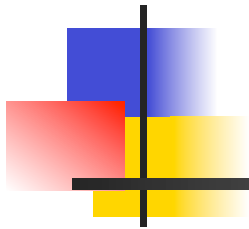


Acesso a Arquivos



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

$$C_{\text{busca_sequencial}} = n$$

onde n é o tamanho do arquivo

- Custo considerando páginas de disco

$$C_{\text{busca_sequencial}} = b$$

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

$$C_{\text{acesso_direto}} = 1$$



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

- 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

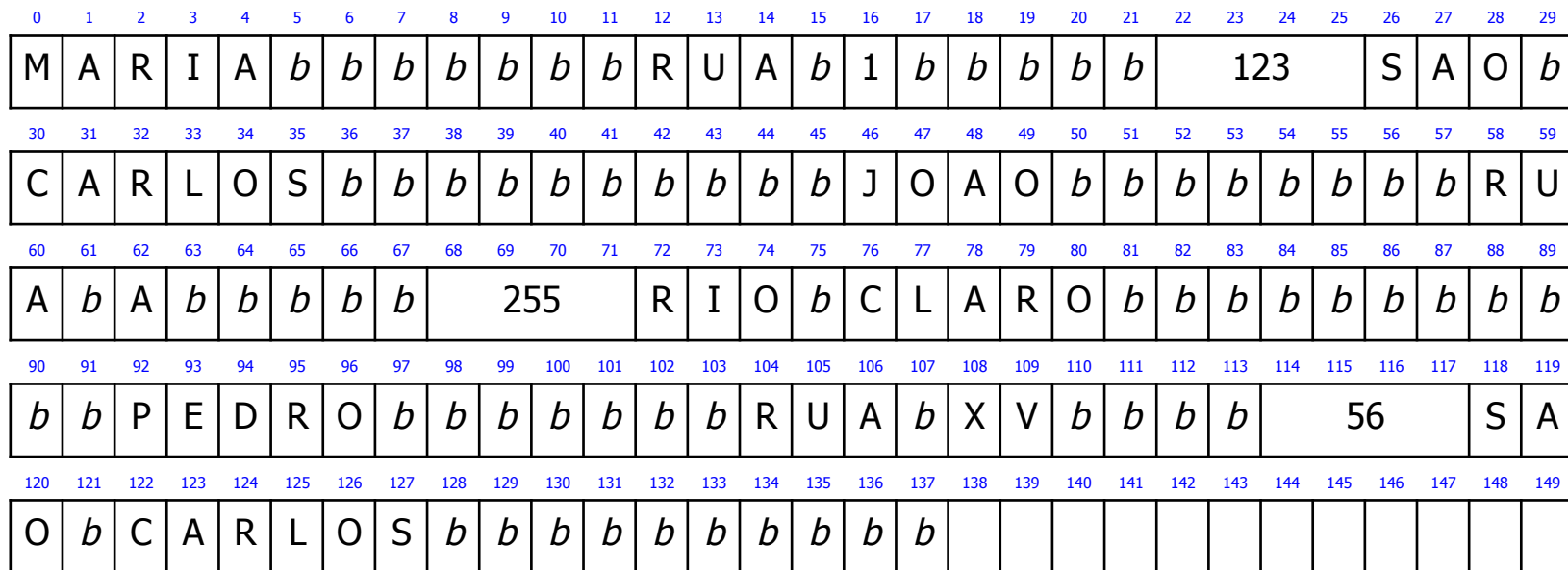
byte offset = RRN x tamanho do registro



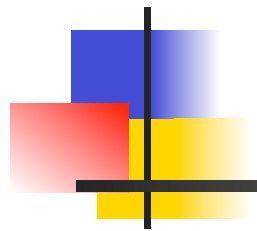
Registros de Tamanho Fixo com Campos de Tamanho Fixo

- 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



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



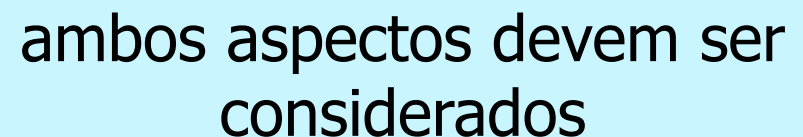
Discussões



Como Escolher a Organização de um Arquivo

- Organização do arquivo
 - registros de tamanho fixo
 - registros de tamanho variável

- Acesso a arquivo
 - busca sequencial
 - acesso direto



ambos aspectos devem ser considerados



Escolha: Registros de Tamanho Fixo

- 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**?