

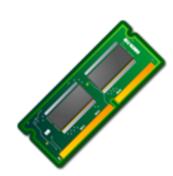
Nội dung

- Giới thiệu
- Tập tin Thư mục
- Đĩa từ
- Cài đặt hệ thống tập tin
- Minh họa một số hệ thống tập tin

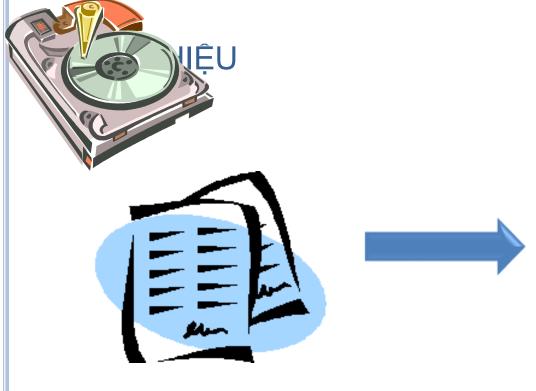
Lưu trữ dữ liệu lớn

- Dữ liệu cần phải lưu lại sau khi kết thúc process
- Nhiều process có thể truy cập dữ liệu cùng lúc









Tìm kiếm thông tin??? Làm sao biết được block nào còn trống??? Quyền hạn???

HỆ THỐNG TẬP TIN

- o cung cấp cơ chế
 - luu trữ
 - truy cập dữ liệu và chương trình trên đĩa
- Đơn vị lưu trữ: tập tin (file)
 - Thư mục là 1 dạng tập tin đặc biệt
- Một số hệ thống tập tin hiện nay:
 - FAT: FAT12, FAT16, FAT32
 - NTFS
 - Ext2, ext
 - Vfat

Nội dung

- o Giới thiệu
- Tập tin Thư mục
- Đĩa từ
- Cài đặt hệ thống tập tin
- Minh họa một số hệ thống tập tin

5

TẬP TIN — PHÂN LOẠI

Loại file	Ví dụ	Ý nghĩa
Thực thi	file.exe	File chứa mã lệnh dùng để load lên bộ nhớ và thực thi
Backup	File.bak	Backup file
Nguồn	File.c	File chứa mã nguồn gồm các dòng code, hàm,
Đối tượng	File.o	File được tổ chức thành các khối được trình liên kết hiểu
Batch	File.sh File.bat	File chứa tập các lệnh
Thư viện	File.dll File.lib	File chứa thư viện các hàm để dùng cho các chương trình
Hình ảnh	File.jpg file.bmp	File hình ảnh được mã hóa bằng các chuẩn JPEG, RLE,
Multimedia	File.mp3 File.wma File.rm	File âm thanh, video,
File text	File.txt	
Nén	File.zip	
Tài liệu	File.pdf File.doc	

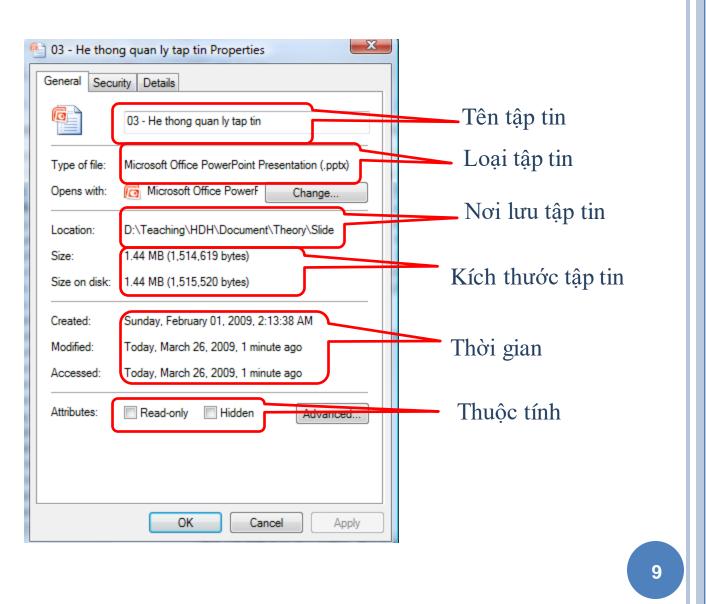
TẬP TIN

Tập tin (file)

- Tập các thông tin liên quan nhau
- Được HĐH ánh xạ trên ổ đĩa vật lý
- Gồm chuỗi các bit, byte, record, ...
- Xác định bằng tên tập tin Example.c

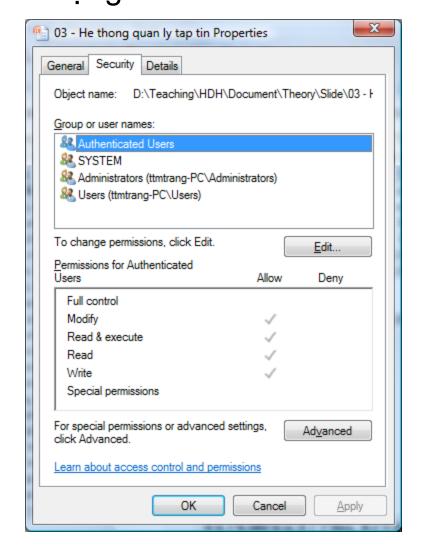
Một số loại file thông dụng

TẬP TIN — THUỘC TÍNH - 1



TẬP TIN — THUỘC TÍNH - 2

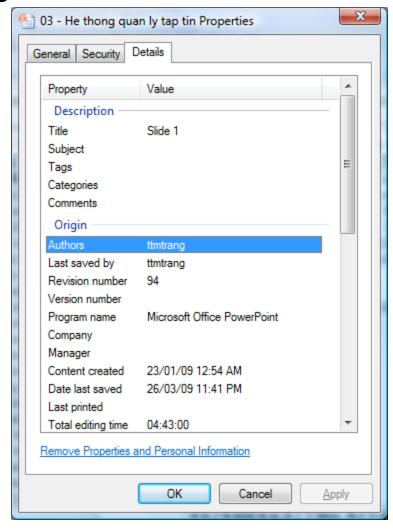
Quyền hạn sử dụng



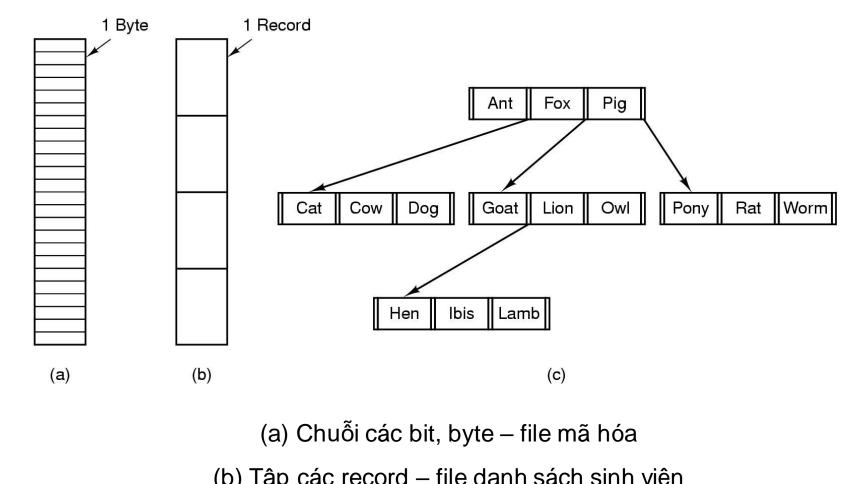
10

TẬP TIN — THUỘC TÍNH - 3

Một số thông tin khác



TẬP TIN - CẦU TRÚC - 1

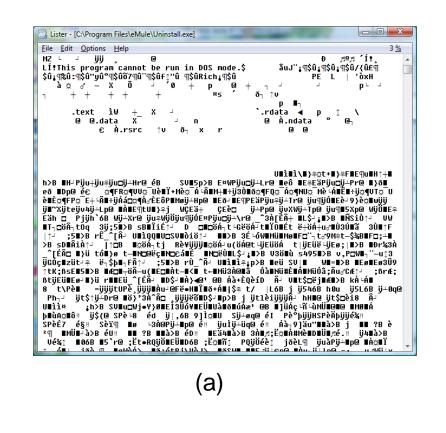


(b) Tập các record – file danh sách sinh viên

(c) Dạng cây - BTree

12

TẬP TIN — CẦU TRÚC - 2



Copyright (C) 1989, 1991 Free Software Foundation, Inc. 59 Temple Place – Suite 330, Boston, MA 02111–1307, USA Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed. Preamble
The licenses for most software are designed to take away your freedom to share and change it. By contrast, the GNU General Public License is intended to guarantee your freedom to share and change free software--to make sure the software is free for all its users. This General Public License applies to most of the Free Software Foundation's software and to any other program whose authors commit to using it. (Some other Free Software Foundation software is covered by the GNU Library General Public License instead.) You can apply it to your When we speak of free software, we are referring to freedom, not price. Our General Public Licenses are designed to make sure that you have the freedom to distribute copies of free software (and charge for this service if you wish), that you receive source code or can get it if you want it, that you can change the software or use pieces of it in new free programs; and that you know you can do these things. To protect your rights, we need to make restrictions that forbid anyone to deny you these rights or to ask you to surrender the rights. These restrictions translate to certain responsibilities for you if you distribute copies of the For example, if you distribute copies of such a program, whether gratis or for a (b)

11

13

(a) File nhị phân – (b) File text

TẬP TIN — THAO TÁC

- Tạo create
- Ghi dữ liệu write
- Đọc dữ liệu read
- Xóa delete
- Mở open
- Đóng close
- Ghi thêm dữ liệu append
- Di chuyển đến 1 khối dữ liệu bất kỳ seek
- Đọc thuộc tính get attr
- 10. Gán thuộc tính set attr
- 11. Đổi tên rename
- 12. Sao chép copy
- 13. Tìm kiếm search
- 14. Liệt kê list, dir

14

TẬP TIN - PHƯƠNG PHÁP TRUY CẬP

- Giả thiết: có 1 tập tin lưu danh sách sinh viên
- Đặt vấn đề: cần đọc thông tin của sinh viên thứ N

Kích thước mỗi record	Giải quyết	Phương pháp
khác nhau	Phải đọc từ đầu	Truy cập tuần tự
Giống nhau	 Tính vị trí logic lưu SV thứ N là p Di chuyển đến vị trí p và đọc 	Truy cập ngẫu nhiên
Khác nhau (Có 1 bảng lưu vị trí lưu mỗi SV)	 Tra bảng Di chuyển đến vị trí p và đọc 	Truy cập index

Nội dung

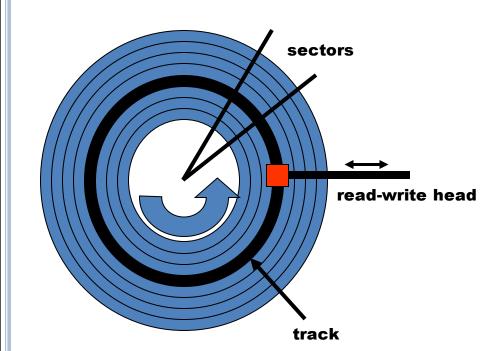
- o Giới thiệu
- o Tập tin Thư mục
- Đĩa từ
- Cài đặt hệ thống tập tin
- Minh họa một số hệ thống tập tin

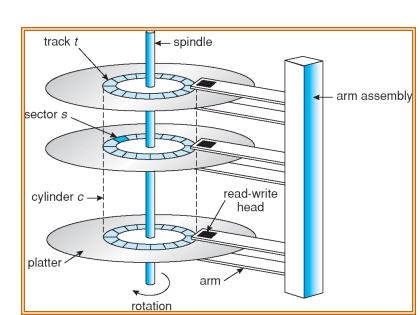
ĐĩA Từ

- Tổ chức đĩa từ
- Thuật toán đọc đĩa
- Phân loại

17

ĐĩA TỪ - CẤU TRÚC - 1







18

ĐĩA Từ - CẤU TRÚC - 2

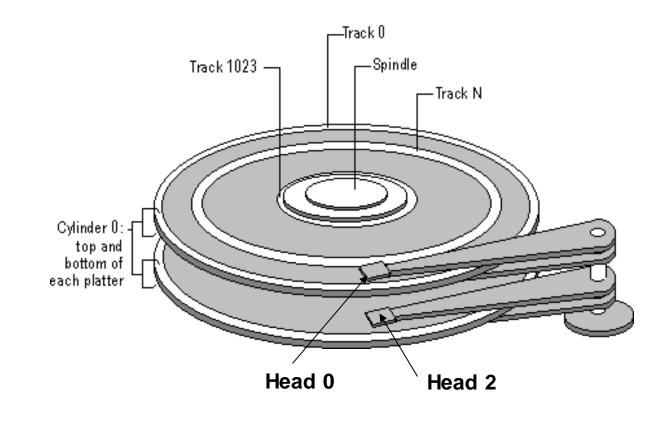
- Cấu trúc vật lý của đĩa từ:
 - Hình tròn, gồm nhiều mặt gọi là head.
 - Mỗi mặt có nhiều đường tròn đồng tâm gọi là track.
 - Trên các đường tròn (track) được chia thành các cung tròn gọi là sector.
 - Tập các track đồng tâm gọi là cylinder
 - Mỗi cung tròn chứa 4096 điểm từ (~ 4096 bit = 512 bytes).
 - Mỗi mặt có 1 đầu đọc để đọc ghi dữ liệu
 - Mỗi lần đọc/ghi ít nhất 1 cung tròn (512B).

ĐĩA TỪ - CẦU TRÚC - 3

- Vị trí của mỗi sector trong đĩa được thể hiện bằng 3 tham số: {sector, track, head}.
 - Head được đánh số từ trên xuống bắt đầu từ 0.
 - Track được đánh số từ ngoài vào bắt đầu từ 0.
 - Sector được đánh số bắt đầu từ 1 theo chiều ngược với chiều quay của đĩa.
- Mỗi lần đọc ghi N sector

19

ĐĩA TỪ - CẤU TRÚC - 4



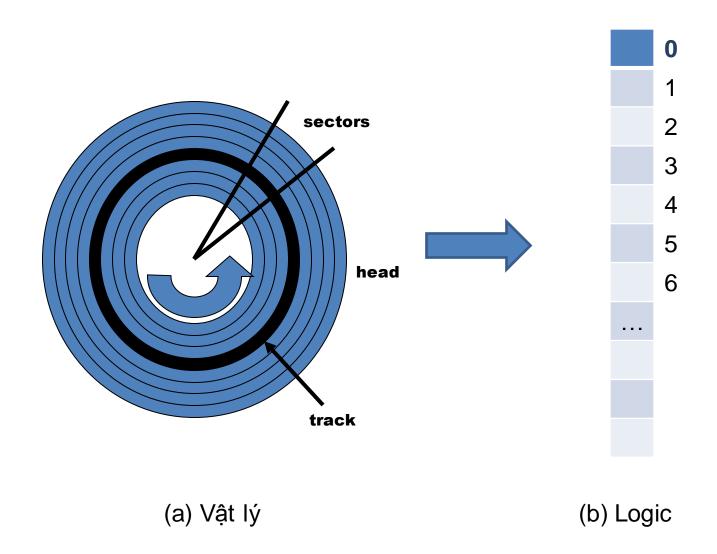
ĐĩA TỪ - DUNG LƯỢNG ĐĨA

- Kích thước đĩa phụ thuộc vào các yếu tố sau:
 - Số mặt từ, head
 - Số track trên mỗi mặt từ
 - Số sector trên mỗi track
 - Kích thước (byte) trên mỗi track.

21

22

ĐĩA Từ - TỔ CHỰC ĐĨA - 1



ĐĩA TỪ - TỔ CHỨC ĐĨA - 2

- Các thông số trên đĩa mềm 1.44MB:
 - 2 head, 80 track/head, 18 sector/track.
 - Dung lượng đĩa = 2 head/disk *80 track/head *18 sector/track = 2880 sector/disk = 0.5 KB/sector * 2880 sector/disk = 1440 KB/disk (~ 1.4MB)
 - Sector logic: 0 đến 2879 và tương ứng với các sector vật lý như sau:
 - Sector 0..17 tương ứng với sector vật lý (1,0,0)..(18,0,0)
 - Sector 18..35 tương ứng với sector vật lý (1,0,1)..(18,0,1)
 - o ...
 - Sector 2879 tương ứng với sector vật lý (18,79,1).

ĐĩA Từ - TỔ CHỰC ĐĨA - 3

Đối từ sector vật lý sang sector logic

ĐĩA Từ - TỔ CHỰC ĐĨA - 4

Đối từ sector logic sang sector vật lý

$$st$$
 : số sectors / track
 th : số tracks / side (head)
 $side$: số tracks / side (head)
 $side$: số lượng head
 $t = l \operatorname{div}(st * side)$
 $t = (l \operatorname{div} st) \operatorname{mod} side$
 $t = (l \operatorname{div} st) \operatorname{mod} side$

25

Đĩa mềm 1.44 MB

- Có 2 head /disk, 80 track /head, 18 sector /track
- Dung lượng đĩa:
 - 2 head/disk * 80 track/head * 18 sector/track = 2880 sector/disk = 0.5 KB/sector * 2880 sector/disk = 1440 KB/disk (~ 1.44 MB)
- Sector logic có chỉ số từ 0 đến 2879 và tương ứng với sector vật lý như sau:

Sector Logic	Sector vật lý (Sector, Track, Head)			
0	(1,0,0)			
1	(2,0,0)			
17	(18, 0, 0)			
18	(1,0,1)			
19	(2,0,1)			
35	(18, 0, 1)			
36	(1, 1, 0)			
37	(2, 1, 0)			
	27			

Bài tập

- Một đĩa cứng có 16 head, mỗi mặt có 684 track, và mỗi track có 18 sector thì sẽ có kích thước là bao nhiêu Megabyte?
- Cho biết sector vật lý (head 0, track 21, sector 6) tương ứng với sector logic nào trên đĩa mềm 1.44MB
- Cho các chỉ số sector logic sau hãy cho biết chỉ số sector vật lý tương ứng:
 - a. 347 b. 348 c. 689 d. 690

l = t*side*st + h*st + s - 1

 $s = (I \bmod st) + 1$ $t = I \operatorname{div} (st * side)$ $h = (I \operatorname{div} st) \operatorname{mod} \operatorname{side}$

26

ĐĩA TỪ - THUẬT TOÁN ĐỌC ĐĨA

- First-Come-First-Serve (FCFS)
- Shortest Seek Time First (SSTF)
- SCAN, C-SCAN
- Look, C-Look

FIRST COME FIRST SERVE - FCFS

- Phục vụ theo thứ tự yêu cầu
- Đơn giản nhưng không đáp ứng tốt dịch vụ

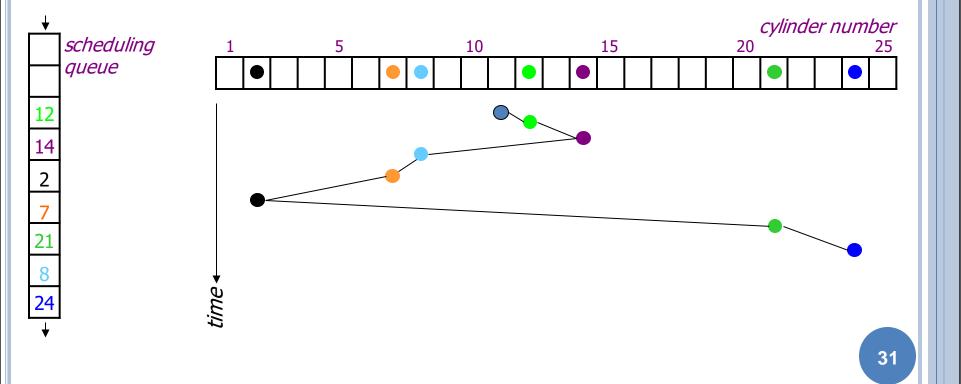
Các khối cần đọc (đầu đọc hiện tại tại vị trí 11):

24 8 21 **7 2** 14 12 cylinder number scheduling queue 30

29

SHORTEST SEEK TIME FIRST - SSTF

- Chọn nhu cầu gần với vị trí hiện hành nhất.
- → Có nhiều yêu cầu chờ ..chờ...và chờ...

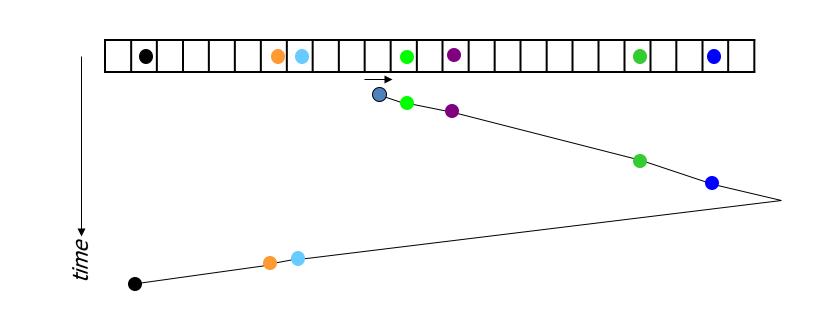


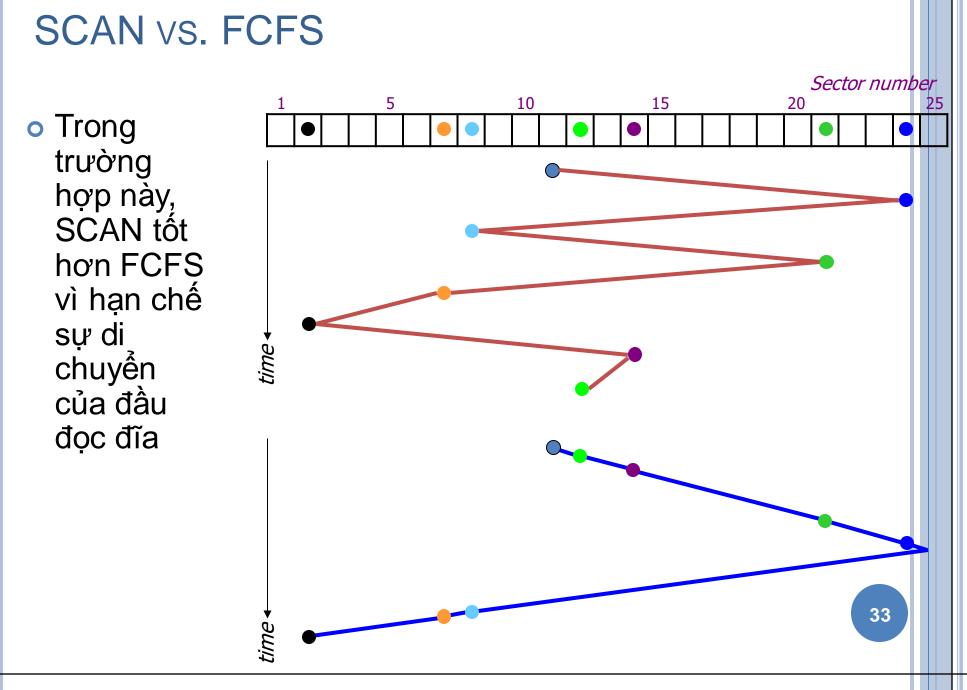
SCAN

- Di chuyển đầu đọc về 1 phía của đĩa đến block xa nhất sau đó di chuyển về phía kia.
- Còn gọi là thuật toán thang máy.

Các khối cần đọc (đầu đọc hiện tại tại vị trí 11):

12 14 2 7 21 8

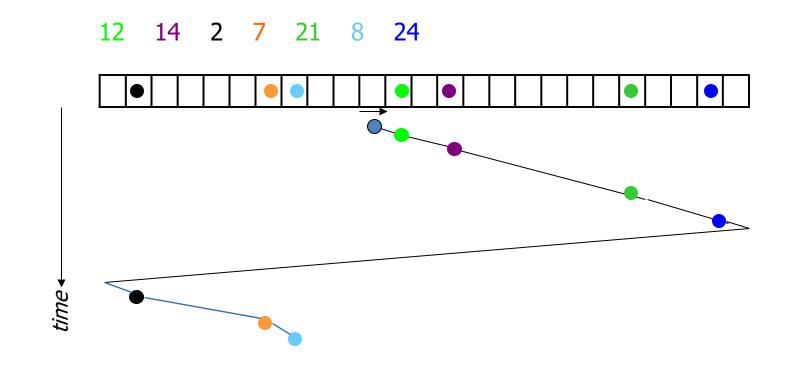




C-SCAN

- Nguyên tắc:
 - Tương tự thuật toán SCAN.
 - Chỉ khác khi di chuyển đến 1 đầu của đĩa thì trở về vị trí bắt đầu của đĩa.

Các khối cần đọc (đầu đọc hiện tại tại vị trí 11):



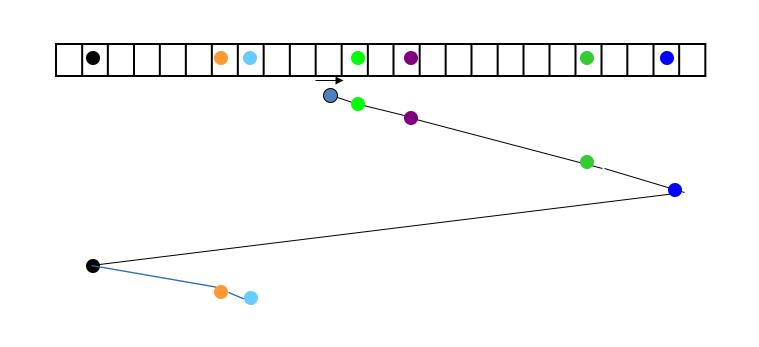
LOOK - C-LOOK

- Nhận xét:
 - Hai thuật toán lập lịch SCAN và C-SCAN luôn luôn di chuyển đầu đọc của đĩa từ đầu này sang đầu kia và di chuyển đến khối cuối cùng ở mỗi hướng.
- Nguyên tắc:
 - Giống SCAN và C-SCAN nhưng chỉ di chuyển đầu đọc đến khối xa nhất chứ không đến cuối.

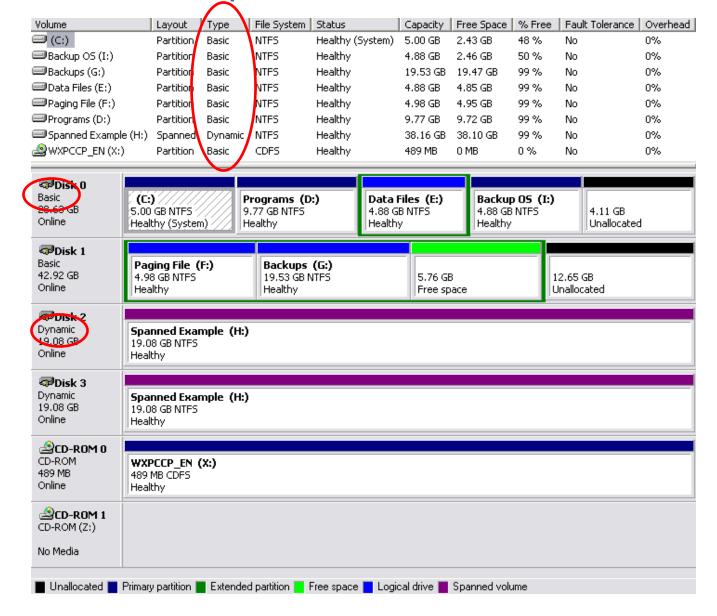
LOOK - C-LOOK

Các khối cần đọc (đầu đọc hiện tại tại vị trí 11):

12 14 2 7 21 8 24



ĐĩA TỪ - PHÂN LOAI

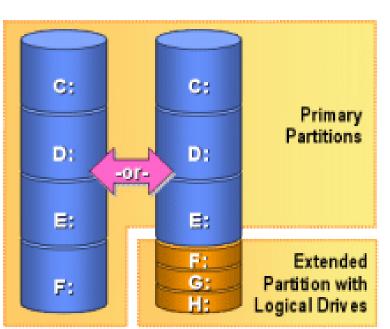


BASIC TYPE

35

37

- Tối đa 4 phân vùng (partition)
- Loại partition
 - Primary
 - Mỗi phân vùng: ấn định 1 ký tự
 - Extended
 - Có thể tạo nhiều logicial drive, ấn định 1 ký tự cho mỗi logical drive

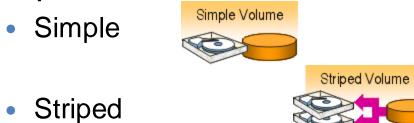


38

DYNAMIC TYPE

- Chia thành nhiều volume
 - Không bị giới hạn số lượng
- Loại Volume

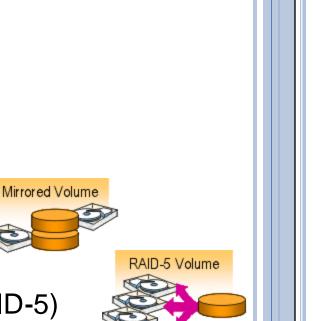




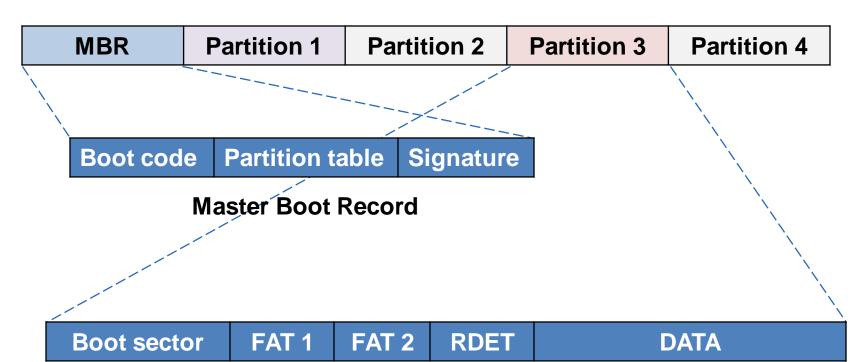
Spanned

Mirrored (RAID-1)

Redundant Array of Independent Disks (RAID-5)



ĐĩA TỪ - CẦU TRÚC



MASTER BOOT RECORD

Address			Doco	Size in		
<u>Hex</u>	<u>Oct</u>	<u>Dec</u>	Description		bytes	
0000	0000	0	Code Area		440 (max. 446)	
01B8	0670	440	Optional Disk	Optional Disk signature		
01BC	0674	444	Usually Nulls	2		
01BE	0676	446	Table of prima (Four 16-byte Partition Table	64		
01FE	0776	510	55h MBR			
01FF	0777	511	signature; 0xAA55 ^[1]		2	
	512					

Nguồn: wikipedia

Bảng mô tả các phân vùng - 1

Offset	Field length (bytes)	Description
0x00	1	status[7] (0x80 = bootable, 0x00 = non-bootable, other = invalid[8])
0x01	3	CHS address of first block in partition. The format is described in the next 3 bytes.
0x01	1	head ^[10]
0x02	1	sector is in bits 5–0[11]; bits 9–8 of cylinder are in bits 7–6
0x03	1	bits 7–0 of cylinder ^[12]
0x04	1	partition type ^[13]
0x05	3	CHS address of last block in partition.[14] The format is described in the next 3 bytes.
0x05	1	head
0x06	1	sector is in bits 5-0; bits 9-8 of cylinder are in bits 7-6
0x07	1	bits 7–0 of cylinder
0x08	4	LBA of first sector in the partition
0x0C	4	number of blocks in partition, in little-endian format

Nguồn: wikipedia

Bảng mô tả các phân vùng - 2

Type:

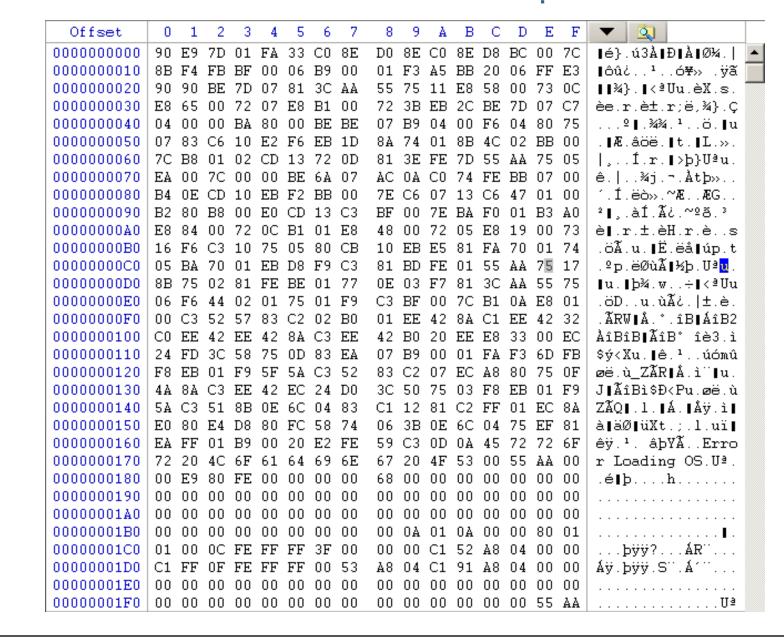
0x07 : Phân vùng chứa "Windows"
0x83 : Phân vùng chứa "Linux"
0x00 : Phân vùng không sử dụng.

Tham khảo thêm:

http://www.win.tue.nl/~aeb/partitions/partition_types-1.html

43

MASTER BOOT RECORD - VÍ Dụ - 1



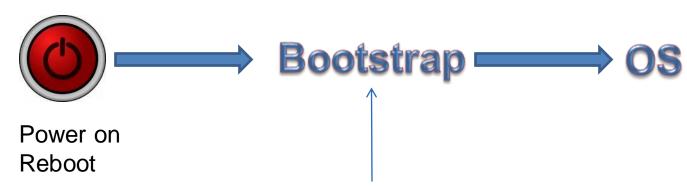
44

Master Boot Record – Ví dụ - 2

Parti	Starting		Ending		Status	Туре	First	Sector#		
tion	Н	Т	S	Н	Т	S			sector	
1	1	0	1	254	1023	63	0x80	0x0C	63	78140097
2	0	1023	1	254	1023	63	0x00	0x0F	78140160	78156225

QUÁ TRÌNH BOOT HỆ THỐNG

- POST (Power-On-Self-Test)
- 2. Tải MBR để đọc thông tin bảng phân vùng.
- 3. Tìm phân vùng "active".
- 4. Chuyển quyền điều khiển về cho đoạn mã chương trình nằm trong Boot Record của phân vùng "active"
- 5. Tải HĐH tại phân vùng "active".



Khởi tạo hệ thống CPU, device controller, main memory, load đoạn code khởi động hđh

46

Nội dung

- o Giới thiêu
- Tập tin Thư mục
- o Đĩa từ
- Cài đặt hệ thống tập tin
- Minh họa một số hệ thống tập tin

CÀI ĐẶT HỆ THỐNG TẬP TIN

- Hệ thống tập tin chứa thông tin gì?
- Thành phần
- Phương pháp cấp pháp vùng nhớ
- Quản lý không gian đĩa trống

45

CÀI ĐẶT HỆ THỐNG TẬP TIN

- Hệ thống tập tin chứa:
 - Cách boot hệ điều hành
 - Tổng số block
 - Block trống
 - Cấu trúc cây thư mục (thư mục, tập tin)

CÀI ĐẶT HỆ THỐNG TẬP TIN

- Các thành phần trong hệ thống tập tin
 - Boot control block
 - Thông tin để boot hđh từ volume này
 - UFS: boot block, NTFS: partition boot sector, FAT: boot sector
 - Volume control block
 - Thông tin chi tiết volume
 - UFS: superblock, NTFS: master file table, FAT: boot sector
 - File control block
 - o Tổ chức: tổ chức các tập tin ntn?
 - o File: thông tin chi tiết của 1 tập tin
 - UFS: inode, NTFS: master file table, FAT: FAT&RDET

51

CÀI ĐẶT HỆ THỐNG TẬP TIN











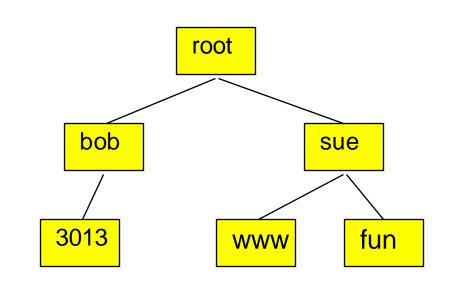




Thư mục

THU MUC - 1

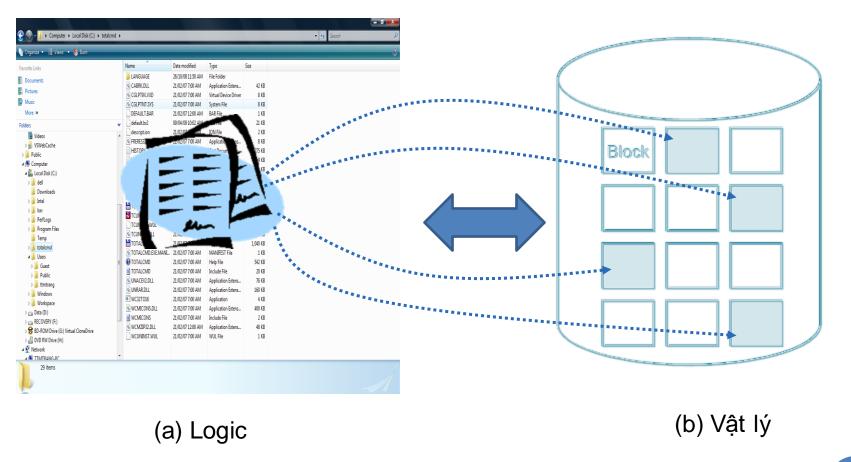
- Là một tập tin đặc biệt
- Giúp cho việc quản lý các tập tin dễ dàng hơn.
 - Gom nhóm các tập tin vào trong các thư mục theo ý nghĩa và mục đích sử dụng của người dùng.
 - Giúp định vị các tập tin 1 cách nhanh chóng.



Thư mục - Đường dẫn (Path)

- Dùng để xác định vị trí lưu tập tin khi hệ thống được tổ chức thành cây thư mục:
 - Đường dẫn tuyệt đối:
 - Ví dụ: "C:\Downloads\software\baigiang.doc"
 - Đường dẫn tương đối:
 - o Ví dụ: "software\baigiang.doc" nếu thư mục hiện hành là "C:\Downloads\"
- Các thư mục đặc biệt:
 - Thư mục hiện hành (.)
 - Thư mục cha (..)

CÀI ĐẶT HỆ THỐNG TẬP TIN



54

53

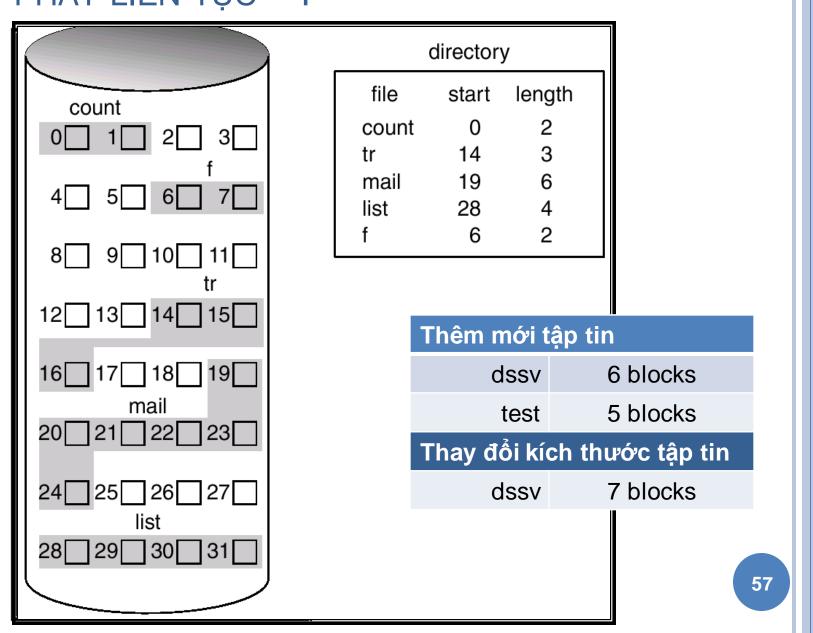
CÀI ĐẶT HỆ THỐNG TẬP TIN

- Làm sao map giữa cây thư mục với các block trên thiết bị lưu trữ
- Mỗi tập tin lưu:
 - Lưu ở block nào?
 - Khi tạo mới, sử dụng block nào?
 - ⇔ cấp phát???

CÁP PHÁT VÙNG NHỚ CHỬA TẬP TIN

- Phương pháp cấp phát:
 - Là cách thức cấp phát vùng nhớ (block) cho tập tin
 - Phương pháp:
 - Cấp phát liên tục
 - Cấp phát bằng danh sách liên kết
 - o Cấp phát bằng chỉ mục

CÁP PHÁT LIÊN TỤC - 1



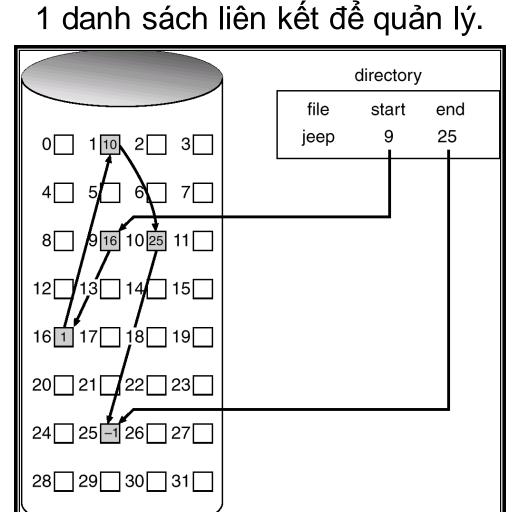
CÁP PHÁT LIÊN TỤC - 2

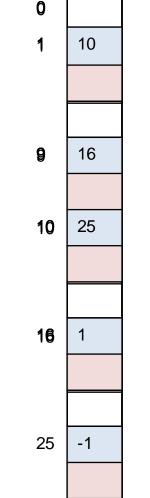
- Cấp phát 1 số block liên tục trên đĩa để lưu trữ nội dung tập tin
- Nhận xét:
 - Đơn giản: chỉ cần quản lý số hiệu khối bắt đầu và tổng số block chiếm bởi tập tin.
 - Truy cập nội dung tập tin nhanh chóng vì các block nằm kề nhau.
 - Gây lãng phí bộ nhớ.
 - Khó khăn khi tập tin mở rộng kích thước.

58

CÁP PHÁT BẰNG DANH SÁCH LIÊN KẾT - 1

 Nội dung tập tin được lưu trữ ở những block không cần liên tục. Các block này được xâu chuỗi tạo thành





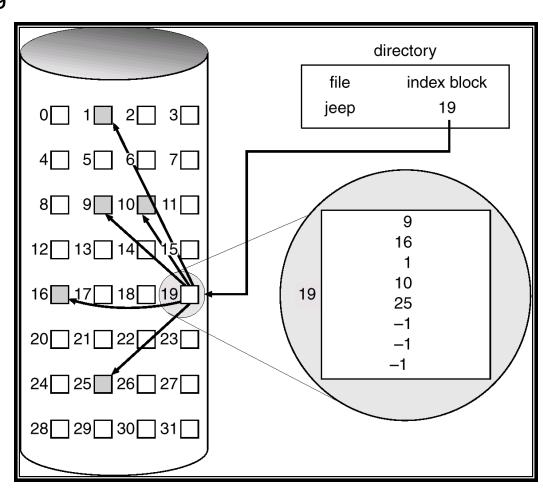
CÁP PHÁT BẰNG DANH SÁCH LIÊN KẾT - 2

- Nhận xét:
 - Đơn giản: Chỉ cần quản lý block bắt đầu.
 - Tận dụng hiệu quả không gian đĩa.
 - Truy cập tập tin lâu hơn vì đầu đọc phải di chuyển nhiều giữa các khối không liên tiếp.
 - Không thể truy cậpngẫu nhiên
 - Khối dữ liệu bị thu hẹp lại vì mỗi khối phải dùng 1 phần để lưu phần liên kết đến khối kế tiếp.

0

CÁP PHÁT BẰNG CHỈ MỤC (INDEX) - 1

- o Mỗi tập tin:
 - Index block: Lưu địa chỉ các block của 1 tập tin bằng 1 mảng



61

59

CÁP PHÁT BẰNG CHỈ MỤC (INDEX) - 2

- Nhận xét:
 - Truy cập ngẫu nhiên
 - Tốn không gian lưu bảng chỉ mục

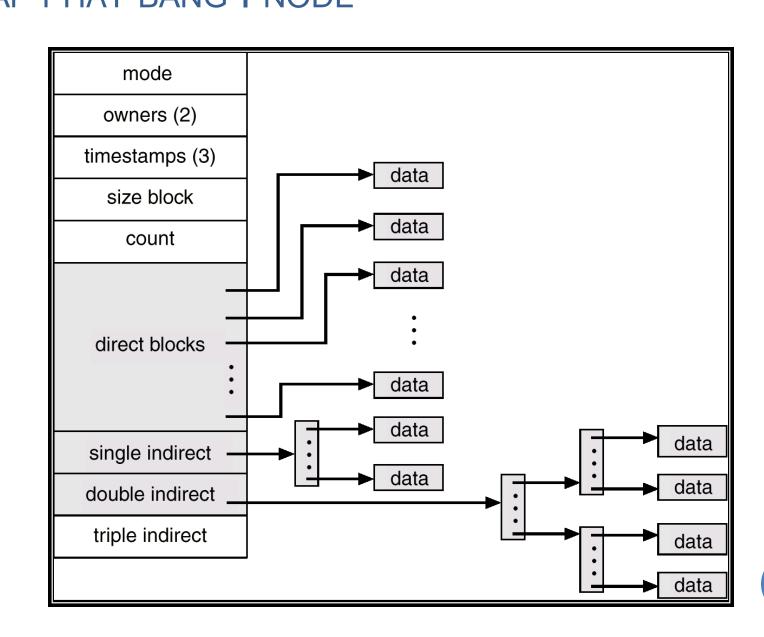
Theo danh sách liên kết	Theo chỉ mục
Truy cập tuần tự	Truy cập ngẫu nhiên
Tốn n*entry lưu địa chỉ block	Tốn 1 block
Không bị giới hạn kích thước	Giới hạn

62

CÁP PHÁT BẰNG CHỈ MỤC (INDEX) - 3

- Phương pháp mở rộng:
 - Chỉ mục kết hợp với danh sách liên kết
 - Liên kết nhiều index block để lưu file lớn
 - VD: dùng entry cuối lưu địa chỉ của block index tiếp theo
 - Chỉ mục đa cấp
 - o Index block cấp 1 lưu danh sách các index block 2,
 - VD: với 2 cấp, mỗi block có1024 entry → quản lý ??? block
 - Chỉ mục kết hợp
 - Sử dụng N entry
 - N-3 entry đầu lưu địa chỉ của các data block
 - Entry ké tiép lưu index block cấp 1
 - Entry kế tiếp lưu index block cấp 2
 - Entry ké tiép lưu index block cấp 3
 - VD: I-node

CÁP PHÁT BẰNG I-NODE



QUẢN LÝ KHÔNG GIAN ĐĨA TRỐNG

- Ghi nhận danh sách các block trống
- Phương pháp:
 - Bit vector
 - Danh sách liên kết (linked list)
 - Nhóm (grouping)
 - Đếm (counting)
 - Bảng đồ không gian (space maps)

BIT VECTOR

- Mỗi block được đại diện bằng 1 bit
 - 0: chứa dữ liệu
 - 1: trống
- Nhận xét:
 - Đơn giản
 - Tính toán nhanh
 - Khi cần kiểm tra → load bit vector lên bộ nhớ

65

66

LINKED — GROUPING - COUNTING

- Linked
 - Các block trống liên kết với nhau: block trống thứ N lưu địa chỉ của block trống thứ N+1
 - Chỉ cần lưu địa chỉ block trống đầu tiên
- Grouping
 - Tương tự linked
 - Lưu địa chỉ của N block trống tiếp theo
- Counting:
 - Với mỗi N block trống liên tiếp: lưu địa chỉ của block trống đầu tiên và số lượng block trống

Nội dung

- o Giới thiêu
- o Tập tin Thư mục
- o Đĩa từ
- o Cài đặt hệ thống tập tin
- Minh họa một số hệ thống tập tin

67

68

MỘT SỐ HỆ THỐNG TẬP TIN

- FAT
- NTFS
- I-node

HỆ THỐNG TẬP TIN FAT - 1

- FAT: File Allocation Table
- Các phiên bản của FAT: FAT12, FAT16, FAT32
 - 12,16,32: Số bít dùng để đánh STT các khối

Boot sector	FAT1	FAT2 (backup	l	ory Other	directories	and files		
0000			0000	0001	0002	0003	0004	0005
0001	0000] [
0002	0003]	empty	empty	/ File1	File1	File1	File2
0003	0004	L						
0004	FFFF]	0006	0007	8000	0009	0010	0011
0005	0006							
0006	0008		File2	File3	File2	empty	empty	empty
0007	FFFF							
0008	FFFF		0012	0013	0014	0015	0016	0017
0009	0000							
]	empty	empty	empty	empty	empty	empty

70

HỆ THỐNG TẬP TIN FAT - 2

	20		S
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

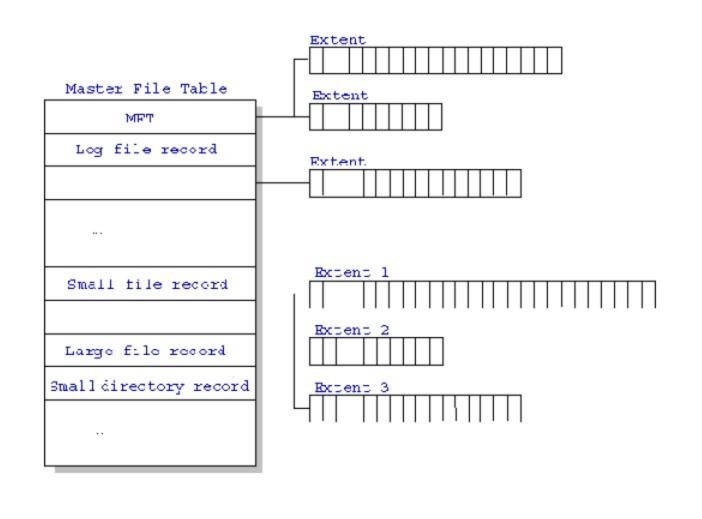
HỆ THỐNG TẬP TIN NTFS - 1

- NTFS: New Technology File System
- Sử dụng MFT (Master File Table).
 - MFT là 1 Metadata file bao gồm 1 danh sách các trường chứa thông tin về mỗi tập tin lưu trữ trên đĩa.
 - Thông tin trong MFT có thể giúp thiết lập các thuộc tính bảo vệ, phục hồi, tìm kiếm, thiết lập quota... cho từng tập tin, thư mục trên đĩa.

Partition boot Master File Table System files File area

69

HỆ THỐNG TẬP TIN NTFS - 2



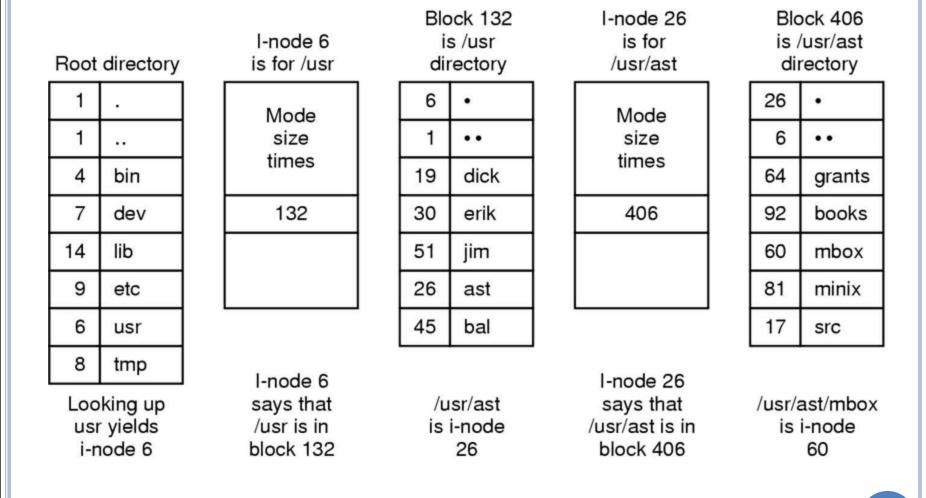
73

HỆ THỐNG TẬP TIN TRÊN UNIX/LINUX: I-NODE - 1 Data blocks **Boot block** Super block inode list mode owners (2) timestamps (3) data size block data count **→** data direct blocks data **→** data single indirect data data double indirect

triple indirect

74

HỆ THỐNG TẬP TIN TRÊN UNIX/LINUX: I-NODE - 2



75