

- TENTATIVA DE REMOÇÃO:
 Busca da folha: $\log_F 2B$
 como não encontra o registro não há mais o que fazer
 logo o custo final será: $\log_F 2B$ operações de E/S
- BUSCA POR INTERVALO:
 Busca da folha de início do intervalo: $\log_F 2B$
 Varredura do intervalo: $\frac{2M}{R}$
 logo o custo final será: $\log_F 2B + \frac{2M}{R}$ operações de E/S

- HEAP + HASH Alternativa 2

- REMOÇÃO:
 Busca no bucket: 1
 Busca no HEAP: 1
 Regravação do HEAP: 1
 Regravação do bucket: 1
 logo o custo final será: 4 operações de E/S
- TENTATIVA DE REMOÇÃO:
 Busca no bucket: 1
 logo o custo final será: 1 operações de E/S
- BUSCA POR INTERVALO:
 Varredura do HEAP: B
 logo o custo final será: B operações de E/S
 uma alternativa considerando chave inteira e que o intervalo $[x, y]$ é pequeno
 seria fazer a busca no bucket e acesso ao HEAP por cada chave do intervalo,
 resultando em um custo aproximado de $2 \times (y - x + 1)$

5. (5 pontos) Considere um arquivo com os seguintes dados:

- registros de tamanho variável ocupando em média 100 bytes cada;
- 2.000.000 (dois milhões) de registros;
- registros armazenados em páginas de 8.192 bytes;
- formato de página conforme Figura 1.

Seja um disco com 512 bytes por setor, 100 setores por trilha, 2.000 trilhas por superfície, 5 pratos de dupla-superfície, tempo médio de *seek* de 10ms e velocidade de rotação dos pratos do disco igual a 10.000rpm.

Pede-se:

- (a) Quantos registros cabem em cada página?

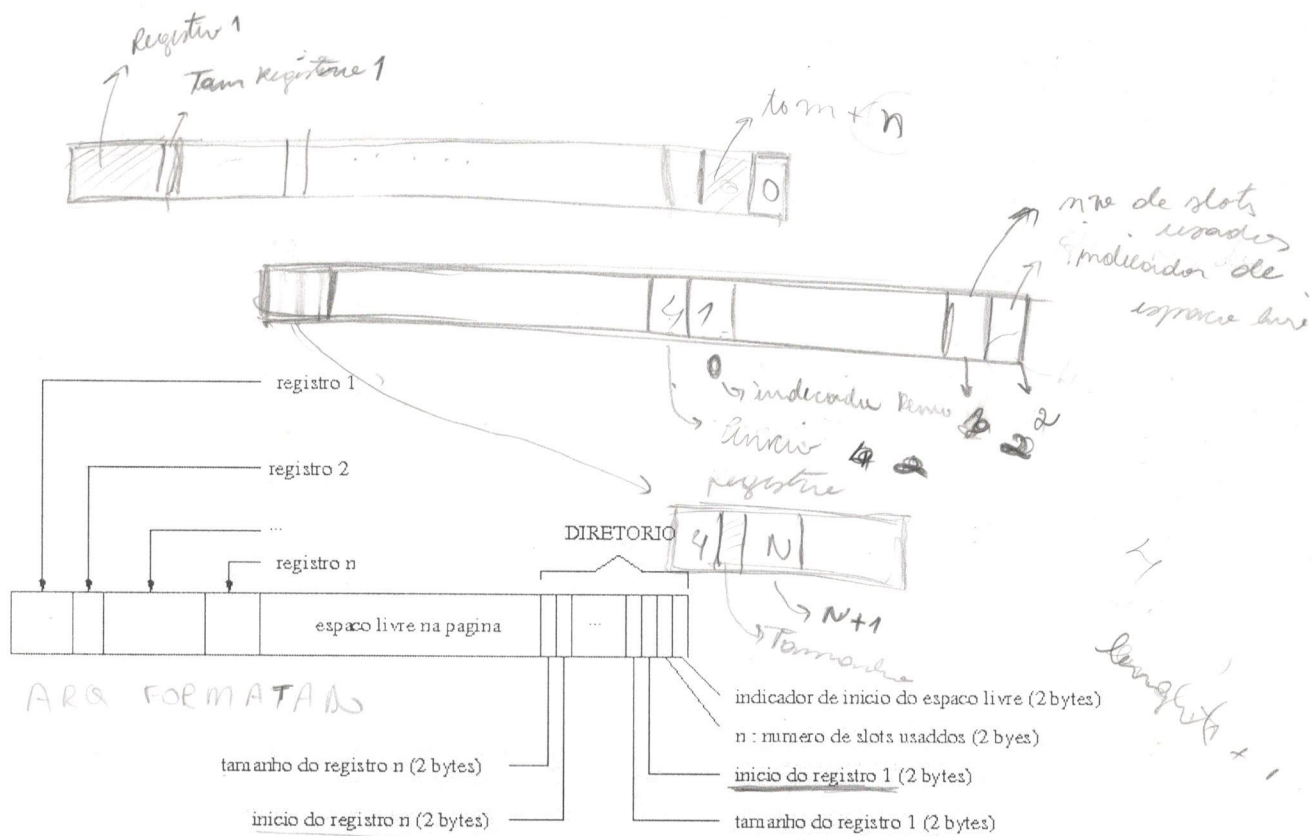


Figura 1: Página para armazenamento de registros de tamanho variável com diretório contendo: indicador do início do espaço livre na página (2 bytes); número de slots usados na página (2 bytes); início do slot com registro na página (2 bytes por slot); e tamanho do registro no slot (2 bytes por slot, sendo que um valor negativo indica que o registro foi removido da página)

$$N(100 + 4) + 2 + 2 \leq 8192$$

$$N \leq 78,73$$

cabem 78 registros em cada página

- (b) Quanto tempo é necessário para ler sequencialmente todo o arquivo considerando que os dados foram distribuídos sequencialmente na mesma trilha e depois no mesmo cilindro?

$$NroPag = \lceil \frac{2.000.000}{78} \rceil = 25.642$$

$$NroPag/Trilhas = \frac{51200}{8.192} = 6,25$$

$$NroTrilhas/Arquivo = \lceil \frac{25.642}{6,25} \rceil = 4.103$$

$$NroCilindros/Arquivo = 411$$

$$TempoSequencial = 411 \times 10 + 4.103 \times 3 + 25.642 \times \frac{6}{6,25} = 41035ms$$

- (c) Recalcule o tempo acima, considerando que todas as cabeças de leitura (uma por superfície) realizam leitura em paralelo e que os dados são alocados de forma ótima.

$$TempoParalelo = 411 \times (10 + 3) + \frac{25.642}{10} \times \frac{6}{6,25} = 7.804ms$$

- (d) Qual o tempo de leitura de todos os registros do arquivo, considerando acesso aleatório, registro por registro?

$$TempoAleatorio = 2.000.000(10 + 3 + \frac{6}{6,25}) = 27.920s = 7h45m$$