

# 16

## Fundamentos de TI

### Aula 16

#### Tópicos Especiais II

#### Banco de Dados

Prof. Dr. Dilermando Piva Jr.

Site Disciplina:  <http://fundti.blogspot.com.br/>

## Sistemas de Arquivos

### Conceitos

Um arquivo é uma estrutura de dados residente em memória auxiliar, que consiste num conjunto de informações estruturadas em unidades de acesso denominadas registros, todos do mesmo tipo e em número indeterminado.

Um registro é cada um dos componentes do arquivo, contendo o conjunto de informações que são tratados de forma unitária.

O registro é constituído por um ou mais elementos denominados campos, que podem ser de tipos diferentes e podem ser compostos por sub-campos.

Registro Físico é aquele que está contido fisicamente em um arquivo, contido em dispositivos de armazenamento de dados e informações.

Registro Lógico contém a estrutura do registro, sendo declarado nas linguagens de programação.

Ficha de Empreadado			
Num. Registro	<input type="text"/>		
Nome	<input type="text"/>		
Endereço	<input type="text"/>		
Cidade	<input type="text"/>	CEP	<input type="text"/>
		Estado	<input type="text"/>
Telefone	<input type="text"/>	E-mail	<input type="text"/>
Data Admissão	<input type="text"/> / <input type="text"/> / <input type="text"/>	Cargo	<input type="text"/>
Salário	<input type="text"/>		

Exemplo de um Registro

### Características dos arquivos

✓ **Residência em suportes de informação externos**

Os arquivos são armazenados em meios magnéticos e óticos também chamados de memórias auxiliares, tais como discos, fitas e CD.

✓ **Independência com relação ao programa que o criou/mantém**

A vida útil de um arquivo não está limitada pela vida útil do programa que o criou e o mantém, podendo também, em diferentes momentos, ser utilizado por outros programas.

✓ **Permanência das informações armazenadas**

Ao contrário das informações armazenadas em memória principal (RAM), as informações contidas em um arquivo não desaparecem quando o computador é desligado.

✓ **Grande capacidade de armazenamento**

A capacidade de armazenamento de informações por um arquivo está limitada apenas pela arquitetura do suporte externo que o contém. Por outro lado, as estruturas de dados residentes em memória principal têm seu tamanho limitado pela capacidade da memória.

### ✓ **Amplas possibilidades de classificações e formas de acesso (consultas)**

Os arquivos, através de *softwares* específicos, podem ter seus registros classificados e acessados de diversas formas, propiciando ao usuário uma rápida e eficiente visualização de seu conteúdo.

### **Classificação dos arquivos conforme seu uso**

Conforme a utilização dos arquivos, estes são classificados em:

#### ✓ **Arquivos Permanentes**

Contém informações que pouco se modificam com o decorrer do tempo. São subdivididos em:

##### ▶ **Arquivos de referência**

As informações contidas nesse tipo de arquivo permanecem praticamente inalteradas por longos períodos, sendo utilizado principalmente como arquivo de consulta. Um arquivo desse tipo pode ser o de uma rede ferroviária, contendo dados sobre número de estações, de trens, linhas, etc..

##### ▶ **Arquivos-mestres**

São arquivos que contém informações que reflete o estado ou situação de uma empresa, entidade ou algum aspecto dela num determinado momento, sendo atualizados periodicamente para adaptá-los a cada nova situação. Os cadastros são bons exemplos desse tipo de arquivo.

##### ▶ **Arquivos históricos**

São gerados a partir das informações dos anteriores, sendo deixados fora de uso para futuros estudos estatísticos, consultas e exigências legais.

#### ✓ **Arquivos Temporários**

São arquivos transitórios. São subdivididos em:

##### ▶ **Arquivos de movimento/transações**

Esses arquivos, de maneira geral, armazenam informações para atualização dos arquivos-mestres. Após atualização dos arquivos-mestres, a partir de arquivos-movimento, este último perde sua validade e podem ser eliminados, ou dependendo do sistema podem ser transformados em arquivos históricos. São arquivos temporários.

##### ▶ **Arquivos de trabalho**

Esses arquivos têm um tempo de vida útil extremamente limitado, normalmente igual à duração de um ciclo de processamento, sendo na maioria das vezes utilizados como auxiliares dos arquivos de movimento, e mesmo dos arquivos-mestres quando estes sofrem processos de classificação de seus registros.

### Organização de arquivos

Dependendo do uso a que se destina e do suporte utilizado, podemos escolher diferentes formas de organizar os registros de um arquivo. As principais organizações são:

#### ✓ Line Sequential ou Seqüencial em linha

Essa forma de organização é típica dos arquivos de armazenamento de textos (.TXT), onde não predomina uma estrutura de dados (registros) mas sim uma estrutura de linhas, onde não identificamos campos nem tamanhos fixos de linhas (registros).

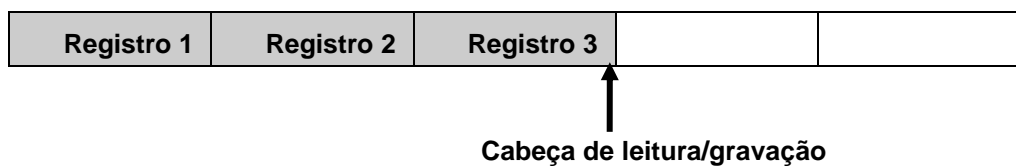
Analisando as linhas contidas em um arquivo desse tipo, encontramos os caracteres de controle OD e OA, *line-feed* e *carriage-return* em cada linha digitada, identificando final de linha ou linha em branco.

#### ✓ Sequential ou Seqüencial

Nesse tipo de organização os registros de um arquivo ocupam posições consecutivas na memória auxiliar, sendo que o computador só pode acessar um registro de cada vez, sempre a partir do primeiro registro.

Em arquivos desse tipo não se fazem operações de gravação quando se está lendo, nem operações de leitura quando se está gravando. Para atualizar registros nesse tipo de arquivo necessário criar novos arquivos.

#### Representação lógica de um arquivo Seqüencial



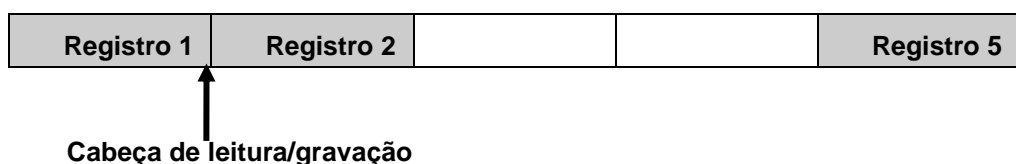
#### ✓ Relative ou Direto/Aleatório

Nesse tipo de organização os registros de um arquivo são gravados e acessados aleatoriamente mediante a sua posição, ou seja, indicando o lugar relativo que ocupam dentro do conjunto de posições possíveis (limitadas ao espaço físico do suporte externo).

Nesta forma de organização pode-se ler e escrever registros em qualquer ordem e em qualquer lugar, sendo uma atribuição do programador o estabelecimento da relação entre a posição que um determinado registro ocupa e o conteúdo da chave relativa do registro (Relative-Key). A gravação seguindo esses critérios é uma tarefa do software gerenciador do arquivo.

A vantagem desse tipo de organização é a rapidez de acesso ao registro, uma vez que não é preciso passar pelos registros anteriores até chegar naquele desejado. O Acesso é direto. Por outro lado, essa organização apresenta como desvantagem o desperdício de espaço físico do suporte externo.

#### Representação lógica de um arquivo Direto



✓ **Index-Sequencial ou Seqüencial Indexado**

Um arquivo com este tipo de organização possui basicamente duas áreas :

▶ **Área de Índices**

É um arquivo seqüencial, criado e mantido pelo software gerenciador do arquivo, no qual cada registro tem uma entrada de ÍNDICE mais uma entrada de ENDEREÇO do registro lógico na área de dados.

O(s) componente(s) do registro lógico, que compõem o índice é predeterminado pelo analista/programadores, e informados ao gerenciador do arquivo através do programa de acesso ao arquivo.

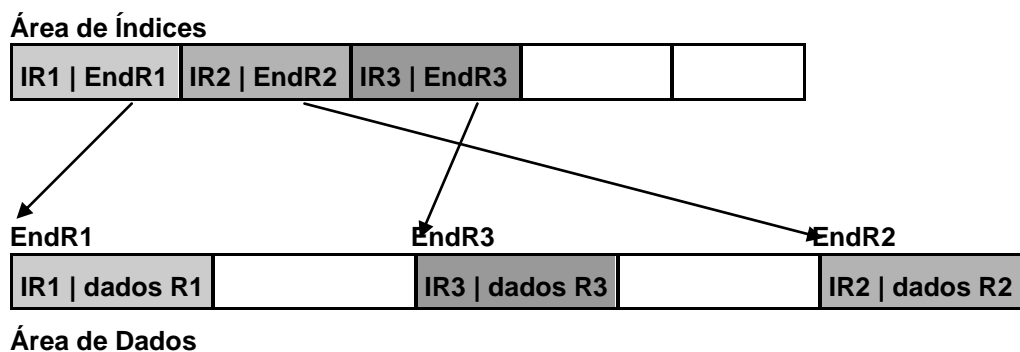
O computador acessa um registro de maneira direta a partir do índice desse registro.

▶ **Área de Dados**

Contém os registro de dados, na forma como são inseridos no arquivo.

O acesso aos registros nessa área é feito através do conteúdo do campo-chave (Record-Key), lendo-se Área de Índices onde é obtendo-se o endereço do registro na área de dados.

**Representação lógica de um Seqüencial Indexado**



IR<sub>n</sub> – Índice do registro n

EndR<sub>n</sub> – Endereço do registro n na área de dados.

A vantagem desse tipo de organização é a rapidez de acesso ao registro. O Acesso é feito a partir da leitura da área de índices, que é muito menor que a área de dados. Por outro lado, essa organização apresenta como desvantagem o desperdício de espaço físico do suporte externo após muitas exclusões físicas de registros.

Arquivos desse tipo, que sofrem muitas atualizações/exclusões devem ser constantemente reorganizados, para evitar desperdícios de espaços (muitos *gaps* no arquivo).

✓ **Tabelas de Dados Relacionadas - Banco de Dados**

**O que é um Banco de Dados ?**

**Um Banco de Dados é um recurso para a manipulação eficiente de um grande conjunto de informações estruturadas e armazenadas de forma organizada, integrada e consistente.**

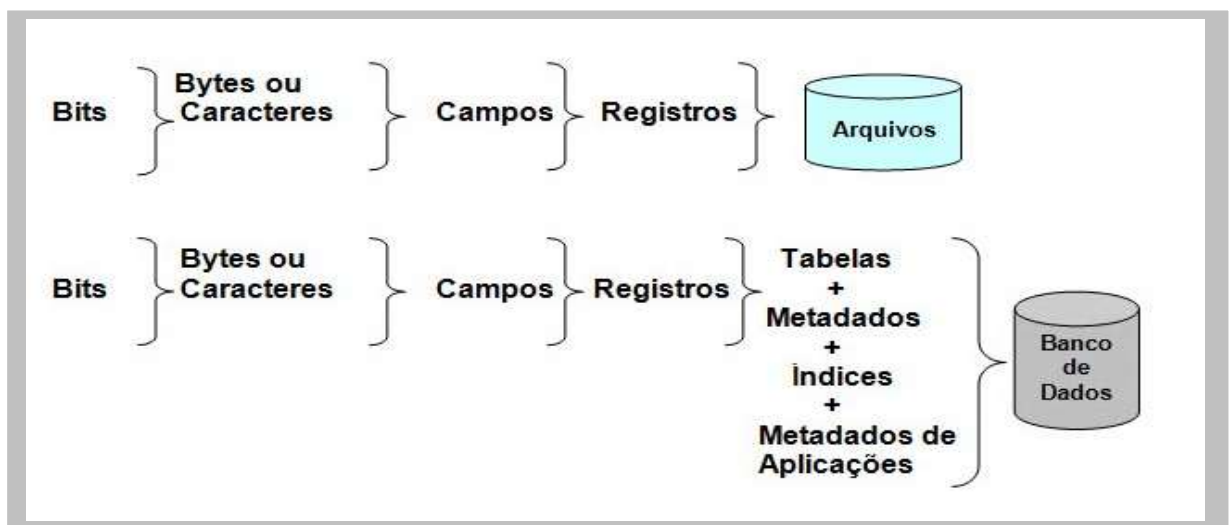
Um Banco de Dados é uma coleção autodescritiva de registros integrados.

O padrão de hierarquia de dados é mostrado na primeira parte da figura abaixo. É comum seguir esse padrão e afirmar que os arquivos são reunidos em Banco de Dados. Embora essa declaração seja verdadeira, ela não define tudo.

Um **Banco de Dados** inclui:

- ▶ **ARQUIVOS** ou **TABELAS de DADOS de usuários**;
- ▶ **ARQUIVOS** ou **TABELAS ÍNDICES**, usados para representar relacionamento entre os dados e também para melhorar o desempenho das aplicações.
- ▶ **METADADOS** ou **DICIONÁRIO DE DADOS**, contendo a descrição dos dados contidos em cada Tabela de Dados; e
- ▶ **METADADOS DA APLICAÇÃO**, tabelas contendo frequentemente dados sobre as aplicações que utilizam as Tabelas de Dados do Banco.

A figura abaixo mostra a relação entre arquivos e Banco de Dados, sendo este último uma evolução dos tipos de arquivos sequencial-indexados.



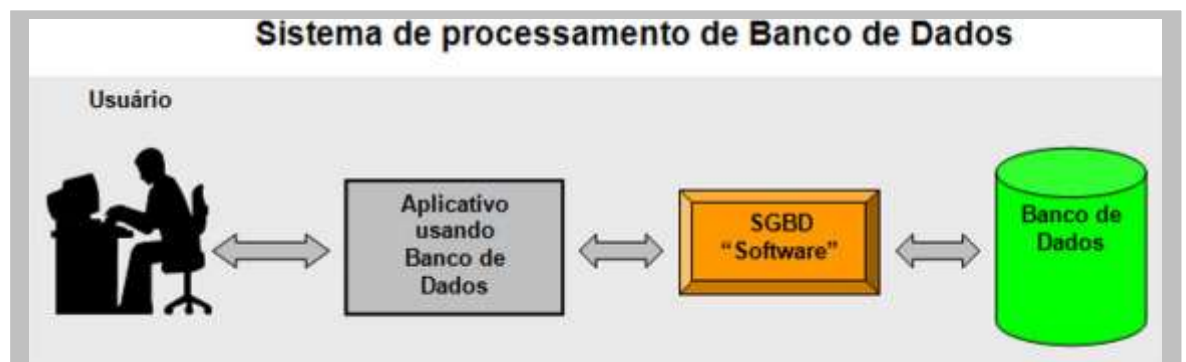
**Relação Arquivos x Banco de Dados**

### O que é um SGBD – Sistema de Gerenciamento de Banco de Dados

O SGBD é um sistema capaz de armazenar e recuperar informações em Banco de Dados.

Antes do advento dos Bancos de Dados, o processamento de arquivos era efetuado pelos Gerenciadores de Arquivos, conforme instruções dadas pela aplicação.

NO caso dos Banco de Dados, **os aplicativos acessam o Banco através do SGBD** (que é uma camada de software), **conforme mostra a figura abaixo**. Esta diferença é significativa porque torna a tarefa de programação da aplicação mais fácil, pois muitas tarefas são delegadas ao SGBD.



### **Operações sobre arquivos**

Dentre as operações realizadas em um arquivo, por um programa utilitário, de sistema operacional ou mesmo de um sistema específico, destacam-se:

✓ **Criação**

Processo de gravação inicial dos registros, num determinado suporte externo. Os registros podem ser originados de dados informados através de um teclado, de outros arquivos ou mesmo internamente de algum processamento.

✓ **Cópia**

Processo de criação de um arquivo novo a partir de outro já existente. A cópia pode ser no mesmo suporte ou num outro diferente, e até mesmo em uma impressora.

✓ **Consulta**

Processo de busca do conteúdo de um ou vários registros de dados, num determinado suporte externo.

✓ **Ordenação ou classificação**

Processo de rearranjar os registros, de forma que fiquem ordenados em relação aos valores de um ou mais campos denominados de campo-chave de ordenação.

✓ **Concatenação**

Processo de obtenção de um arquivo a partir de dois ou mais arquivos de entrada, de mesma estrutura, onde figurem todos os registros do primeiro arquivo de entrada, e na seqüência, todos os registros do segundo, e assim por diante.

✓ **Interseção**

Processo de obtenção de um arquivo a partir de dois ou mais arquivos de entrada, de mesma estrutura, onde figurem no arquivo de saída somente os registros comuns aos de entrada.

✓ **Junção ou intercalação**

Processo de obtenção de um arquivo a partir de dois ou mais arquivos de entrada, de mesma estrutura e ordenados em relação a um ou mais campos-chave, onde figurem no arquivo de saída todos os registros de ambos arquivos de entrada, mantendo a ordenação.

✓ **Atualização**

Processo de modificação de um arquivo-mestre, por exemplo através de um arquivo de movimento/transações. As formas de atualização incluem: Inclusão, Alteração, Exclusão e Bloqueio de registros.

✓ **Reorganização**

Processo de reposicionamento dos registros de um arquivo que sofreu sucessivas alterações, de forma que sejam ocupados os possíveis espaços livres intermediários resultantes de remoções de registros, obtendo-se dessa forma uma otimização de memória tanto auxiliar como principal quando esse arquivo estiver sendo processado.

✓ **Remoção**

Processo de eliminação total de um arquivo, quando este não se faz mais necessário, liberando espaço no suporte externo.

✓ **Partição**

Processo de decomposição de um arquivo em dois, de acordo com alguma característica dos seus registros. Por exemplo, particionar um arquivo de alunos, segundo o conteúdo do campo IDADE, obtendo dois arquivos, sendo que um contém somente os alunos maiores de idade, e outro contendo os menores.

### ✓ **Criptografia**

Processo de codificação do conteúdo dos registros, conforme determinados padrões de software, visando preservar a segurança do arquivo.

### ✓ **Compressão**

Processo de substituição de caracteres repetitivos, efetuado por algoritmos apropriados, de proporcionando economia de espaço em disco.

## **Segurança e Controles dos arquivos**

Dados e informações armazenados em arquivos podem ser de grande importância para organizações e usuários, sendo que a perda ou destruição delas pode significar um desastre para seus proprietários.

Diante disso torna-se necessário estabelecer determinadas políticas e medidas de proteção e controle que reduzam ou afastem totalmente esses riscos. Da mesma forma outras medidas de proteção devem ser tomadas para garantir a privacidade e confidencialidade dos conteúdos dos arquivos, de forma que estes não possam ser acessados por pessoas não autorizadas, ou para outros propósitos que não aqueles para os quais foi criado.

**A proteção dos arquivos deve ser efetuada considerando dois aspectos:**

**a) Proteção contra falhas do sistema ou do suporte.**

**b) Proteção contra uso indevido e/ou não autorizado.**

### ✓ **Proteção contra Falhas do Sistema ou do suporte**

A proteção dos arquivos e também das operações de atualização realizadas sobre eles deve ser efetiva, de forma que, se em determinado momento ocorrer uma falha no sistema ou quebra do suporte externo, os dados e informações possam ser recuperados tal como se encontravam antes da falha.

Os seguintes procedimentos, são alguns dos que devem ser adotados para proteção contra falhas de sistema e/ou suporte:

#### ▶ **Cópias de Segurança – (Backup/Restore)**

Periodicamente, através de software específicos, são efetuadas cópias de segurança (backup) dos principais arquivos, guardando-se essas cópias em lugar seguro, para serem utilizadas (restauradas) em caso de perda ou danificação do arquivo original.

O Backup pode ser classificado em:

- **Backup Total**

Consiste na cópia de todos os arquivos sempre que for iniciado o processo Backup.

- **Backup Incremental**

Consiste em copiar somente os arquivos, de um conjunto de arquivos, que sofreram modificação após seu último backup.

- **Backup Seletivo**

Consiste em efetuar cópias somente de determinados arquivos.



### ▶ Registros de operações

Consiste em registrar todas as operações realizadas com os arquivos periodicamente, uma vez por dia, por exemplo.

### ▶ Replicação de suporte (RAID)

A tecnologia RAID (Redundant Arrays of Inexpensive Disks) significa Conjuntos Redundantes de Discos Baratos.

Essa tecnologia, originária de *Mainframes*, utiliza um conjunto de discos rígidos para gravar simultaneamente os dados em mais de um disco. Se um disco falhar, o sistema operacional poderá recuperar os dados do disco de reserva. Dessa forma, quanto mais discos (mais redundância), mais segurança tem o sistema.

O RAID pode também acelerar o acesso ao disco, pois estando as cópias armazenadas em mais de um disco, as solicitações por dados podem ser atendidas simultaneamente, acelerando o acesso.

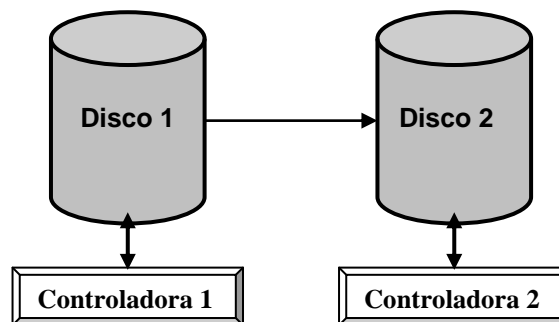
**Dentre as tecnologias RAID destacamos:**

#### ▪ RAID-1 ( Mirror )

Duplica uma partição e transfere a cópia para outro disco físico. Essa estratégia é a maneira mais simples de proteger um disco simples de falhas.

O espelhamento é uma forma de backup contínuo.

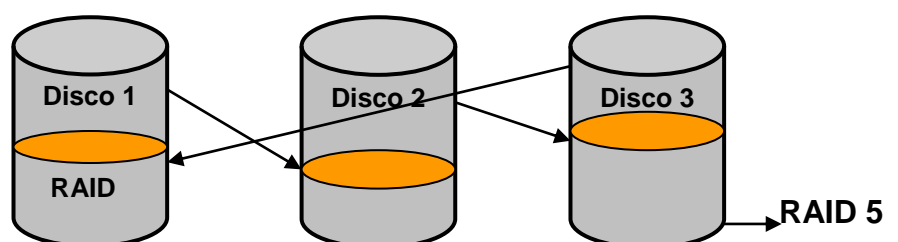
RAID-1 torna-se mais eficiente quando usada a Duplexação, que consiste em uma placa controladora de disco adicional para a Segunda unidade de disco.



#### ▪ RAID-5 ( Striping com paridade )

É atualmente a abordagem mais popular para projetos de proteção contra falhas.

Suporta um mínimo de 3 discos e máximo de 32. Cerca de 30% de cada disco é utilizado pelo sistema de proteção de dados, onde são gravados os dados dos outros discos, utilizando-se para isso *softwares* específicos.



- **Proteção contra uso indevido e/ou não autorizado**

A funcionalidade dos arquivos consiste na possibilidade de acesso a seus dados por parte de diferentes programas ou usuários. Isto implica na possibilidade dos dados serem danificados por erro ou serem utilizados de forma ilegítima por usuários não autorizados.

Dentre as medidas que devem ser adotadas para proteger arquivos contra uso indevido e/ou acesso não autorizado, destacamos:

- **Criptografia**

Consiste no encobrimento dos dados e informações através de algum método cifrado (software).

A Criptografia é realizada pela conversão dos dados em um código muito complexo que é extremamente difícil de ser quebrado.

- **Conta e Senha**

A cada usuário atribuem-se uma Conta e uma Senha para acesso aos ambientes informacionais e conseqüentemente aos arquivos. Liga-se à essa Conta/Senha o nível de acesso (possibilidade de Leitura/Gravação/Alteração/Exclusão, etc) desse usuário aos arquivos.