

Escalonamento de Disco

Marcelo Johann

Introdução

- Disco magnético talvez seja o mais importante dispositivo de E/S
 - Gerência de memória (área de swap)
 - Sistema de arquivos (arquivos + diretórios)
- Prós e contras:
 - Meio barato de armazenamento permanente
 - Se comparado com processador e memória apresenta uma velocidade de acesso muito lenta
- Necessidade:
 - Otimizar o desempenho do disco visando aumentar a sua largura de banda, tempo de resposta e variância
 - ex.: Bufferização, caches, configurações RAID

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 2

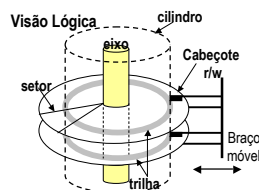
Disco rígido



Fisicamente:
2 pratos, 4



Velocidade linear constante



IBM Model DKLA-24320
8944 cilindros
15 cabeçotes
63 setores por trilha

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 3

Acesso a dados

- Menor unidade de transferência é um bloco lógico (n setores)
- Acessar dado implica em localizar trilha, superfície e setor
- Dois métodos:
 - Método CHS (*Cylinder, Head, Sector*)
 - Método LBA (*Linear Block Addressing*)
 - Tradução do L-CHS (*Logical*) para P-CHS (*Physical*)
- Discos modernos endereçam blocos lógicos sequencialmente
 - Conversão de um bloco lógico para sua localização física
 - Não é um mapeamento direto por haver setores fisicamente defeituosos e pelo número de setores por trilha não ser constante
 - Cilindros que possuem mesmo número de setores (zonas)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 4

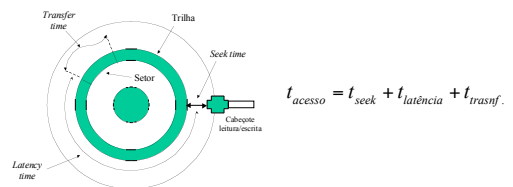
Fatores de desempenho do disco

- Largura de banda (MB/sec)
- Para ler/escrever dados é necessário que o cabeçote de leitura/escrita esteja **posicionado** na trilha e no início do setor desejados
- Três tempos envolvidos:
 - Tempo de **busca** (*seek time*)
 - Tempo necessário para posicionar o cabeçote na trilha desejada
 - Tempo de **latência** rotacional
 - Tempo necessário para atingir o início do setor a ser lido/escrito
 - Tempo de **transferência**
 - Tempo para escrita/leitura efetiva dos dados

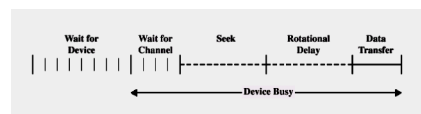
INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 5

Temporização de acesso ao disco



$$t_{\text{acesso}} = t_{\text{seek}} + t_{\text{latência}} + t_{\text{transf.}}$$



INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 6

Tempo de busca (*seek*)

- Influenciado pelo tempo de acionamento, aceleração e deslocamento do cabeçote até a trilha desejada
- Não é linear em função do número do trilhas
 - Tempo de identificação da trilha (confirmação)
- Na realidade existem três tempos de *seek*
 - *seek time*: tempo de deslocamento até uma determinada trilha
 - *head switch time*: tempo para acionar o cabeçote de leitura/escrita
 - *cylinder time*: tempo de deslocamento da trilha *i* para a trilha *i+1*
- Tempo médio de *seek*
 - Dado fornecido pelo fabricante
 - e.g.; 5 a 10 ms (tecnologia anos 2000)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 7

Tempo de latência rotacional

- Definido pela velocidade de rotação do motor
 - e.g. (anos 2000):
 - discos rígidos (5400 rpm a 10000 rpm);
 - unidades de *floppy* (300 rpm a 600 rpm)
- Considera-se o tempo médio
 - Não se sabe a posição relativa do cabeçote com a do setor a ser lido
 - Metade do tempo de uma rotação
 - e.g.; 3 ms para um disco de 10000 rpm (6 ms uma rotação)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 8

Tempo de transferência

- Tempo de transferência de dados de/para disco depende da velocidade de rotação

$$T = \frac{b}{rN} \quad \text{OU} \quad T = \frac{60b}{RN}$$

T = tempo de transferência
b = número de bytes a serem transferidos
N = número de bytes em uma trilha
r = velocidade de rotação, nro de rotações por segundo (RPS)
R = velocidade de rotação, nro de rotações por minuto (RPM)

- Tempo médio de acesso é dado por:

$$T_{\text{acesso}} = t_{\text{seek_médio}} + \frac{1}{2r} + \frac{b}{rN} \quad \text{OU} \quad T_{\text{acesso}} = t_{\text{seek_médio}} + \frac{30}{R} + \frac{60b}{RN}$$

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 9

Exemplo: disco IBM deskstart 180 GXP

- **Bytes per sector:** 512 bytes
- **Data zones:** 27
- **Sector per track:** 536 to 1092
- **Platters:** 3
- **Tracks per cylinder:** 6
- **Cylinders:** 70.553
- **Sectors (total):** 361.882.080
- **Total size:** 185.283.264.960 B
- **Seek time**
 - Track to track: 1.1 ms
 - Average: 8.5 ms
 - Full stroke: 15 ms
- **Latency time:** 4.7 ms
- **Buffer size:** 8 MB
- **Rotation speed:** 7200 RPM

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 10

Desempenho de acesso a disco

- Depende de três fatores:
 - Tempo de *seek*
 - Latência rotacional
 - Tempo de transferência
- Blocos a serem acessados sequencialmente devem estar próximos
 - Problema da alocação de dados no disco (estudaremos mais adiante)
 - Objetivo é minimizar a distância a ser percorrida pelo cabeçote de leitura e escrita do disco

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 11

Exemplo

- Acessar um arquivo de dados de 1.3 Mbytes armazenado em disco com as seguintes características:
 - $T_{\text{seek_médio}} = 10 \text{ ms}$, 10000 rpms, 512 bytes por setor, 320 setores por trilha
- **Caso I:** Acesso seqüencial (2560 setores = 8 trilhas x 320 setores)
 - $T_{\text{acesso}} = 10 \text{ ms} + 8 \times (3 + 6) \text{ ms} = 0.082 \text{ s}$
 - (Obs: supondo trilhas vizinhas, despreza-se o tempo de *seek*)
- **Caso II:** Acesso randômico (leitura na base um setor por vez)
 - $T_{\text{acesso}} = 2560 \times (10 \text{ ms} + 3 \text{ ms} + 0.01875 \text{ ms}) = 33.328 \text{ s}$

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 12

Escalonamento do disco

- Problema:
 - Requisições provêm de diferentes processos e existem rajadas de requisições (teoria das filas):
 - fila de requisições
 - Atendimento pode “quebrar” a ordem de acesso a um arquivo mesmo com um bom mapeamento físico
- Solução: reordenar as requisições para otimizar o tempo de busca (*seek*) e latência rotacional
- Objetivos:
 - Aumentar a taxa de transferência (rendimento)
 - Reduzir o tempo médio de resposta
 - Ser justo no atender a requisições dos processos (variância)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 13

Políticas de escalonamento do disco

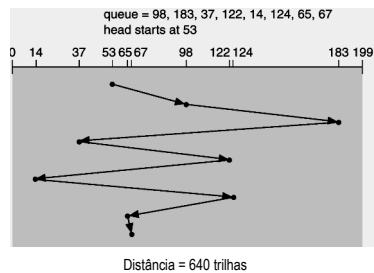
- Otimização de busca
 - FIFO ou FCFS
 - SSTF (*Shortest Seek Time First*)
 - Scan (elevador) e suas variações
- Otimização rotacional
 - SLTF (*Shortest Latency Time First*)
 - SPTF (*Shortest Positioning Time First*)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 14

FCFS - *First Come First Served*

- Acessa na ordem em que as requisições são solicitadas



INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 15

Prós e contras

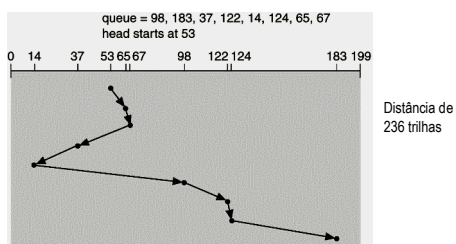
- Prós:
 - Simples de executar
 - Justo: requisições são atendidas na ordem
- Contras:
 - Padrão de busca aleatório (movimentação mecânica)
 - Comportamento ruim sob carga altas (fila de requisição grande)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 16

SSTF - *Shortest Seek Time First*

- Seleciona a requisição com o menor tempo de *seek* em relação a posição atual do cabeçote de leitura/escrita



INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 17

Prós e contras

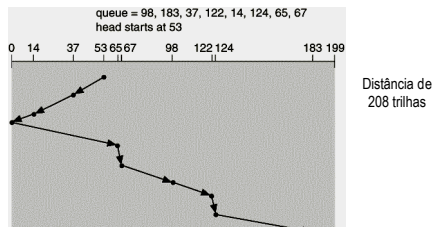
- Prós:
 - Redução do tempo de busca (maior rendimento)
 - Tempo médio tende a ser mais baixo
- Contra:
 - Não garante justiça (“fura” a fila)
 - Postergação indefinida
 - Variância alta (ruim para sistema interativos, aceitável para sistemas batch)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 18

SCAN (algoritmo do elevador)

- Atende requisições em uma direção preferencial
 - Muda de direção ao atingir os cilindros mais interno ou mais externo
 - Definição da direção preferencial (fixa ou em base na fila → critério SSTF)



INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 19

Prós e contras

- Prós
 - Oferece bons tempos médios de resposta
 - Bom rendimento
 - Variância menor que o SSTF
- Contra
 - Não justo
 - Trilhas das extremidades são "visitadas" menos frequentemente que as trilhas internas

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 20

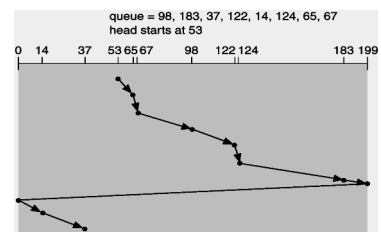
C-SCAN (1)

- Variação do algoritmo de SCAN
- Procedimento é idêntico ao do algoritmo SCAN porém as requisições são atendidas apenas em um sentido da varredura
 - Compensar o fato que, se a leitura ocorresse nos dois sentidos da varredura, os setores próximo ao centro seriam acessados em um tempo médio menor
- Ao final da varredura o cabeçote é reposicionado no início do disco
- Fornece uma visão lógica onde o disco é tratado como uma fila circular
 - Oferece um tempo médio de acesso mais uniforme que o algoritmo SCAN

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 21

C-SCAN (2)



Varição C-LOOK: não necessita esperar até chegar ao extremo do disco

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 22

Otimização rotacional: estratégias combinadas

- Discos mais novos o tempo de seek está na ordem de grandeza da latência rotacional
- SLTF – *Shortest Latency Time First*
 - Reordena o atendimento de requisições de um mesmo cilindro em função do atraso rotacional mais curto.
- SPTF – *Shortest Positioning Time First*
 - Reordena requisições considerando a soma do tempo de seek com a latência
 - Variação: *Shortest Access Time First (SATF)*
 - Reordena as requisições considerando a soma do tempo de seek e de transferência
- Ambas são implementadas organizando a fila de requisições por blocos (n setores) de acordo com a política usada (SLTF ou SPTF)

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 23

Próxima aula...

Sistemas de Arquivos

INF01142 - Sistemas Operacionais I N - Marcelo Johann - 2010/2

Aula 21 : Slide 24