



Outros tipos de índice

Leandro C. Cintra

M.C.F. de Oliveira

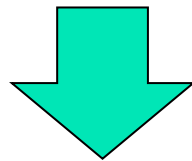
Thiago A. S. Pardo

Cristina D. A. Ciferri



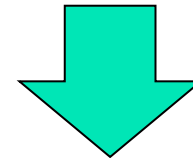
Pesquisa

Quais os dados do CD de
código **DG18807** ?



geralmente usado
internamente por um
programa, mas
raramente pelo usuário
de forma direta

Quais os dados do CD
Symphony No. 9 de
Beethoven ?



consulta típica de um
usuário



Acesso por múltiplas chaves

- Acesso a registros se faz por **chaves secundárias**
 - e não por **chaves primárias**
- Solução
 - cria-se um índice que relaciona uma chave secundária à chave primária (e não diretamente ao registro)
 - **índice secundário**



Acesso por múltiplas chaves

- Passo 1
 - pesquisar o **índice de chave secundária** para encontrar a chave primária relacionada
- Passo 2
 - usar a chave primária para pesquisar o **índice de chave primária** para encontrar o *byte offset* (ou RRN) do registro no arquivo de dados
- Passo 3
 - recuperar o registro desejado



Acesso por múltiplas chaves

Beethoven	ANG3795
Beethoven	DG139201
Beethoven	DG18807
Beethoven	RCA2626
Corea	WAR23699
Dvorak	COL31809
Prokofiev	LON2312
Rimsky	MER75016
Springsteen	COL38358
Sweet Honey In The	FF245

arquivo de índice secundário

ANG3795	167
COL31809	353
COL38358	211
DG139201	396
DG18807	256
FF245	442
LON2312	32
MER75016	300
RCA2626	77
WAR23699	132

arquivo de índice primário

Acesso por múltiplas chaves

Composer index	
<i>Secondary key</i>	<i>Primary key</i>
BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRINGSTEEN	COL38358
SWEET HONEY IN THE R	FF245

arquivo de índice secundário 1
- índice de compositor -

Title index	
<i>Secondary key</i>	<i>Primary key</i>
COQ D'OR SUITE	MER75016
GOOD NEWS	FF245
NEBRASKA	COL38358
QUARTET IN C SHARP M	RCA2626
ROMEO AND JULIET	LON2312
SYMPHONY NO. 9	ANG3795
SYMPHONY NO. 9	COL31809
SYMPHONY NO. 9	DG18807
TOUCHSTONE	WAR23699
VIOLIN CONCERTO	DG139201

arquivo de índice secundário 2
- índice de título -



Acesso por múltiplas chaves

- Índices permitem muito mais do que simplesmente melhorar o tempo de busca por um registro
- Múltiplos índices secundários
 - permitem manter diferentes visões dos registros em um arquivo de dados
 - permitem combinar chaves associadas e, desse modo, fazer buscas que combinam visões particulares



Inserção

- Inserção de um registro no arquivo de dados
 - inserção da entrada correspondente no arquivo de índice primário
 - inserção da entrada correspondente em cada arquivo de índice secundário
- Campo chave
 - deve ser armazenado em sua forma canônica no índice secundário
- Duplicação de chave no índice secundário
 - chaves duplicadas devem ser mantidas agrupadas e ordenadas



Remoção

delete all references

- Remoção de um registro no arquivo de dados
 - remoção da entrada correspondente no arquivo de índice primário
 - remoção da entrada correspondente em cada arquivo de índice secundário
- Processo altamente custoso
 - o custo cresce com o aumento do número de índices secundários
 - índices são mantidos ordenados
 - remoção física requerendo deslocamento dos registros



Remoção

delete some references

- Remoção de um registro no arquivo de dados
 - remoção da entrada correspondente no arquivo de índice primário
- Características
 - manutenção da entrada correspondente em cada arquivo de índice secundário
 - pesquisa no índice de chave primária acusará que o registro foi removido



Remoção

delete some references

- Vantagem

- economia de tempo quando existem vários índices secundários, principalmente se eles são mantidos em disco

- Desvantagens

- crescimento do tamanho dos índices secundários com o decorrer das remoções
- necessidade de reorganização periódica dos índices secundários
- queda de desempenho na pesquisa devido à existência de falsos candidatos



Atualização

- Alteração do valor da chave secundária
 - reordenação do índice secundário
- Alteração do valor da chave primária
 - reordenação do índice primário
 - atualização dos índices secundários
 - reordenação dos índices secundários se houver repetição da chave secundária
- Alteração dos demais campos
 - não afeta nenhum dos índices

e se o tamanho
do registro
mudar?



Busca usando múltiplas chaves

- Uma das aplicações mais importantes das chaves secundárias é localizar conjuntos de registros do arquivo de dados usando uma ou mais chaves
- Pode-se fazer uma **busca em vários índices e combinar (AND,OR,NOT) os resultados**
- Exemplo: encontre todos os registros de dados tal que
 - **composer = "BEETHOVEN" AND title = "SYMPHONY NO. 9"**



Busca usando múltiplas chaves

Composer index	
<i>Secondary key</i>	<i>Primary key</i>
BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRINGSTEEN	COL38358
SWEET HONEY IN THE R	FF245

Title index	
<i>Secondary key</i>	<i>Primary key</i>
COQ D'OR SUITE	MER75016
GOOD NEWS	FF245
NEBRASKA	COL38358
QUARTET IN C SHARP M	RCA2626
ROMEO AND JULIET	LON2312
SYMPHONY NO. 9	ANG3795
SYMPHONY NO. 9	COL31809
SYMPHONY NO. 9	DG18807
TOUCHSTONE	WAR23699
VIOLIN CONCERTO	DG139201



Busca usando múltiplas chaves

Composer index	
<i>Secondary key</i>	<i>Primary key</i>
BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRINGSTEEN	COL38358
SWEET HONEY IN THE R	FF245

Title index	
<i>Secondary key</i>	<i>Primary key</i>
COQ D'OR SUITE	MER75016
GOOD NEWS	FF245
NEBRASKA	COL38358
QUARTET IN C SHARP M	RCA2626
ROMEO AND JULIET	LON2312
SYMPHONY NO. 9	ANG3795
SYMPHONY NO. 9	COL31809
SYMPHONY NO. 9	DG18807
TOUCHSTONE	WAR23699
VIOLIN CONCERTO	DG139201



Busca usando múltiples claves

Composer index	
<i>Secondary key</i>	<i>Primary key</i>
BEETHOVEN	ANG3795
BEETHOVEN	DG139201
BEETHOVEN	DG18807
BEETHOVEN	RCA2626
COREA	WAR23699
DVORAK	COL31809
PROKOFIEV	LON2312
RIMSKY-KORSAKOV	MER75016
SPRINGSTEEN	COL38358
SWEET HONEY IN THE R	FF245

Title index	
<i>Secondary key</i>	<i>Primary key</i>
COQ D'OR SUITE	MER75016
GOOD NEWS	FF245
NEBRASKA	COL38358
QUARTET IN C SHARP M	RCA2626
ROMEO AND JULIET	LON2312
SYMPHONY NO. 9	ANG3795
SYMPHONY NO. 9	COL31809
SYMPHONY NO. 9	DG18807
TOUCHSTONE	WAR23699
VIOLIN CONCERTO	DG139201

Busca usando múltiplas chaves

Composer index		Title index	
Secondary key	Primary key	Secondary key	Primary key
BEETHOVEN	ANG3795	COQ D'OR SUITE	MER75016
BEETHOVEN	DG139201	GOOD NEWS	FF245
BEETHOVEN	DG18807	NEBRASKA	COL38358
BEETHOVEN		M	RCA2626
COREA		LON2312	
DVORAK		ANG3795	
PROKOFIEV		COL31809	
RIMSKY-KORSAK		DG18807	
SPRINGSTEEN	COL38358	TOUCHSTONE	WAR23699
SWEET HONEY IN THE R	FF245	VIOLIN CONCERTO	DG139201

Matching: tira vantagem da ordenação das chaves associadas a uma chave secundária
→ algoritmos importantes para isso