

CAPÍTULO II - ORGANIZAÇÃO DE ARQUIVOS

1. Método De Acesso	07
2. Organização Seqüencial	09
3. Organização Serial.....	10
4. Organização Indexada	11

CAPÍTULO II

ORGANIZAÇÃO DE ARQUIVOS

O tema organização de arquivos refere-se a forma como os registros são armazenados em um arquivo baseado em computador. Confunde-se com MÉTODO DE ACESSO, que consiste na forma como esses podem ser recuperados. A organização do arquivo determina os métodos de acesso que podem ser utilizados na recuperação dos registros, mas tratam-se de coisas distintas.

Apesar de este ser um assunto muito abrangente e com muitas variantes em termos de abordagem, trataremos de apenas três tipos de organização (SEQÜENCIAL, SERIAL E INDEXADA) e seus respectivos métodos de acesso. Essa escolha baseia-se na necessidade de discutirmos alguns conceitos essenciais para o estudo do modelo Relacional de banco de dados, que constitui o objeto principal desse texto.

1. MÉTODOS DE ACESSO

Para recuperarmos um registro em um arquivo, podemos utilizar acesso **SEQÜENCIAL** ou **DIRETO**.

O método **SEQÜENCIAL** de acesso é o mais tradicional e consiste em efetuar a leitura dos registros, um após o outro, comparando o ARGUMENTO DE PESQUISA, com o valor do campo CHAVE (primária ou secundária) no registro corrente, até encontrar os registros desejados ou o final do arquivo.

exemplo:

PROGRAMA Y

INÍCIO....

.

Repita até fim

ler registro

chave secundária (campo chave)

SE NOME = "JOSÉ"

ENTÃO IMPRIMIR

argumento de pesquisa

Fim repita (volte a ler)

.

FIM DO PROGRAMA

O método **DIRETO** consiste em recuperar o(s) registro(s) desejado(s), sem a necessidade de efetuar a leitura dos registros que o(s) antecede(m), o que pode ser feito através de um **ÍNDICE** (que abordaremos no item organização indexada) ou com o auxílio de um algoritmo de **RANDOMIZAÇÃO** que localiza o registro, calculando a posição ocupada pelo registro no disco, com base no valor do argumento de pesquisa, que deve ser um campo numérico.

Em ambos os casos, a localização do registro ocorre a cargo do gerenciador de arquivos, de maneira transparente para o programador, que só precisa escolher a organização adequada para o arquivo e fornecer no programa o argumento de pesquisa.

exemplo:

PROGRAMA Z

INÍCIO....

.

.

ABRIR ARQUIVO ALUNO INDEXADO POR NOME

.

NOME="JOSÉ" → argumento de pesquisa

LOCALIZAR REGISTRO → acesso direto (indexado)

SE ENCONTROU REGISTRO

ENTÃO IMPRIMIR

.

.

FIM DO PROGRAMA

2. ORGANIZAÇÃO SEQÜENCIAL

A ORGANIZAÇÃO SEQÜENCIAL caracteriza-se pela existência de uma CHAVE DE ORDENAÇÃO. Essa chave determina a ordem em que os registros são armazenados e pode ser SIMPLES ou COMPOSTA por dois ou mais campos. Geralmente, coincide com a chave primária, mas não obrigatoriamente.

A organização seqüencial somente permite o ACESSO SEQÜENCIAL.

A figura a seguir apresenta um arquivo com ORGANIZAÇÃO SEQÜENCIAL e CHAVE PRIMÁRIA(MATRICULA) distinta da CHAVE DE ORDENAÇÃO (NOME - ordem alfabética).

ARQUIVO ALUNO

chave primária **chave de ordenação**

↙ ↙

MATRICULA	NOME	ENDEREÇO	DT_NASC
001	Ana	SQS 308 ...	23/08/78
003	José	QND 14	25/09/70
002	José	SQN 410 ...	10/08/85
005	Maria	GAMA	05/04/76
.	.	.	.
.	.	.	.

3. ORGANIZAÇÃO SERIAL

Nesta forma de organização os registros são armazenados de acordo com a ordem de inclusão. o arquivo não possui chave de ordenação, portanto não existe preocupação com a ordem de armazenamento dos registros. No entanto, é sempre recomendável o arquivo possua uma chave primária.

A organização serial somente permite o ACESSO SEQÜENCIAL. Não deve ser utilizada em processos de exclusão e alteração de registros na modalidade batch (atualização em lote), pois degrada a performance.

É muito utilizada em processos de inclusão de registros onde não haja preocupação em manter a seqüência dos mesmos (“pools” de digitação). É também empregada no arquivo de dados que serve de base para a organização indexada, que estudaremos no próximo item.

A figura a seguir apresenta um arquivo com ORGANIZAÇÃO SERIAL. Note que ele não possui CHAVE DE ORDENAÇÃO.

ARQUIVO ALUNO

 **chave primária**

MATRICULA	NOME	ENDEREÇO	DT_NASC
005	Maria	SQS 308 ...	23/08/78
003	José	QND 14	25/09/70
002	Ana	SQN 410 ...	10/08/85
001	José	GAMA	05/04/76
.	.	.	.

4. ORGANIZAÇÃO INDEXADA

Nesta forma de organização, os registros são armazenados em um arquivo de dados com organização serial e para cada campo (ou combinação deles) através do qual se deseja obter acesso direto (indexado) deve-se criar um arquivo de índice (processo de indexação).

Um mesmo arquivo de dados pode possuir diversos arquivos de índice a ele associados. Porém, apesar da flexibilidade para a criação de índices, esse recurso deve ser utilizado com critério, pois a manutenção de muitos índices pode degradar a performance no processo de atualização do arquivo. Ou seja, ganha-se na consulta on-line, mas pode-se perder na atualização de dados.

O arquivo de índice é composto basicamente por duas colunas. A primeira corresponde ao campo utilizado no processo de indexação (endereço lógico) e a segunda armazena um valor (endereço físico) que serve como referência, para que o gerenciador de arquivos localize o registro no disco magnético.

Os registros dos arquivos índice são ordenados pelo endereço lógico. Portanto, se utilizarmos um algoritmo de leitura seqüencial em um arquivo indexado por nome, por exemplo, obteremos os registros em ordem alfabética, mesmo sendo o arquivo de dados um arquivo serial. Ou seja prevalece a ordem do índice. Porém nesse exemplo, a performance a performance do arquivo indexado seria menor, se comparada a de um arquivo seqüencial por nome.

Sempre que um arquivo índice for referenciado por um programa, ele será carregado para memória principal, o que torna desprezível o tempo de busca dos registros nesse arquivo. Além disso, o algoritmo utilizado na busca é o de pesquisa binária, o que reduz ainda mais o tempo.

Os índices constituídos com base no valor da chave primária ou candidata são conhecidos como ÍNDICES PRIMÁRIOS e os demais como ÍNDICES SECUNDÁRIOS.

Em resumo, a organização indexada é formada pela combinação de pelo menos um arquivo de dados e um ou mais arquivos de índice.

A figura a seguir apresenta o cenário da ORGANIZAÇÃO INDEXADA.

ARQUIVO ALUNO

ÍNDICE PRIMÁRIO

MATR	TSL
001	220
002	321
003	231
005	110
.	

ÍNDICE SECUNDÁRIO

NOME	TSL
Ana	321
José	220
José	231
Maria	110
.	331

TRILHA, SETOR E LADO DO DISCO

(endereço físico)

chave primária (endereço lógico)

TSL	MATR	NOME	ENDEREÇO	DT_NASC
110	005	Maria	SQS 308 ...	23/08/78
231	003	José	QND 14	25/09/70
321	002	Ana	SQN 410 ...	10/08/85
220	001	José	GAMA	05/04/76
331