

---

# GBC053–Gerenciamento de Banco de Dados

## Introdução à Organização de Arquivos

### Métodos de Acesso/Índices

### Parte 1 - Panorâmica

Ilmério Reis da Silva

ilmerio arroba ufu.br

*MS Teams:* GBC053.2021.2

UFU/FACOM

---

# *Organização de Arquivos - OBJETIVO*

---

***Entender como os dados podem ser organizados no espaço em disco para alcançar bom desempenho e segurança nas operações com SGBDs***

# *Fundamentos de Organização de Arquivos*

---

## *Características de dados em memória externa*

- *Unidade de acesso é página (ex: 4KB, 8KB)*
- *Custo dominante é I/O, cujo tempo gasto é fixo para qualquer página acessada aleatoriamente*
- *Acesso sequencial diminui tempo pois diminui o número médio de seeks por registro*

# Organização de Arquivos - SUBCAMADAS

---

## *Interação entre subcamadas do SGBD*

### *O Método de acesso a arquivo:*

- *recebe do processador de consultas*
  - ✓ *requisição de registro, índice, ou arquivo*
- *gerencia páginas alocadas para cada arquivo*
- *manipula registros*
- *gerencia espaço disponível nas páginas*
- *solicita páginas ao Gerenciador de buffer pool que:*
  - ✓ *verifica se página está no buffer*
  - ✓ *se necessário, solicita página ao Gerenciador de espaço em disco que*
    - *faz acesso ao disco e transmite/recebe páginas para/do buffer pool.*

# Organização de Arquivos - Definição

---

***Def.: Organização de arquivos é o método de dispor os registros em um arquivo armazenado em disco***

***Exemplo: tuplas do esquema empregado(nome, idade, salario) serão armazenadas em registros de um arquivo em disco podendo ser organizados por:***

- *Opção 1: ordem cronológica de inserção*
- *Opção 2: ordem de nome*
- *Opção 3: ordem de idade*
- *Opção 4: aleatoriamente*
- *Etc.*

*OBS: nome e idade seriam “chaves de busca”.*

---

# *Tipos de Organização de Arquivos*

---

***Arquivo não ordenado (heap file):*** registros distribuídos aleatoriamente nas páginas

***Arquivo ordenado (sorted file):*** registros armazenados em ordem de uma chave de busca

***Arquivo indexado:*** registros armazenados de acordo com uma estrutura de índice

- ***Hash:  $H : \text{KeySet} \rightarrow \text{BucketSet}$***
- ***ÁRVORE***

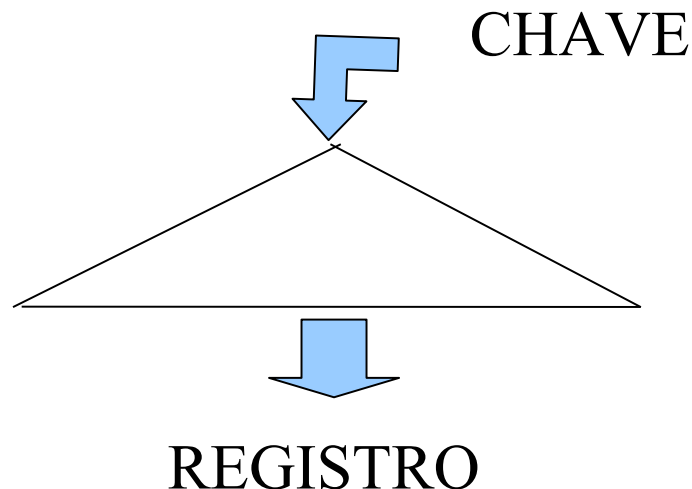
***Índice:*** estrutura auxiliar que auxilia as operações de busca/alteração/remoção de registros em arquivos

# Índices - Características

---

## *Principais Características de Índices*

- *chave de busca, simples ou composta*
- *estrutura facilita busca baseada na chave*
- *facilita atualização e remoção aleatória*
- *múltiplos índices causam sobrecarga na inserção, atualização e remoção*



# Índices – Alternativas de Dados no Índice

---

## Principais Características de Índices

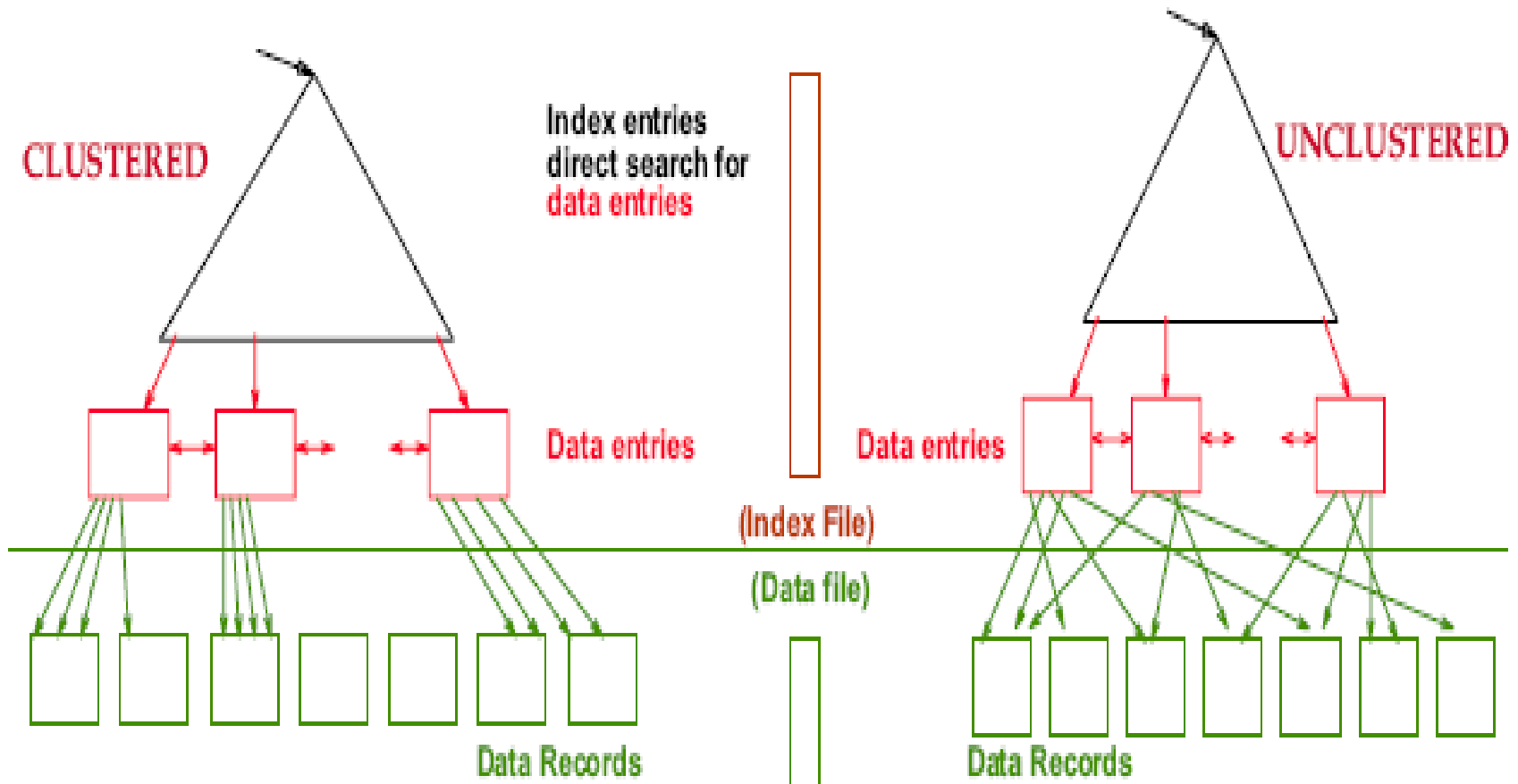
- **ALTERNATIVAS PARA DADOS ASSOCIADOS À CHAVE – “DATA ENTRY”**
  - (1) *<registro>*: este é o arquivo indexado
  - (2) *<k + rid>* : chave mais um identificador de registro - *rid=(pageId, slotId)*
  - (3) *<k + lista\_rids>*: chave mais uma lista de identificadores, usado p/ índices com chaves duplicadas
- *temos no máximo um índice com alternativa (1)*
- *agrupado / não agrupado*
- *denso / esperso*
- *primário / secundário*



# Índices – Agrupado e Não Agrupado

## Principais Características de Índices

- *Índice agrupado e não agrupado, Alternativa 2*



# Índices – Exemplo de índice agrupado

## Principais Características de Índices

- *Exemplo de índice agrupado - Alternativa 2*

2000	●	→	Paulo, 44, 2000
2000	●	→	Pedro, 35, 2000
2000	●	→	Carlos, 44, 2000
2500	●	→	José, 40, 2500
3000	●	→	João, 35, 3000
3500	●	→	Ilmério, 40, 3500
3500	●	→	Rodrigo, 40, 3500
4000	●	→	Maria, 30, 4000
4000	●	→	Sara, 35, 4000
5000	●	→	Sabrina, 31, 5000

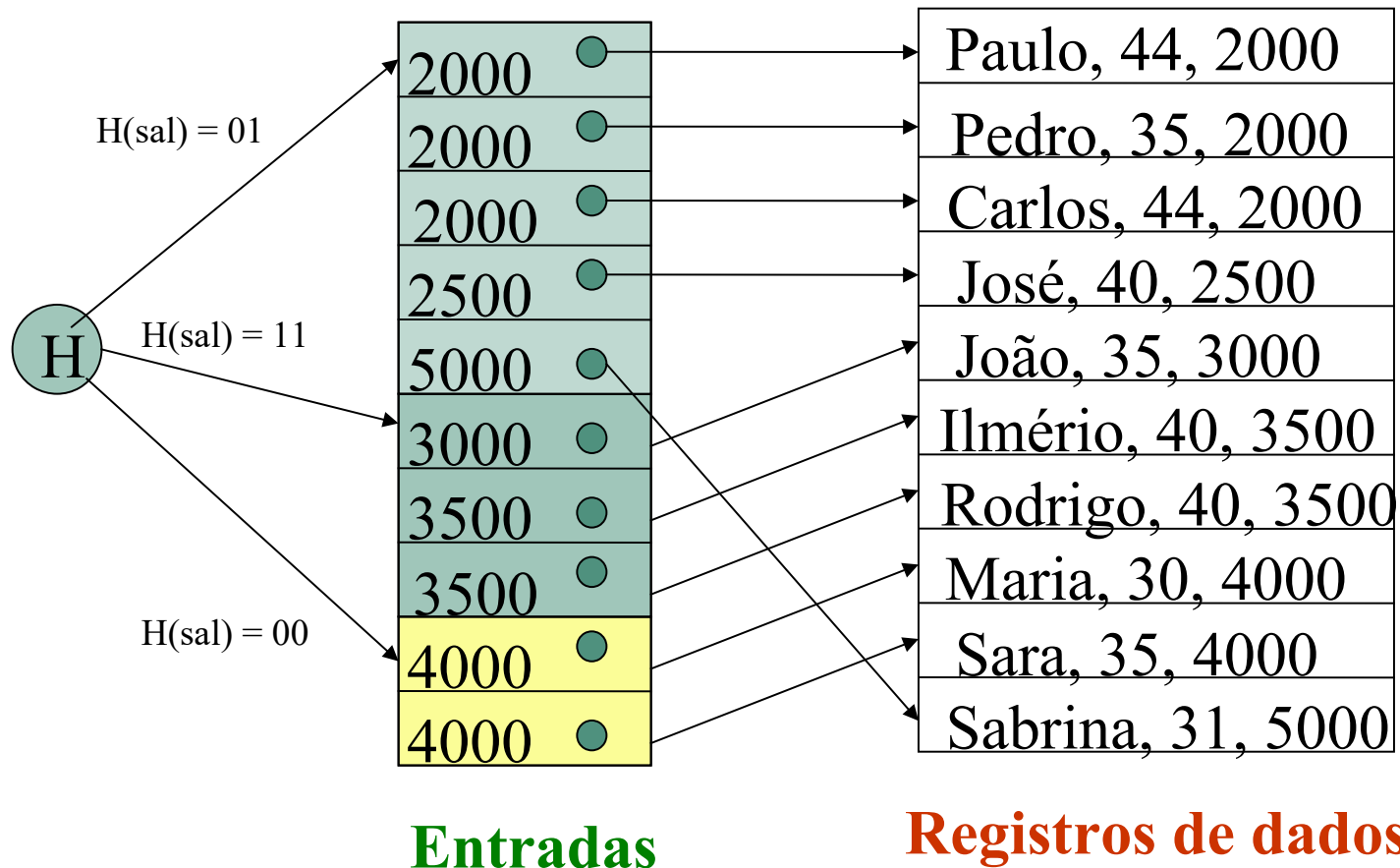
**Entradas**

**Registros de dados**

# Índices – exemplo de índice não agrupado

## Principais Características de Índices

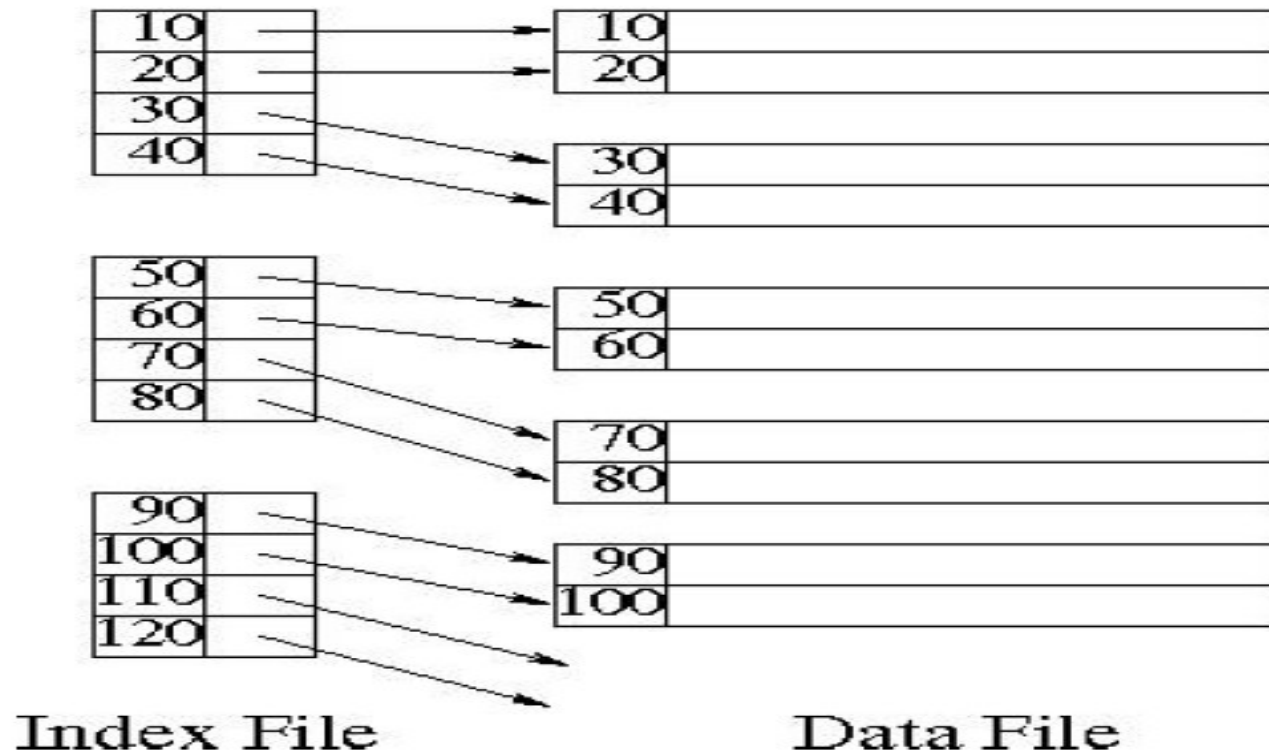
- Exemplo de índice não agrupado - Alternativa 2



# Índices – índice denso

## Principais Características de Índices

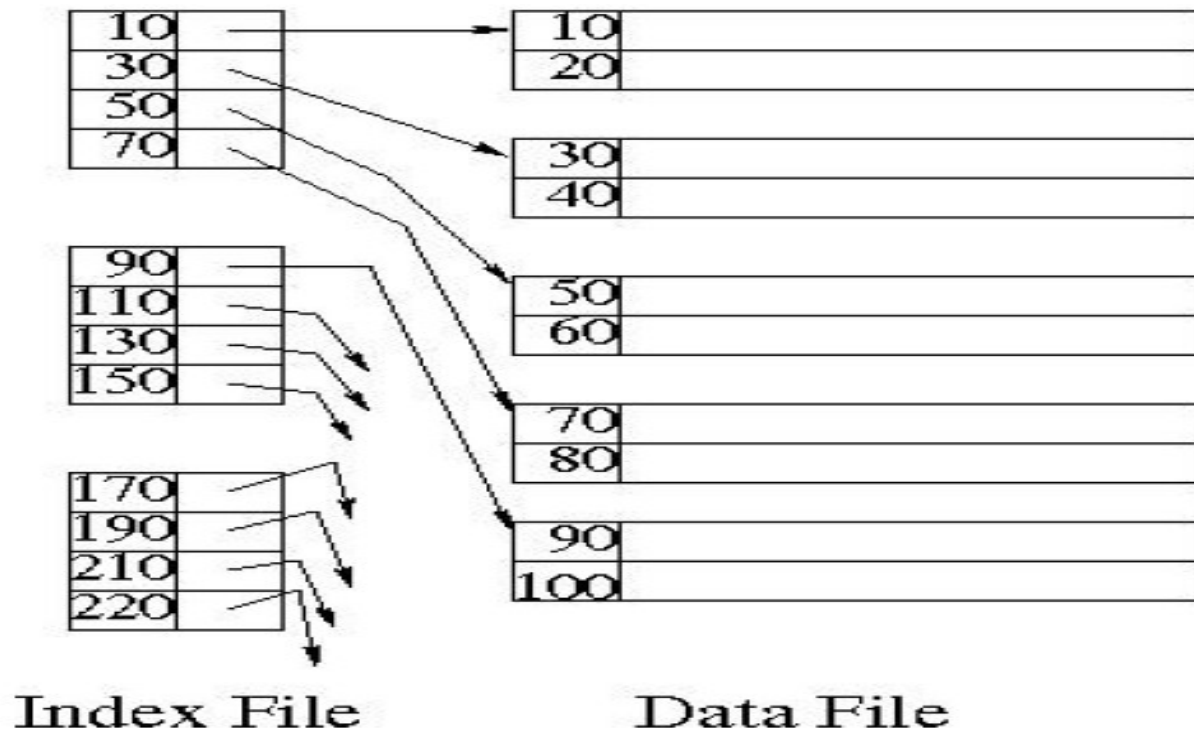
- **Índice denso:** se para cada valor  $k$  da chave de busca existe uma entrada  $(k, rid)$



# Índices – índice esparsos

## Principais Características de Índices

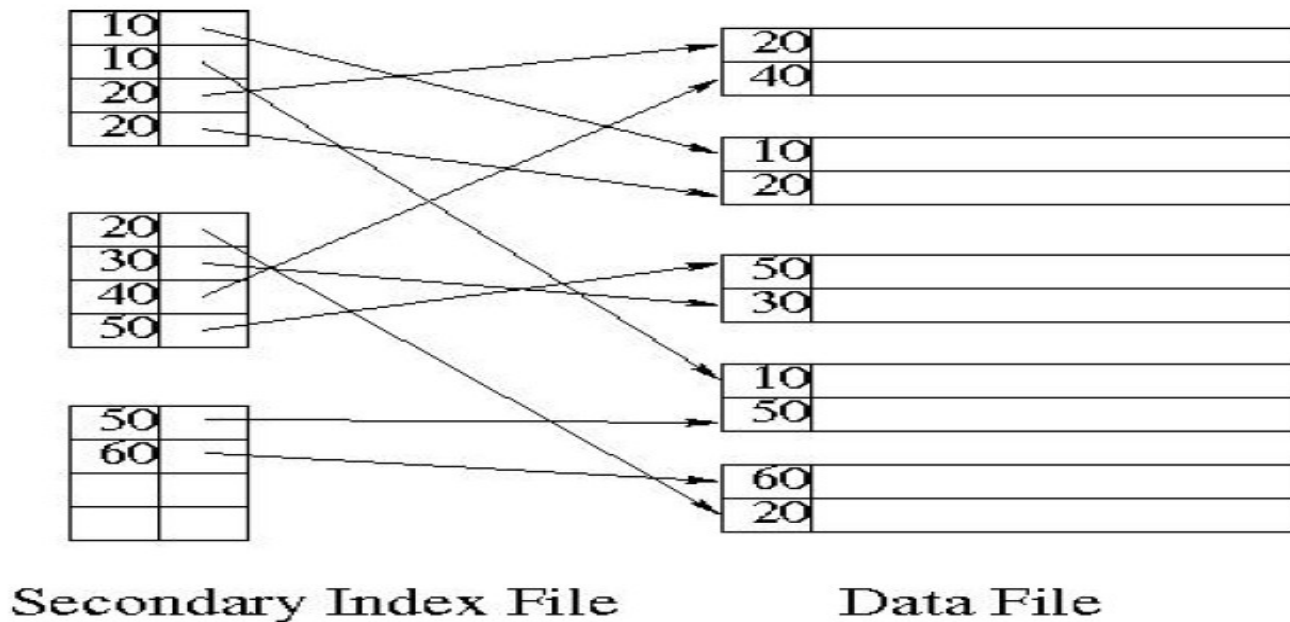
- *Índice esparsos = não denso (tem que ser agrupado)*



# Índices – Secundário

## Principais Características de Índices

- *Índice Primário : chave do índice é chave primária do arquivo*
- *Índice Secundário: chave pode ocorrer mais de uma vez*



---

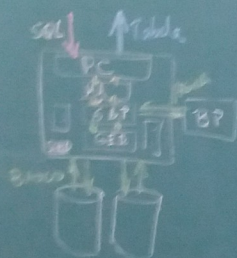
*Organização de Arquivos*  
*Parte 2 – Uma abordagem comparativa*  
*Após estudo das Estruturas de Dados em*  
*Disco*



# RESUMO NO QUADRO



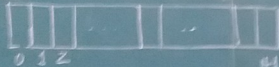
SBD



## ORGANIZAÇÂO DE ARQUIVOS

COMO DISPOR OS DADOS NO ARQUIVO DE DADOS?

### ARQUIVO



### EXEMPLO

EMPREGADO (NOME, SALARIO, DATA\_NASCIMENTO, etc.)

ISRT (JOÃO, 5000, 10/01/81)  
ONDE?

## 1) ARQUIVO NÃO ORDEMADO (RAW FILE)

• Não tem o espaço usado por cada registro

## 2) ARQUIVO ORDEMADO

- POR NOME; ou
- POR SALARIO; ou
- POR DATA\_NASCIMENTO; ou
- POR NOME + DATA\_NASCIMENTO; ou
- POR ... (CÓDIGO FUNÇÃO + DATA\_NASC)
- POR ...

## 3) ARQUIVO INDEXADO

