Sistemas de Banco de Dados

Fundamentos em Bancos de Dados Relacionais

Wladmir Cardoso Brandão
www.wladmirbrandao.com

ARMAZENAMENTO EM MEMÓRIA

Armazenamento em Memória

BDs são armazenados fisicamente em meios (mídias) de armazenamento computacional

- Meios de armazenamento formam uma **hierarquia**, em que dados residem e transitam, sendo que a hierarquia reflete a *distância* do meio à CPU
  - Memória primária → *próxima* e operada diretamente pela CPU
  - Memória secundária → *distante* e não operada pela CPU
  - Memória terciária → *muito distante* e não operada pela CPU
- Programas residem e são executados em memória primária
- BDs são geralmente grandes e persistem em memória secundária
- SGBD transfere partes do BD entre memórias de acordo com a necessidade

Armazenamento em Memória

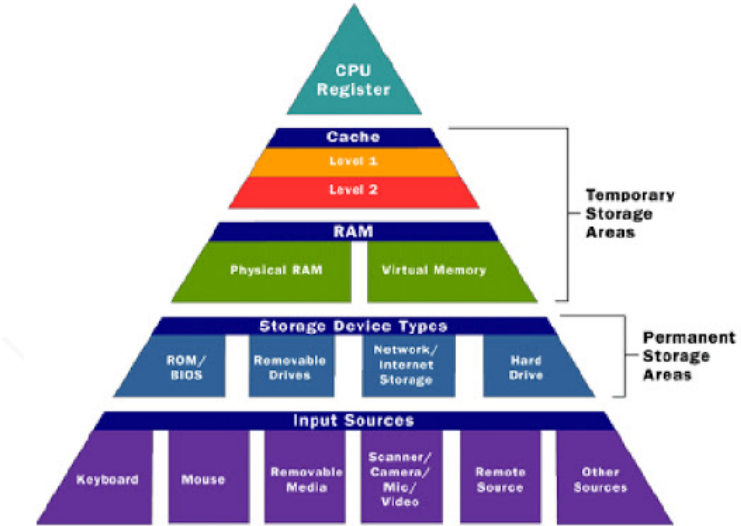
Existe uma correlação entre capacidade de armazenamento, velocidade de transferência e custo em meios de armazenamento

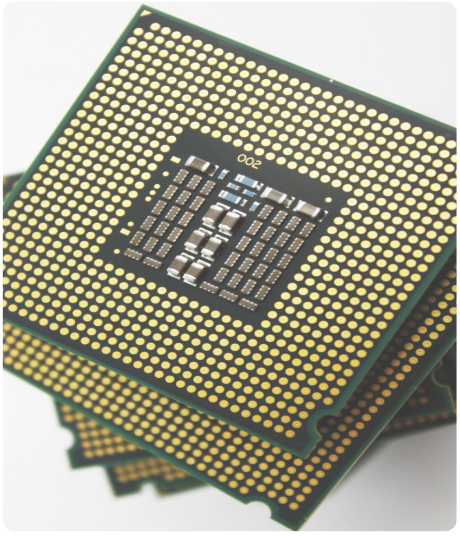
- Capacidade de armazenamento → quantidade de dados (bytes) que podem ser armazenados na memória
- Velocidade de transferência → quantidade de dados (bits) que podem ser transferidos de ou para a memória por unidade de tempo (segundo)
- Custo → unidade monetária ($) por quantidade de dados (bytes) que podem ser armazenados na memória

Correlação:

- $>$ capacidade $\Rightarrow$ $<$ velocidade
- $>$ velocidade $\Rightarrow$ $>$ custo

Hierarquia de Memória





Registrador

Memória eletrônica

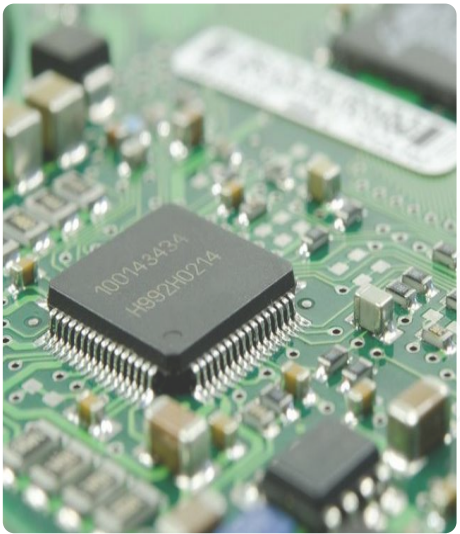
Interna da CPU

Rápida $\rightarrow \approx$ 60 Tbps

Pequena $\rightarrow$ centenas de bytes

Cara $\rightarrow >$ 500 R$/MB

Utilizada para execução de instruções de programa



Cache

Memória eletrônica

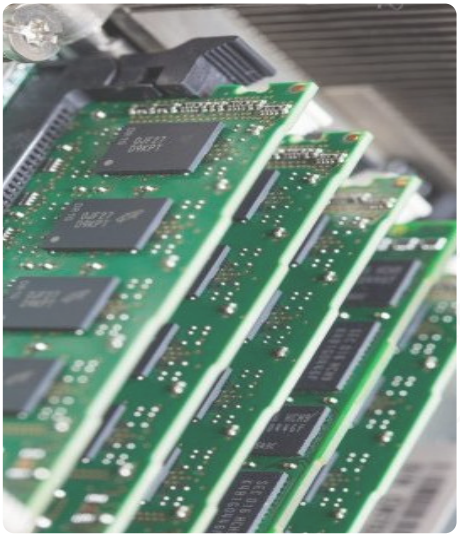
Vários níveis → L0 a L4

Rápida → L1 ≈ 6 Tbps

Pequena → L4 ≈ 128 MB

Cara → L0 > 100 R$/MB

Acelera a execução de instruções de programa (pré-busca e *pipelining*)



RAM

Memória eletrônica

Acesso aleatório

Rápida $\rightarrow \approx$ 80 Gbps

Pequena $\rightarrow$ dezenas de GB

Cara $\rightarrow \approx$ 0,05 R$/MB

Utilizada para manter instruções de programa e dados temporários

Flash

Memória eletrônica

Resistente e durável

Rápida $\rightarrow \approx$ 5 Gbps

Média $\rightarrow$ alguns TB

Barata $\rightarrow \approx$ 0,0007 R$/MB

Utilizada para manter dados de maneira persistente

HD

Memória magnética

Discos em alta rotação

Lenta $\rightarrow \approx$ 100 Mbps

Grande $\rightarrow$ dezenas de TB

Barata $\rightarrow \approx$ 0,0002 R$/MB

Utilizada para manter dados de maneira persistente

Fita

Memória magnética removível

Acesso sequencial

Lenta $\rightarrow \approx$ 2 Mbps

Grande $\rightarrow$ PB (jukebox)

Barata $\rightarrow \approx$ 0,00003 R$/MB

Utilizada para manter dados pouco mutáveis e acessados de maneira persistente, como *backups*

Óptica

Memória removível

Discos ópticos

Lenta $\rightarrow \approx$ 20 Mbps

Grande $\rightarrow$ PB (jukebox)

Barata $\rightarrow \approx$ 0,0001 R$/MB

Utilizada para manter dados pouco mutáveis e de acesso sequencial de maneira persistente, como multimídia

Hierarquia de Memória

Comparativo entre diferentes tipos de memória:

Tipo	Nome	Velocidade (bps)	Capacidade	Custo (R$/MB)	Volátil
CPU	Registrador	60T	KB	500	sim
Primária	Cache	6T	MB	100	sim
Primária	RAM	80G	GB	0,05	sim
Secundária	Flash	5G	TB	0,0007	não
Secundária	HD	100M	TB	0,0002	não
Terciária	Óptico	20M	PB	0,0001	não
Terciária	Fita	2M	PB	0,00003	não

Os valores de velocidade, capacidade e custo são estimativas, a fim de fornecer uma ordem de grandeza. Estimativas foram baseadas em memórias disponíveis atualmente, podendo variar de acordo com a tecnologia e o fabricante

Armazenamento em Memória

Em sistemas de banco de dados, os dados são efetivamente armazenados em diferentes tipos de memória de acordo com sua natureza

- Transientes → persistem em memória por um período limitado de tempo, apenas durante a execução do programa
- Persistentes → permanecem em memória por longos períodos de tempo, sendo acessados e processados repetidamente durante esse período

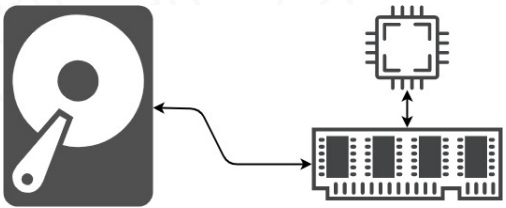
SGBDs devem ser capazes de gerenciar eficientemente a transferência de dados transientes e permanentes entre memórias

- No **projeto físico**, DBAs e projetistas devem escolher as melhores técnicas de organização de dados para garantir equilíbrio entre custo e desempenho, atendendo aos requisitos funcionais e operacionais do BD

Armazenamento em Memória

Aplicações tipicamente necessitam de apenas uma pequena parte do BD de cada vez para processamento, sendo responsabilidade do SGBD garantir que:

1. A parte seja transferida da memória secundária para a primária
2. A CPU processe os dados em memória primária adequadamente
3. Os dados processados sejam transferidos de volta à memória secundária

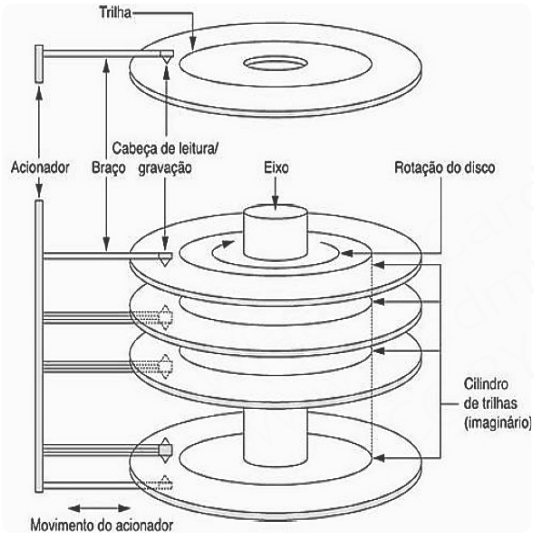


Armazenamento em Memória

Tipicamente BDs são armazenados de maneira permanente em discos magnéticos

- BDs são muito grandes para caberem inteiramente em memória primária, com capacidade limitada de armazenamento
- Custo de armazenamento em memória primária é muito alto
- Memórias terciárias tem grande capacidade de armazenamento e baixo custo, mas são muito lentas e frequentemente demandam intervenção manual (*off-line*)
- Discos magnéticos apresentam excelente relação custo-benefício, ainda mais vantajosa que outros tipos de memória secundária

Disco Magnético (HD)



Acesso aleatório

Múltiplas superfícies

Armazenamento em TRILHAS

Trilhas divididas em BLOCOS

Tamanho do bloco é fixado na formatação do HD e não pode ser trocado dinamicamente

Transferências entre memória primária e HD ocorrem em unidades de bloco

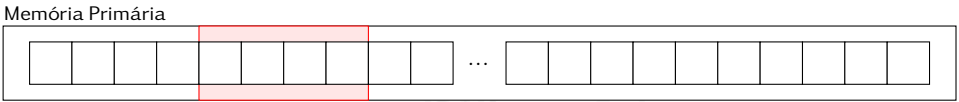
Disco Magnético (HD)

Bloco (página) → unidade mínima de transferência de dados entre disco e memória primária

- Tamanho fixado na formatação, geralmente entre 512B a 8KB, que não pode ser alterado dinamicamente
- Separados nas trilhas por **lacunas** de tamanho fixo que incluem dados de controle, como ponteiro para o bloco subsequente
- Pode ser acessado aleatoriamente pelo seu endereço de hardware, denominado **endereço de bloco**
- Hardware controladores de disco usam o endereço do bloco para transferir o bloco do disco para um *buffer* em memória primária

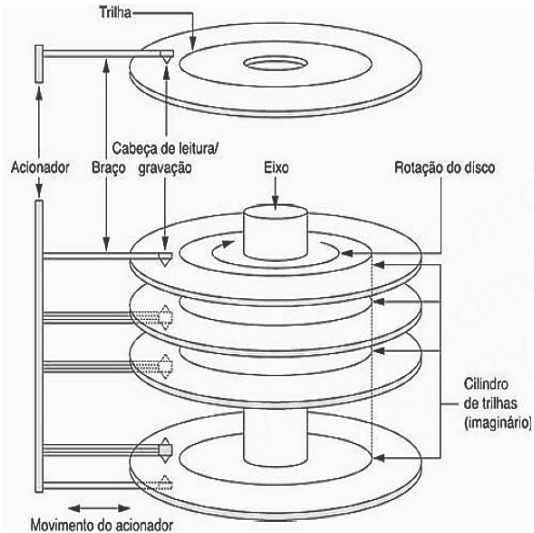
Disco Magnético (HD)

Buffer → área reservada contígua em memória primária



- Controladores de disco usam o endereço de bloco e de *buffer* para realizar a transferência do bloco de disco para a memória primária
  - Leitura (Input) → bloco é copiado para *buffer*
  - Escrita (Output) → *buffer* é copiado para bloco

HD: Leitura e Escrita (I/O)



1) Controlador recebe endereços de bloco e *buffer*
2) Controlador comanda acionador a movimentar braço para posicionar cabeça na trilha do endereço de bloco
3) Discos giram até o ponto de leitura e escrita
4) Dados são copiados de ou para *buffer*

HD: Leitura e Escrita (I/O)

Tempo de Transferência → tempo necessário para transferir um bloco entre disco e memória primária

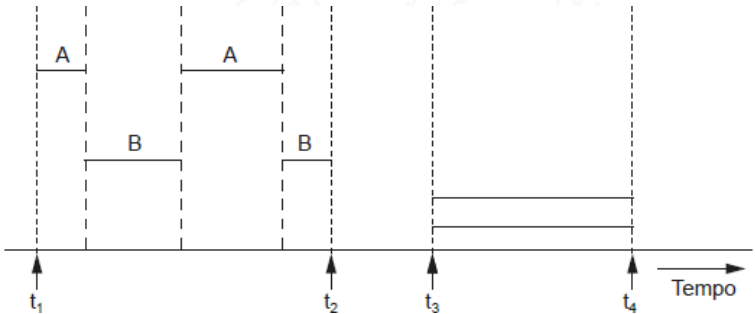
- Tempo de Busca → tempo necessário para posicionar a cabeça de leitura e escrita na trilha do endereço de bloco
- Tempo de Latência → ou atraso rotacional é o tempo necessário para o disco girar até o ponto de leitura e escrita
- Tempo de Transferência de Bloco → tempo necessário para os dados serem copiados de ou para o *buffer* em memória primária

Transferência de Bloco ≪ Busca + Latência

- Transferir múltiplos blocos consecutivos na mesma trilha ou cilindro elimina tempos de busca e latência acumulados, tornando a transferência mais eficiente

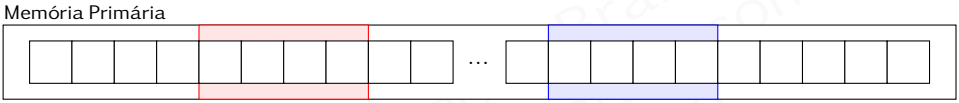
Buffering de Blocos → técnica que reserva vários *buffers* em memória primária para agilizar a transferência de blocos do disco

- Controladores de disco e CPUs podem operar de forma independente e paralela usando *buffers* diferentes



HD: Leitura e Escrita (I/O)

Duplo Buffering → uso de dois *buffers* para leitura ou gravação em disco



- Enquanto o controlador de disco transfere dados de ou para um *buffer*, a CPU processa dados no outro *buffer*
- Permite leitura ou gravação contínua em blocos consecutivos
- Elimina tempos de busca e latência para todas as transferências de bloco, com exceção da primeira
- Dados ficam prontos para processamento mais rapidamente, reduzindo ociosidade da CPU e, consequentemente o tempo de espera das aplicações

HD: Alocação de Blocos

A forma como os blocos são alocados em disco impacta o desempenho de I/O

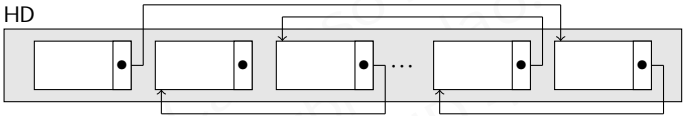
- Alocação Contígua → blocos consecutivos em disco

  - Rápido I/O com *duplo buffering*
  - Difícil expansão, podendo resultar em múltiplas realocações em caso de alteração dos dados

HD: Alocação de Blocos

A forma como os blocos são alocados em disco impacta o desempenho de I/O

- Alocação por Ligação → cada bloco contém um ponteiro para o próximo

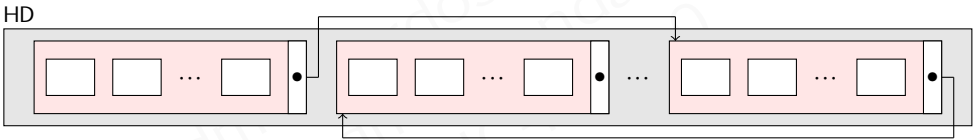


  - Facilita expansão
  - I/O mais lento pela impossibilidade de uso de *duplo buffering*

HD: Alocação de Blocos

A forma como os blocos são alocados em disco impacta o desempenho de I/O

- Alocação por Segmento → agrupa blocos consecutivos em segmentos (*clusters*) e cada segmento contém um ponteiro para o próximo segmento

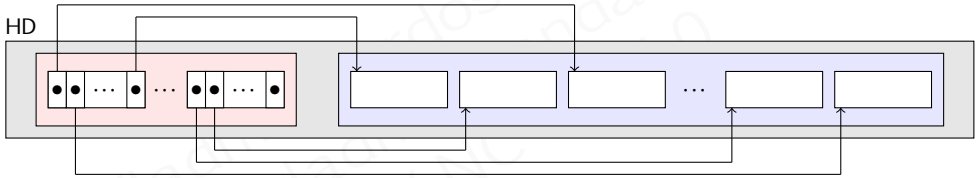


  - Combinação de alocação contígua e por ligação
  - Torna *duplo buffering* viável em um segmento, agilizando I/O
  - Facilita expansão, reduzindo o número de realocações em caso de alteração dos dados

HD: Alocação de Blocos

A forma como os blocos são alocados em disco impacta o desempenho de I/O

- Alocação Indexada → blocos especiais de índice são criados contendo ponteiros para blocos de dados



  - Rápido I/O com busca sendo efetuada em blocos de índice, que podem ter alocação contígua ou por segmento (*duplo buffering*)
  - Fácil expansão, com realocações ocorrendo em blocos de índice

Referências Bibliográficas

[1] Elmasri, Ramez; Navathe, Sham. *Fundamentals of Database Systems*. 7ed. Pearson, 2016.

[2] Silberschatz, Abraham; Korth, Henry F.; Sudarshan, S. *Database System Concepts*. 6ed. McGraw-Hill, 2011.

[3] Date, Christopher J. *An Introduction to Database Systems*. 8ed. Pearson, 2004.