút i-node) để tìm ra các khối đĩa.

Lưu trữ các thuộc tính của file



Hình (a): Mỗi entry chứa tên, các địa chỉ đĩa và thuộc tính của file.

Hình (b): Mỗi entry chỉ chứa tên file và số hiệu nút chỉ số. Địa chỉ đĩa và thuộc tính của file được lưu trong nút chỉ số.

5.4. Ví dụ về các hệ thống file

- Hệ thống File MS-DOS
- Hệ thống File UNIX V7

Hệ thống File MS-DOS

- Hệ thống file MS-DOS là hệ thống file đầu tiên sử dụng trên IBM PC và các hệ điều hành Windows từ đời đầu đến Windows Vista.
- MS-DOS kiểm soát việc cấp phát các khối đĩa cho file thông qua bảng phân phối file (FAT) trong bộ nhớ chính.

Hệ thống File MS-DOS

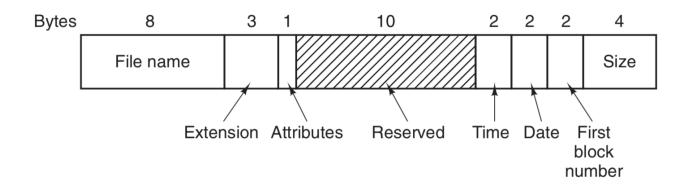
Hệ thống file FAT của MS-DOS đã trải qua ba phiên bản: FAT-12, FAT-16, FAT-32.

Block size	FAT-12	FAT-16	FAT-32
0.5 KB	2 MB		
1 KB	4 MB		
2 KB	8 MB	128 MB	
4 KB	16 MB	256 MB	1 TB
8 KB		512 MB	2 TB
16 KB		1024 MB	2 TB
32 KB		2048 MB	2 TB

Kích thước cực đại của một phân khu ứng với các kích thước khối khác nhau. Các ô trống tương ứng với các trường hợp không tồn tại.

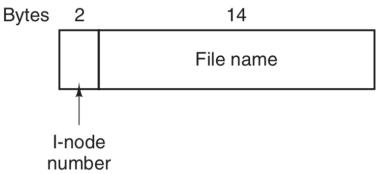
Hệ thống File MS-DOS

• Định dạng của một entry thư mục của MS-DOS:



Hệ thống File UNIX V7

- Hệ thống file Unix là một cấu trúc cây, xuất phát từ thư mục gốc, cùng các kết nối tạo thành đồ thị không chu trình có hướng. Tên file tối đa 14 ký tự và có thể chứa bất cứ ký tự ASCII nào ngoại trừ "/" và null.
- Thư mục trong Unix chứa các entry tương ứng với các file trong thư mục. Vì UNIX sử dụng I-node, nên nội dung của entry rất đơn giản chỉ gồm tên file và số hiệu i-node của file đó.



Hết Chương 6