

# Armazenamento em Disco, Estruturas básicas de Arquivos e Hashing

INF 220-Banco de Dados I - Prof. Jugurta Lisboa Filho

Apresentação elaborada por Glauber Costa (Estágio em Ensino 2011 – PPGCC/UFV)

# Agenda

---

- ▶ Objetivos
- ▶ Introdução
- ▶ Estrutura geral do SBGD
- ▶ Estruturas de arquivos e de armazenamento
- ▶ Dispositivos de Armazenamento
- ▶ Medidas de Desempenho
- ▶ Gerenciamento do Buffer
- ▶ Organização de Arquivos
- ▶ Sistemas de Arquivos
- ▶ Organização de Registros em Arquivos
- ▶ Armazenamento do dicionário de dados

# Objetivos

---

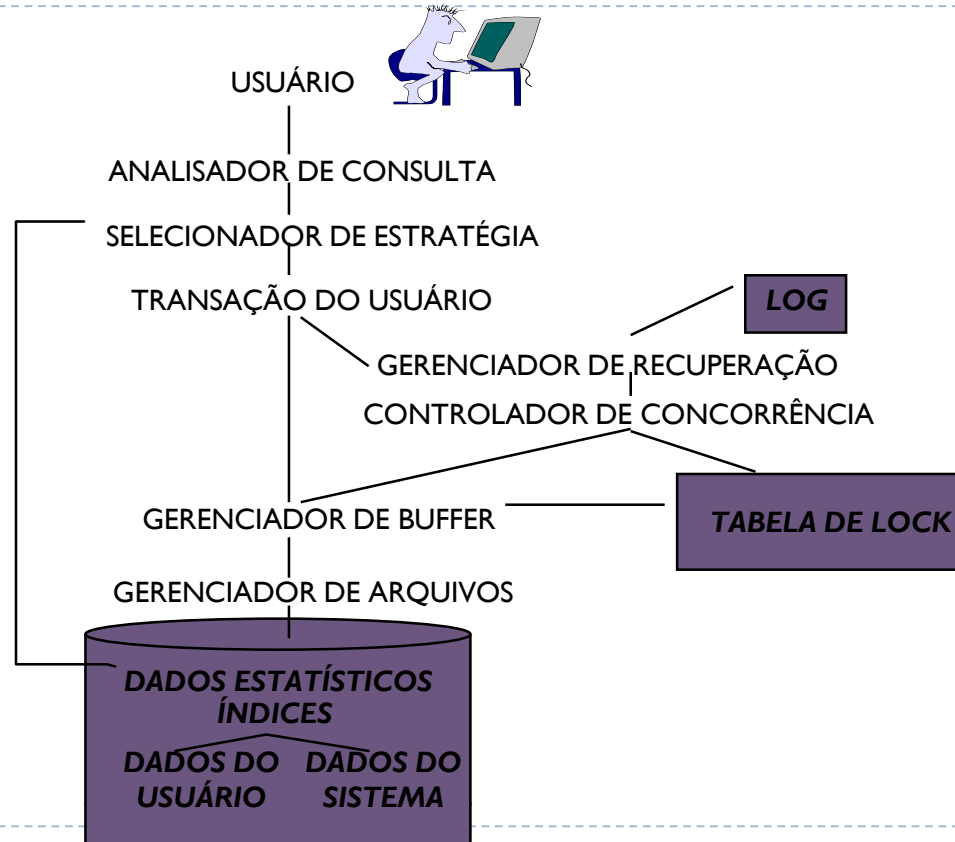
- ▶ Compreender a forma como os dados são armazenados na memória secundária e como este armazenamento afeta o desempenho das aplicações de banco de dados.

# Introdução

---

- ▶ Banco de Dados:
  - ▶ Commodity!
- ▶ Diferencial:
  - ▶ A satisfação do usuário é diretamente proporcional ao **desempenho** do banco de dados.
- ▶ Fatores que influenciam no desempenho:
  - ▶ Projeto bem elaborado
  - ▶ Sistema de hard/swf adequado ao porte do banco de dados em questão.
- ▶ Em suma:
  - ▶ Projetista de BD que não domina o ambiente onde o este está inserido elabora projetos menos eficazes.

# ESTRUTURA GERAL DO SGBD



# ESTRUTURA GERAL DO SGBD

---

## ▶ Componentes funcionais

### ▶ Gerenciador de arquivos

- ▶ Gerencia a alocação do espaço em disco.

### ▶ Gerenciador de buffer

- ▶ Responsável pela transferência de informação entre o disco e a memória principal.

### ▶ Interpretador de consultas

- ▶ Traduz da linguagem de consulta para uma linguagem de mais baixo nível.

### ▶ Seletor de estratégia

- ▶ Identifica a melhor estratégia para executar a consulta.

# ESTRUTURA GERAL DO SGBD

---

- ▶ Componentes funcionais (Continuação)

- ▶ Gerenciador de autorização e integridade

- ▶ Verifica permissão de acesso e integridade.

- ▶ Gerenciador de recuperação

- ▶ Mantém o BD em um estado consistente.

- ▶ Controlador de concorrência

- ▶ Resolve conflitos de acesso.

# ESTRUTURA GERAL DO SGBD

---

## ▶ Estruturas de Dados

- ▶ Arquivos de dados
- ▶ Dicionário de dados
- ▶ Índices - permitem acesso mais rápido
- ▶ Dados estatísticos - armazenam informações sobre os dados do BD



# ESTRUTURA DE ARQUIVOS E DE ARMAZENAMENTO

---

## ▶ Mídias disponíveis

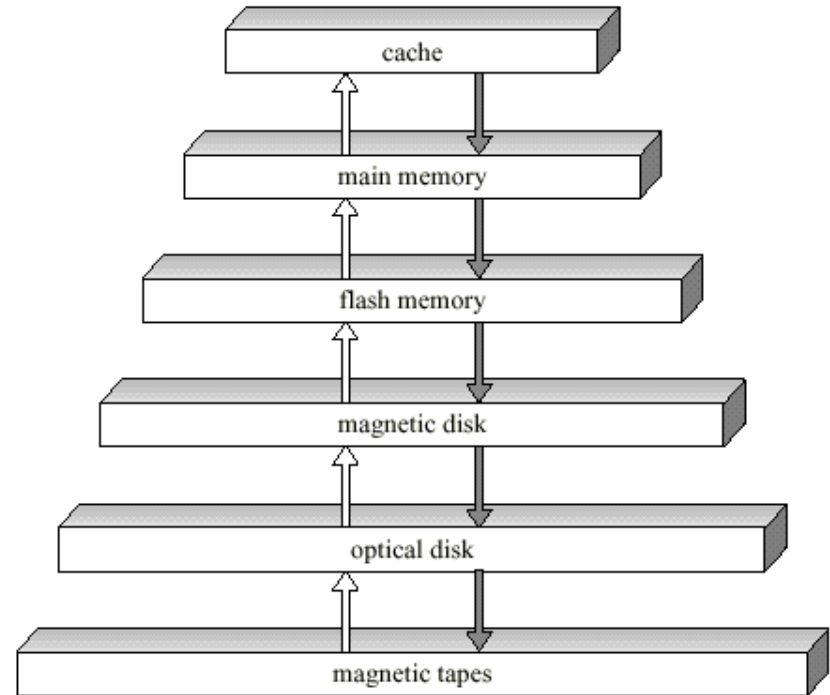
- ▶ **CACHE:** forma + rápida e cara. Memória pequena, gerenciada pelo SO/hardware.
- ▶ **MEMÓRIA PRINCIPAL:** utilizada para dados na memória. Memória volátil.
- ▶ **MEMÓRIA FLASH:** apenas de leitura eletronicamente apagável e programável (EEPROM). Tão rápida como a principal. Escrita complicada (apaga tudo e reescreve).
- ▶ **DISCO MAGNÉTICO:** principal mídia para armazenamento de BD. Memória de acesso direto (não sequencial). Não volátil.
- ▶ **ARMAZENAMENTO ÓTICO:** CD Rom + popular. (*WORM* - *write-once, read-many*).
  - ▶ Sistema *jukebox* - contém unidades com vários discos que podem ser carregados automaticamente.
- ▶ **FITA MAGNÉTICA:** não volátil. Usada para back-up e armazenamento de dados. Acesso sequencial. Alta capacidade (5GB). Baixo custo.

# ESTRUTURA DE ARQUIVOS E DE ARMAZENAMENTO

---

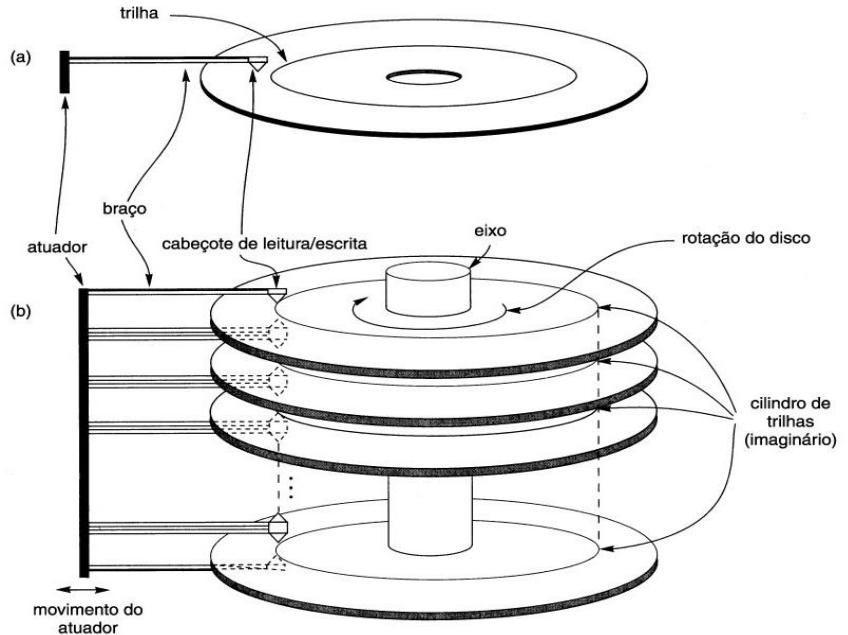
## ► Hierarquia dos meios físicos de armazenamento

- Memória principal: +rápida, mas volátil (cache e principal)
- Memória secundária: não volátil, tempo de acesso moderado (flash, disco)
- Memória terciária: não volátil, acesso lento (fita e disco ótico)



# Dispositivos de Armazenamento

- ▶ **Disco** - feito de metal ou vidro coberto por material magnético.
- ▶ **Trilhas**: círculo concêntrico onde são guardadas as informações
- ▶ **Setor** : menor unidade de informação para E/S.
- ▶ **Cilindro**: conjunto de trilhas com o mesmo diâmetro em diferentes discos.
- ▶ **Braço**: e suporte do braço
- ▶ **Cabeçote**: flutua a microns da superfície

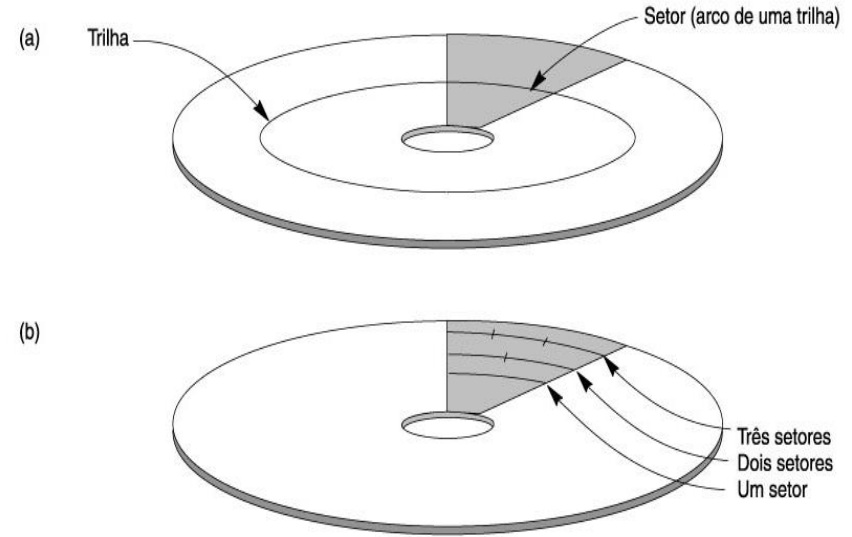


(a) Um disco de face única com hardware de leitura/escrita.  
(b) Um conjunto de discos com hardware de leitura/escrita.

# Dispositivos de Armazenamento

## ► Organização dos setores:

- Uma parte da trilha que se estende por um ângulo fixo. A medida que se afasta do centro, o “arco” armazena mais informações.
- Divisão com que o ângulo do setor aumente a medida que o setor se afaste do centro, permitindo uma densidade informacional uniforme.



(a) Setores formados por um ângulo fixo.

(b) Setores que mantêm uma densidade uniforme de gravação.

# Medidas de desempenho

---

## ▶ TEMPO DE ACESSO:

- ▶ compreendido entre um pedido de leitura/gravação até início da transferência. Consiste de:
  - Tempo de PROCURA: reposicionamento do braço sobre a trilha
  - Tempo de LATÊNCIA rotacional: aguarda o posicionamento sobre o setor.

## ▶ TAXA DE TRANSFERÊNCIA DE DADOS

- ▶ Taxa na qual dados são recuperados ou armazenados no disco. (ex.: 1 a 5 Mb por segundo)

## ▶ TEMPO DE OCORRÊNCIA DE FALHA

- ▶ Tempo que se espera que o disco funcione sem apresentar falhas. (ex.: de 3 a 91 anos)

# Gerenciamento de Buffer

---

- ▶ Buffer

- ▶ Área de memória principal usada na transferência de dados entre o disco e a memória.

- ▶ Otimização de Acesso a Dados:

- ▶ Minimizar o número de blocos a serem acessados (Clustering)
  - ▶ Manter o maior número de blocos possível na memória

- ▶ Gerenciador do Buffer:

- ▶ Recebe requisição de bloco
  - ▶ Se o bloco já está na memória
    - OK
  - ▶ Se não:
    - Primeiro Caso: busca o bloco do disco e armazena em um buffer vazio
    - Segundo Caso: Não há buffer vazio  $\Rightarrow$  remoção de blocos

# POLÍTICAS DE SUBSTITUIÇÃO DE BUFFER

---

- ▶ Política de remoção de páginas em SOs
  - ▶ Normalmente é usada LRU (*least recently used*) menos recentemente utilizada.
- ▶ Gerenciador do Buffer utiliza técnicas mais sofisticadas
  - ▶ Em programas de uso genérico não é possível prever quais blocos serão referenciados.
  - ▶ Um SBD, muitas vezes, é capaz de antecipar quais blocos serão necessários para executar uma operação.
  - ▶ Ex.: A expressão **empréstimo [X] cliente** pode ser resolvida pelo algoritmo:  
    **para** cada tupla E de empréstimo **faça**  
        **para** cada tupla C de cliente **faça**  
            **se** E[nomecli] = C[nomecli] **então**  
                inclui na relação resultante
  - ▶ Blocos contendo as tuplas da relação empréstimo podem ser descartados imediatamente após a última tupla ser processada => LRU não é eficiente.

# Organização de Arquivos

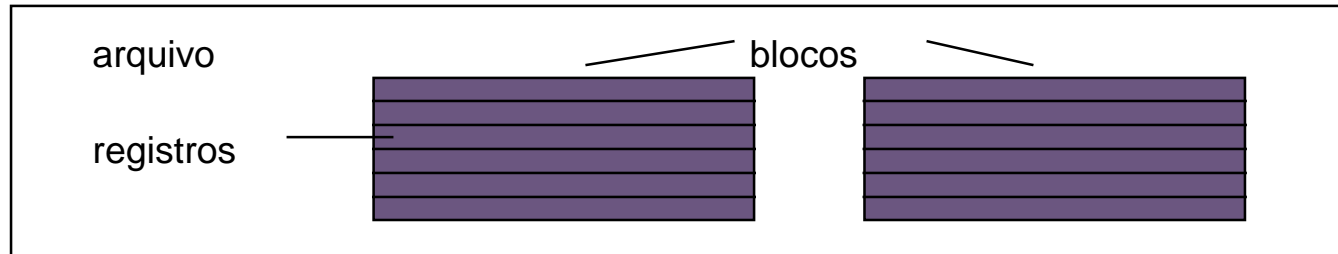
---

## ▶ Bloco

- ▶ unidade de memória usada na transferência entre o disco e a memória principal.

## ▶ Registro

- ▶ Um arquivo é organizado como uma sequência de registros.
- ▶ Registros podem ter tamanho fixo ou variável.





# Organização de Arquivos

## ▶ Arquivos com registro de tamanho fixo

Ex: TYPE Aparelhos = RECORD

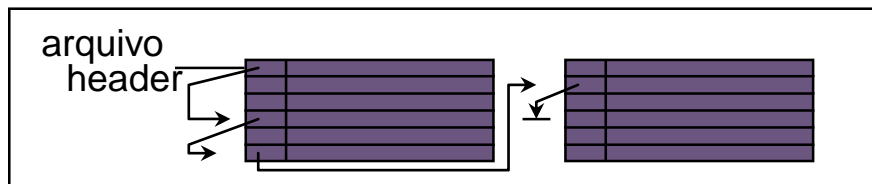
NumAp : INTEGER;

Tipo : CHAR(15);

Idade: INTEGER;

Dono : CHAR(15);

END



## ▶ Problemas:

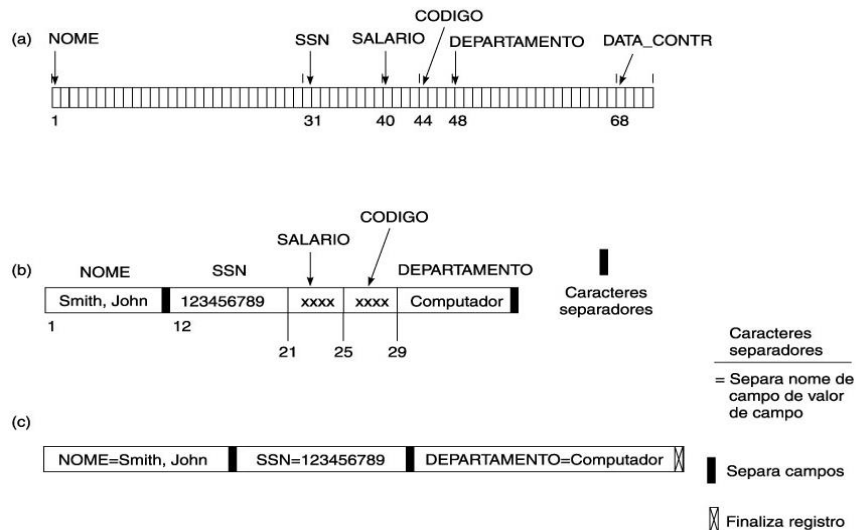
- ▶ Remoção de um registro - o espaço precisa ser reaproveitado por outro registro.
- ▶ Tamanho do bloco normalmente não é múltiplo do tamanho do registro
  - um registro é dividido em dois blocos
  - sobra espaço em um bloco

## ▶ Soluções:

- Introduzir marcador de registro deletado: busca de registros vagos
- Lista de registros vagos

# Sistema de Arquivos

- ▶ Arquivos com registros de tamanho fixo
- ▶ Arquivos com registros de tamanho variável
  - ▶ Aparecem nas seguintes situações:
    - ▶ Armazenamento de diferentes tipos de registros em um arquivo
    - ▶ Campos com tamanho variável
    - ▶ Campos repetidos



- (a) Registro de tamanho fixo com seis campos e tamanho de 71 bytes.
- (b) Um registro com dois campos de tamanho variável e três campos de tamanho fixo.
- (c) Um registro de tamanho variável com três tipos de caracteres separadores.

# Organização de Registros em Arquivos

- ▶ Organização de arquivo pode ser de diversas formas:
  - ▶ HEAP - qualquer registro pode ser colocado em qualquer lugar vago. Um arquivo para cada relação.
  - ▶ SEQUENCIAL - os registros são armazenados em ordem sequencial, com base no valor da chave primária.
  - ▶ CLUSTERING - registros de diferentes relações podem ser agrupados no mesmo bloco/arquivo, de forma que uma operação de E/S seja executada de forma eficiente.

Brighton	A-217	750	
Downtown	A-101	500	
Downtown	A-110	600	
Mianus	A-215	700	
Perryridge	A-102	400	
Perryridge	A-201	900	
Perryridge	A-218	700	
Redwood	A-222	700	
Round Hill	A-305	350	



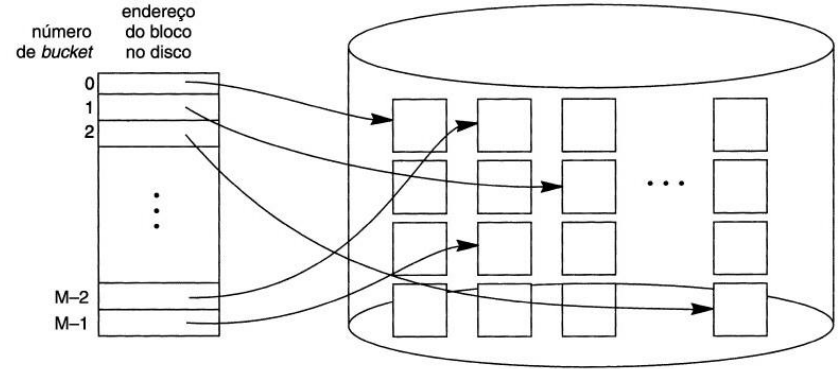
Hayes	Main	Brooklyn
Hayes	A-102	
Hayes	A-220	
Hayes	A-503	
Turner	Putnam	Stamford
Turner	A-305	

(a) Organização Sequencial

(b) Estrutura de clustering

# Organização de Registros em Arquivos

- ▶ **HASHING** - uma função hashing é calculada sobre um atributo. O resultado da função especifica em qual bloco o registro será gravado.



Correspondência entre números de *bucket* e endereços de blocos de disco.

# Armazenamento do Dicionário de Dados

---

- ▶ **Informações necessárias:**
  - ▶ Nome das relações
  - ▶ Nomes dos atributos de cada relação
  - ▶ Domínios dos atributos
  - ▶ Nomes e definições das visões
  - ▶ Restrições de integridade
  - ▶ Usuários autorizados
- ▶ **Além disso muitos sistemas armazenam dados estatísticos sobre as relações:**
  - ▶ Número de tuplas
  - ▶ Método de armazenamento
  - ▶ Nome do índice
  - ▶ Atributos indexados
  - ▶ Tipo do índice

---

▶ Em muitos SGBD, dados do dicionário são armazenados dentro do BD, com os dados.

# Bibliografia

---

- ▶ Lisboa Filho, J. – Notas de aula – Universidade Federal de Viçosa.
- ▶ Elmasri, R.; Navathe, S. – Sistemas de Banco de Dados – 6ª. Edição. São Paulo - 2011