



Universidade Tecnológica Federal do Paraná  
Departamento Acadêmico de Computação  
Bacharelado em Ciência da Computação

# Sistemas Operacionais

## **Discos Rígidos**

**Prof. Rodrigo Campiolo**  
**Prof. Rogério A. Gonçalves**

29/10/19

# Introdução

- Discos Rígidos (Hard Disks – HD) ou Discos Magnéticos são dispositivos de E/S para armazenamento persistente de dados.



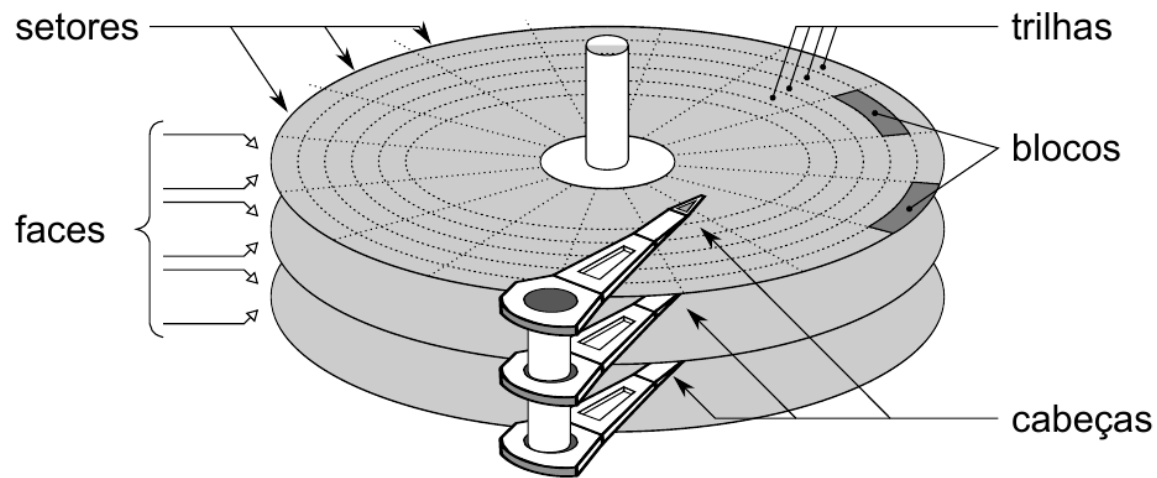
Fonte: <https://spectrum.ieee.org>

# Estrutura do Disco

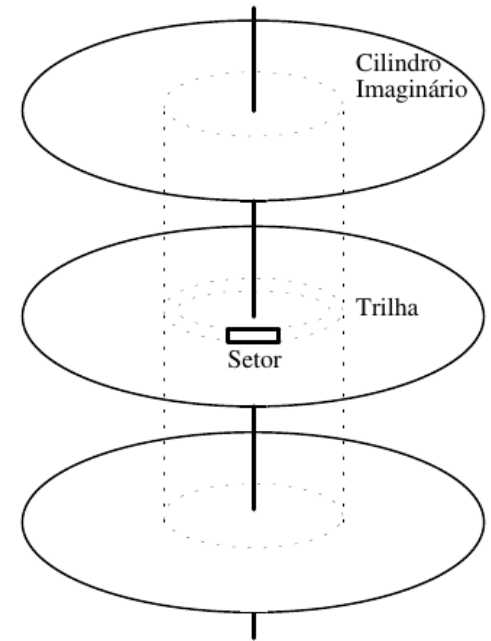
---

- Composto por discos metálicos sobrepostos que giram em rotações aproximadas entre 5000 a 15000 rpms.
- Possuem uma **cabeça de leitura** para cada face e as faces são divididas em **trilhas** e **setores**. O conjunto de trilhas sobreposto em diferentes discos formam os **cilindros**.
- Setor(es) em uma trilha formam um **bloco lógico** (unidade básica de leitura e escrita).

# Estrutura do Disco



Fonte: Maziero



Fonte: Oliveira

# Estrutura do Disco

---

- Unidades de disco são endereçadas como agregados unidimensionais de *blocos lógicos*.
- Cada bloco lógico é a menor unidade de transferência.
- O agregado unidimensional de blocos lógicos é mapeado nos setores do disco sequencialmente.
  - Setor 0 é o primeiro setor da primeira trilha no cilindro mais externo.
  - Mapeamento prossegue na ordem por essa trilha, depois o restante das trilhas nesse cilindro, e depois pelo restante dos cilindros de fora para dentro.

# Tempo de Acesso

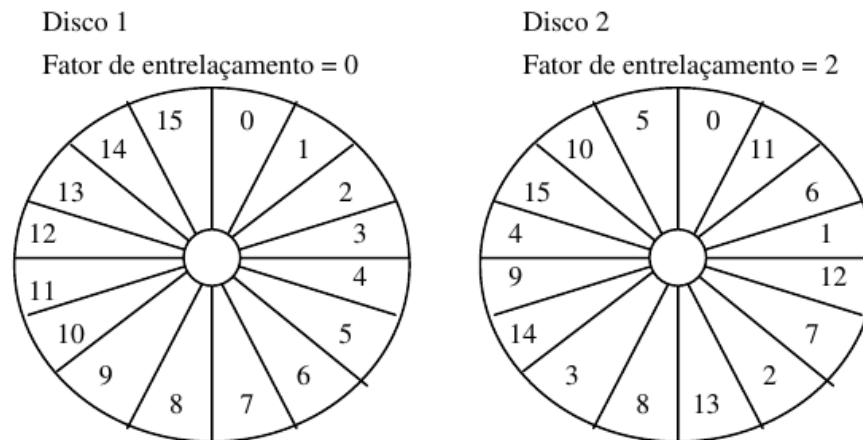
- Tempo de acesso é dado por:

$$t_{\text{acesso}} = t_{\text{busca}} + t_{\text{latência}} + t_{\text{transferência}}$$

- $t_{\text{busca}}$**  (*seek time*): posicionar o cabeçote no cilindro para o acesso.
- $t_{\text{latência}}$**  (*latency time*): posicionar o cabeçote no setor para o acesso.
- $t_{\text{transferência}}$**  (*transfer time*): tempo de transferência dos dados.

# Entrelaçamento

- Entrelaçamento (*interleaving*): técnica para numerar os setores de forma não contígua em uma trilha. Usa-se um espaçamento fixo entre eles.



**Figura:** Exemplo de entrelaçamento para trilha com 16 setores.

Fonte: Oliveira

# Escalonamento de Disco

---

- Para otimizar o tempo de acesso:
  - Minimizar os movimentos do cabeçote de leitura e escrita.
  - Maximizar o número de bytes transferidos (*throughput*)
- O **escalonamento de disco** consiste em algoritmos para selecionar e organizar um conjunto de requisições de acesso ao para otimizar o tempo de acesso.



# Escalonamento de Disco

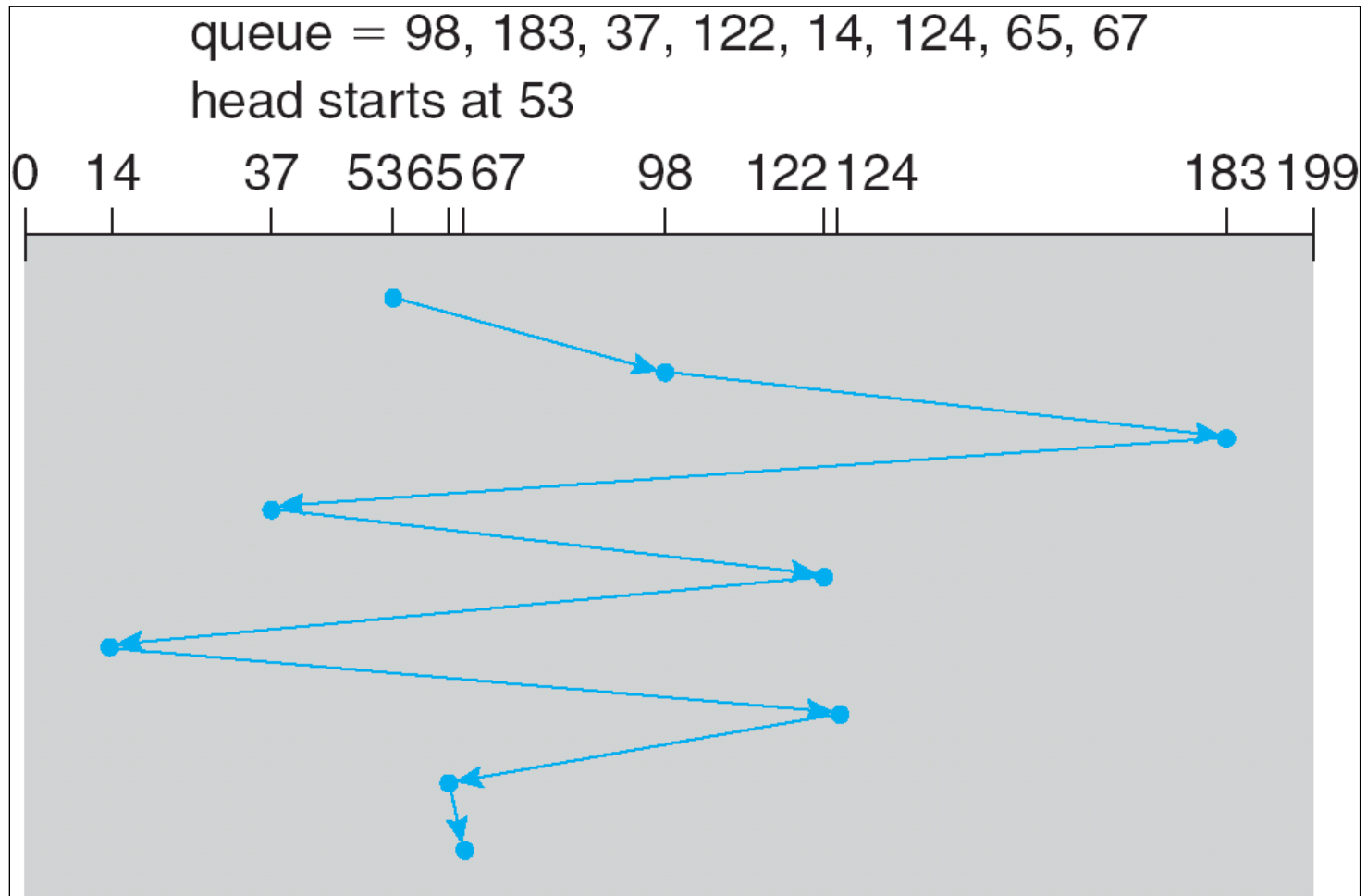
---

- Existem vários algoritmos para escalonar o atendimento das solicitações de E/S de disco.
- Ilustramos com uma fila de solicitação (0-199).

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

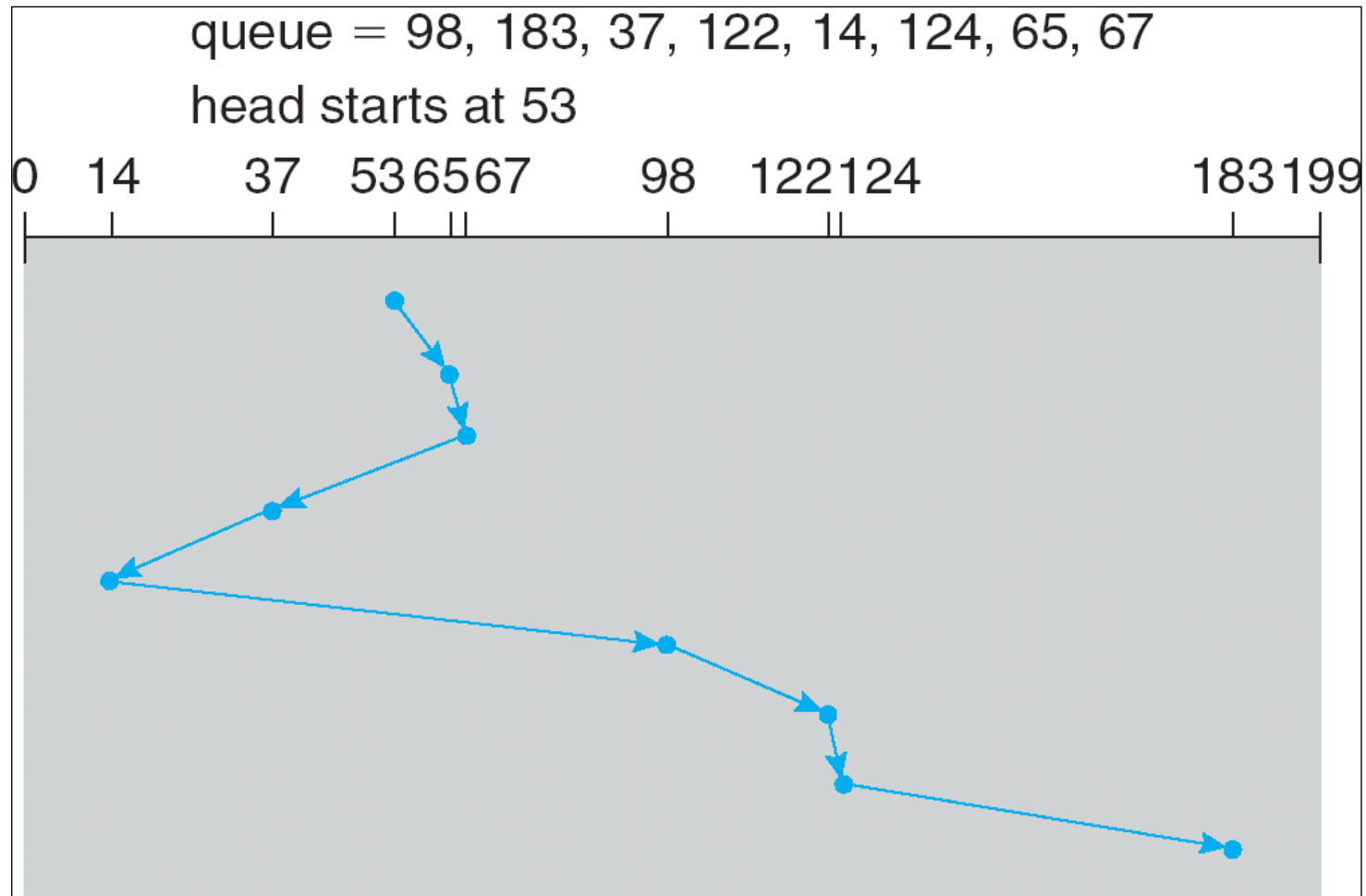
Posição inicial do ponteiro na cabeça: 53

# FCFS (First-Come First-Served)



Fonte: Silberschatz

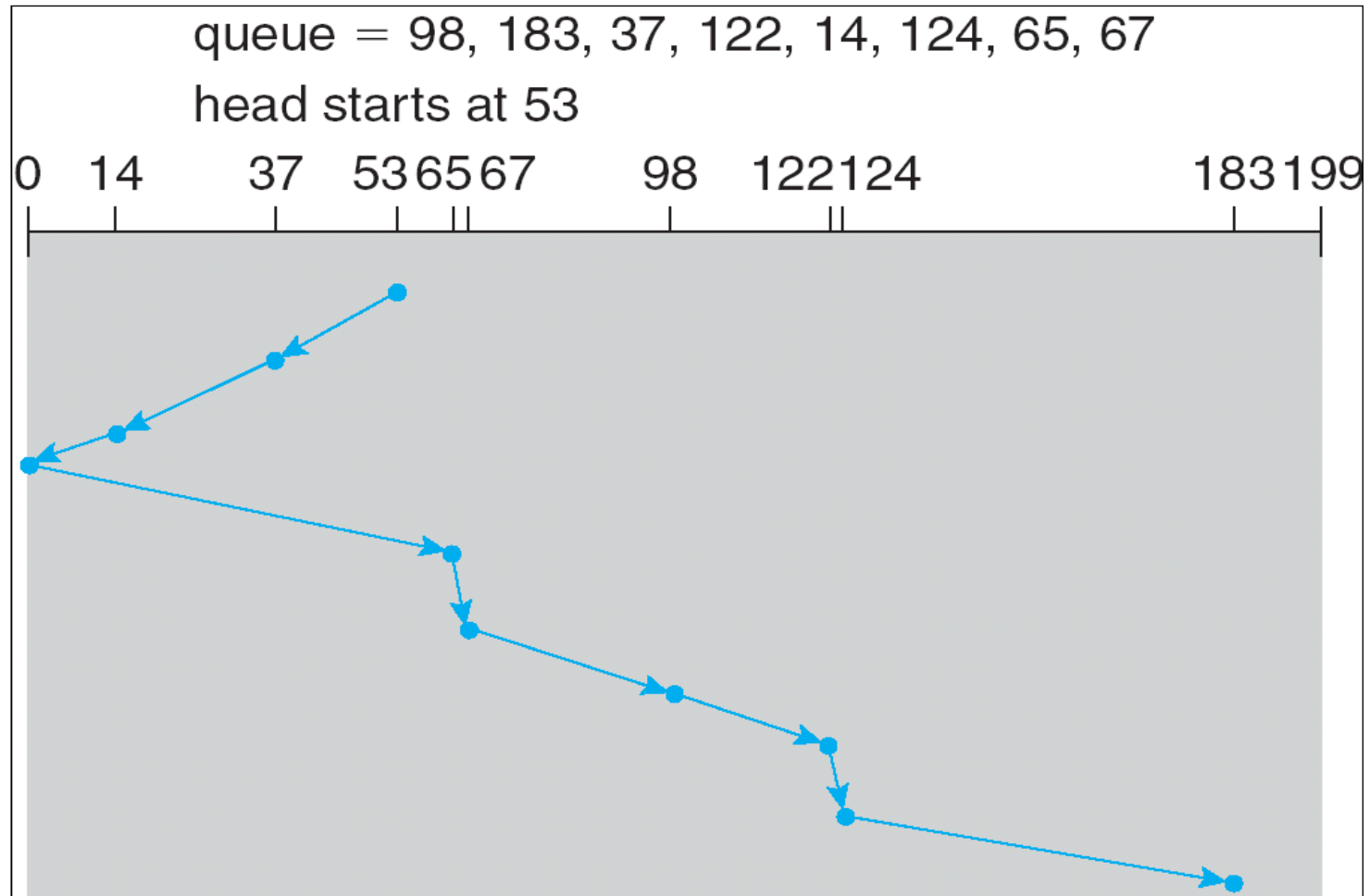
# SSTF (Shortest Seek Time First)



Movimento total da cabeça: 236

Fonte: Silberschatz

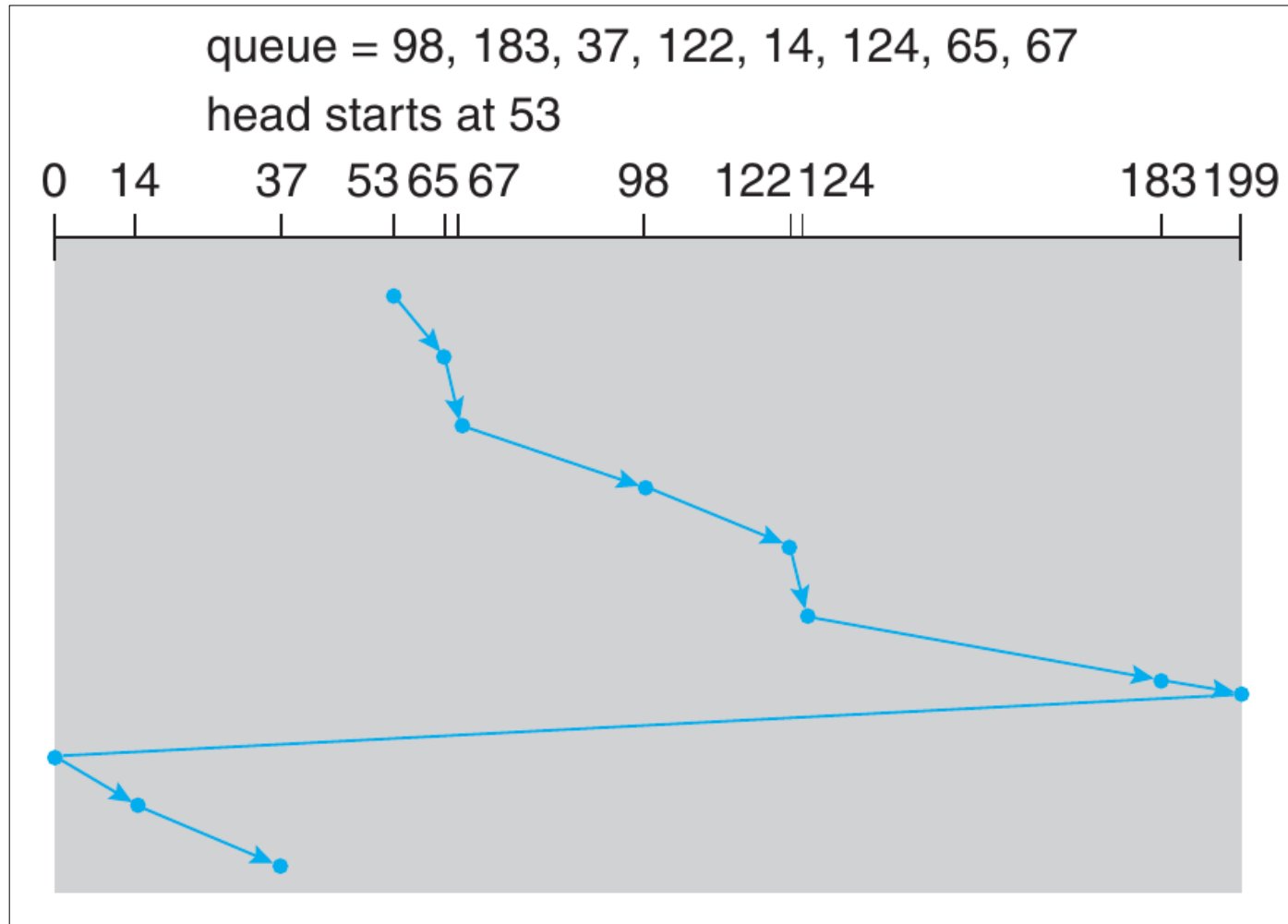
# SCAN (Alg. do Elevador)



Movimento total da cabeça: 236

Fonte: Silberschatz

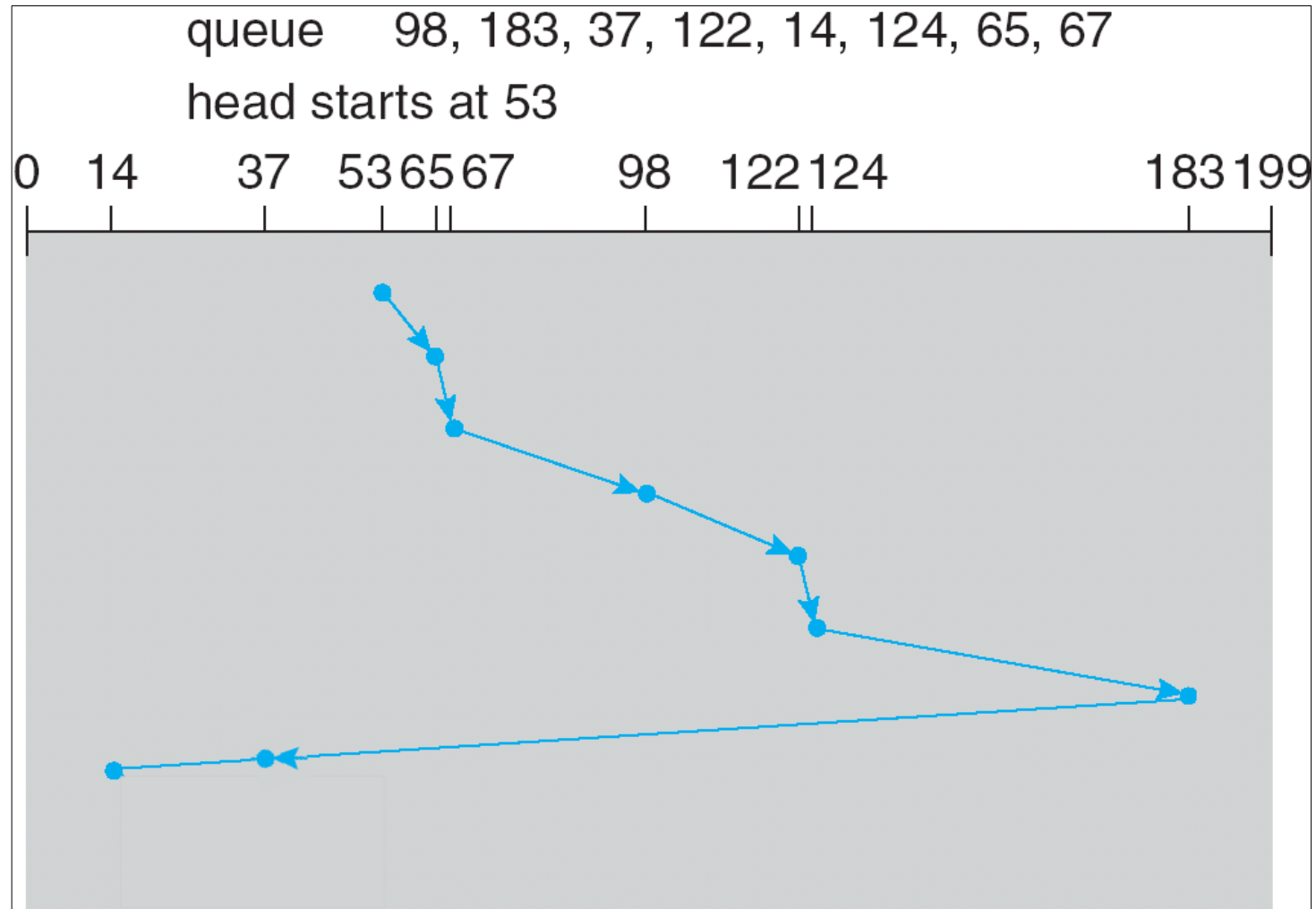
# C-SCAN (Alg. do Elevador)



Movimento total da cabeça: 382

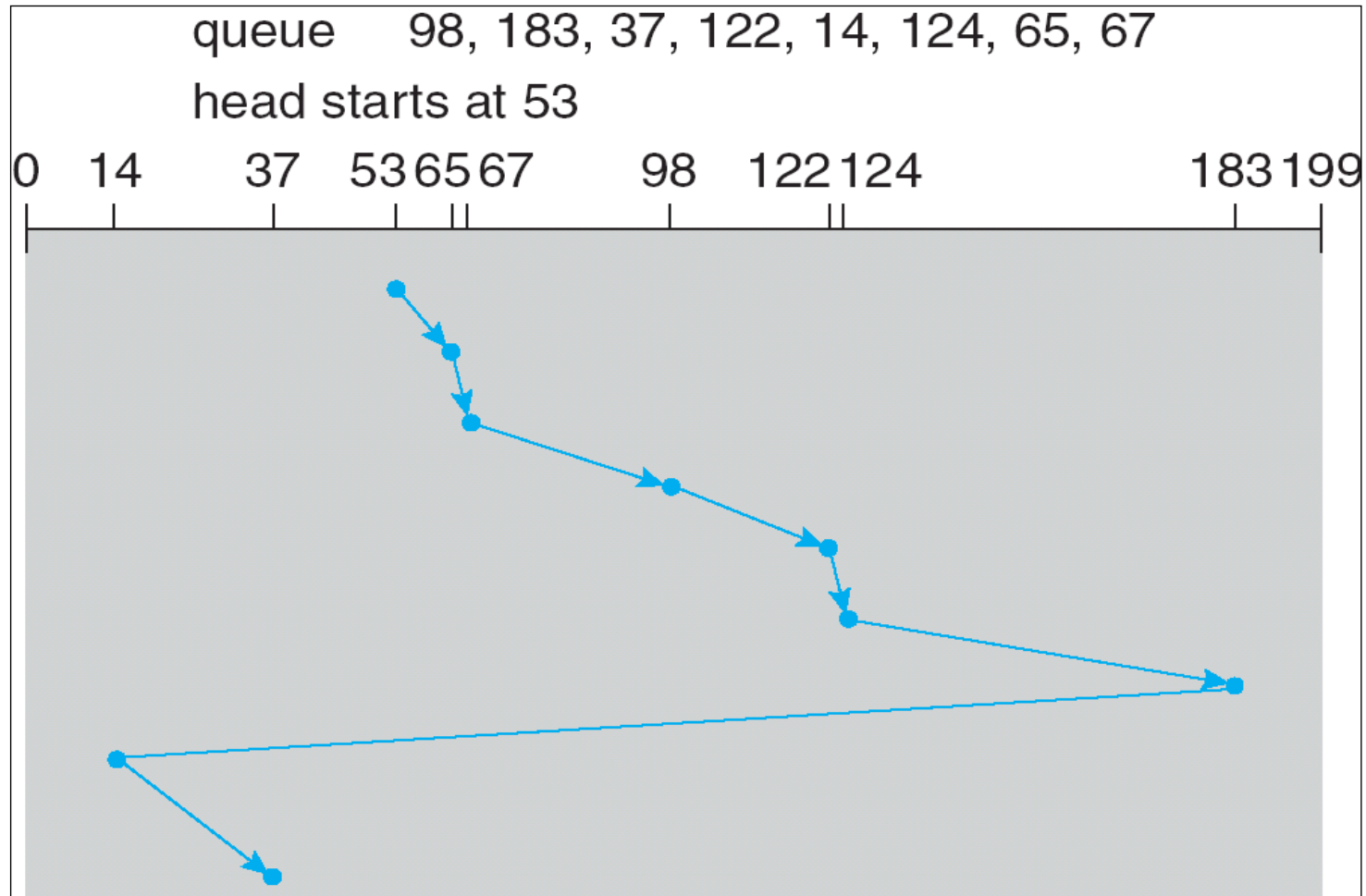
Fonte: Silberschatz

# LOOK (Alg. do Elevador)



Movimento total da cabeça: 299

# C-LOOK (Alg. do Elevador)



Movimento total da cabeça: 322

Fonte: Silberschatz

# Escalonamento de Disco - Síntese

- SSTF é comum e tem um apelo natural
- SCAN e C-SCAN funcionam melhor para sistemas que têm cargas pesadas sobre o disco
- O desempenho depende do número e tipo de solicitações.
- Requisições para 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 seja substituído por um algoritmo diferente, se necessário.
- SSTF ou C-LOOK é uma escolha razoável para o algoritmo padrão.



# Atividades

---

- Analise o desempenho de cada algoritmo de escalonamento de disco para a seguinte situação:

**Posição Inicial:** 70

**Ordem Req.:** 198, 1, 6, 183, 38, 71, 72, 3, 140, 50

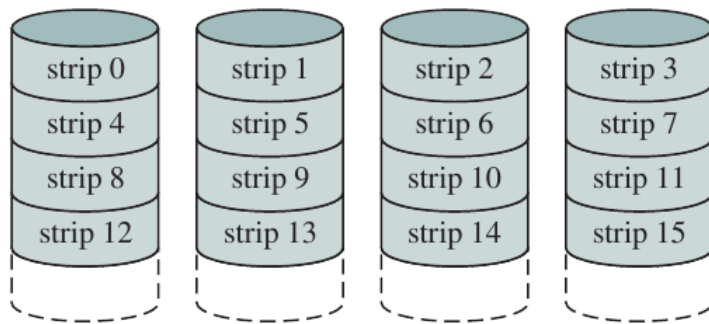
# RAID

---

- RAID – *Redundant Array of Inexpensive Disks*
  - Conjunto de discos redundantes vistos pelo SO como um único disco lógico.
  - Desempenho, disponibilidade e tolerância a falhas.
  - Podem ser implementados em hardware ou software.
  - Diferentes níveis RAID.

# RAID 0

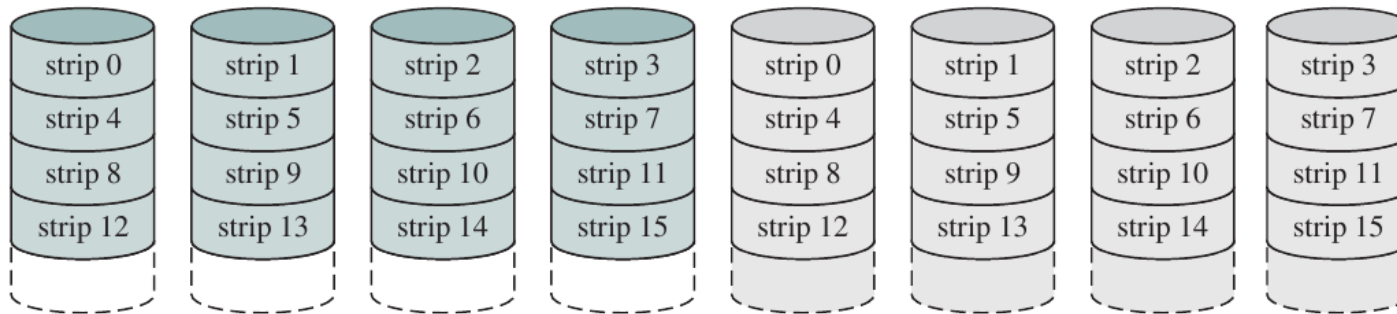
- Dados divididos em diferentes discos.
- Cada parte é denominada de *strip*.
- Técnica denominada de *stripping*.
- Vantagem: paralelismo, desempenho.



**Figura:** RAID 0 (não redundante) Fonte: Stallings

# RAID 1

- Dados espelhados em diferentes discos.
- Técnica denominada de *mirroring*.
- Disponibilidade e tolerância a falhas.
- Desvantagem: custo.

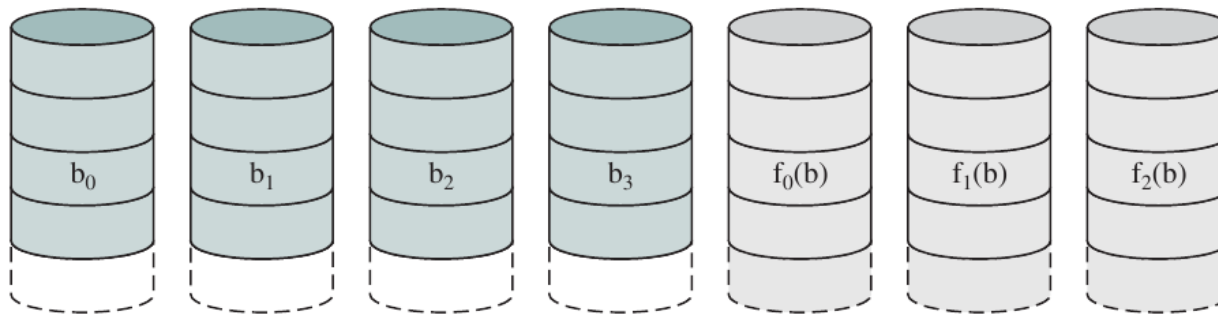


**Figura:** RAID 1 (espelhamento)

Fonte: Stallings

# RAID 2

- Strips são pequenos (byte ou palavra).
- Tipicamente é usado *Hamming Code* para calcular a paridade.
- Discos sincronizados para E/S.

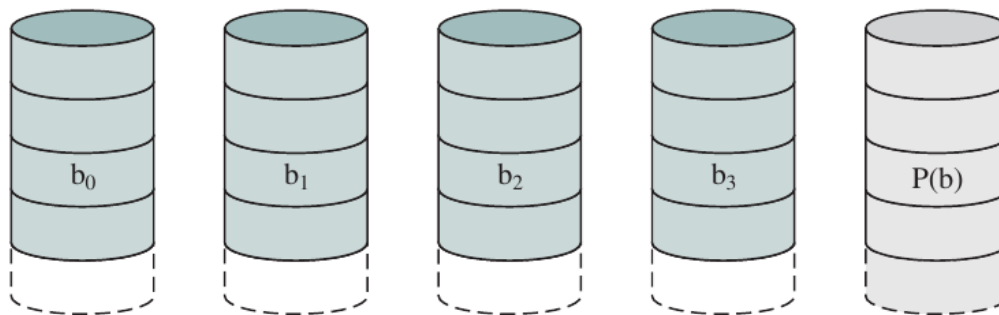


**Figura:** RAID 2 (paridade – código de correção)

Fonte: Stallings

# RAID 3

- Similar ao RAID 2.
- Calcula paridade de um conjunto de bits de mesma posição em diferentes discos.
- Apenas um disco de redundância.
- Possível reconstruir dados de um disco.

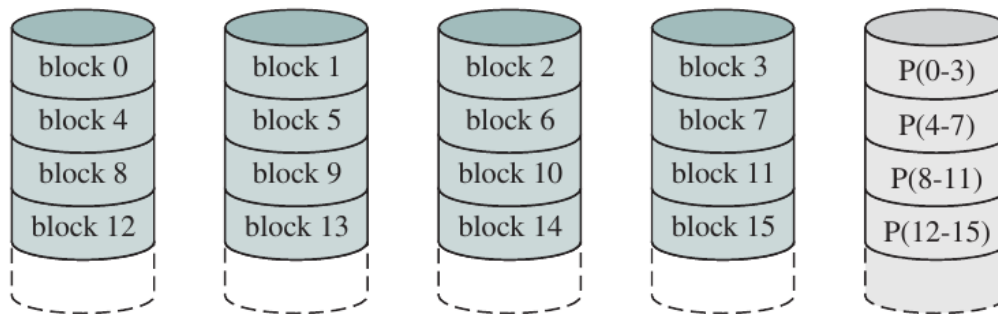


**Figura:** RAID 3 (paridade – código de detecção)

Fonte: Stallings

# RAID 4

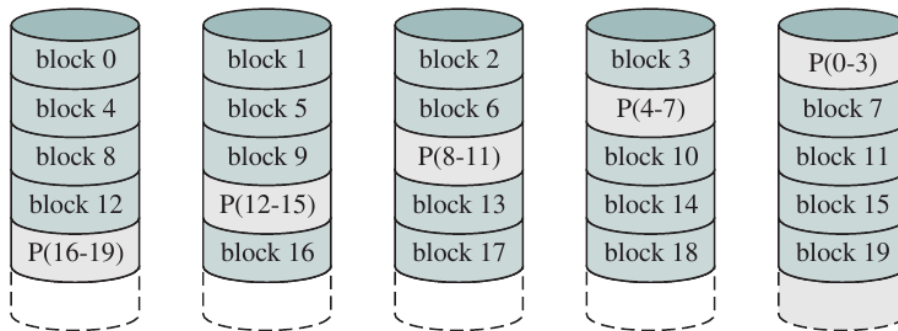
- Acesso independente por discos.
- Strips são maiores.
- Paridade dos blocos.



**Figura:** RAID 4 (paridade de blocos) Fonte: Stallings

# RAID 5

- Similar ao RAID 4.
- Paridade distribuída entre diferentes discos.



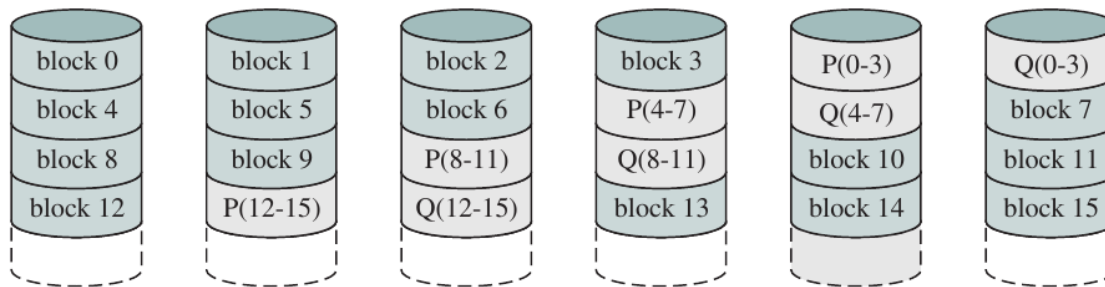
**Figura:** RAID 5 (paridade de blocos distribuída)

Fonte: Stallings



# RAID 6

- Faz uso de dois esquemas de cálculo de paridade.
- Desvantagem é o custo para atualizar as paridades na escrita.



**Figura:** RAID 6 (redundância dual)

Fonte: Stallings

# RAID - Síntese

Category	Level	Description	Disks Required	Data Availability	Large I/O Data Transfer Capacity	Small I/O Request Rate
Striping	0	Nonredundant	$N$	Lower than single disk	Very high	Very high for both read and write
Mirroring	1	Mirrored	$2N$	Higher than RAID 2, 3, 4, or 5; lower than RAID 6	Higher than single disk for read; similar to single disk for write	Up to twice that of a single disk for read; similar to single disk for write
Parallel access	2	Redundant via Hamming code	$N+m$	Much higher than single disk; comparable to RAID 3, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
	3	Bit-interleaved parity	$N+1$	Much higher than single disk; comparable to RAID 2, 4, or 5	Highest of all listed alternatives	Approximately twice that of a single disk
Independent access	4	Block-interleaved parity	$N+1$	Much higher than single disk; comparable to RAID 2, 3, or 5	Similar to RAID 0 for read; significantly lower than single disk for write	Similar to RAID 0 for read; significantly lower than single disk for write
	5	Block-interleaved distributed parity	$N+1$	Much higher than single disk; comparable to RAID 2, 3, or 4	Similar to RAID 0 for read; lower than single disk for write	Similar to RAID 0 for read; generally lower than single disk for write
	6	Block-interleaved dual distributed parity	$N+2$	Highest of all listed alternatives	Similar to RAID 0 for read; lower than RAID 5 for write	Similar to RAID 0 for read; significantly lower than RAID 5 for write

**Figura:** Comparação entre os diferentes níveis de RAID. Fonte: Stallings

# Referências

---

- OLIVEIRA, R. S. et al. **Sistemas Operacionais**. 4ª Edição e Slides online. Bookman. 2010.
- SILBERSCHATZ, A.; GALVIN, P. B.; GAGNE, G. **Fundamentos de Sistemas Operacionais**. 9. ed. LTC, 2015.
- MAZIERO, C. **Sistemas Operacionais: Conceitos e Mecanismos**. Online. 2019.
- Stallings, William. **Operating systems. Internals and design principles**, 7th edition, 2011.