

Nguyên lý hệ điều hành

Nguyễn Hải Châu
Khoa Công nghệ Thông tin
Trường Đại học Công nghệ



Các hệ thống lưu trữ

Cấu trúc đĩa
Lập lịch đĩa
Quản lý đĩa
Quản lý không gian swap
Cấu trúc RAID
...



Cấu trúc đĩa

- Các ổ đĩa được đánh địa chỉ như một mảng lớn, 1 chiều với mỗi phần tử là một khối logic – đơn vị truyền nhận nhỏ nhất
- Mảng một chiều nói trên được ánh xạ vào các sector đĩa một cách tuần tự
 - Sector 0 là sector đầu tiên trên rãnh đầu tiên của track nằm ngoài cùng trên đĩa
 - Quá trình ánh xạ theo thứ tự chỉ số: sector, track, cylinder (từ ngoài vào trong)



Lập lịch đĩa (1)

- HĐH cần sử dụng phần cứng một cách hiệu quả - với đĩa: thời gian truy cập nhanh và băng thông lớn
- Thời gian truy cập bị ảnh hưởng bởi:
 - *Seek time* (thời gian dịch đầu đọc): Thời gian chuyển đầu đọc đến cylinder chứa sector cần truy cập
 - *Rotational latency* (Độ trễ quay): Thời gian chờ đĩa quay để đầu đọc gặp sector cần truy cập



Lập lịch đĩa (2)

- *Seek time*: Càng nhỏ càng tốt
 - Tương đương: Khoảng cách dịch đầu đọc càng nhỏ càng tốt
- Băng thông đĩa là tổng số byte đã được truyền chia cho tổng thời gian giữa yêu cầu đầu tiên và thời gian hoàn thành lần truyền dữ liệu cuối cùng



Lập lịch đĩa (3)

- Có nhiều thuật toán lập lịch đĩa
- Chúng ta minh họa với dãy các yêu cầu (Giả sử đĩa có 200 track từ 0-199):

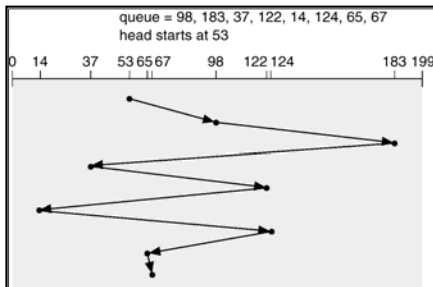
98, 183, 37, 122, 14, 124, 65, 67

Đầu đọc đang nằm ở cylinder 53



FCFS (First come first serve)

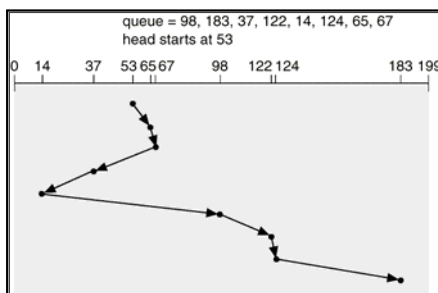
Tổng số bước di chuyển của đầu đọc là: 640 cylinder.



SSTF: Shortest seek time first

- Yêu cầu có seek time nhỏ nhất tính từ vị trí hiện tại của đầu đọc
- Lập lịch SSTF là một dạng của lập lịch SJF có thể gây ra một số yêu cầu không bao giờ được phục vụ (starvation)
- Ví dụ minh họa: Tổng số bước di chuyển của đầu đọc là 236 cylinder.

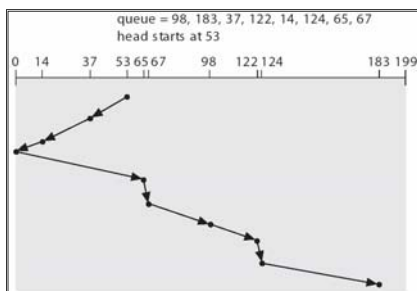
Ví dụ SSTF



SCAN

- Đầu đọc của đĩa di chuyển từ một phía (ví dụ bên ngoài hoặc bên trong đĩa) sang phía kia để phục vụ các yêu cầu đọc, sau đó di chuyển ngược lại... quá trình này lặp đi lặp lại
- Phương thức hoạt động tương tự thang máy nên thuật toán này còn được gọi là thuật toán thang máy (*elevator algorithm*)
- Ví dụ minh họa: Đầu đọc phải dịch chuyển 208 cylinder.

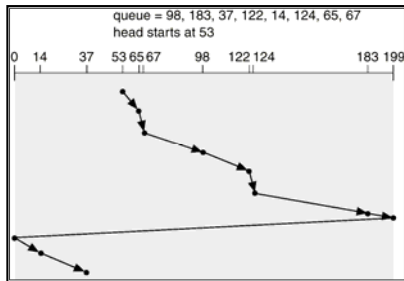
Ví dụ SCAN



C-SCAN

- Đầu đọc chuyển từ một phía (trong/ngoài) sang phía kia và phục vụ các yêu cầu. Khi sang đến phía kia, đầu đọc quay trở lại nhưng trong khi quay trở lại không phục vụ yêu cầu nào.
- C-SCAN xem các cylinders như một danh sách vòng

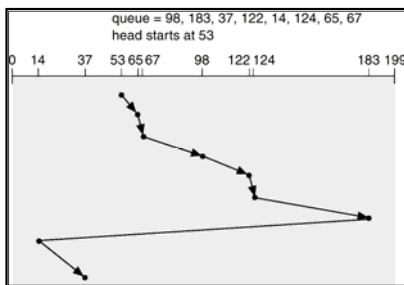
Ví dụ C-SCAN



C-LOOK

- Là một trường hợp của C-SCAN
- C-SCAN: Đầu đọc chuyển giữa cylinder 0 và n (cylinder cuối)
- C-LOOK: Đầu đọc chuyển giữa c_{min} và c_{max} trong đó c_{min} là cylinder có số thứ tự nhỏ nhất trong số các yêu cầu; c_{max} là cylinder có số thứ tự nhỏ nhất trong số các yêu cầu
- C-LOOK giảm quãng đường di chuyển của đầu đọc so với C-SCAN

Ví dụ C-LOOK



Chọn thuật toán lập lịch đĩa

- SSTF là phổ biến và “tự nhiên”
- SCAN và C-SCAN thực hiện tốt với các hệ thống đọc ghi đĩa nhiều
- Hiệu năng nói chung phụ thuộc vào số lượng và tính chất của các yêu cầu truy cập đĩa
- Yêu cầu đọc ghi đĩa có thể bị ảnh hưởng bởi phương pháp cấp phát tệp

Chọn thuật toán lập lịch đĩa

- Các thuật toán lập lịch đĩa nên được cài đặt như một module độc lập của HĐH để dễ thay thế khi cần thiết
- SSTF hoặc LOOK có thể chọn là thuật toán ngầm định

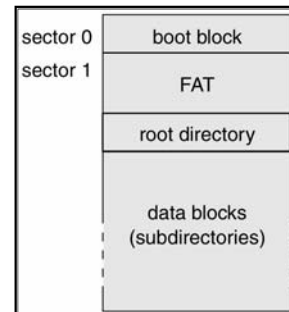
Quản lý đĩa

- *Low-level format* hoặc *physical format* — Chia đĩa thành các sector để bộ điều khiển đĩa (disk controller) có thể đọc/ghi
- Để lưu tệp lên đĩa, HĐH cần ghi cấu trúc dữ liệu lên đĩa:
 - HĐH chia đĩa thành các *partition* (phân vùng) – mỗi partition là một nhóm các cylinder
 - HĐH thực hiện *logical formatting* hay tạo hệ thống tệp.

Tổ chức đĩa

- Đĩa cứng có boot block để khởi tạo hệ thống:
 - bootstrap được lưu trong ROM
 - *Bootstrap loader* là một chương trình nhỏ nằm trên boot block của đĩa cứng.
- Ví dụ bootstrap loader:
 - Linux: GRUB, LILO
 - Windows: NTLDR
- Với các sector hỏng: Phương pháp *sector sparing*

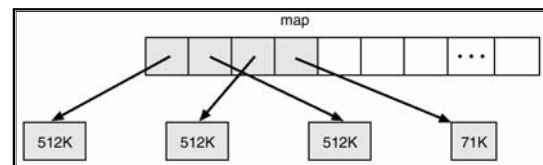
Tổ chức đĩa của MS-DOS



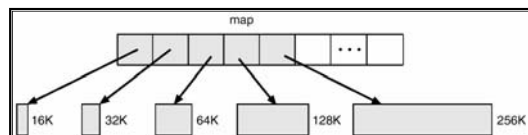
Quản lý không gian swap

- Không gian swap được xem như một phần mở rộng của bộ nhớ trong và nằm trên đĩa
- Không gian swap có thể nằm trên hệ thống tệp hoặc trên một partition riêng
- Quản lý không gian swap
 - UNIX BSD 4.3 cấp phát swap khi tiến trình bắt đầu thực hiện và lưu các segment: text và data
 - Nhân dùng *swap maps* để quản lý việc sử dụng không gian swap.

4.3 BSD Text-Segment Swap Map



4.3 BSD Data-Segment Swap Map



Cấu trúc RAID

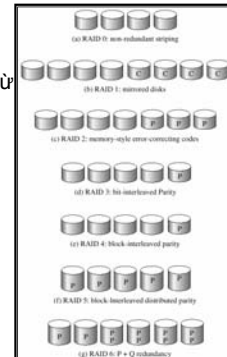
- **RAID** – Hệ thống lưu trữ sử dụng nhiều ổ đĩa để tăng độ tin cậy (**reliability**) thông qua sự dư thừa (**redundancy**).
- Có 6 mức RAID.

RAID (tiếp)

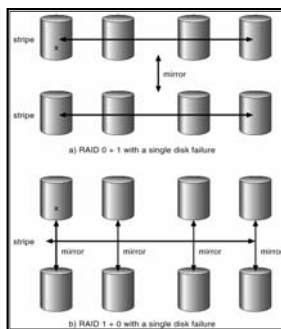
- Disk striping: Sử dụng một tập đĩa như một đĩa
- RAID cải thiện hiệu năng và độ tin cậy của hệ lưu trữ bằng cách lưu trữ có dư thừa
 - *Mirroring/shadowing*: Mỗi đĩa có một bản copy (duplicate).
 - *Block interleaved parity*: Mức độ dư thừa ít hơn mirroring.

Các mức RAID

- Xem chi tiết các mức RAID trong giáo trình từ trang 471 đến 475



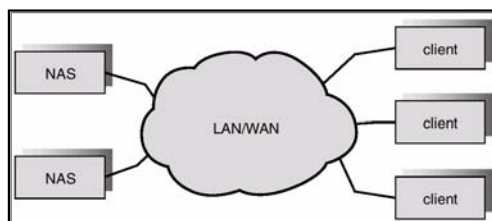
RAID (0 + 1) và (1 + 0)



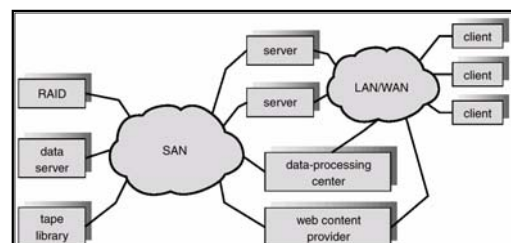
Kết nối đĩa

- Có thể kết nối đĩa theo 2 cách:
 1. **Host attached** thông qua cổng vào/ra
 2. **Network attached** thông qua kết nối mạng

Network-Attached Storage



Storage-Area Network



Cài đặt hệ lưu trữ ổn định

- Write-ahead log scheme requires stable storage.
- Để cài đặt hệ lưu trữ ổn định:
 - Replicate information on more than one nonvolatile storage media with independent failure modes.
 - Update information in a controlled manner to ensure that we can recover the stable data after any failure during data transfer or recovery.

Các thiết bị tertiary

- Thuật ngữ: Tertiary storage
- Đặc tính của tertiary storage là giá rẻ
- Nói chung, tertiary storage được làm để *tháo dỡ*
- Ví dụ: Đĩa mềm, CD-ROM, USB

Đĩa tháo được

- Đĩa mềm: Đĩa phủ từ nằm trong vỏ bảo vệ
 - Hầu hết các đĩa mềm chứa được khoảng 1 MB; một số đĩa sử dụng công nghệ tương tự có thể chứa 1GB.
 - Tốc độ đĩa mềm nhanh nhưng dễ mất dữ liệu do hỏng bề mặt (tiếp xúc)

Đĩa tháo được

- Đĩa quang-từ ghi dữ liệu trên đĩa nhựa cứng bề mặt phủ vật liệu từ.
 - Nhiệt laser được sử dụng để khuếch đại từ trường yếu để ghi 1 bit
 - Ánh sáng laser light is được sử dụng để đọc dữ liệu (hiệu ứng Kerr).
 - Đầu đọc đĩa loại này xa bề mặt đĩa hơn là đầu đọc đĩa từ → giảm hỏng do xước, va chạm.
- Đĩa quang không sử dụng vật liệu từ mà dùng vật liệu đặc biệt có thể bị biến đổi do tia laser

Đĩa WORM

- WORM ("Write Once, Read Many Times") Ghi một lần, đọc nhiều lần
- Đĩa có 3 lớp: Hai lớp nhựa ở 2 mặt, ở giữa là một lớp phim mỏng chế tạo từ nhôm
- Để ghi 1 bit: Ổ đĩa dùng tia laser đốt một lỗ nhỏ trên lớp phim nhôm.
- Rất bền và tin cậy
- Ví dụ: CD-ROM, DVD-ROM

Băng

- Băng rẻ hơn đĩa, nhưng truy cập ngẫu nhiên chậm hơn
- Băng thường được dùng cho các ứng dụng không yêu cầu truy cập nhanh. Ví dụ: Lưu trữ dữ liệu, backup
- Các hệ thống băng từ lớn sử dụng robot để thay băng: Chuyển băng giữa các ổ băng và thư viện băng
 - stacker – Thư viện băng nhỏ (một vài băng)
 - silo – Thư viện băng lớn (vài nghìn băng)

Các vấn đề của HĐH

- Nhiệm vụ chính của HĐH là quản lý các thiết bị vật lý và cung cấp một máy ảo cho ứng dụng (thông qua trừu tượng hóa)
- Với đĩa cứng có hai mức trừu tượng hóa:
 - Thiết bị – Một mảng các khối dữ liệu
 - Hệ thống tệp – HĐH phục vụ các yêu cầu truy cập đĩa (qua cơ chế hàng chờ/lập lịch) từ nhiều ứng dụng

Tốc độ truy cập đĩa

- Hai yếu tố ảnh hưởng đến tốc độ truy cập đĩa: băng thông bandwidth và độ trễ (latency).
- Băng thông được đo bằng byte/second.
 - Sustained bandwidth: Tốc độ trao đổi dữ liệu trung bình trong một lần đọc/ghi lớn (tổng số byte/thời gian)
 - Effective bandwidth – Băng thông trung bình của toàn bộ các lần vào/ra bao gồm **seek / locate**...

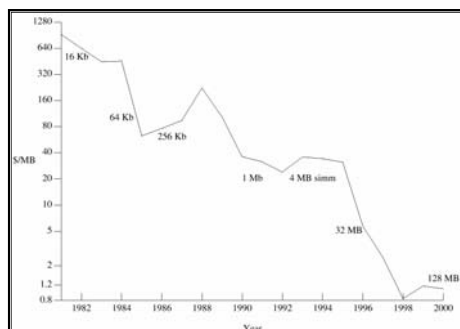
Tốc độ truy cập đĩa

- Độ trễ truy cập: Thời gian cần để định vị dữ liệu trên đĩa.
 - Độ trễ cho đĩa: Chuyển đầu đọc đến cylinder cần thiết + độ trễ quay (thường < 35ms)
 - Độ trễ băng: Cần tua băng → Độ trễ từ vài chục đến vài trăm giây

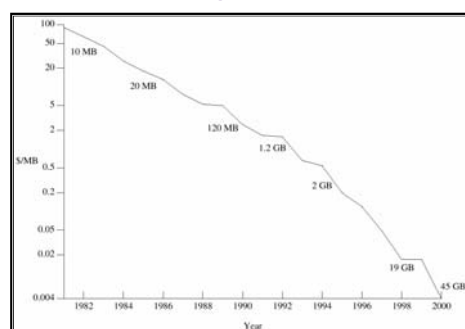
Độ tin cậy

- Đĩa cứng có độ tin cậy cao hơn băng hoặc đĩa tháo được
- Lưu trữ trên đĩa quang tin cậy hơn đĩa từ hoặc băng
- Đầu đọc đĩa cứng hỏng → mất dữ liệu
- Đầu đọc băng, CD... hỏng không gây mất dữ liệu

Giá 1MB DRAM từ 1981 đến 2000



Giá 1MB đĩa cứng từ 1981 đến 2000



Giá 1MB băng, từ 1984 đến 2000

