Acesso a Arquivos



Cristina D. A. Ciferri

Thiago A. S. Pardo Leandro C. Cintra

M.C.F. de Oliveira

Moacir Ponti Jr.



Formas de Acesso

- Busca sequencial
 - lê o arquivo registro a registro
- Acesso direto
 - realiza um seek direto para o início do registro desejado e lê o registro imediatamente

Desempenho da pesquisa (busca) em disco: número de acessos a disco

Custa da Busca Sequencial - Pior Caso -

Custo sem considerar páginas de disco

 $C_{busca_sequencial} = n$ onde n é o tamanho do arquivo

Custo considerando páginas de disco

C_{busca_sequencial} = b onde b é o número de páginas de disco que contêm os registros



Busca Sequencial

- Vantagens
 - fácil de programar
 - requer estruturas de arquivos simples
 - pode ser aplicada a qualquer arquivo
- Desvantagem
 - pode ser ineficiente



Busca Sequencial

- Exemplos de utilidade
 - busca por uma cadeia específica em um arquivo ASCII
 - busca em arquivos com poucos registros
 - até 10 registros
 - busca em arquivos pouco pesquisados
 - arquivos armazenados em armazenamento terciário

Custo do Acesso Direto

 Um único acesso traz o registro, independentemente do tamanho do arquivo

$$C_{acesso_direto} = 1$$

Formas de Prover o Acesso Direto

- Uso do RRN
 - para registros de tamanho fixo
- Uso de um arquivo índice separado
 - obrigatório para registros de tamanho variável
 - também pode ser utilizado para registros de tamanho fixo



- RRN (relative record number)
 - usado para registros de tamanho fixo
 - fornece a posição relativa de cada registro dentro do arquivo
 - permite calcular o byte offset no qual cada registro começa

byte offset = RRN x tamanho do registro



Registros de Tamanho Fixo com Campos de Tamanho Fixo

- Exemplo
 - Registros de tamanho fixo
 - tamanho de 46 bytes
 - Campos de tamanho fixo
 - nome: string de 12 caracteres (12 bytes)
 - rua: string de 10 caracteres (10 bytes)
 - número: inteiro (4 bytes)
 - cidade: string de 20 caracteres (20 bytes)

Registros de Tamanho Fixo com Campos de Tamanho Fixo

_			_		_		_																						
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
M	Α	R	Ι	Α	b	b	b	b	b	b	b	R	U	Α	b	1	b	b	b	b	b		12	23		S	Α	0	b
	_ ` `																										- 1		
30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59
	Α	R	1	0	S	Ь	b	b	b	b	b	b	b	Ь	b	J	0	Α	0	b	b	b	b	b	b	b	b	R	U
	, ,																	<u> </u>											oxdot
60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89
Α	b	Α	b	b	b	Ь	b		25	55		R	Ι	0	b	С	L	Α	R	0	b	b	b	b	b	b	b	b	Ь
	D		D				D			,,		11	*				_		_ '\		D	D	D		D	D	D		
90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119
Ь	b	Р	Е	D	R	0	b	b	b	b	b	b	h	R	U	Α	b	Х	٧	Ь	b	b	b			6		S	Α
	D	Г	_				D	D		D	D	D		'\	U			^	v		D	D	D		J	U		٦	$ \cap $
120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149
	b		Λ	Ъ			(h	6	h	h	h	<i>h</i>	<i>h</i>	<i>h</i>	<i>h</i>	h												
0	b	С	Α	R	L	0	S	b	b	b	b	b	b	b	b	b	b												
$\overline{}$					L														L										ш

registro	RRN	Byte Offset
primeiro (MARIA)	0	$0 \times 46 = 0$
segundo (JOAO)	1	1 x 46 = 46
terceiro (PEDRO)	2	$2 \times 46 = 92$





- Organização do arquivo
 - registros de tamanho fixo
 - registros de tamanho variável
- Acesso a arquivo
 - busca sequencial
 - acesso direto

ambos aspectos devem ser considerados



- Vantagem
 - RRN pode ser usado para prover acesso direto aos registros
- Problema
 - escolha do tamanho dos registros
 - dependente do tamanho dos campos

Escolha: Registros de Tamanho Variável

- Principal resultado da análise
 - registros têm tamanhos muito diferentes
- Problemas
 - como acessar esses registros diretamente?
 - como manter o acesso direto a esses registros eficientemente?
 - como superar as limitações das linguagens de programação?