

Banco de Dados – Atividade 13 – Organização de Arquivos e Hashing
Prof. Dr. Aparecido V. de Freitas

1. Um arquivo tem 20.000 registros de ALUNO de tamanho fixo. Cada registro tem os seguintes campos:

- ✓ Nome (30 bytes)
- ✓ CPF (9 bytes)
- ✓ Endereço (40 bytes)
- ✓ Telefone (9 bytes)
- ✓ DataNascimento (8 bytes)
- ✓ Sexo (1 byte)
- ✓ DepPrinc (4 bytes)
- ✓ DepSec (4 bytes)
- ✓ TipoAluno (4 bytes, INTEGER)
- ✓ TituloAcademico (3 bytes)

Um byte adicional é usado como marcador de exclusão. O arquivo é armazenado em um disco com os seguintes parâmetros:

- ✓ Tempo médio para transferência de um bloco do disco para a memória = 1ms;
- ✓ Tempo médio de posicionamento (busca) no bloco (Seek Time) = 30ms;
- ✓ Atraso rotacional = 12,5 ms.

A) Calcule o tamanho do Registro **R** em bytes.

B) Considere que o bloco tenha 4K de tamanho (4096 bytes). Calcule o fator de Bloco **Bfr** e o número **b** de blocos de arquivo, considerando uma organização **não espalhada** (vide dicas abaixo).

C) Calcule o tempo médio necessário para localizar um registro ao realizar uma pesquisa linear no arquivo se:

- i. Os blocos do arquivo forem armazenados consecutivamente e o **buffering** duplo for utilizado;
- ii. Os blocos de arquivo não forem armazenados de maneira consecutiva.

D) Suponha que o arquivo esteja ordenado por **CPF**. Ao realizar uma Pesquisa Binária, calcule o tempo necessário para procurar um registro, dado seu valor de **CPF**.

Dicas:

- ✓ Em geral, um registro **R** pode não dividir o bloco exatamente. Por exemplo, o tamanho do registro é 80 e o tamanho do bloco é 100. Para aproveitar esse espaço livre, pode-se utilizar a organização espalhada (spanned). Um ponteiro no final do primeiro bloco aponta para o bloco que contém o restante do registro. Nesse caso, os registros podem se espalhar por mais de um bloco. Caso isso não ocorra, a organização é dita não espalhada.
- ✓ Quando o tempo exigido para se processar um bloco de registros na memória é menor que o tempo exigido para se ler o próximo bloco e preenchê-lo na memória, pode-se empregar a técnica de **buffering** dupla. Enquanto a CPU processa um bloco, o processador de E/S está lendo e transferindo o próximo bloco para a memória. O buffering duplo permite a leitura e a gravação contínua de dados em blocos de disco consecutivos, o que elimina o tempo de busca e o atraso rotacional para todas as transferências de bloco, com exceção da primeira.

