

NGUYÊN LÝ HỆ ĐIỀU HÀNH

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

Chương 6: Hệ thống file

- Giới thiệu
- File
- Thư mục
- Thực hiện hệ thống file
- Ví dụ về các hệ thống file

6.1. Giới thiệu

- Ba yêu cầu cơ bản đối với các thiết bị lưu trữ thông tin.
 - Nó phải có khả năng lưu trữ được một lượng rất lớn thông tin.
 - Thông tin không được phép mất đi khi tiến trình kết thúc.
 - Phải cho phép nhiều tiến trình truy nhập thông tin cùng một lúc.
- Giải pháp thông thường cho tất cả các vấn đề này là lưu trữ thông tin trên bộ nhớ ngoài (đĩa từ, băng từ, đĩa quang,) theo từng đơn vị gọi là **file**/tập tin

6.1. Giới thiệu

- File là một kỹ thuật trừu tượng. Nó cung cấp khả năng lưu trữ thông tin trên đĩa, và có thể đọc lại sau đó.
- Đối với người dùng:
 - file là đơn vị nhớ logic, việc lưu trữ và truy cập file không phụ thuộc vào thiết bị lưu trữ cụ thể.
- Đối với hệ điều hành:
 - đảm nhiệm việc ánh xạ file lên các thiết bị lưu trữ.
- Thành phần của hệ điều hành làm nhiệm vụ quản lý file được gọi là **hệ thống file (file system)**
- Hệ thống file gồm 2 phần:
 - Các file: Chứa dữ liệu/chương trình của người dùng/hệ thống
 - Cấu trúc thư mục: Cung cấp các thông tin về file

6.1. Giới thiệu

- Một số vấn đề liên quan đến **hệ thống file (file system)** gồm:
 - Khái niệm file
 - Cấu trúc file
 - Kiểu file
 - Truy cập file
 - Các thuộc tính của file
 - Các thao tác với file
 - Tổ chức thực hiện hệ thống file

6.2. File



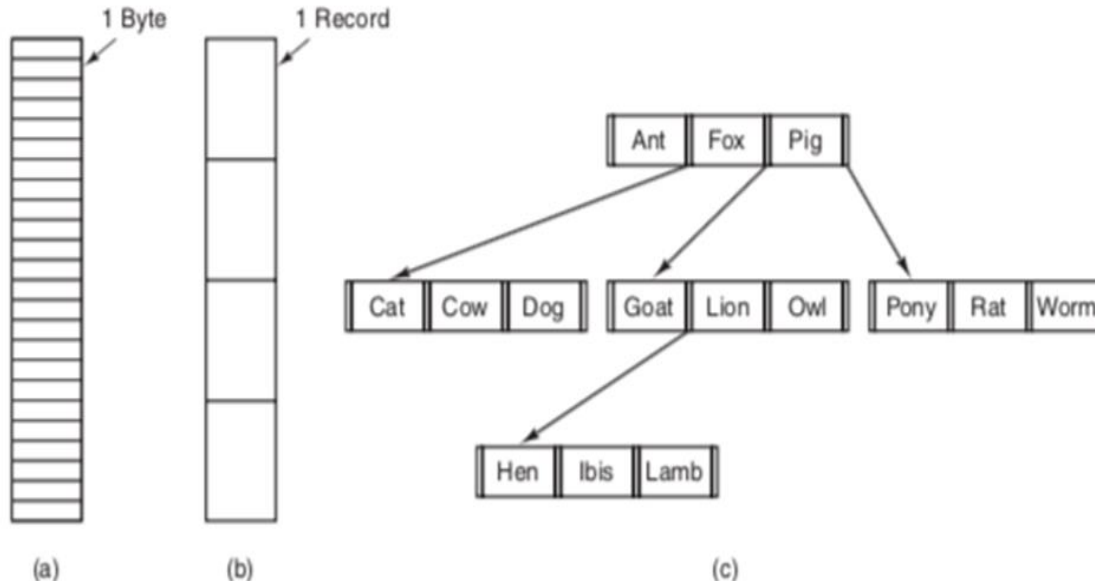
- File là gì?
- Cấu trúc file
- Kiểu file
- Truy cập file
- Các thuộc tính của file
- Các thao tác với file

File là gì?

- Từ cách tiếp cận lưu trữ:
 - File là một tập hợp các thông tin có quan hệ với nhau được ghi lên bộ lưu trữ thứ cấp.
- Từ cách nhìn của người sử dụng:
 - File là một đơn vị lưu trữ logic của thông tin trên máy tính ở đó người sử dụng có thể truy xuất.

Cấu trúc file

- File có thể cấu trúc theo một số cách sau:
 - chuỗi byte không có cấu trúc,
 - một chuỗi các bản ghi,
 - 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:
 - Là file 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 được tiến hành theo thứ tự tùy ý.

Các thuộc tính của file

- **Tên file (name):** Chuỗi kí tự (vd: hello.c)
 - Tên file là thông tin lưu ở dạng người dùng có thể đọc
 - 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:
 - ◆ Tên file
 - ◆ Phần mở rộng là mô tả về file và giúp hệ điều hành nhận ra phần mềm để mở file.
 - ◆ Tên file và phần mở rộng được ngăn cách bởi dấu chấm

Các thuộc tính của file (tiếp)

- **Định danh (Identifier):** Thẻ xác định duy nhất một file
- **Kiểu (Type):** Dùng cho hệ thống hỗ trợ nhiều kiểu file
- **Vị trí (Position):** Trỏ tới thiết bị và vị trí của file trên đó
- **Kích thước (Size):** Kích thước hiện thời/ tối đa của file
- **Bảo vệ (Protection):** Điều khiển truy nhập: Ai có thể đọc/ghi..
- **Thời gian (Time):** Thời điểm tạo, sửa đổi, sử dụng cuối ...

Các thao tác với file

- **Create** – tạo file
- **Write** – ghi vào file
- **Read** – đọc file
- **Open** – mở file
- **Close** – đóng file
- **Delete** – xóa 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

Chương 6: Hệ thống file

- Giới thiệu
- File
- **Thư mục**
- Thực hiện hệ thống file
- Ví dụ về các hệ thống file

6.3. 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 logic, các thuộc tính của file được lưu trong cấu trúc dữ liệu: **Bản ghi file**
- Các bản ghi file được lưu trong **Thư mục file**
- 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:

- Cách 1:

- ♦ Toàn bộ thuộc tính của file được lưu trong thư mục, bản thân file chỉ chứa dữ liệu. Ví dụ: MS-DOS

- Cách 2:

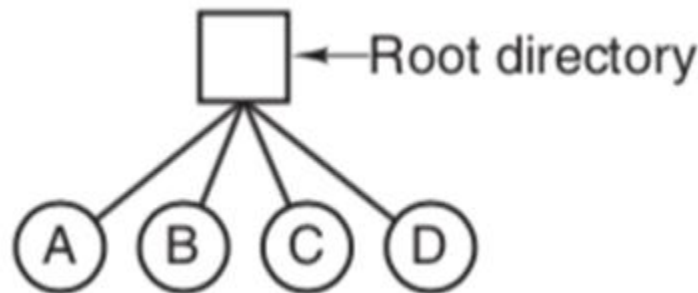
- ♦ Thư mục chỉ lưu thông tin tối thiểu cần thiết cho việc tìm vị trí của file trên đĩa.
- ♦ Các thuộc tính file còn lại được lưu trữ luôn cùng với dữ liệu của file.
- ♦ Ví dụ: thư mục sử dụng trong hệ thống file EXT2 và EXT3 của Linux

Cấu trúc hệ thống thư mục

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

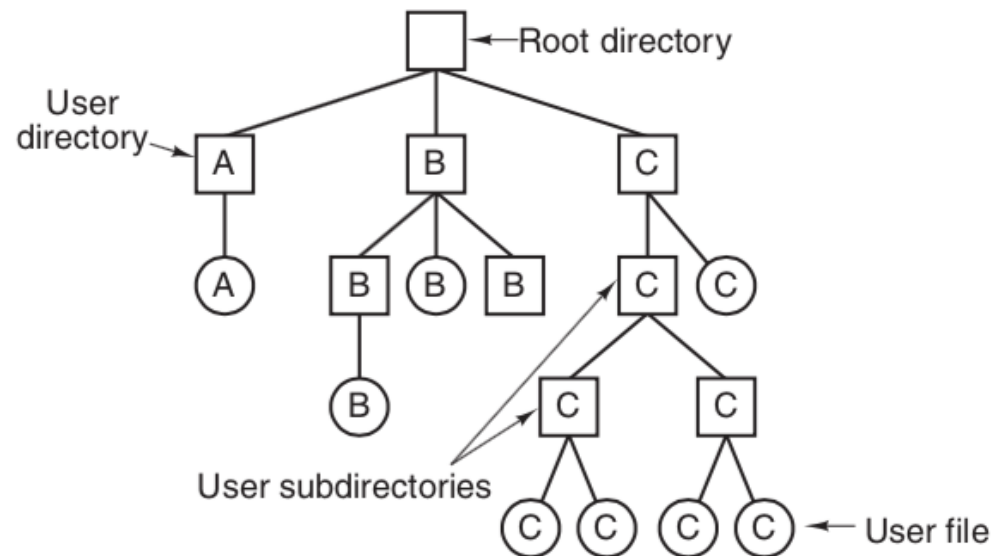
Hệ thống thư mục đơn cấp

- 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 tùy ý các thư mục con cung cấp công cụ mạnh mẽ để tổ chức công việc.



Đường dẫn

- Thông tin về vị trí của file được gọi là đường dẫn
- Khi hệ thống file được tổ 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	<code>\usr\ast\mailbox</code>
UNIX	<code>/usr/ast/mailbox</code>

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

Chương 6: Hệ thống file

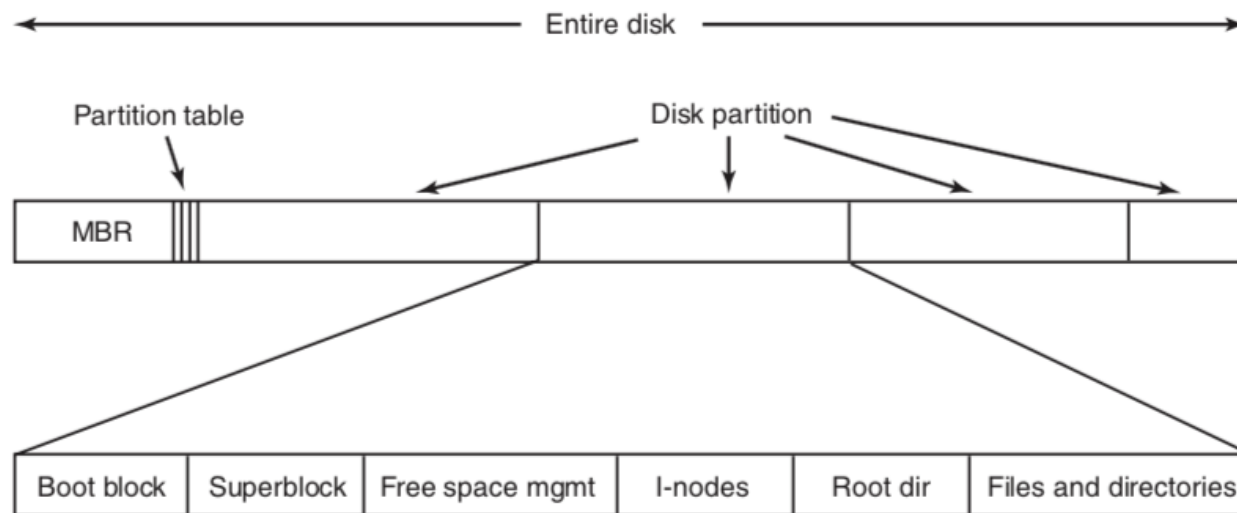
- Giới thiệu
- File
- Thư mục
- **Thực hiện hệ thống file**
- Ví dụ về các hệ thống file

6.4. Thực hiện hệ thống file

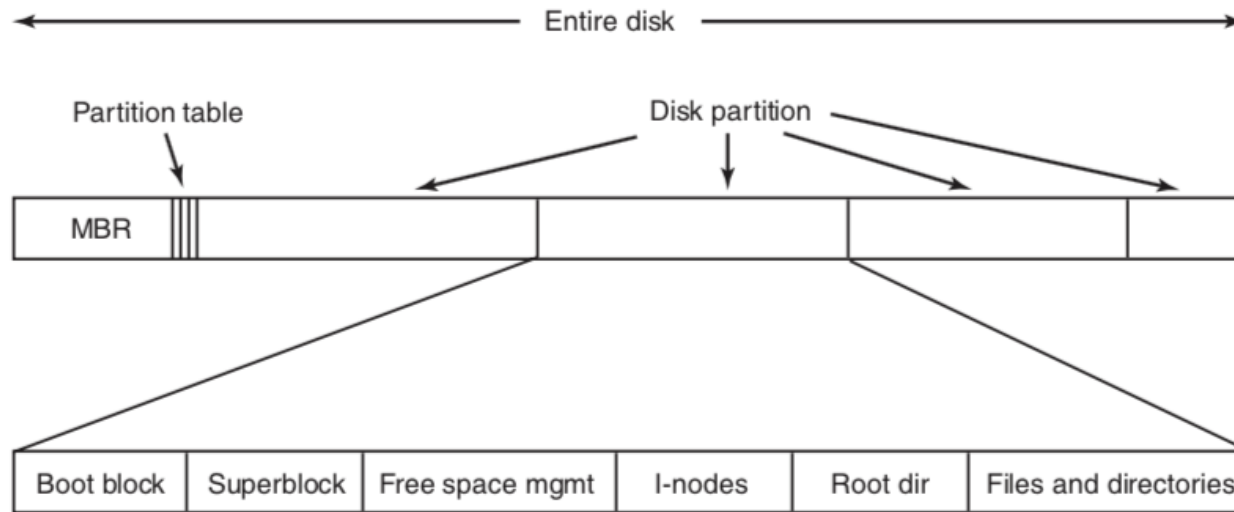
- Cấu trúc hệ thống file
- Tổ chức thực hiện file
- Tổ chức thực hiện thư mục

Cấu trúc hệ thống file

- Hệ thống File được lưu trữ trong ổ đĩa.
- Ổ 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 bộ nhớ

Tổ chức thực hiện file

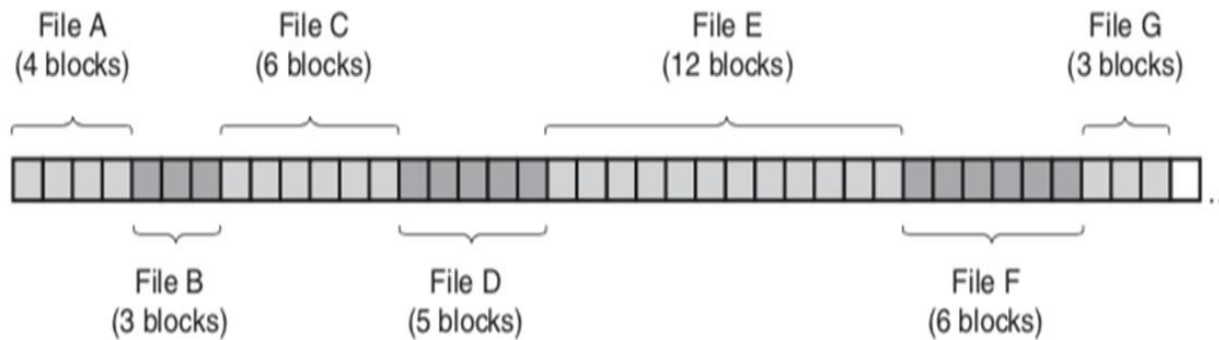
- Số lượng các sector trên một phân khu đĩa cứng là rất lớn, khó quản lý.
- Người ta thường nhóm nhiều sector trên phân khu đĩa lại thành một block (còn gọi là khối đĩa, nhóm sector, cluster, ...)
- Block là đơn vị nhỏ nhất để cấp phát đĩa

Tổ chức thực hiện file

- File được lưu 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

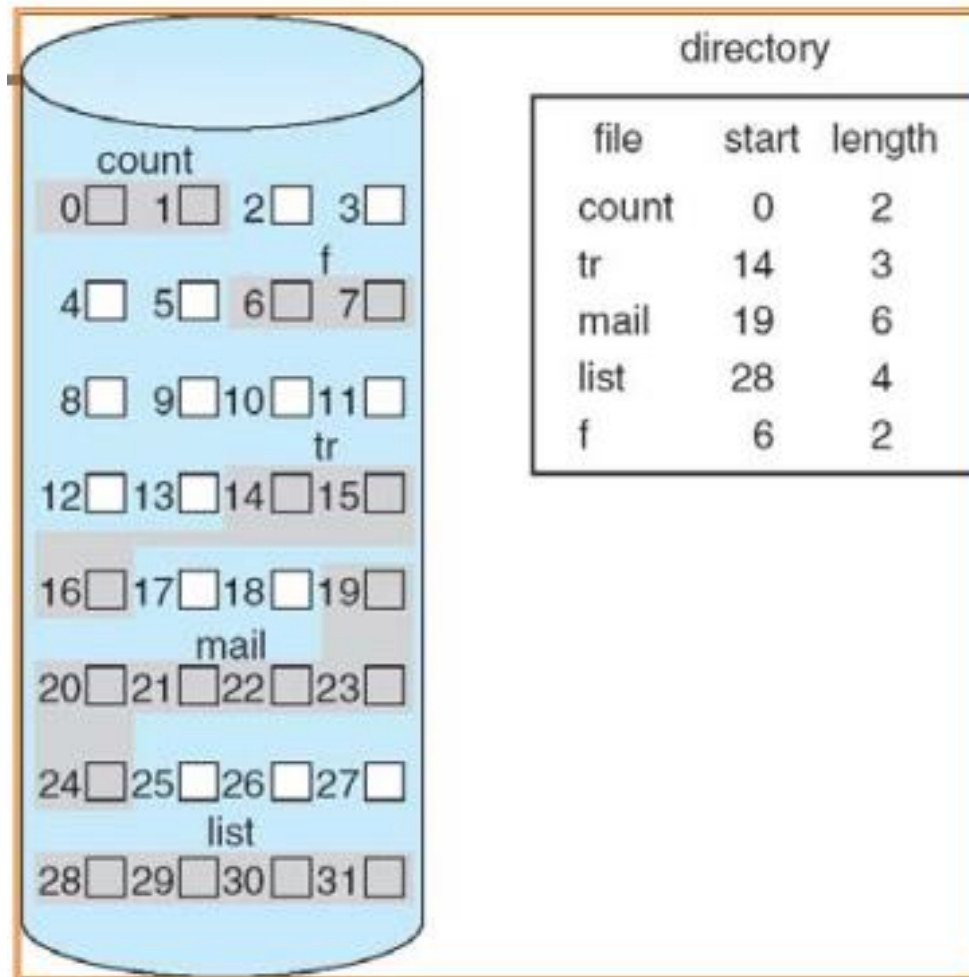
- Mỗi file được chứa vào một dãy liên tiếp các khối đĩa.
- Vị trí file trên đĩa được xác định bởi vị trí khối đầu tiên và độ dài hoặc số khối mà file đó chiếm.



Thư mục

Tên file	Bắt đầu	Độ dài
File A	0	4
File B	4	3
File C	7	6
...		

Ví dụ cấp phát liên tục



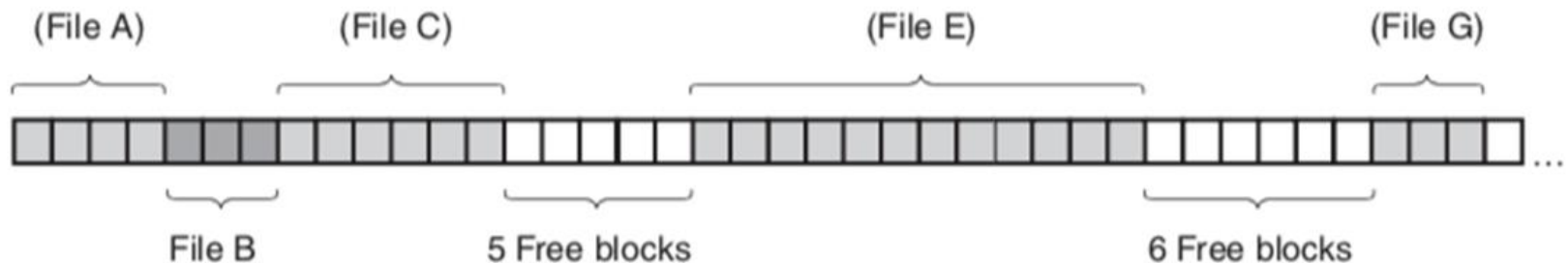
Cấp phát liên tục

- **Ưu điểm:**

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

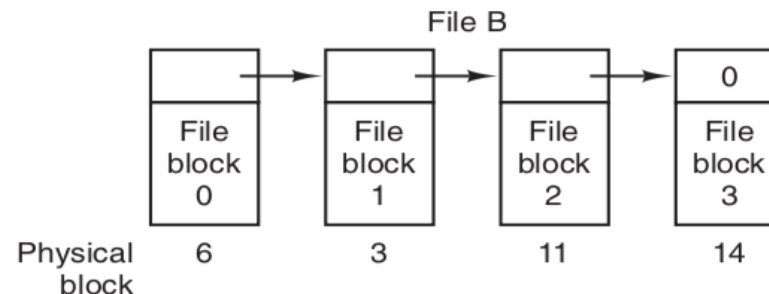
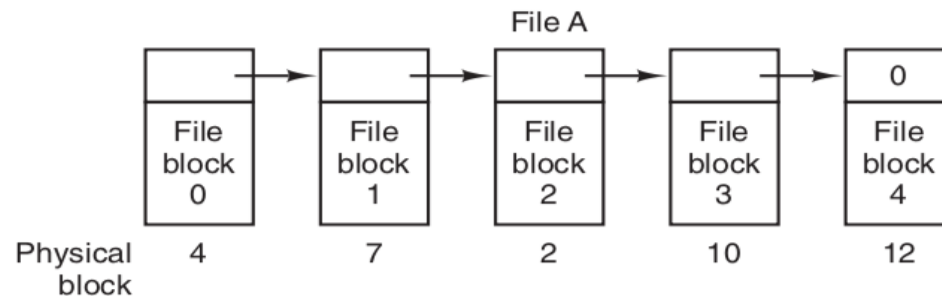
- **Nhược điểm:**

- 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

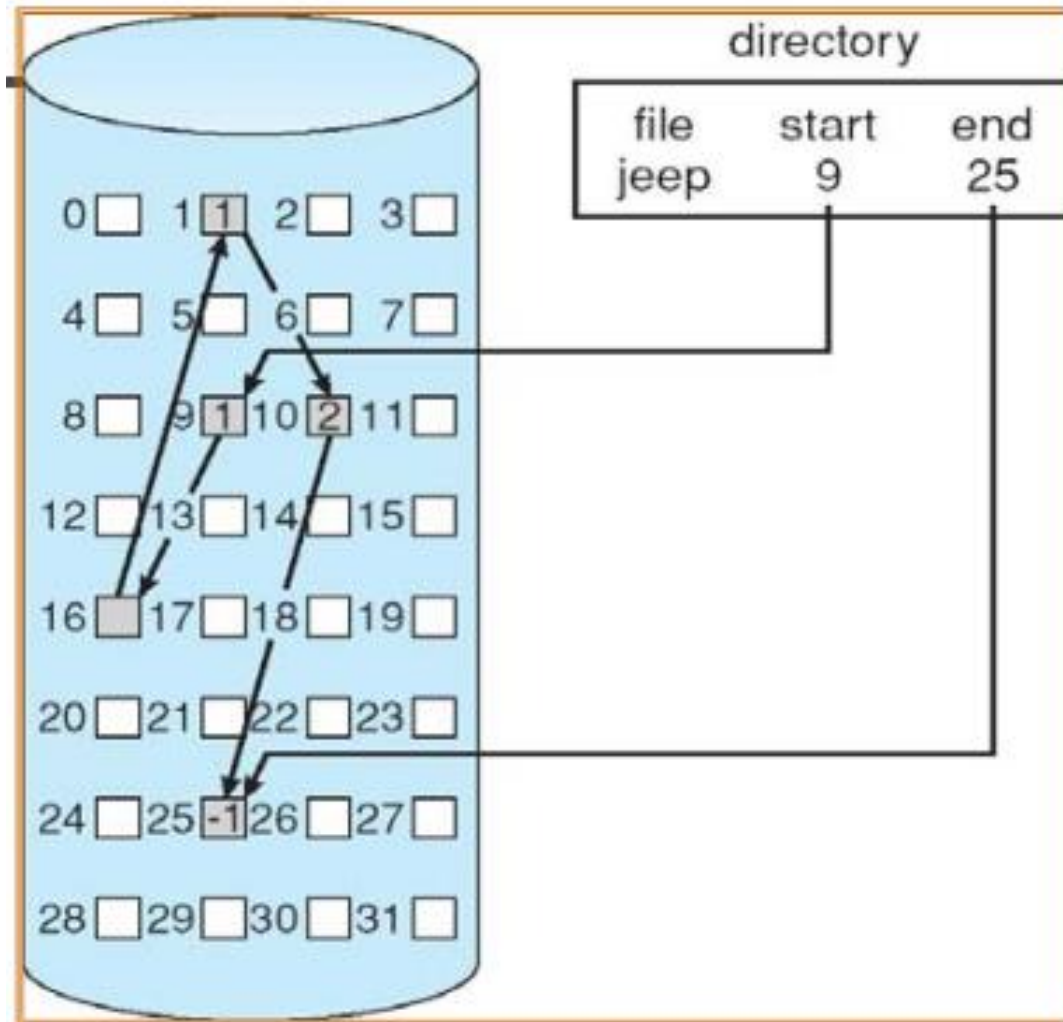


Cấp phát bằng danh sách liên kết

- File được lưu vào các khối nhớ không liên tiếp.
- Trong mỗi khối nhớ:
 - Từ nhớ đầu tiên của mỗi khối là con trỏ, trỏ tới khối tiếp theo
 - Phần còn lại của khối là dữ liệu



Ví dụ cấp phát bằng danh sách liên kết



Cấp phát bằng danh sách liên kết

■ Ưu điểm:

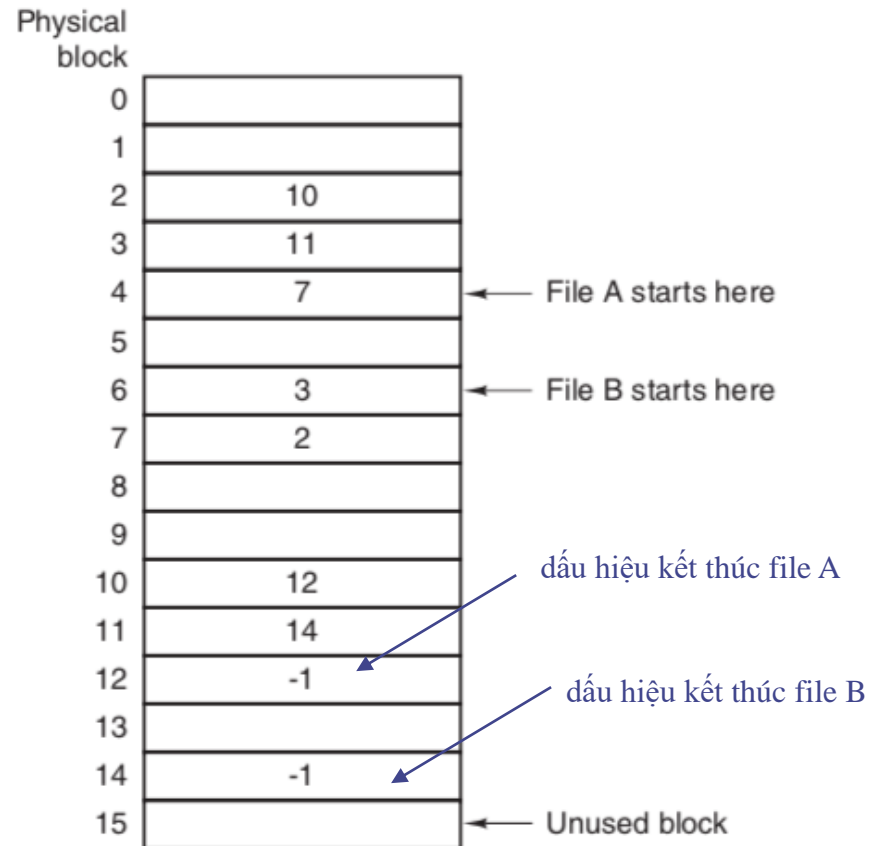
- Không xảy ra lãng phí không gian nhớ do phân mảnh ổ đĩa.
- Dễ dàng quản lý thư mục: chỉ cần lưu địa chỉ của block đầu tiên và cuối cùng.

■ 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*.
- Các con trỏ định vị này được lưu vào một bảng FAT (File Allocation Table) trong bộ nhớ chính.



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

Ví dụ cấp phát bằng danh sách liên kết sử dụng bảng nhớ

thư mục

Tên file	...	6
----------	-----	---

Bảng nhớ

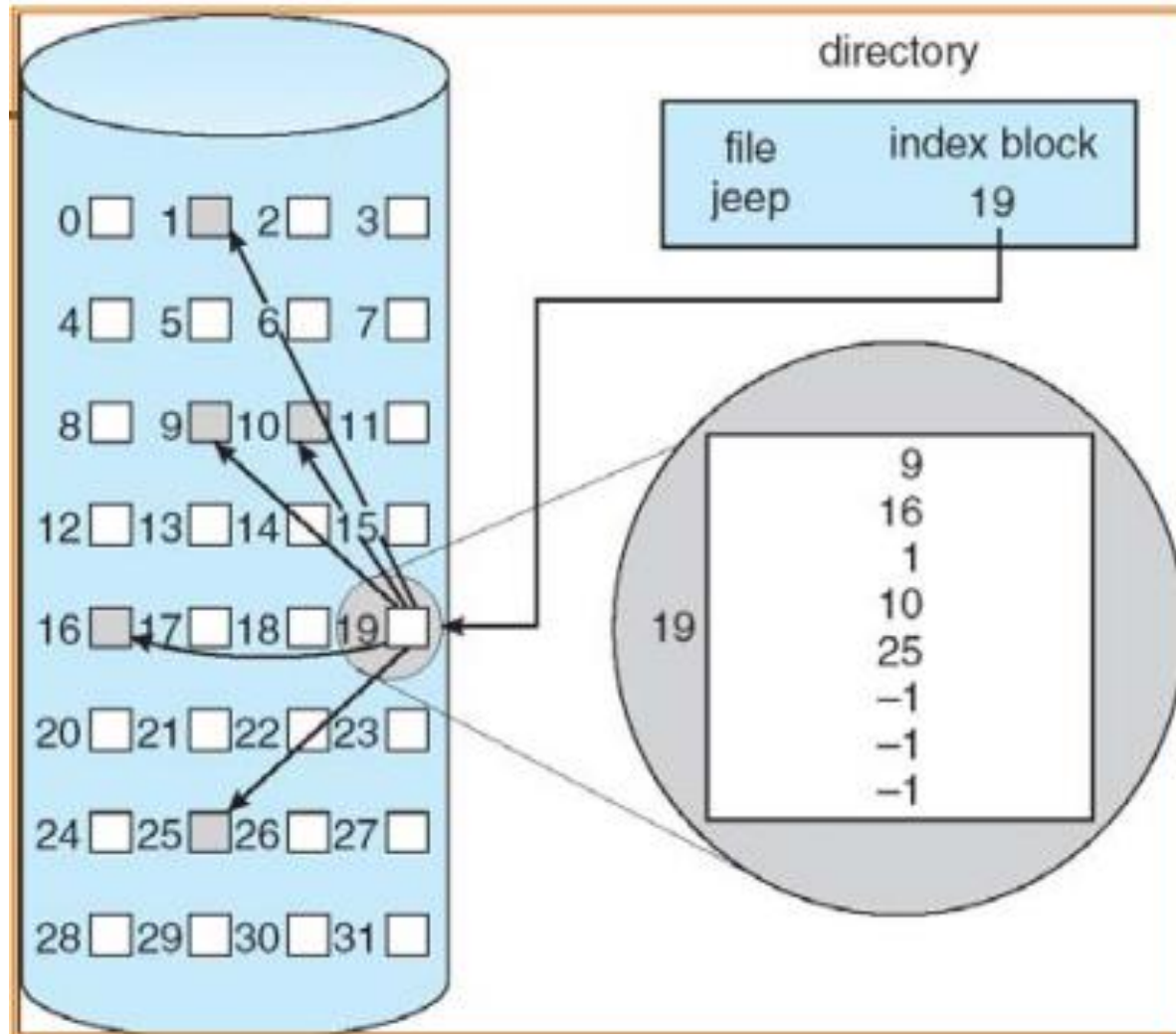
0	1	2	3	4	5	6	7	8	9	10	...
			9			3			-1		...

```
graph TD; D[Tên file | ... | 6] --> M6[6 | 3]; M3[3 | 9] --> M9[9 | -1]; M6 --> M9;
```

Cấp phát sử dụng nút chỉ số (i-node)

- Mỗi file có một khối chỉ số (index – block hay i-node) chứa danh sách các khối dữ liệu của file.
- i-node là một danh sách chứa 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.

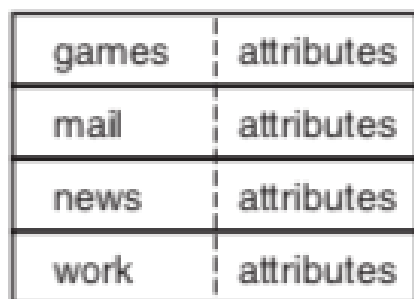
Ví dụ về cấp phát sử dụng nút chỉ số



Tổ chức thực hiện thư mục

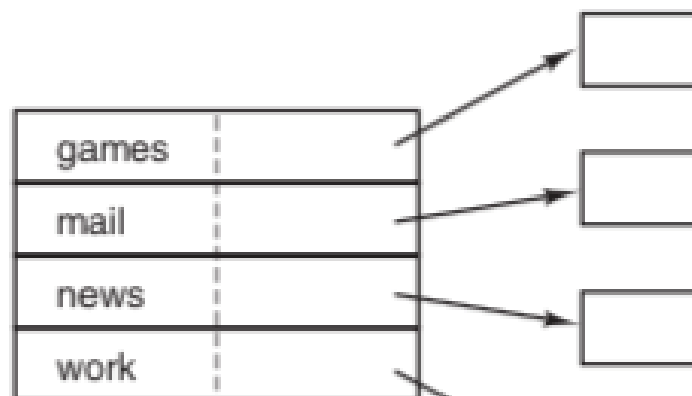
- Trong bất cứ hệ thống nào, chức năng chính của hệ thống thư mục là ánh xạ tên file sang các thông tin cần thiết để định vị dữ liệu của file trên đĩa .
 - Hệ điều hành sẽ sử dụng đường dẫn mà người dùng cung cấp để xác định entry (khoản mục) của bảng thư mục.
 - Entry thư mục sẽ cung cấp thông tin (với 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



games	attributes
mail	attributes
news	attributes
work	attributes

(a)



(b)

Data structure
containing the
attributes

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 (i-node).

Chương 6: Hệ thống file

- Giới thiệu
- File
- Thư mục
- Thực hiện hệ thống file
- Ví dụ về các hệ thống file

6.5. 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

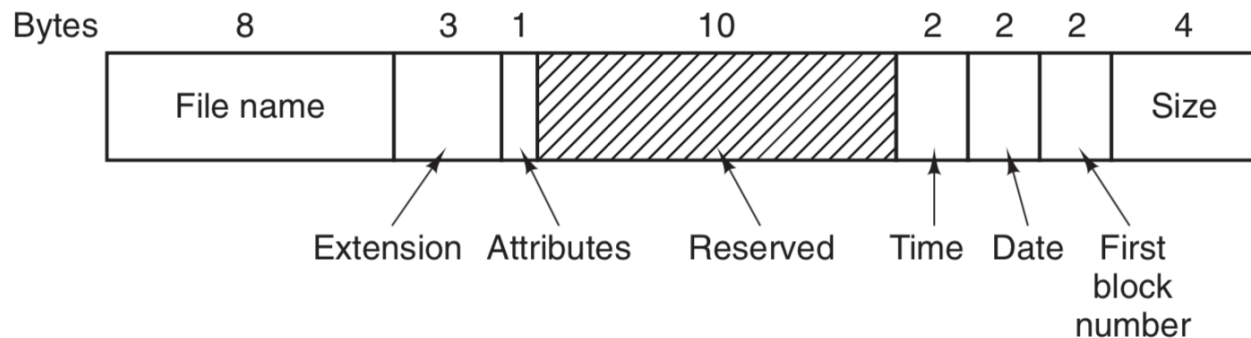
- Hệ thống file FAT của MS-DOS đã trải qua ba phiên bản: FAT-12, FAT-16, FAT-32.

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.*

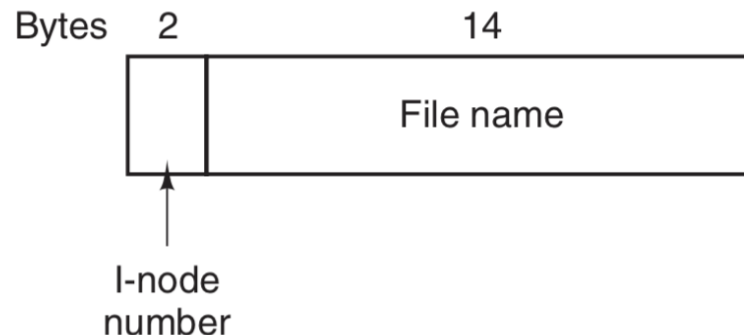
Hệ thống File MS-DOS

- 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