

Apontamentos de BD

João Aragonez

Conteúdo

| | | |
|----------|---|----------|
| 1 | Conceitos Iniciais | 5 |
| 1.1 | Sistemas de informação | 5 |
| 1.2 | Sistemas de Gestão de Bases de Dados (SGBD) | 5 |
| 1.2.1 | Vantagens dos SGBD's | 6 |
| 1.2.2 | Desvantagens dos SGBD's | 6 |
| 1.3 | Modelos e Níveis de Abstração nos SI | 6 |
| 1.4 | Modelos de Dados | 7 |
| 1.5 | Arquitetura dos SGBD | 7 |
| 1.6 | Conceção de Bases de Dados | 8 |
| 1.7 | Utilizadores de Bases de Dados | 8 |

Capítulo 1

Conceitos Iniciais

1.1 Sistemas de informação

Definição 1 (Sistemas de Informação). Consiste na área que estuda as atividades de pendor estratégico, operacional e de gestão subjacentes à recolha, processamento, armazenamento, distribuição e uso de informação e de tecnologias associadas, tanto pela sociedade como por organizações.

Também é comum definir SI como a interação entre tecnologia e processos de negócio, mais concretamente, a gestão de 3 componentes fundamentais: **dados, tecnologia e pessoas**.

Entre outros, menciona-se os seguintes tipos de sistemas de informação:

- *ERP* (*Enterprise Resource Planning*);
- SIG (Sistemas de Informação Geográfica);
- Sistemas de *office automation*;
- Sistemas de *Business Intelligence*;
- Sistemas Especialistas;
- *WWW* (*World Wide Web*).

1.2 Sistemas de Gestão de Bases de Dados (SGBD)

Definição 2 (Base de Dados). Consiste em nada mais que conjuntos de dados interligados.

Definição 3 (Sistema de Gestão de Bases de Dados). Consiste numa ferramenta de software desenhada para a manutenção e gestão de bases de dados

Dado que os sistemas operativos atuais se encontram munidos de um sistema de ficheiros, perfeitamente capazes de lidar com o armazenamento de informação, surge a seguinte questão: *porquê usar um SGBD?* A verdade é que os sistemas de informação apresentam necessidades comuns que não são cobertas por sistemas de ficheiros. Assim, os SGBD têm por objetivo realizar:

- Controlo de redundância;
- Segurança e controlo de acessos, dada a heterogeneidade de utilizadores e de dados;
- Persistência de dados;
- Oferecer múltiplas interfaces para diferentes tipos de utilizadores;
- Representar relações complexas;
- Assegurar constrangimentos de integridade sobre os dados;
- Realizar controlo de concorrência, por forma a manter os dados consistentes;
- Permitir que uma grande quantidade de iterrações (*queries*) possam ser feitas sobre os dados sem necessidade de programação adicional;
- Garantir tolerância a faltas (e.g., realizando *backups*).

1.2.1 Vantagens dos SGBD's

- **Independência dos dados:** encapsulando o modo real de representação e armazenamento dos dados, os SGBD's disponibilizam uma visão abstrata dos dados.
- **Acesso Eficiente aos Dados:** os SGBD incorporam técnicas para armazenamento e recolha eficiente dos dados;
- **Integridade dos dados e segurança:** os SGBD garantem a aplicação de restrições de integridade no acesso e manipulação de dados;
- **Capacidade de administração dos dados:** é possível mudar a representação dos dados por forma a minimizar a redundância e melhorar o armazenamento de forma totalmente transparente ao utilizador;
- **Acesso Concorrente e Recuperação de Falhas:** existe suporte à concorrência no acesso aos dados, garantido um efeito semelhante a um acesso sequencial;
- **Redução do tempo de desenvolvimento de aplicações:** disponibiliza uma interface de alto nível para os dados e funções de acesso comuns, sendo para além disso uma componente da aplicação que não necessita de ser verificada.

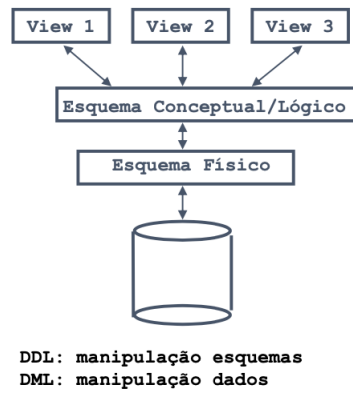
1.2.2 Desvantagens dos SGBD's

- **Overhead demasiado elevado:** requer investimento em hardware, software e formação no uso destes sistemas;
- **Tratamento demasiado geral:** Dependendo da aplicação, os mecanismos de segurança, controlo de concorrência, integridade e de recuperação de faltas podem não ser suficientes;
- **Desadequados a sistemas com requisitos de tempo-real;**
- **Desadequados a bases de dados simples/imutáveis ou sem concorrência de acessos;**
- **Desadequados a certos tipos de dados, como texto.**

1.3 Modelos e Níveis de Abstração nos SI

Num SGBD, os dados podem ser descritos segundo diversos modelos, que correspondem a diferentes níveis de abstração acerca da sua representação/armazenamento:

- **Modelo Conceptual** (ou esquema externo), que descreve como os utilizadores vêem os dados. Permite particularizar o acesso aos dados através de **Vistas** - conjuntos de registos visíveis para grupos específicos de utilizadores e apenas computados quando necessário (i.e., não são explicitamente armazenados). Este nível permite **independência dos dados lógicos**, pois alterações ao esquema lógico requerem unicamente redefinição de vistas, pelo que o utilizador não se dará conta de eventuais extensões e modificações das estruturas de dados.
- **Modelo Lógico** (ou esquema conceptual), que corresponde à estrutura lógica dos dados (e.g., relações existentes no modelo relacional). Este nível permite **independência dos dados físicos**, pois a organização física nada influi sobre o esquema lógico dos dados.
- **Modelo Interno** (ou esquema físico), que especifica os detalhes de armazenamento das relações (e.g., definição de tipos de ficheiros a utilizar e de índices).

Figura 1.1: Modelo *ANSI/SPARC*

1.4 Modelos de Dados

Definição 4 (Modelo de Dados). Coleção de conceitos para descrever dados, relacionamentos, semântica de dados e restrições.

Definição 5 (Esquema). Descrição de uma coleção específica de dados à luz de um dado modelo de dados (i.e., o resultado da aplicação de um modelo de dados um conjunto de dados específico).

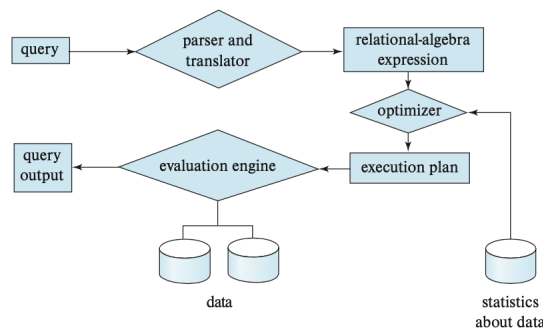
Entre outros modelos de dados, destacam-se o **Modelo Relacional**, o **Modelo Entidade-Associação**, o **Modelo Baseado em Objetos**, **Modelos de Dados Semi-Estruturados** (como *XML/JSON*), ou os **Modelos em Rede e Hierárquicos** (não usados atualmente).

Contudo, o modelo de dados mais amplamente difundido nos SGBD é o **modelo relacional**, cujos conceitos fundamentais são a **relação** (i.e., um tuplo de atributos) e o **esquema**, que corresponde à especificação do nome da relação e do nome e tipo dos seus atributos.

Numa fase mais inicial do desenvolvimento de bases de dados, podem-se usar **Modelos Semânticos de Dados**, passíveis de serem diretamente traduzidos para o modelo relacional. O exemplo mais paradigmático destes modelos é o **Modelo Entidade-Associação**.

1.5 Arquitetura dos SGBD

As arquiteturas dos SGBD procuram, por um lado, maximizar a **eficiência e escalabilidade**, mais concretamente, acelerando as interrogações sobre os dados. A figura abaixo exhibe as fases que compõem o processamento de uma *query*: **análise e tradução**, **otimização** e **avaliação**.

Figura 1.2: Processamento de uma *query*

Por outro lado, procuram maximizar a **concorrência e a robustez**, existindo um **gestor de transações** para lidar com questões de concorrência, bem como um **gestor de recuperação** e um **gestor de locks**.

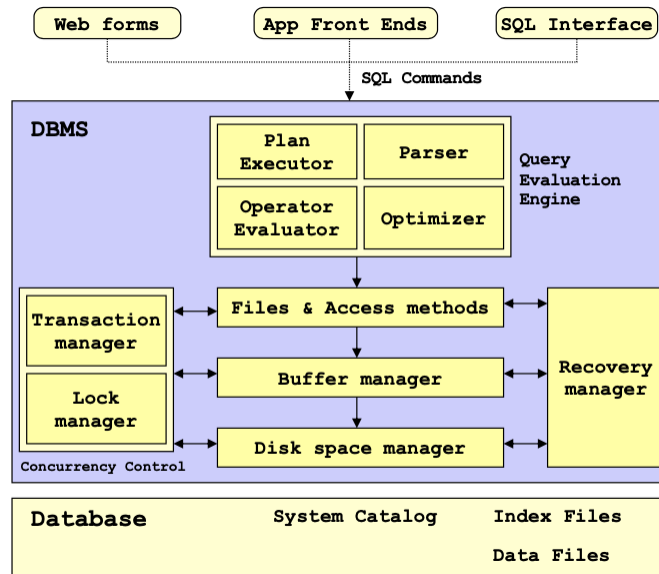


Figura 1.3: Arquitetura de um SGBD

1.6 Conceção de Bases de Dados

O processo de conceção de um base de dados incide inicialmente no **desenho lógico**, i.e., sobre o **esquema** a adotar. Para esta decisão contribuem fatores associados ao **Negócio** (como determinar quais os atributos mais relevantes para o domínio em questão), bem como fatores de **Engenharia**, como definição de esquemas e distribuição dos atributos por estes.

1.7 Utilizadores de Bases de Dados

- Implementadores de Bases de Dados;
- Utilizadores das aplicações;
- Programadores de aplicações ao definirem o modelo lógico do sistema de informação;
- **DBA** (*Database Administrators*), que concebem e mantêm a base de dados, em termos de desenho físico e lógico, segurança e configuração dos mecanismos de disponibilidade e recuperação.