

Sistemas de Banco de Dados

Fundamentos em Bancos de Dados Relacionais

Wladmir Cardoso Brandão

www.wladmirbrandao.com

Material distribuído sob licença CC BY-NC-ND 4.0

Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International



INDEXAÇÃO DE NÍVEL ÚNICO



Construção de índices para acelerar recuperação de registros de arquivos

ÍNDICE → Caminho alternativo de acesso a registros de um arquivo

PROFESSOR						ÍNDICE	
CPF	Nome	Sexo	Salario	Departamento		Ponteiro	Departamento
12345678900	Roberto Machado	M	1200.00	1	#00	#00	1
21345678900	Carlos A. Martins	M	3200.00	1		#10	2
32145678900	Ana Maria Freitas	F	7500.00	2	#10
...		#FF	N
52345678902	Luiz A. Barbosa	M	5300.00	N	#FF		

- ▶ Arquivo de índice → arquivo adicional ao arquivo de dados (indexado)
- ▶ Contém campos de indexação provenientes do arquivo indexado
- ▶ Registro → campos de indexação + ponteiro para bloco que armazena o registro do arquivo indexado referenciado pelos campos de indexação



ÍNDICE DE NÍVEL ÚNICO → arquivo ordenado pelo campo de indexação

- DENSO → um registro de índice para cada registro no arquivo indexado

PROFESSOR						ÍNDICE	
CPF	Nome	Sexo	Salario	Departamento		Ponteiro	Departamento
12345678900	Roberto Machado	M	1200.00	1	#00	#00	1
21345678900	Carlos A. Martins	M	3200.00	1		#00	1
32145678900	Ana Maria Freitas	F	7500.00	2	#10	#10	2

- ESPARSO → registros de índice para alguns registros no arquivo indexado

PROFESSOR						ÍNDICE	
CPF	Nome	Sexo	Salario	Departamento		Ponteiro	Departamento
12345678900	Roberto Machado	M	1200.00	1	#00	#00	1
21345678900	Carlos A. Martins	M	3200.00	1			
32145678900	Ana Maria Freitas	F	7500.00	2	#10	#10	2



ÍNDICE ESPARSO COM registros de tamanho fixo

PROFESSOR					
CPF	Nome	Sexo	Salario	Departamento	
12345678900	Roberto Machado	M	1200.00	1	#00
12345678901	Manuela Costa	F	2700.00	3	
21345678900	Carlos A. Martins	M	3200.00	1	#10
32145678900	Ana Maria Freitas	F	7500.00	2	
52345678902	Luiz A. Barbosa	M	5300.00	3	#20

ÍNDICE	
Ponteiro	CPF
#00	12345678900
#10	21345678900
#20	52345678902

- ▶ Campo de indexação → referencia chave primária do arquivo indexado
- ▶ Demanda arquivo indexado ordenado pela chave primária
- ▶ Um registro de índice para cada bloco do arquivo indexado
 - ▶ Primeiro registro de cada bloco do arquivo indexado (âncora do bloco) encontra-se referenciado por um registro no arquivo de índice



Para um arquivo indexado de PROFESSOR, com 10.000 registros de tamanho fixo de 185B, ordenado pela chave primária CPF e armazenado em um disco com blocos de 4KB, teremos:

- ▶ Fator de Bloco $\rightarrow F = \left\lfloor \frac{4KB}{185B} \right\rfloor = \left\lfloor \frac{4 \times 1.024B}{185B} \right\rfloor \approx \lfloor 22,14 \rfloor = 22$
- ▶ # Blocos $\rightarrow B = \left\lceil \frac{10.000}{22} \right\rceil \approx \lceil 454,54 \rceil = 455$
- ▶ Espaço $\rightarrow S = 455 \times 4KB = 1.820KB \approx 1,77MB$

Pesquisas nesse arquivo demandarão acessos a blocos de disco:

- ▶ Pela chave primária $\rightarrow A = \lceil \log_2 455 \rceil \approx \lceil 8,83 \rceil = 9$
- ▶ Por outro campo $\rightarrow A = 455$



Para um índice primário criado sobre a chave primária CPF de 11B, onde o ponteiro de bloco ocupa 16B, teremos:

- ▶ Fator de Bloco $\rightarrow F_P = \left\lfloor \frac{4KB}{11B+16B} \right\rfloor = \left\lfloor \frac{4 \times 1.024B}{27B} \right\rfloor \approx \lfloor 151,70 \rfloor = 151$
- ▶ # Blocos $\rightarrow B_P = \left\lceil \frac{455}{151} \right\rceil \approx \lceil 3,01 \rceil = 4$
- ▶ Espaço $\rightarrow S_P = 4 \times 4KB = 16KB$

Pesquisas nesse arquivo só podem ser realizadas pelo campo de indexação e demandarão acessos a blocos de disco:

- ▶ $A_P = \lceil \log_2 4 \rceil = 2$
- ▶ +1 acesso para recuperar o registro no arquivo indexado



Arquivo de índice é significativamente menor que arquivo indexado, ocupando menos blocos em disco

- ▶ Esparsidade → menos registros no arquivo de índice
- ▶ Tamanho → registros de índice menores que registros do arquivo indexado
- ▶ Ordenação → arquivo ordenado com pesquisa de complexidade logarítmica
 - ▶ Pesquisa binária no arquivo de índice para encontrar a chave procurada
 - ▶ Acesso direto ao registro de dados através do endereço de bloco recuperado a partir do índice

Operações de atualização no arquivo indexado podem envolver reordenação de registros do próprio arquivo indexado e do arquivo de índice



Ou **ÍNDICE CLUSTERING**, é um **ÍNDICE ESPARSO** com registros de tamanho fixo

PROFESSOR					
CPF	Nome	Sexo	Salario	Departamento	
12345678900	Roberto Machado	M	1200.00	1	#00
21345678900	Carlos A. Martins	M	3200.00	1	
32145678900	Ana Maria Freitas	F	7500.00	2	#10
12345678901	Manuela Costa	F	2700.00	3	
52345678902	Luiz A. Barbosa	M	5300.00	3	#20

ÍNDICE	
Ponteiro	Departamento
#00	1
#10	2
#10	3

- ▶ Campo de indexação → referencia campo de agrupamento (chave não exclusiva) do arquivo indexado
- ▶ Demanda arquivo indexado ordenado pelo campo de agrupamento
- ▶ Um registro de índice para cada valor distinto no campo de agrupamento
 - ▶ Bloco de primeira ocorrência de um valor distinto no campo de agrupamento encontra-se referenciado por um registro no arquivo de índice



No arquivo indexado de PROFESSOR, para um índice de agrupamento criado sobre o campo DEPARTAMENTO de 8B, contendo 200 valores distintos, onde o ponteiro de bloco ocupa 16B, teremos:

- ▶ Fator de Bloco $\rightarrow F_C = \left\lfloor \frac{4KB}{8B+16B} \right\rfloor = \left\lfloor \frac{4 \times 1.024B}{24B} \right\rfloor \approx \lfloor 170,66 \rfloor = 170$
- ▶ # Blocos $\rightarrow B_C = \left\lceil \frac{200}{170} \right\rceil \approx \lceil 1,17 \rceil = 2$
- ▶ Espaço $\rightarrow S_C = 2 \times 4KB = 8KB$

Pesquisas nesse arquivo só podem ser realizadas pelo campo de indexação e demandarão acessos a blocos de disco:

- ▶ $A_C = \lceil \log_2 2 \rceil = 1$
- ▶ +1 acesso para recuperar o registro no arquivo indexado



ÍNDICE DENSO com registros de tamanho fixo

PROFESSOR						ÍNDICE	
CPF	Nome	Sexo	Salario	Departamento		Ponteiro	Departamento
12345678900	Roberto Machado	M	1200.00	1	#00	#00	1
12345678901	Manuela Costa	F	2700.00	3		#10	1
21345678900	Carlos A. Martins	M	3200.00	1	#10	#10	2
32145678900	Ana Maria Freitas	F	7500.00	2		#00	3
52345678902	Luiz A. Barbosa	M	5300.00	3	#20	#20	3

- ▶ Campo de indexação → referencia campo não ordenado do arquivo indexado
- ▶ Não demanda ordenação no arquivo indexado
- ▶ Um registro de índice para cada registro do arquivo indexado
 - ▶ Bloco de cada registro do arquivo indexado encontra-se referenciado por um registro no arquivo de índice



No arquivo indexado de PROFESSOR, para um índice secundário criado sobre o campo DEPARTAMENTO de 8B, onde o ponteiro de bloco ocupa 16B, teremos:

- ▶ Fator de Bloco $\rightarrow F_S = \left\lfloor \frac{4KB}{8B+16B} \right\rfloor = \left\lfloor \frac{4 \times 1.024B}{24B} \right\rfloor \approx \lfloor 170,66 \rfloor = 170$
- ▶ # Blocos $\rightarrow B_S = \left\lceil \frac{10.000}{170} \right\rceil \approx \lceil 58,82 \rceil = 59$
- ▶ Espaço $\rightarrow S_S = 59 \times 4KB = 236KB$

Pesquisas nesse arquivo só podem ser realizadas pelo campo de indexação e demandarão acessos a blocos de disco:

- ▶ $A_S = \lceil \log_2 59 \rceil \approx \lceil 5,88 \rceil = 6$
- ▶ +1 acesso para recuperar o registro no arquivo indexado



Arquivo de índice tipicamente menor que arquivo indexado

- ▶ Tamanho → registros de índice menores que registros do arquivo indexado
- ▶ Ordenação → arquivo ordenado com pesquisa binária
- ▶ Flexibilidade → múltiplos índices sobre um mesmo arquivo
 - ▶ Não demandam arquivo indexado ordenado
 - ▶ Somente um índice primário ou de agrupamento por arquivo indexado, por demandarem ordenação
- ▶ Desempenho → proporcionalmente, maior ganho em tempo de pesquisa
 - ▶ Para índices primário ou de agrupamento, tem-se a opção de pesquisa binária tanto no arquivo indexado, quanto no arquivo índice



Comparativo entre diferentes tipos de índices de nível único

Tipo	Esparsidade	Arquivo Indexado	Quantidade Registros
Primário	Esparso	Ordenado	Número de blocos do arquivo indexado
Agrupamento	Esparso	Ordenado	Número de valores distintos no campo de agrupamento
Secundário	Denso	Qualquer	Número de registros do arquivo indexado



- [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.