

Exercícios

Considere o arquivo de dados a seguir, *'entry-sequenced'*, com registros de tamanho variável. Cada registro é descrito por 3 campos, separados pelo marcador '|':

Chave primária Título Autor

Obs. antes de cada registro informo o seu byte offset e após cada registro informo o número de bytes: essas informações (em vermelho) são dadas para facilitar, mas elas não fazem parte do arquivo de dados.

Arquivo de Dados (inicial)

...

100	L21 The Hunger Games Suzanne Collins	36
137	A05 The Diary of a Young Girl Anne Frank	40
178	L36 Little Women Louisa May Alcott	34
213	F80 The Fault in Our Stars John Green	37
251	L06 The Lightning Thief Rick Riordan	36
288	A75 The Giver Lois Lowry	24
313	C89 Let it Snow John Green	26
340	F04 City of Bones Cassandra Clare	33
374	L94 Paper Towns John Green	26
401	F01 A Wrinkle in Time Madeleine L'Engle	39
440	L49 An Old Fashioned Girl Louisa May Alcott	43
481	F11 Divergent Veronica Roth	27
509	A12 Clockwork Angel Cassandra Clare	35
545	M13 Vampire Academy Richelle Mead	36
579	L14 Looking for Alaska John Green	33
613	C49 Will's Wonder Book Louisa May Alcott	40

Pede-se:

1. Esquematize o conteúdo do índice primário.
2. Esquematize o conteúdo do índice secundário por autor, supondo que está sendo adotada a estratégia de acoplamento fraco
3. Esquematize o índice secundário por autor, supondo que é mantido como um arquivo de lista invertida
4. Mostre a configuração do arquivo de dados e dos índices primário e secundário (ambas as versões) após a inserção dos seguintes registros, nessa ordem.

M99 The Transfer Veronica Roth	31
C58 The Initiate Veronica Roth	31
A74 Norwegian Wood Haruki Murakami	35

5. Mostre a configuração do arquivo de dados e dos índices primário e secundário (ambas as versões) após a remoção dos registros de chave C89, L94 e C49 nesta ordem. Suponha que está sendo adotada uma estratégia de remoção lógica e a estratégia 'delete some references'.

6. Mostre a configuração do arquivo de dados após a remoção dos registros de chaves C89, L94 e C49 nesta sequência, supondo que está sendo usada uma estratégia de remoção lógica para reaproveitamento dos *slots* vagos pelos registros, na qual será utilizada a estratégia *first-fit* de reaproveitamento.

7. Suponha que em seguida às remoções do item 6 acima, é inserido o registro abaixo.

C43|Kafka on the Shore|Haruki Murakami 38

Mostre como ficam o arquivo de dados e os índices primário e secundário (nas duas versões).

8. Refaça os itens 6 e 7 assumindo que a estratégia de reutilização adotada é a *best-fit*.

9. Refaça os itens 6 e 7 assumindo que a estratégia de reutilização adotada é a *worst-fit*.

10. Repita os itens 1 a 7 supondo agora que o arquivo é *entry sequenced*, mas organizado em registros de tamanho fixo (45 bytes), e campos de tamanho variável (separados pelo delimitador '|'). Antes de cada registro informe o seu RRN (essa informação não é parte do conteúdo do arquivo).

Arquivo de Dados com registros de tamanho fixo (inicial)

...

01	L21 The Hunger Games Suzanne Collins-----
02	A05 The Diary of a Young Girl Anne Frank----
03	L36 Little Women Louisa May Alcott-----
04	F80 The Fault in Our Stars John Green-----
05	L06 The Lightning Thief Rick Riordan-----
06	A75 The Giver Lois Lowry-----
07	C89 Let it Snow John Green-----
08	F04 City of Bones Cassandra Clare-----
09	L94 Paper Towns John Green -----
10	F01 A Wrinkle in Time Madeleine L'Engle----
11	L49 An Old Fashioned Girl Louisa May Alcott--
12	F11 Divergent Veronica Roth-----
13	A12 Clockwork Angel Cassandra Clare-----
14	M13 Vampire Academy Richelle Mead-----
15	L14 Looking for Alaska John Green-----
16	C49 Will's Wonder Book Louisa May Alcott----

Suponha um arquivo com 5.000.000 de registros, em que os registros têm tamanho fixo de 64 bytes. Ele está armazenado em um HD (hipotético) com as seguintes características:

- ◆ n° de bytes por setor: 1.024
- ◆ n° de setores por trilha: 32
- ◆ n° de trilhas por cilindro: 12
- ◆ n° de cilindros: 1.332
- ◆ tempo de rotação: 6.000rpm
- ◆ latência (*delay* de rotação): 1ms
- ◆ tempo de busca (*seek*): 8ms

11. Considerando o disco e o arquivo acima:

- Qual a capacidade nominal do disco?
- Qual o tamanho total do arquivo?
- Quantos cilindros o arquivo ocupa (supondo que há espaço livre suficiente para ele ser armazenado em um único *extent*)?
- Nessa situação, qual o tempo necessário para transferir todo o conteúdo do arquivo (acesso sequencial)?
- Porque é interessante manter o arquivo em um único *extent*, i.e., uma única área contígua?

12. Supondo que o disco opera com páginas de 8.192 bytes (8K) ao fazer a alocação dos arquivos, e que não há espaço livre suficiente para armazenar o arquivo em um único *extent*.

- Quantas páginas esse arquivo ocupa?
- Nessa situação, qual o tempo necessário para transferir todo o conteúdo do arquivo (acesso sequencial)?