

Estrutura de Armazenamento em Massa

- Estrutura do disco
- Escalonamento de Disco
- Gerenciamento de Disco
- Gerenciamento do Swap-Space
- Estrutura RAID
- Implementação do Armazenamento Estável
- Estrutura do Armazenamento Terciário

Estrutura do Disco

- Discos são endereçados com um grande array unidimensional de *blocos* lógicos, onde o bloco lógico é a menor unidade de transferência.
- Esse array unidimensional é mapeado sequencialmente em setores do disco.
 - ◆ Setor 0 é o primeiro setor da primeira trilha do cilindro mais externo.
 - ◆ Mapeamento procede em ordem através da trilha, depois o resto nas trilhas daquele cilindro e depois o resto em outros cilindros, dos mais externos para os mais internos.

Escalonamento do Disco

- O Sistema Operacional é responsável pelo uso eficiente do hardware — para os discos isto significa ter um rápido tempo de acesso e largura de banda de disco.
- Tempo de acesso tem dois componentes principais:
 - ◆ *Tempo de Busca (Seek time)* é o tempo para que o disco mova a cabeça para o cilindro contendo o setor desejado.
 - ◆ *Latência Rotacional (Rotational latency)* é o tempo adicional de espera para que o disco rotacione até o setor desejado esteja sob a cabeça do disco.
- Minimizar tempo de busca
- Tempo de busca \approx distancia da busca
- Largura de banda é o número total de bytes transferidos dividido pelo tempo total time entre a primeira requisição e o término da última transferência.

Escalonamento do Disco

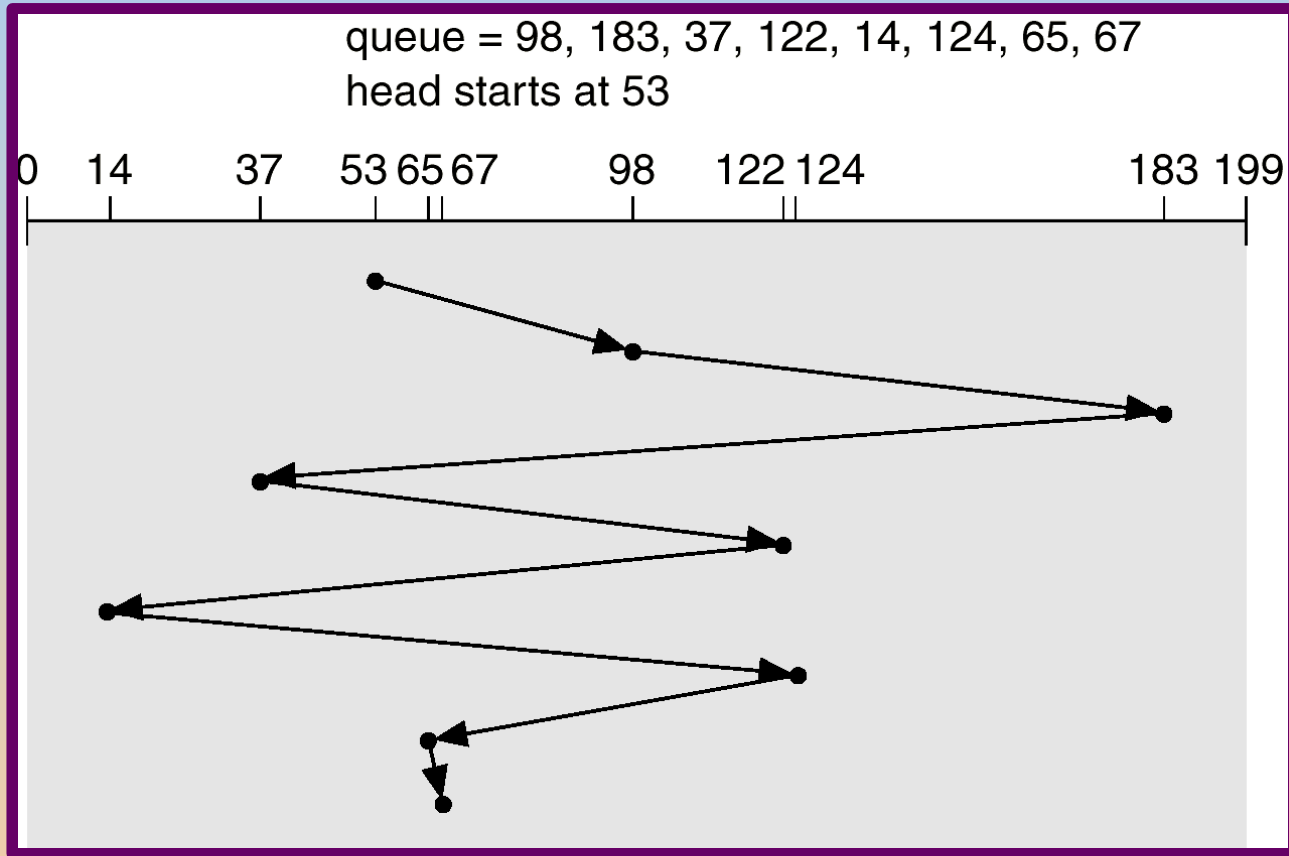
- Vários algoritmos existem para escalonar o atendimento das requisições de E/S.
- Ilustramos com a seguinte fila de requisições (0-199)

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

Inicialment a cabeça aponta para o cilindro 53

FCFS

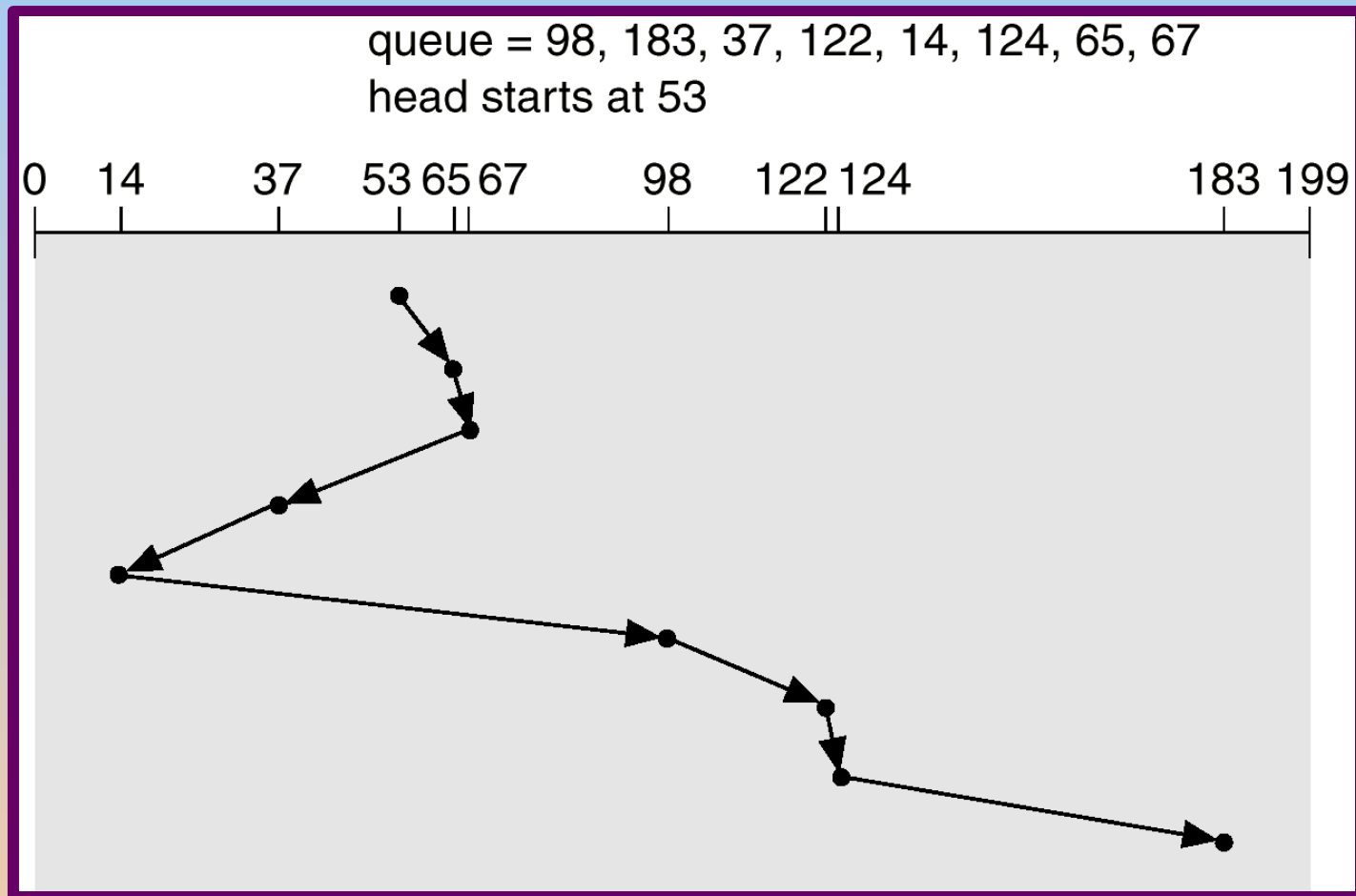
A ilustração mostra um total de movimentações da cabeça de 640 cilindros.



SSTF

- Seleciona a requisição com o menor tempo de busca a partir da posição atual da cabeça.
- O escalonamento SSTF é uma forma de escalonamento SJF, por isso pode levar a inanição (starvation) de algumas requisições.
- A seguinte ilustração mostra um total de movimentação da cabeça de 236 cilindros.

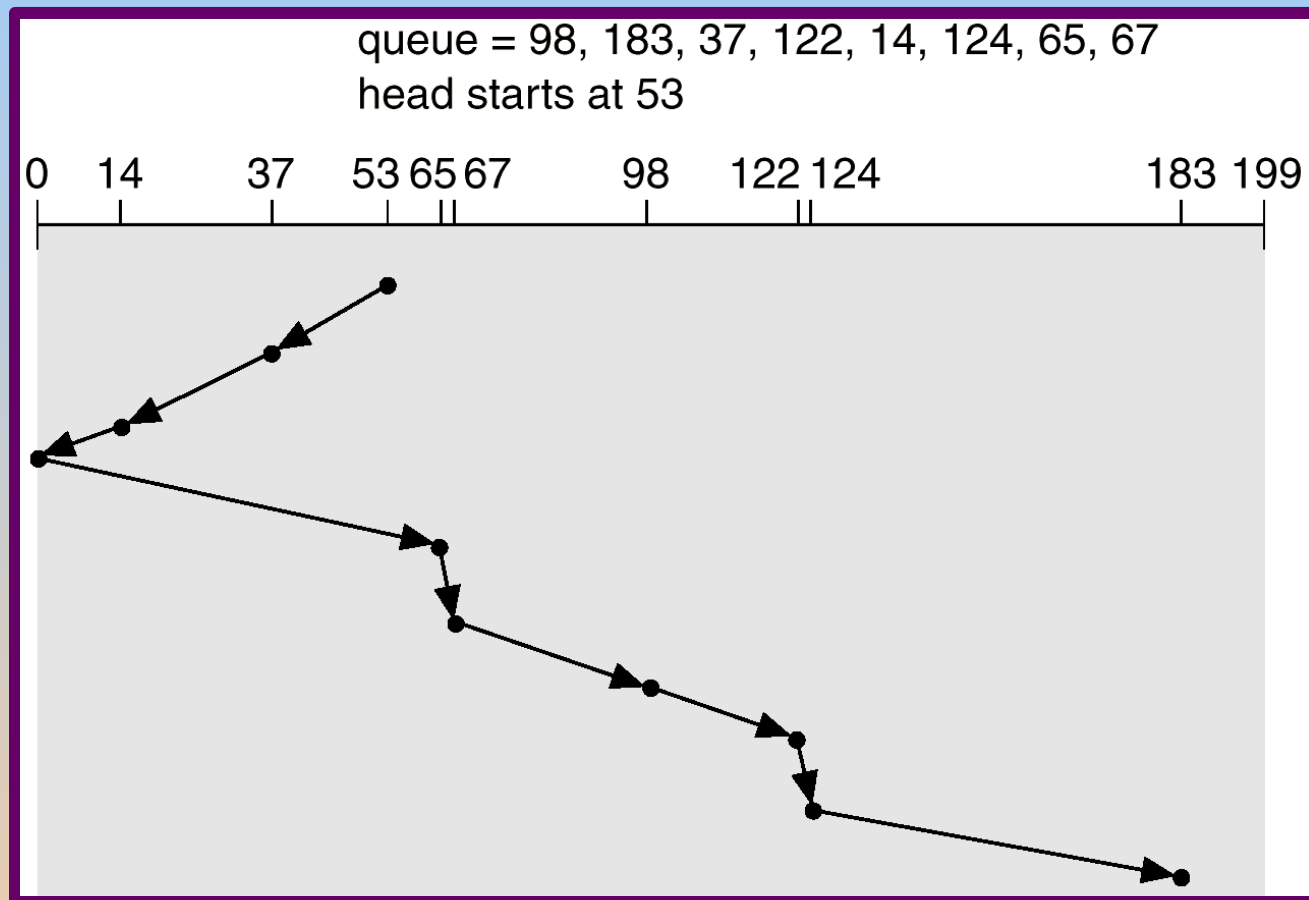
SSTF (Cont.)



SCAN

- O braço do disco inicia em um dos extremos do disco e se move em direção ao outro extremo atendendo as requisições até que ele chegue ao ponto final. Desse ponto ele muda direção e continua o atendimento das requisições até alcançar o outro extremo.
- Chamado de algoritmo do elevador.
- A ilustração mostra um total de movimentações da cabeça de 208 cilindros.

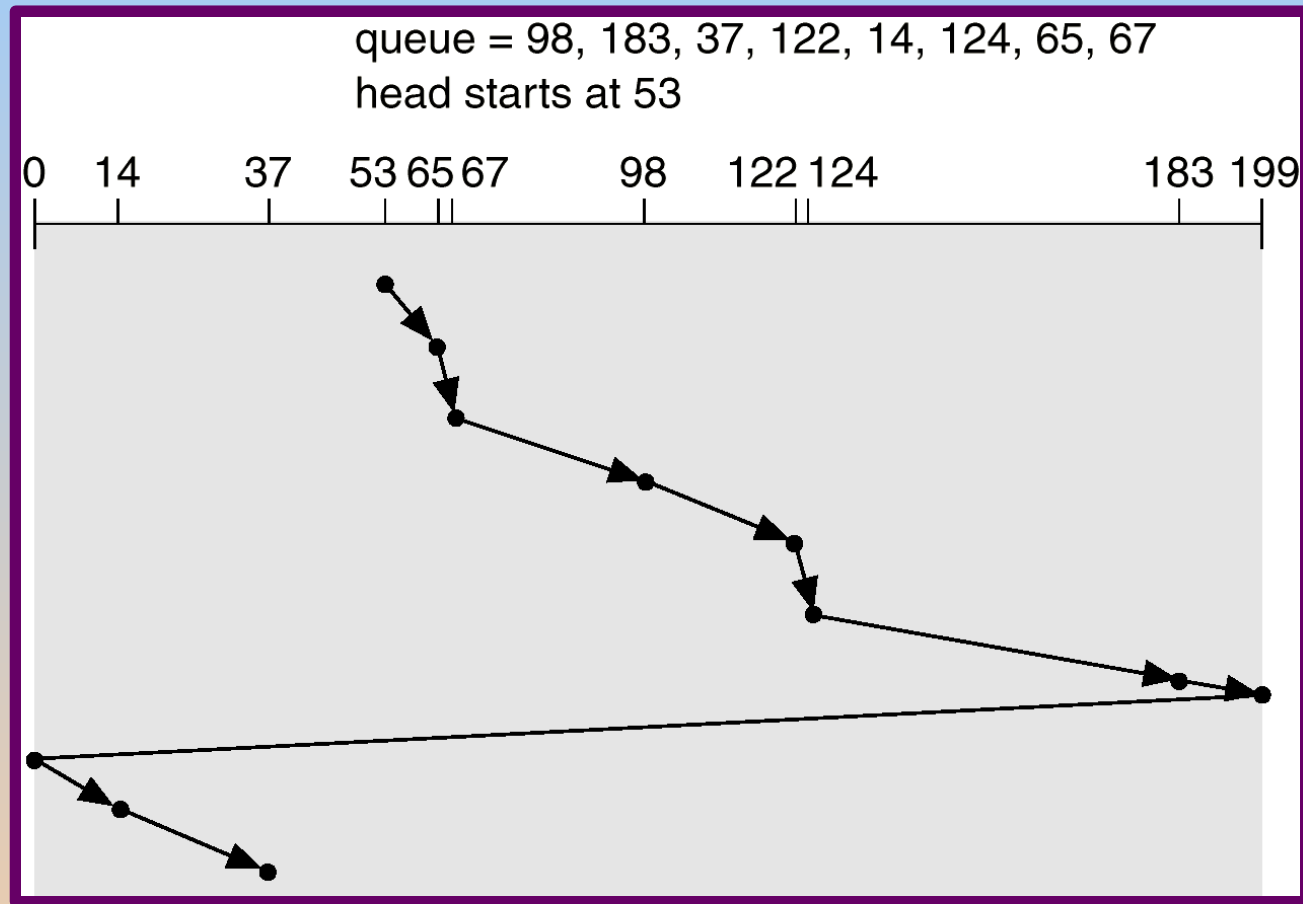
SCAN (Cont.)



C-SCAN

- Fornece um tempo de espera mais uniforme que o SCAN.
- A cabeça move de um extremo ao outro do disco atendendo as requisições no caminho. Quando chega, volta imediatamente ao início do disco sem atender nenhuma requisição no trajeto.
- Trata os cilindros como uma lista circular e sai do último para o primeiro cilindro.

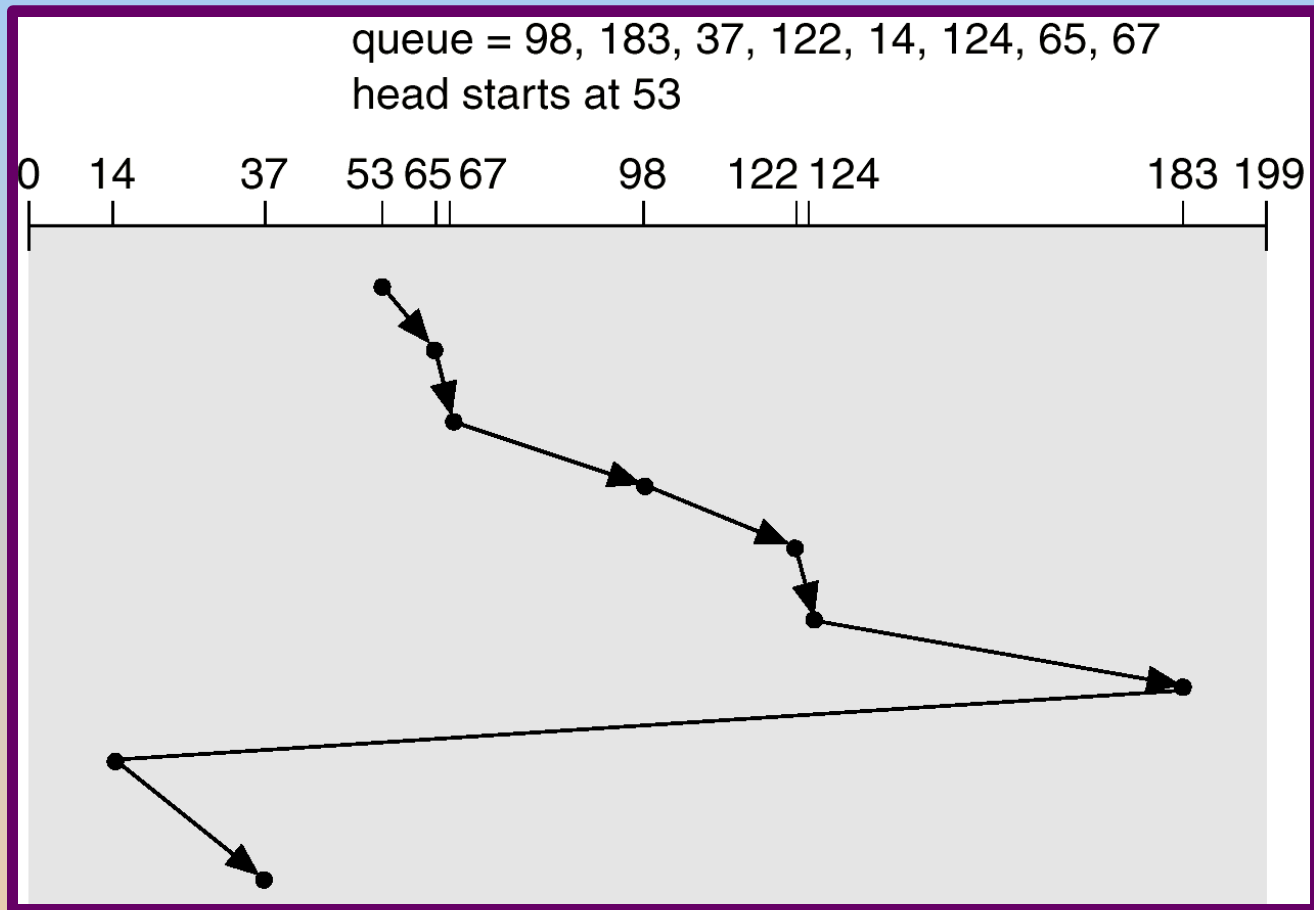
C-SCAN (Cont.)



C-LOOK

- Versão do C-SCAN
- O braço se move apenas o suficiente até a última requisição em cada direção, depois troca a direção imediatamente sem ir até o final do disco.

C-LOOK (Cont.)



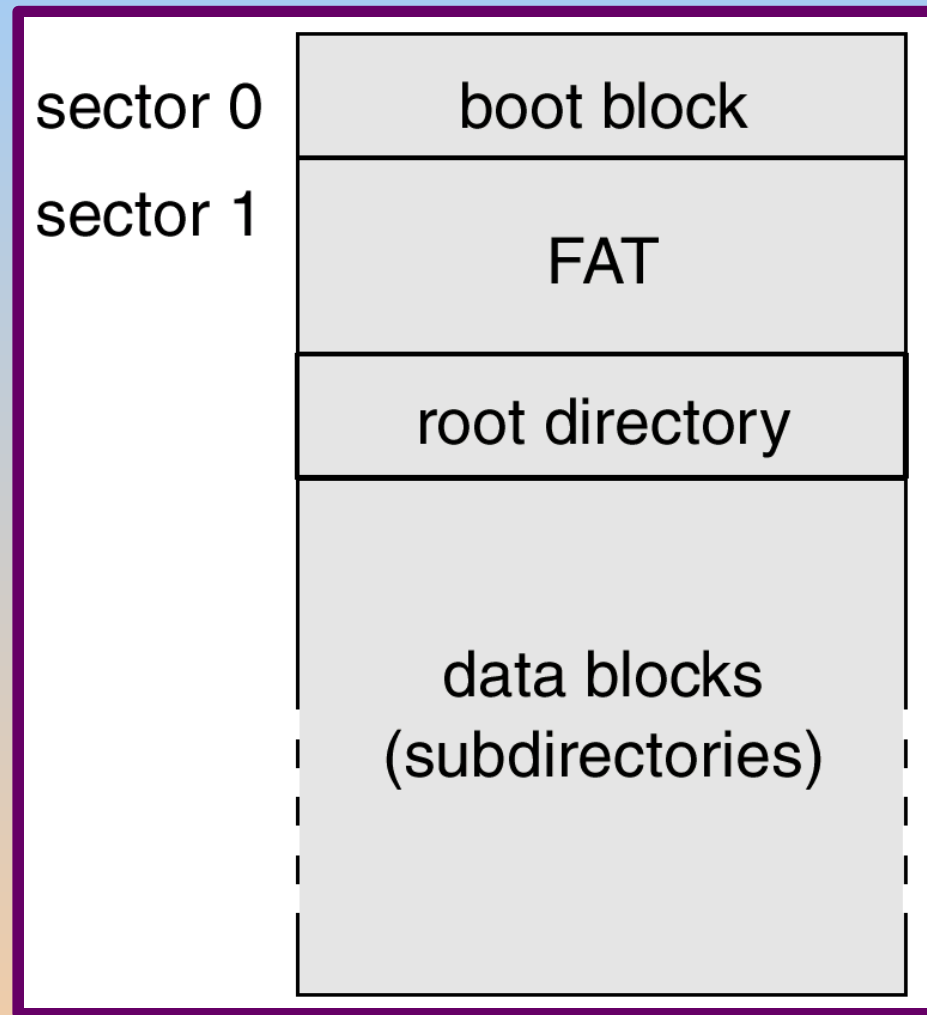
Selecionando um Algoritmo de Escalonamento de Disco

- SSTF é comum e tem um apelo natural
- SCAN e C-SCAN tem melhor desempenho em sistemas que possuem uma carga pesada sobre o disco.
- O desempenho depende do número e do tipo de requisições.
- Requisições para o serviço de disco podem ser influenciadas pelo método de alocação de arquivo.
- O algoritmo de escalonamento de disco deve ser escrito como um módulo separado do sistema operacional, permitindo que possa ser trocado por outro algoritmo se necessário.
- Tanto SSTF ou LOOK são escolhas razoáveis para algoritmo default.

Gerenciamento do Disco

- *Formatação de baixo nível ou formused to handle bad blocks* — Divisão de um disco em setores que a controladora do disco possa ler e escrever.
- Para usar um disco para gravar arquivos o sistema operacional ainda necessita registrar as estruturas de diretório no disco.
 - ◆ *Particionar* o disco em um ou mais grupos de cilindros.
 - ◆ *Formatação lógica* ou “fazer um sistema de arquivos”.
- Bloco de Boot inicia o sistema
 - ◆ O *bootstrap* é armazenado na ROM.
 - ◆ *Programa Bootstrap loader*
- Métodos tais como *sector sparing* são usados para manipular blocos ruins.

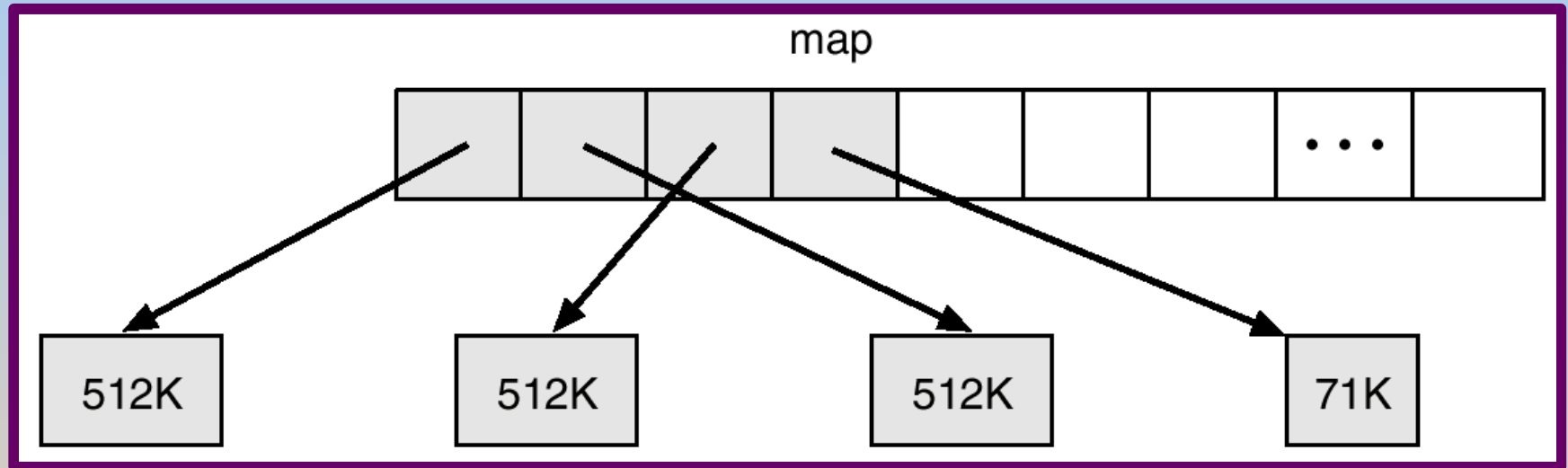
MS-DOS Disk Layout



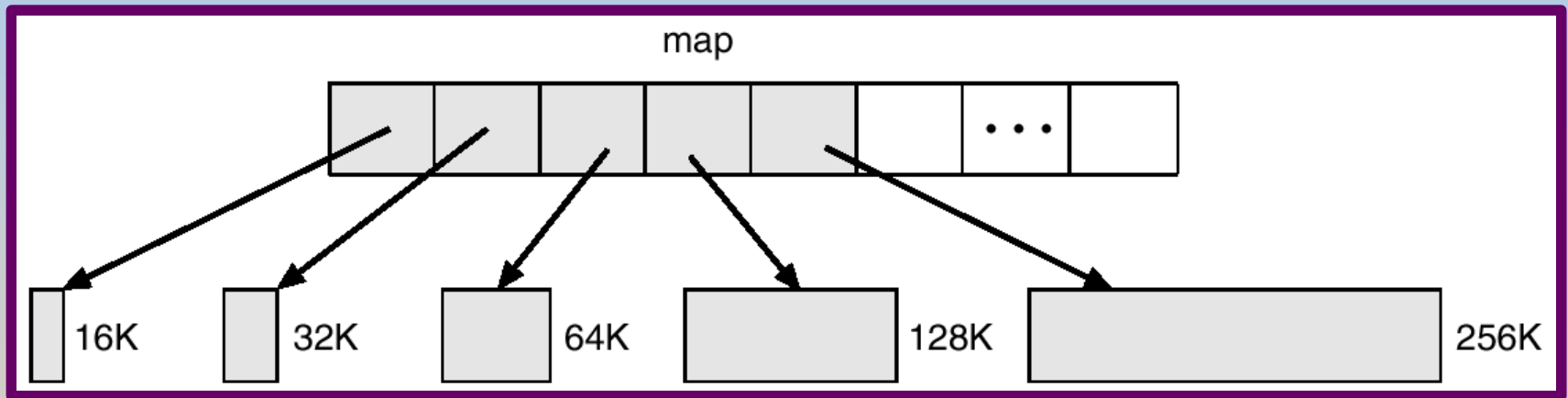
Gerenciamento do Swap-Space

- Swap-space — Memória Virtual usa o espaço em disco como uma extensão da memória principal.
- Swap-space pode estar inserido em um sistema de arquivos normal ou, mais comunmente, pode estar em uma partição separada do disco.
- Gerencia do Swap-space
 - ◆ 4.3BSD aloca swap space quando o process inicia; mantém o *segmento de texto* (o programa) e o *segmento de dados*.
 - ◆ Kernel usa *mapas de swap para monitorar o uso do espaço*.
 - ◆ Solaris 2 aloca espaço de swap somente quando uma página é retirada da memória física, não quando a página é criada.s

4.3 BSD Text-Segment Swap Map



4.3 BSD Data-Segment Swap Map



RAID Structure

- **RAID** – múltiplos discos oferecem confiabilidade (**reliability**) via redundância (**redundancy**).
- RAID pode ser organizado em seis níveis diferentes.

RAID (cont)

- Várias melhorias nas técnicas de utilização dos discos envolvem o uso de múltiplos discos trabalhando cooperativamente.
- Uma faixa de disco requer um grupo de discos como uma única unidade de armazenamento.
- Esquemas RAID melhoram a confiabilidade e o desempenho do armazenamento ao gravar dados redundantes.
 - ◆ *Espelhamento* mantém uma duplicata de cada disco.
 - ◆ *Paridade de blocos intercalados* usa muito menos redundância.

Dispositivos de Armazenamento Terciário

- Custo mais baixo é a característica que define o armazenamento terciário.
- Geralmente, o armazenamento terciário é construído com mídias removíveis
- Exemplos comuns são disquetes, CDs e DVDs, além de outros tipos.

Tarefas do Sistema Operacional

- As principais tarefas do SO são gerenciar dispositivos físicos e apresentar uma abstração de máquina virtual às aplicações
- Para os discos, o SO oferece duas abstrações:
 - ◆ Raw device – um array de blocos de dados.
 - ◆ Sistema de Arquivo – o SO enfileira e escalona requisições intercaladas de várias aplicações.

Interface da Aplicação

- A maioria dos SOs manipula os discos removíveis quase como discos fixo
- As fitas são apresentadas como mídia de armazenamento bruto, isto é, uma aplicação não abre um arquivo na fita, ela abre todo o drive de fita com um dispositivo bruto.
- Usualmente o drive de fita é reservado para uso exclusivo da aplicação.
- Já que o SO não fornece serviços de sistema de arquivo, a aplicação deve decidir como usar o array de blocos.
- Já que toda aplicação faz suas próprias regras em como organizar uma fita, uma fita cheia de dados geralmente só pode ser usada pelo programa que a criou.

Nomes de Arquivos

- A tarefa de nomeação de arquivos em mídias removíveis é especialmente difícil quando queremos escrever dados em um cartucho removível conectado em um computador e depois usá-la em outro computador.
- Os S.O.s atuais geralmente deixam o problema do espaço de nomes não resolvido para mídia removível e depende das aplicações e usuários resolverem como acessar e interpretar os dados.
- Alguns tipos de mídia removível (ex., CDs) estão padronizados que os computadores os usam da mesma forma.

Hierarchical Storage Management (HSM)

Gerenciamento Hierárquico

- Um sistema hierárquico de armazenamento estende além da memória primária e secundária para incorporar o armazenamento terciário — geralmente implementado como um jukebox robótico de fitas ou discos removíveis
- Extensão do sistema de arquivos
 - ◆ Arquivos pequenos e frequentemente usados ficam no disco
 - ◆ Arquivos grandes, antigos e inativos vão para jukebox.
- HSM geralmente encontrada em centros de supercomputação e outras instalações com grande volume de dados.