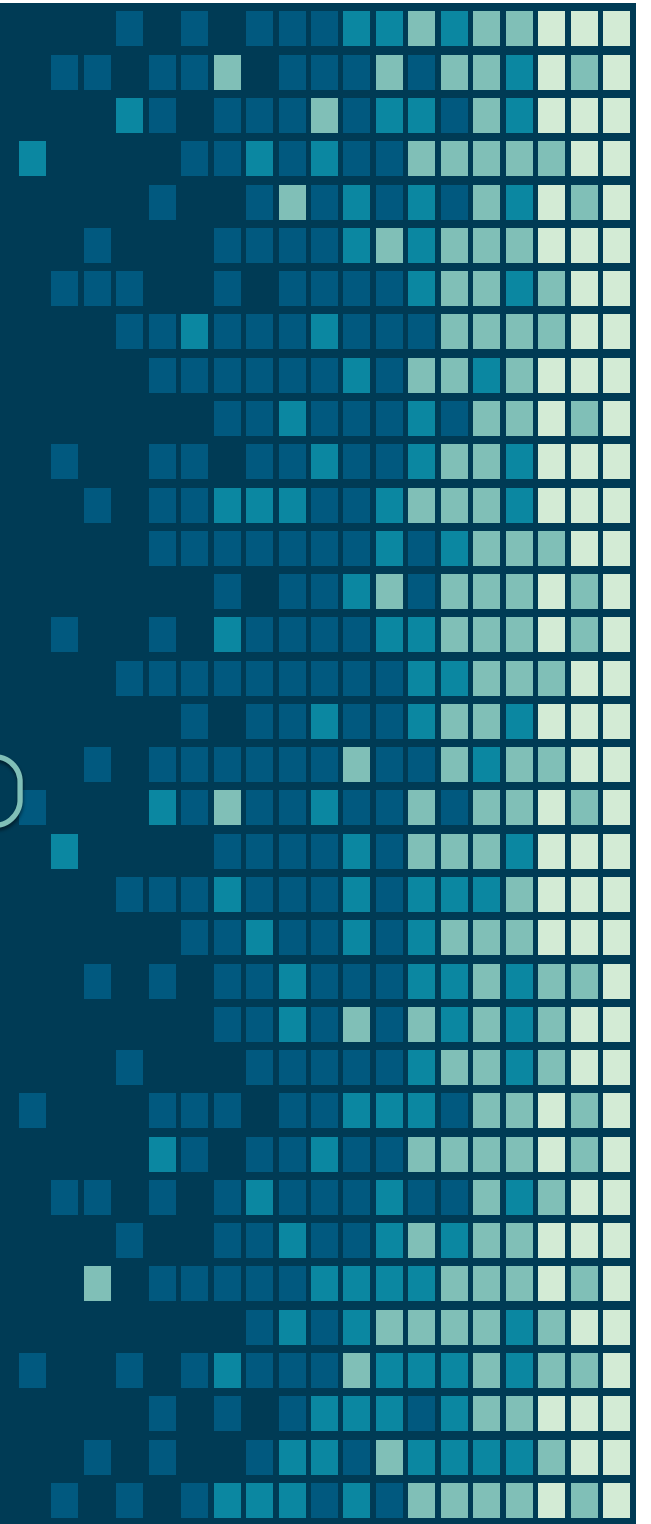


Sistemas de Armazenamento Físico de dados

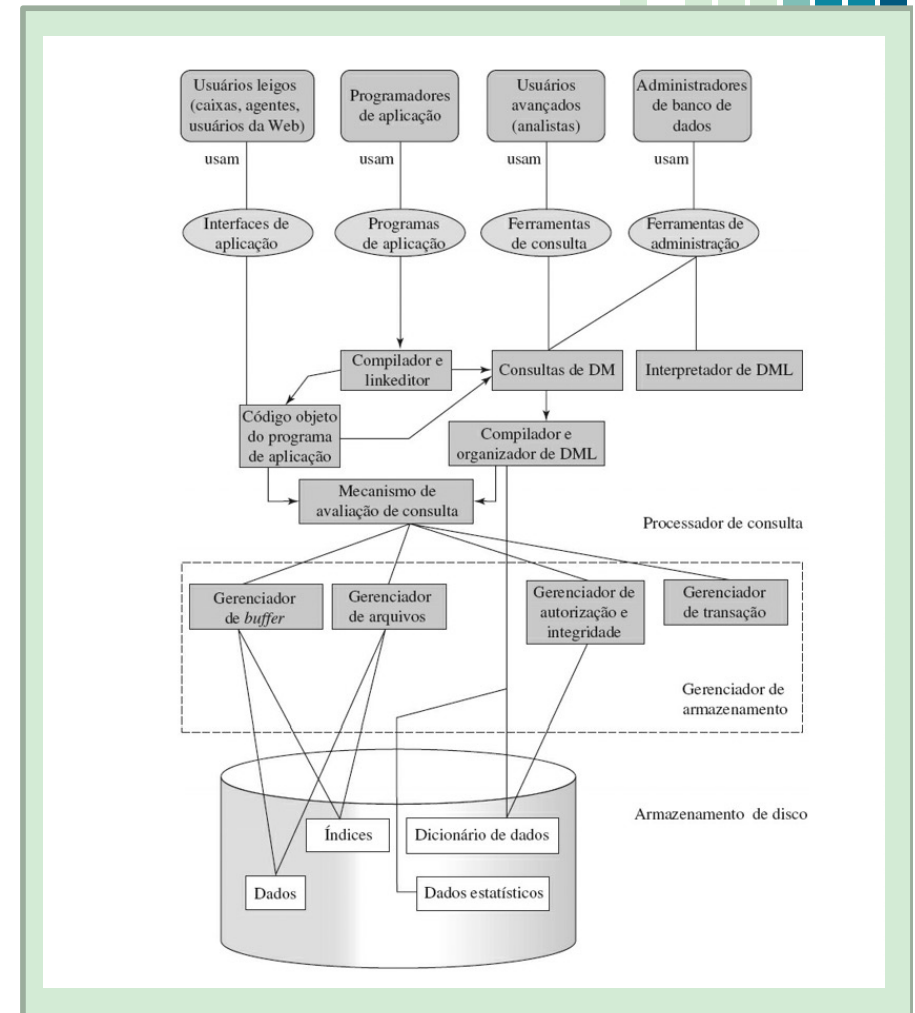


Objetivos

- ❖ Que tipos de memória existem num computador?
- ❖ Quais são as características físicas dos discos rígidos e das fitas e como que afetam o design de sistemas de bancos de dados?
- ❖ O que são os sistemas RAID de memória e quais são as suas vantagens?
- ❖ Como é que um SGBD registra o espaço em disco?
- ❖ Como é que um SGBD lê e modifica os dados em disco? Qual é o significado de um bloco enquanto unidade de armazenamento e transferência de dados?
- ❖ Como é que um SGBD cria e mantém arquivos de registros? Como é que os registros estão organizados em blocos, e como estão os blocos organizados dentro de um arquivo?

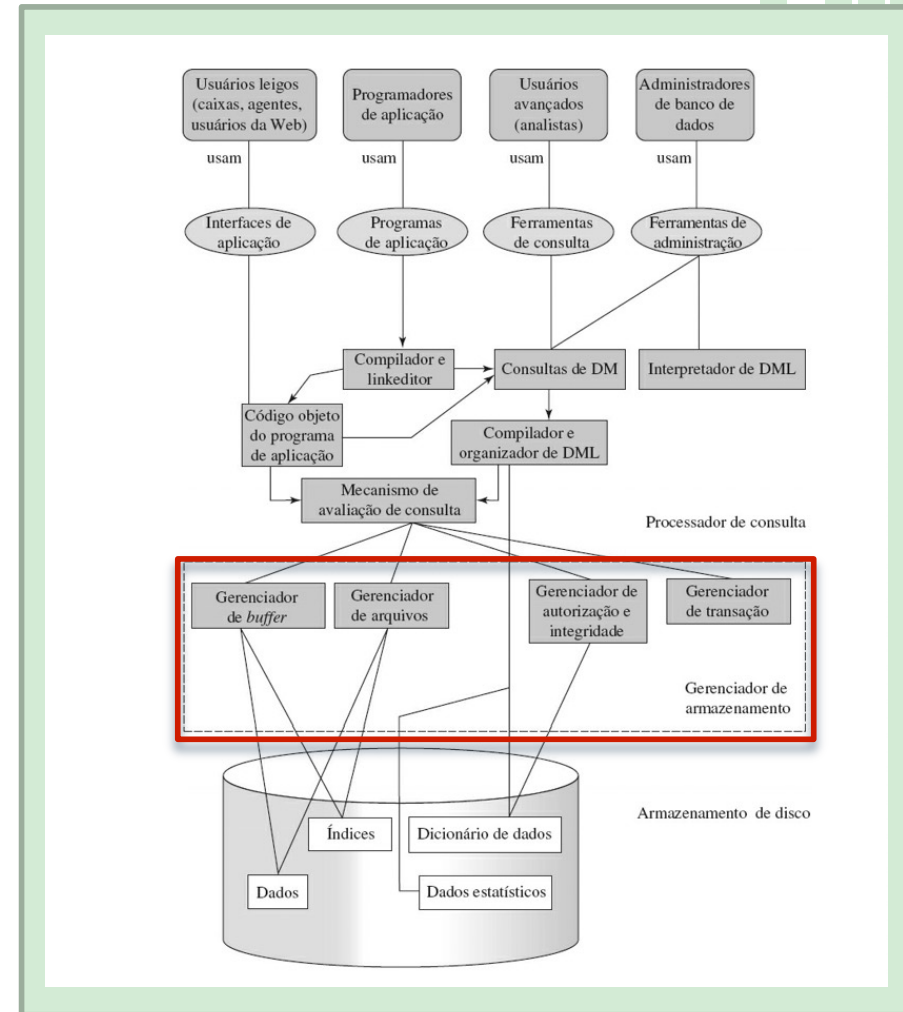
Estrutura Simplificada de um SGBD

- ✓ Bancos de Dados armazenam grandes quantidades de dados por períodos longos de tempo em meios de armazenamento secundário
- ✓ Um **SGBD** provê geralmente várias opções para organização física dos dados.
- ✓ **Projeto Físico** de Banco de Dados:
 - ✓ Busca determinar o melhor tipo de organização dos dados, dentre todas as possíveis, para uma determinada aplicação;
 - ✓ Especificar o Modelo Físico de Banco de Dados, levando em consideração o Modelo de Dados lógico e informações sobre volumes, acessos e necessidade de disponibilidade
 - ✓ Visando garantir uma implementação com ótima performance
 - ✓ Assegurando aspectos como padronização, portabilidade, disponibilidade e capacidade de recuperação tempestiva dos dados.
- ✓ Cada sistema tem as suas próprias particularidades.

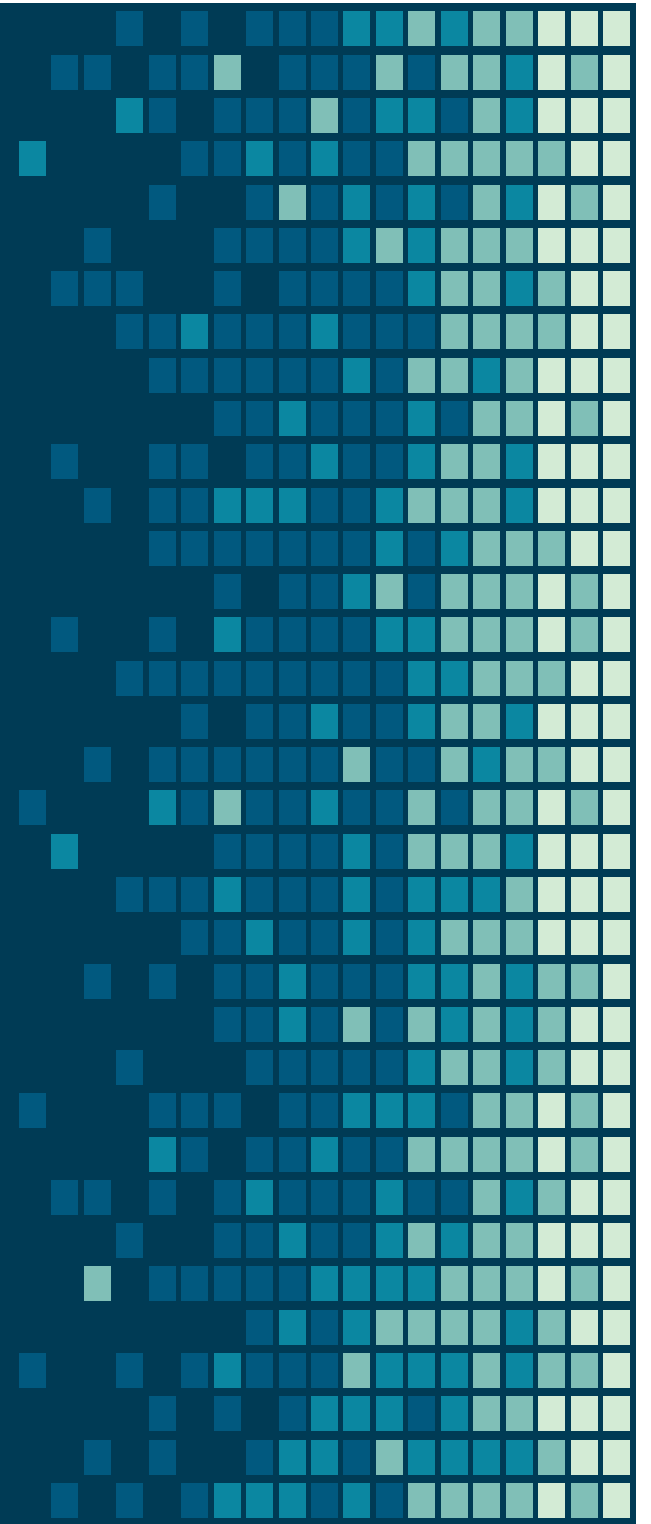


Discos e Arquivos

- ✓ Um SGBD guarda informação em discos.
- ✓ Este fato tem grandes implicações no projeto de um SGBD
 - ✓ **READ**: transferência de dados do disco para a memória principal (RAM).
 - ✓ **WRITE**: transferência de dados da RAM para o disco.
- ✓ Ambas são operações de custo elevado em termos de tempo e espaço em memória, de modo que devem ser planejadas cuidadosamente!



Meios Físicos de Armazenamento



Aspectos a Serem Considerados

- ❖ Velocidade com a qual um dado pode ser acessado
- ❖ Custo para ler e armazenar cada unidade de dado
- ❖ Disponibilidade
 - ❖ Perda de dados em caso de falha no sistema
 - ❖ Falhas físicas nos dispositivos de armazenamento

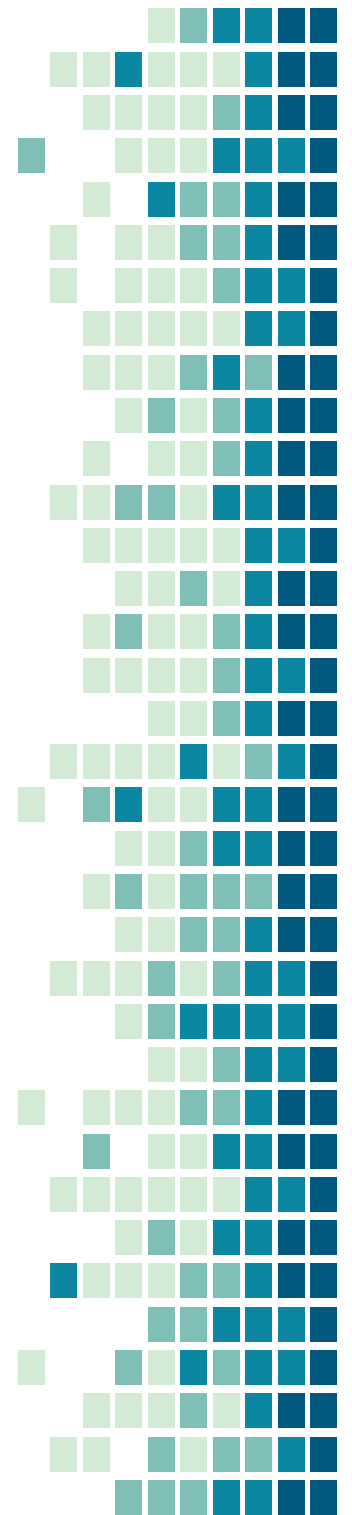
Visão Geral dos Meios Físicos de Armazenamento

1. Voláteis

- ❖ Memória RAM (Principal)
- ❖ Memória Cache

2. Não-voláteis

- ❖ Memória secundária
 - Disco magnético
 - Disco ótico
 - Memória flash
- ❖ Memória terciária
 - Fita



Por que não armazenar tudo em memória principal?

- ✓ Custo muito alto



- ✓ Memória volátil = capacidade de não reter a informação quando a energia elétrica é desligada

Hierarquia de Armazenamento

1. Armazenamento Primário:

Memória Principal + Caches de Memória:

acessado diretamente pela CPU, acesso rápido, custo alto

2. Armazenamento Secundário ou On-Line:

Discos magnéticos, óticos e memória flash

armazena a base de dados em si, acesso lento

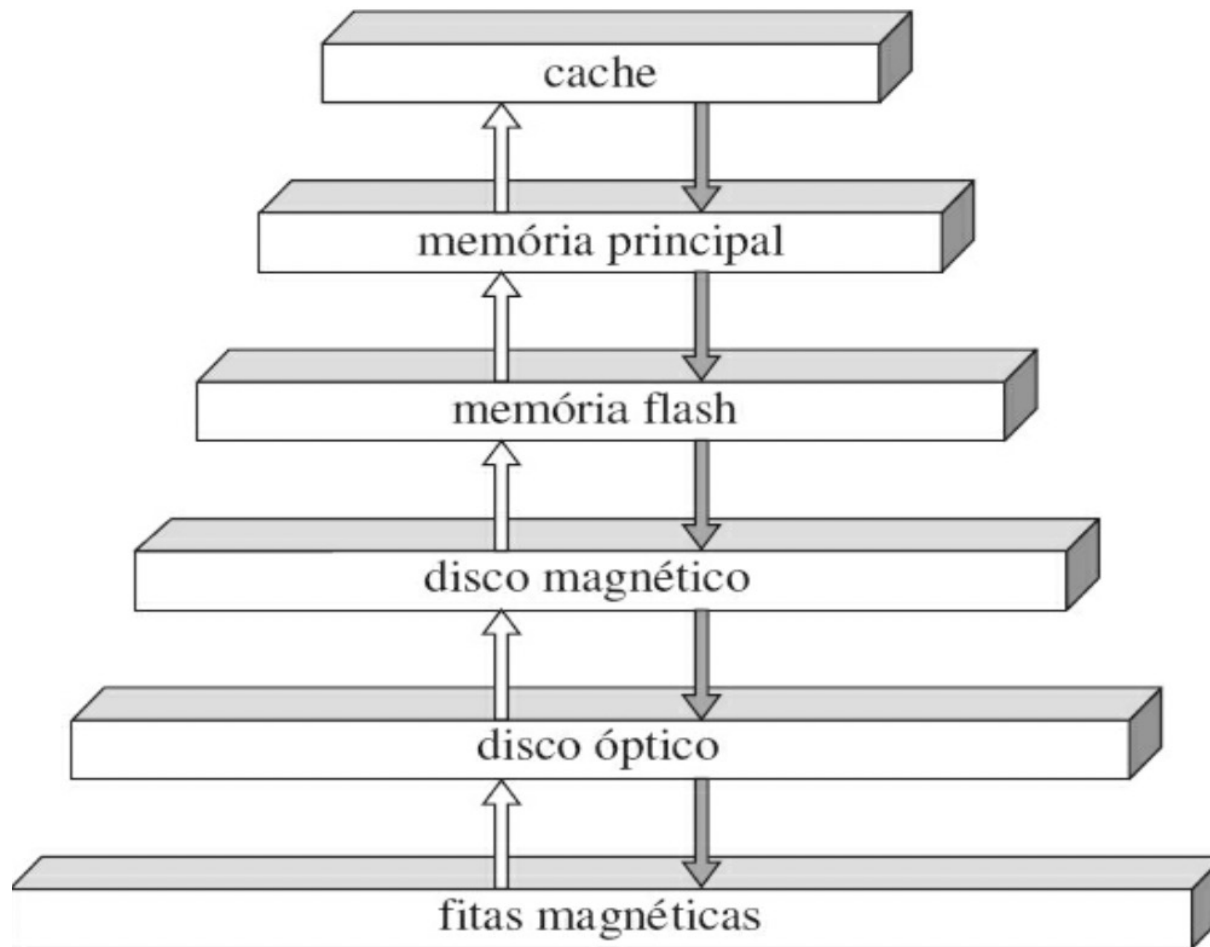
dados são copiados nos meios de armazenamento primário para serem processados e depois reescritos novamente

3. Armazenamento Terceário ou Off-Line:

Fitas = para versões antigas da base de dados (ou backups)



Hierarquia de Armazenamento

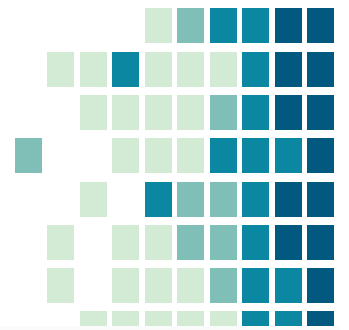


Discos Magnéticos

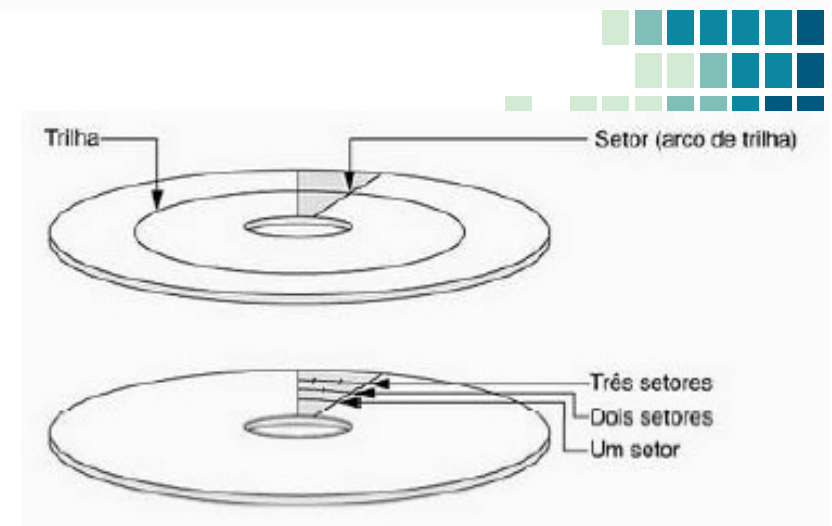
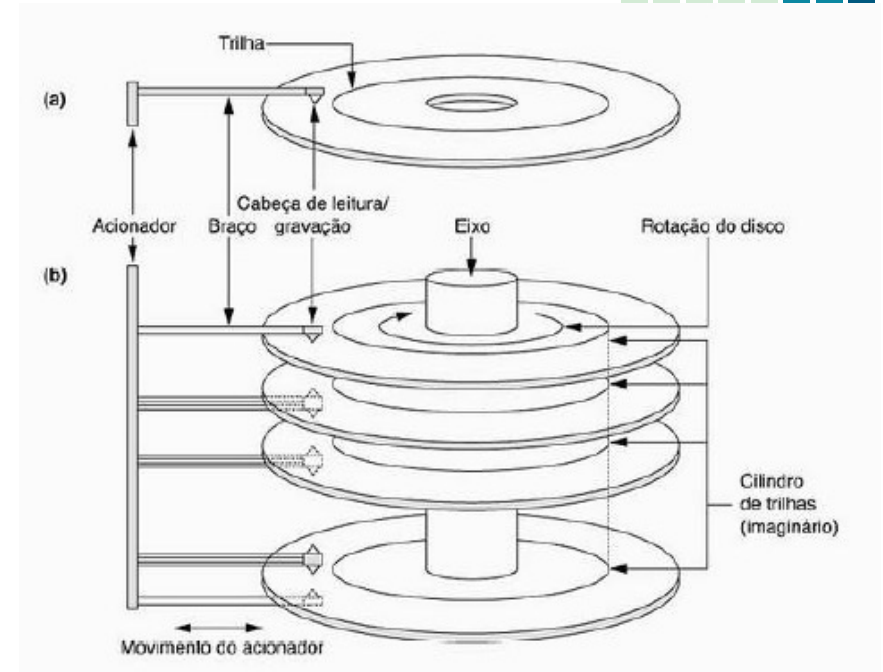
- ❑ Principal dispositivo de memória secundária.
- ❑ Principal vantagem com relação às fitas:
 - ❑ **acesso aleatório (randômico) x sequencial.**
- ❑ Dados são armazenados e devolvidos em unidades chamadas **blocos** ou **páginas**.
- ❑ Ao contrário da RAM, o tempo para devolver um bloco de um disco varia com a sua localização em disco, o que tem grande impacto no desempenho de um SGBD.



Componentes de um disco

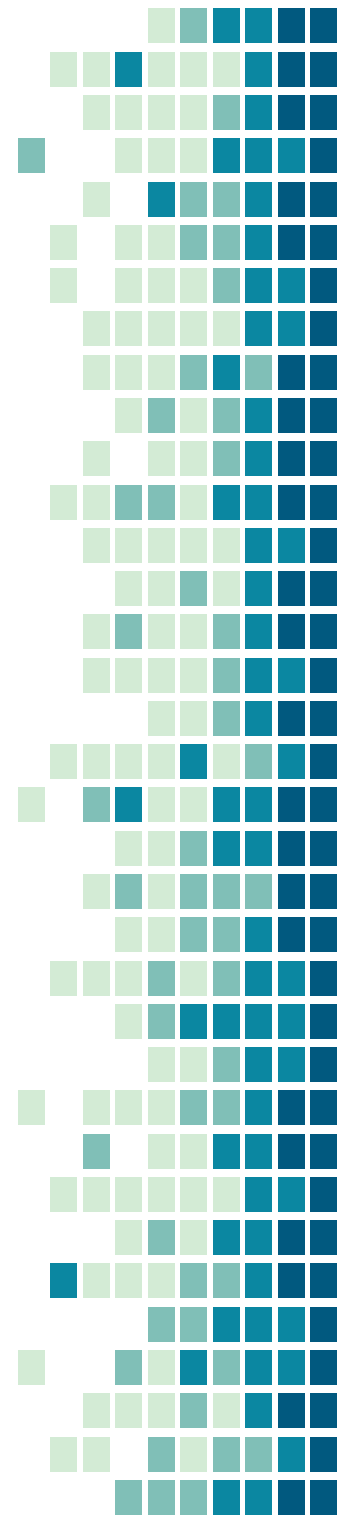


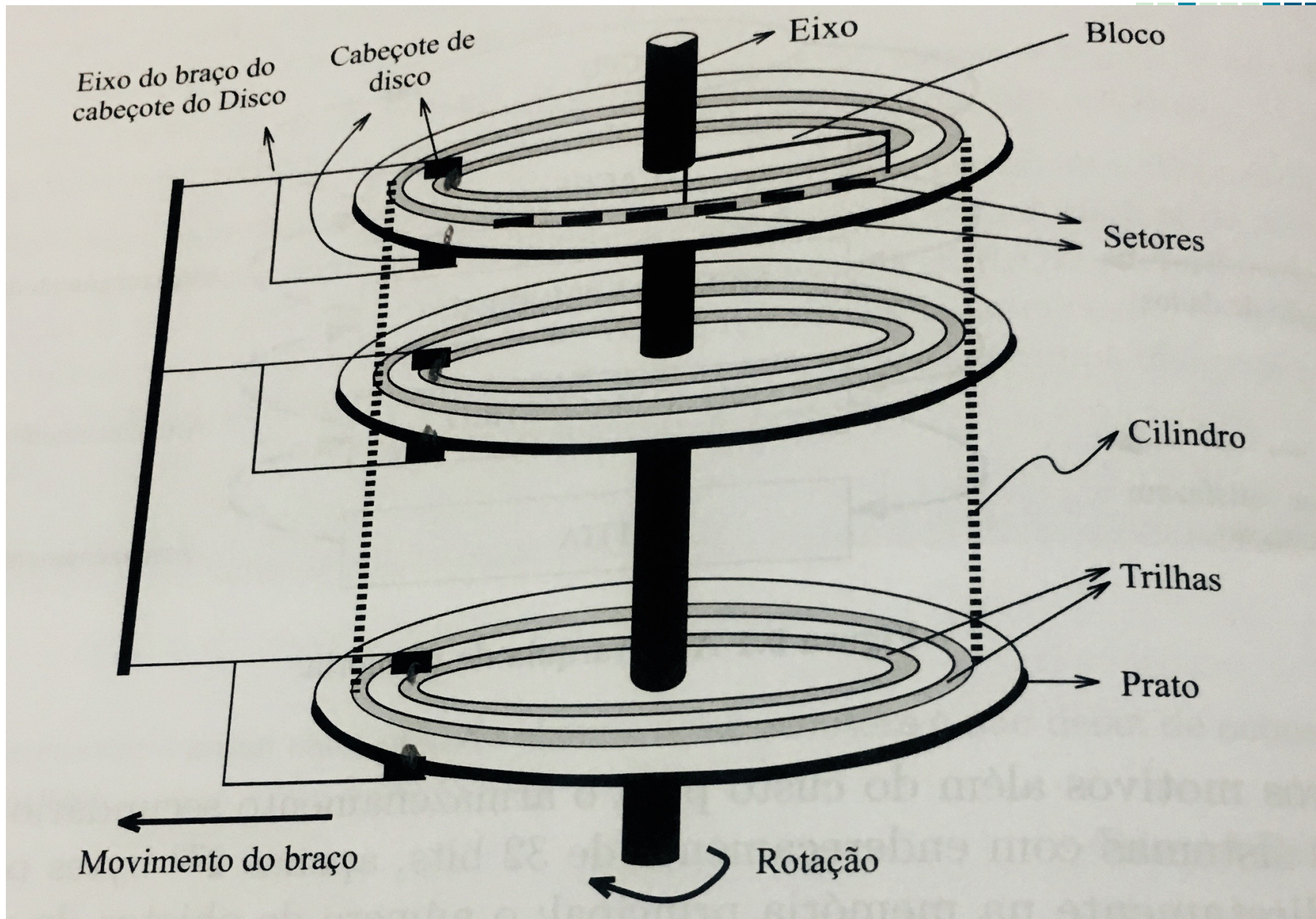
- Os blocos são armazenados em trilhas **trilhas** que formam cada prato (disco)
 - Até 152000 trilhas por superfície
- Cada trilha é dividida em vários **setores** (menor unidade de informação lida/escrita)
 - 1 setor = 32 até 4096 bytes.
 - Tamanho padrão é 512 bytes.
- Os **pratos** giram. Por exemplo, 5400 rpm.
- A agulha move-se para dentro ou para fora de modo a posicionar a cabeça sobre a trilha desejada.
- As trilhas de mesmo diâmetro formam cilindro (imaginário!).
- Só uma cabeça lê/escrive em cada instante.
- Tamanho do **bloco** é múltiplo do tamanho do setor (que é fixo).
 - Geralmente de 512 8.192 bytes.



Otimização de Acesso de Blocos de Disco

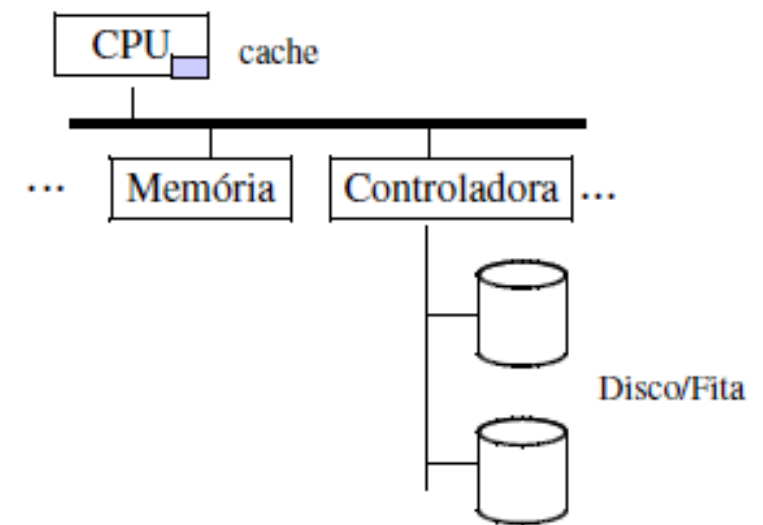
- ✓ **Bloco** = é uma sequência contígua de bytes de uma única trilha de um prato
 - ✓ Dados são transferidos do disco para a memória principal em blocos
 - ✓ Os tamanhos dos blocos variam de 512 bytes a vários kb
 - ✓ Blocos menores – mais transferências do disco
 - ✓ Blocos maiores – mais espaço desperdiçado
 - ✓ O tamanho mais comum varia de 4 a 16 kbytes





Subsistema de discos

- ❑ Um controlador de discos, comumente embutido na unidade de disco, controla o disco e o interliga ao sistema de computação, fazendo interface entre o disco e a memória RAM
 - ❑ Recebe e realiza os comandos de leitura/escrita de dados
 - ❑ Remapeamento de setores ruins
- ❑ ATA, SATA, SCSI (mais usada para conectar discos a computadores pessoais e estações de trabalho)



Medidas de Desempenho de Disco



1. Capacidade

2. **Tempo de Acesso** – acesso de leitura ou escrita requer três passos:

- Tempo de procura (seek): posicionamento do braço na trilha correta. De 3 a 10 ms.
- Tempo de atraso (latência rotacional): espera até o setor desejado seja rotacionado até a cabeça de leitura/escrita. 5400 to 15000 r.p.m.
- Tempo de transferência: transferência dos bits armazenados no setor que está ao alcance da cabeça. 25 a 100 Mb por segundo
 - Bloco: unidade de transferência

3. **Confiabilidade**

- Tempo médio para a ocorrência de falhas: média de tempo que se pode esperar que o disco trabalhe sem que ocorra falhas. De 3 a 5 anos.

Exemplo – Seagate

Tabela 16.2 Especificações de discos típicos de alto nível da Seagate. (a) Seagate Enterprise Performance 10K HDD – 1200 GB.

Especificações	1.200 GB
Número do modelo SED	ST1200MM0017
Número do modelo SED FIPS 140-2	ST1200MM0027
Nome do modelo	Enterprise Performance 10K HDD v7
Interface	6 Gb/s SAS
Capacidade	
Formatada com 512 bytes/setor (GB)	1.200
Taxa de transferência externa (MB/s)	600
Desempenho	
Velocidade do eixo (RPM)	10 K
Latência média (ms)	2,9
Taxa de transferência sustentada do diâmetro externo ao interno (MB/s)	204 a 125
Cache multissegmentado (MB)	64
Configuração/confiabilidade	
Discos	4
Cabeças	8
Erro de leitura não recuperável por bits lidos	1 por 10E16
Taxa de falha anual (AFR)	0,44%
Físico	
Altura (pol./mm, máxima)	0,591/15,00
Largura (pol./mm, máxima)	2,760/70,10
Profundidade (pol./mm, máxima)	3,955/100,45
Peso (lb/kg)	0,450/0,204

Fonte: cortesia da Seagate Technology

(continua)

Tabela 16.2 (b) Características internas da unidade de discos Seagate de 300 GB a 900 GB.

(Continuação)

	ST900MM0006	ST600MM0006	ST450MM0006	ST300MM0006	
	ST900MM0026	ST600MM0026	ST450MM0026	ST300MM0026	
	ST900MM0046	ST600MM0046	ST450MM0046	ST300MM0046	
	ST900MM0036				
Capacidade da unidade	900	600	450	300	GB (formatados, arredondado)
Cabeças de leitura/gravação de dados	6	4	3	2	
Bytes por trilha	997,9	997,9	997,9	997,9	KBytes (média, valores arredondados)
Bytes por superfície	151.674	151.674	151.674	151.674	MB (não formatados, arredondado)
Trilhas por superfície (total)	152	152	152	152	KTrilhas (acessíveis ao usuário)
Trilhas por polegada	279	279	279	279	KTPI (média)
Pico de bits por polegada	1.925	1.925	1.925	1.925	KBPI
Densidade de área	538	538	538	538	Gb/pol. ²
Velocidade de rotação do disco	10K	10K	10K	10K	rpm
Latência de rotação média	2,9	2,9	2,9	2,9	ms