



Introdução aos Sistemas de Bases de Dados

Orlando Belo

*Departamento de Informática
Escola de Engenharia
Universidade do Minho*

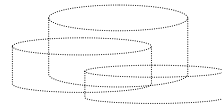
Índice

- Introdução ao armazenamento de informação.
- Sistemas de ficheiros convencionais.
- Sistemas de bases de dados.
- Arquitectura funcional de um sistema de bases de dados.
- Perfis de utilização de um sistema de bases de dados.
- Características, funções e componentes de um sistema de gestão de bases de dados.
- Linguagens das bases de dados.
- A arquitectura ANSI/SPARC.
- Independência de dados.
- Arquitecturas multiutilizador.



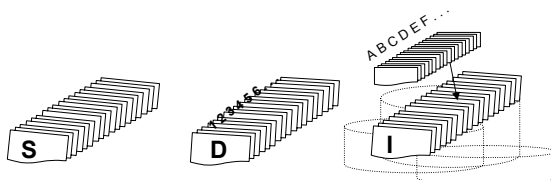
Introdução ao armazenamento de informação

- Informação e dados.
- A necessidade de armazenar informação.
- Sistemas de armazenamento de informação.
- A necessidade de organizar a informação.
 - Gestão e armazenamento dos apontamentos de um curso.
- As bases de dados.
- Os sistemas de gestão de bases de dados.
- Aplicações de sistemas de bases de dados.
 - Hipercursos, bancos, multibanco, comércio electrónico, serviços académicos, etc.



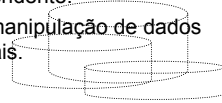
Sistemas de ficheiros convencionais

- Um Sistema Baseado em Ficheiros é uma colecção de programas que realizam um conjunto de serviços específicos para os utilizadores do sistema e em que cada um deles define e gere o seu próprio conjunto de dados.
- Diferentes organizações e acessos à informação implicariam diferentes tipos de ficheiros:
 - Sequencial.
 - Directo.
 - Indexado.

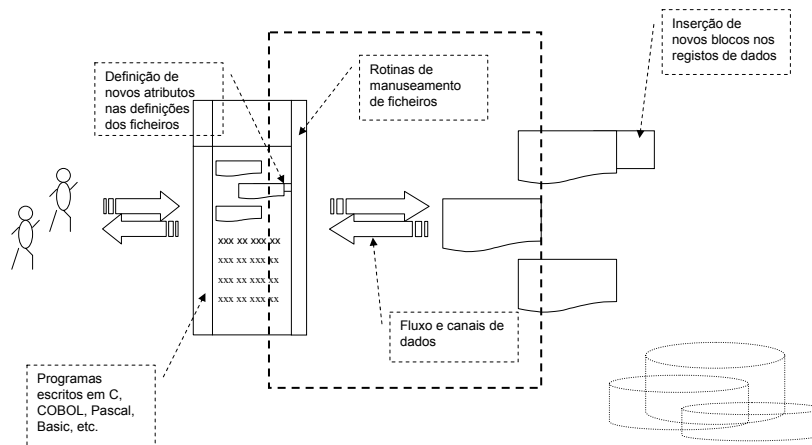


Limitações dos ficheiros

- Separação e isolamento dos dados.
 - Duplicação de dados.
 - Dependência dos dados.
 - Formatos dos ficheiros incompatíveis.
 - Sistemas de interrogações (queries) fixos.
 - Grande dependência do programador das aplicações.
 - Proliferação das aplicações.
- O "porquê" destas limitações...
- A definição dos dados está integrada na própria estrutura dos programas em vez de estar guardada separadamente e de forma independente.
 - Por não existir mecanismos de controlo no acesso e na manipulação de dados além daquele que é imposto pelos programas aplicacionais.

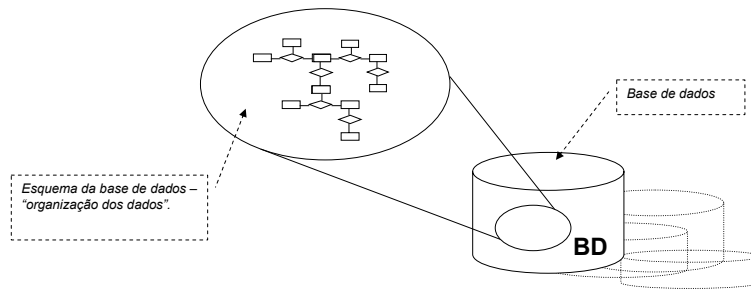


Definição e acesso aos dados em ficheiros



Bases de dados

- Uma Base de Dados (BD) é, em termos gerais, uma colecção de dados partilhados, assim como a descrição de cada um desses dados, que estão logicamente relacionados organizados de acordo com os requisitos de informação de uma organização.



2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

7

Informação de uma base de dados

- Uma base de dados armazena informação relacionada com:
 - Os dados operacionais de uma organização.
 - A descrição dos dados – catálogo do sistema, dicionário de dados ou metadados (dados acerca dos dados).
- Uma base de dados pode ser vista como uma colecção de registos integrados auto-descritivos; esta característica garante a independência dos dados relativamente aos programas que os manipulam.

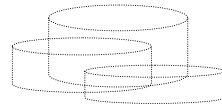


2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

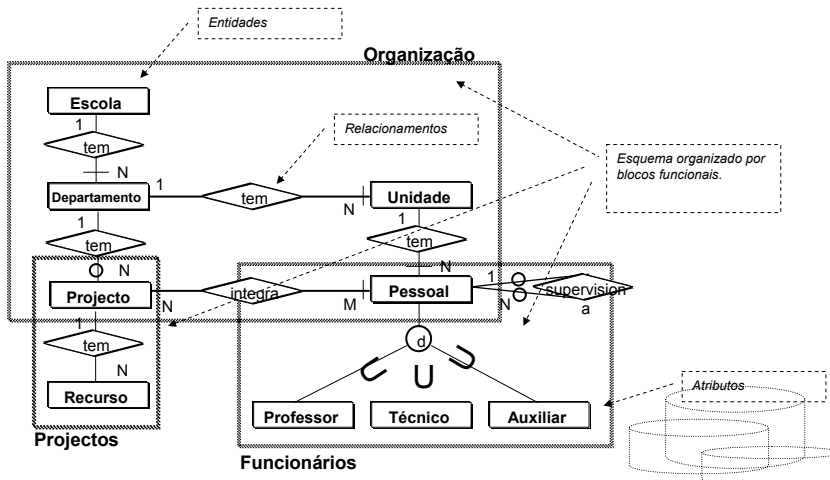
8

Abstracção de dados

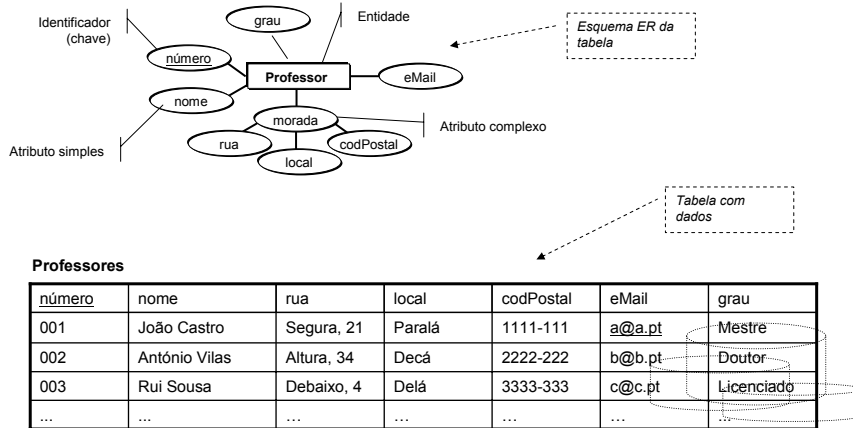
- As bases de dados possibilitam a modificação da definição interna de um objecto sem que os utilizadores desse objectos sejam afectados.
- As estruturas dos dados são armazenadas na própria base de dados, estando assim separadas dos programas aplicativos que manipulamos dados.
 - (análise das implicações de operações de inserção, modificação e remoção de metadados).



Esquema de uma base de dados



Esquema e instaciações de uma entidade



2003, O.Belo, *Sistemas de Bases de Dados, Introdução*.

11

Sistemas de gestão de bases de dados

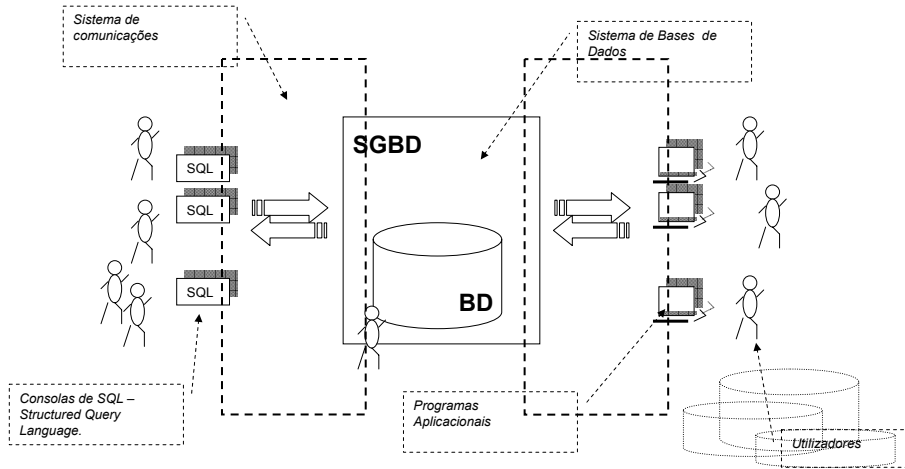
- Um Sistema de Gestão de Bases de Dados (SGBD) é um sistema de software que permite aos seus utilizadores definir, criar e manter uma BD, assim como providencia formas de acesso controladas a essa BD.
- Um SGBD é a peça de software que assegura a interacção entre os utilizadores ou entre os programas aplicativos e a base de dados.
- Em termos gerais um SGBD disponibiliza as seguintes funcionalidades:
 - Permite aos utilizadores definirem a base de dados.
 - Disponibiliza os meios para executar operações de inserção, actualização, remoção ou consulta sobre os dados armazenados na base.
 - Assegura os mecanismos para o controlo de acessos à base de dados – segurança, integridade, concorrência, recuperação e controlo de catálogos.



2003, O.Belo, *Sistemas de Bases de Dados, Introdução*.

12

Um sistema de bases de dados



2003, O. Belo, *Sistemas de Bases de Dados, Introdução.*

13

Linguagens de um sistema de bases de dados

- Um SGBD disponibiliza o seguinte conjunto de funcionalidades:
 - Uma Linguagem de Definição de Dados (LDD) – “Data Description Language”.
 - Uma Linguagem de Manipulação de Dados (LMD) – “Data Manipulation Language”.
 - Uma linguagem para a definição e implementação de regras para a gestão e controlo das bases de dados, Linguagem de Controlo de Dados (LCD) – “Data Control Language”.

```
CREATE TABLE Alunos (  
  Número INTEGER PRIMARY KEY,  
  Nome VARCHAR(50) NOT NULL,  
  Data_Nascimento DATETIME,  
  Distrito VARCHAR(15),  
  Curso CHAR(3) FOREIGN KEY REFERENCES Cursos(Código));
```

Instrução de
definição de dados

Instrução de
manipulação de
dados

SELECT *

```
FROM Alunos  
WHERE Número = '19829';
```



2003, O. Belo, *Sistemas de Bases de Dados, Introdução.*

14

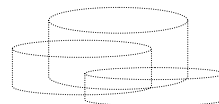
Componentes de um SBD

- Hardware.
 - Plataformas computacionais – computadores pessoais, computadores de grande porte, redes de computadores, etc.
- Software.
 - SGBD, sistemas operativos, programas de rede, utilitários, etc.
- Dados.
 - Dados operacionais, e catálogo do sistema.
- Procedimentos.
 - Instruções e regras que regulam o desenho e a utilização de uma base de dados – entrada no sistema, arranque do sistema, cópias de segurança, permissões de utilização, actualizações de metadados, recuperação de dados, etc.
- Utilizadores.
 - Diferentes tipos de utilizadores do sistema – administradores, arquitectos, programadores de aplicações e utilizadores finais.



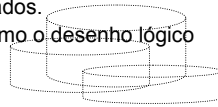
Unidades de Armazenamento

| | 2^x | byte |
|------------------|----------|-----------------------------------|
| <i>Kilobyte</i> | 2^{10} | 1,024 |
| <i>Megabyte</i> | 2^{20} | 1,048,576 |
| <i>Gigabyte</i> | 2^{30} | 1,073,741,824 |
| <i>Terabyte</i> | 2^{40} | 1,099,511,627,776 |
| <i>Petabyte</i> | 2^{50} | 1,125,899,906,842,620 |
| <i>Exabyte</i> | 2^{60} | 1,152,921,504,606,850,000 |
| <i>Zettabyte</i> | 2^{70} | 1,180,591,620,717,410,000,000 |
| <i>Yottabyte</i> | 2^{80} | 1,208,925,819,581,336,886,706,176 |



Perfis de utilização de uma BD

- Administradores
 - Dados
 - Gestão dos recursos de dados, planeamento da BD, desenvolvimento e manutenção de padrões, políticas e procedimentos, e desenho conceptual e lógico da BD.
 - Bases de Dados.
 - Implementação física da BD, projecto e implementação do desenho físico, controlo de segurança e de integridade, manutenção do sistema operacional e análise e optimização do desempenho do sistema.
- Arquitectos de Sistemas de Dados.
 - Este perfil está organizado em duas vertentes:
 - Arquitectos da base de dados lógica, que identificam os dados os seus relacionamentos e as restrições a aplicar sobre os dados.
 - Arquitectos da base de dados física, que decidem como o **desenho lógico** da base de dados pode ser realizado fisicamente.



Perfis de utilização de uma BD

- Programadores de Aplicações.
 - Profissionais que desenvolvem as aplicações requeridas com base em especificações fornecidas por analistas do sistema.
- Utilizadores.
 - Estes são os “clientes” das BDs e podem ser classificados em:
 - Especialistas (experts), utilizadores familiarizados com a estrutura da BD e dos serviços disponibilizados por um SGBD.
 - Regulares (naïves), utilizadores que acedem à BD através de aplicações específicas, restringindo-se o seu conhecimento a operações muito elementares.



Vantagens de um SGBD

- Controlo da redundância de dados.
- Consistência de dados.
- Mais informação para o mesmo volume de dados.
- Partilha de dados.
- Melhoria na integridade de dados.
- Melhoria em questões de segurança.
- Aplicação de standards.
- Economia de escala.
- Balanceamento em conflitos de requisitos.
- Melhoria no acesso aos dados e na resposta do sistema.
- Aumento de produtividade.
- Melhoria na manutenção.
- Aumento de concorrência.
- Melhoria nos serviços de cópias de segurança e de serviços de recuperação.



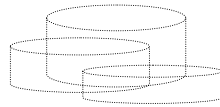
Desvantagens de um SGBD

- Sistemas mais complexos.
- Dimensão.
- Custo (adicionais de hardware e de conversão).
- Desempenho.
- Vulnerabilidade do sistema.

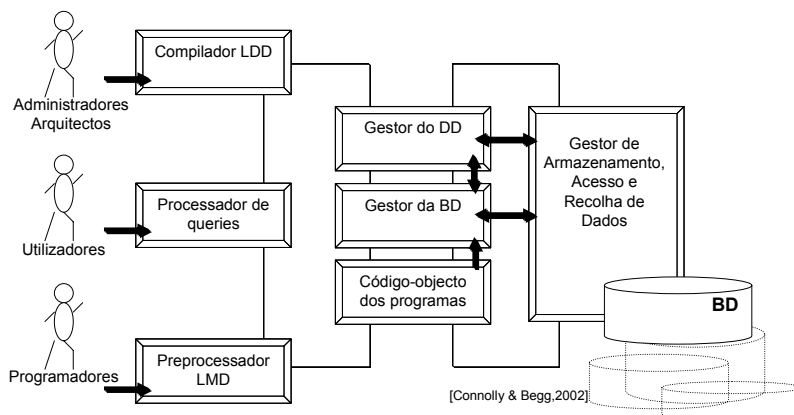


Funções básicas de um SGBD

- Cood apresentou oito serviços que um SGBD, de qualquer tipo e dimensão, deveria ser capaz de prestar:
 - Armazenamento, actualização e obtenção de dados.
 - Um catálogo do sistema acessível aos utilizadores.
 - Suporte transaccional.
 - Controlo de concorrência.
 - Serviços de recuperação.
 - Serviços de autorização e controlo de acessos.
 - Suporte para comunicação de dados.
 - Serviços de integridade.
- e adicionalmente:
 - Serviços para a promoção da independência de dados.
 - Serviços utilitários.

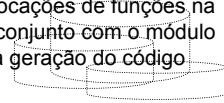


Componentes de um SGBD



Componentes de um SGBD

- Processador de interrogações.
 - Transforma as *queries* em instruções directamente interpretáveis pelo gestor da base de dados.
- Gestor de bases de dados.
 - Actua como interface entre as aplicações e as *queries* dos utilizadores, verificando se os pedidos executados e analisando os esquemas externos e conceptuais de forma a determinar quais os registos conceptuais poderão satisfazer os pedidos executados. De seguida, comunicará com o gestor de ficheiros para satisfazer os pedidos.
- Gestor de ficheiros.
 - Manipula os ficheiros de dados e gere o seu armazenamento em disco.
- Preprocessador da linguagem de manipulação de dados.
 - Converte as instruções de manipulação de dados em invocações de funções na linguagem nativa do sistema – *host language*. Actua em conjunto com o módulo de processamento de interrogações de forma a garantir a geração do código apropriado.

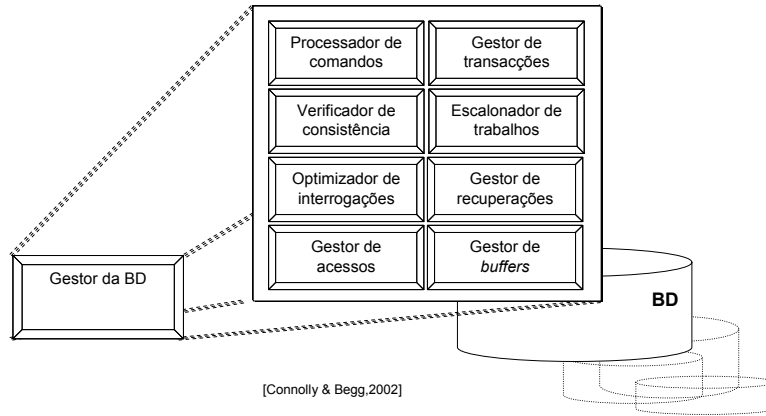


Componentes de um SGBD

- Compilador da linguagem de descrição de dados.
 - Converte as instruções de descrição de dados num conjunto de tabelas, contendo os metadados respectivos, que serão depois armazenadas no catálogo do sistema. A informação de controlo associada será armazenada nos cabeçalhos dos ficheiros de dados.
- Gestor do catálogo do sistema.
 - Gere os acessos ao catálogo do sistema e trata da sua manutenção.



O Gestor da Base de Dados



2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

25

Componentes do Gestor da BD

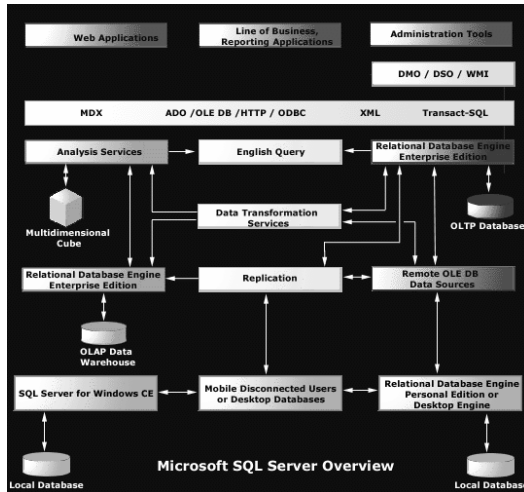
- Processador de comandos
- Verificador de consistência
- Optimizador de interrogações
- Gestor de acessos
- Gestor de transacções
- Escalonador
- Gestor de recuperações
- Gestor de "buffers"



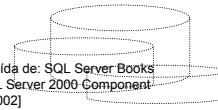
2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

26

A Arquitectura Global do SQL Server



[Figura extraída de: SQL Server Books OnLine, SQL Server 2000 Component Overview, 2002]



2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

27

Arquitecturas Multiutilizador para SGBD

- Teleprocessamento
 - Arquitectura tradicional para sistemas multiutilizador
 - Configuração típica
 - Processamento centralizado
 - Terminais sem autonomia e capacidade de processamento
 - Carga excessiva na máquina central
 - Arquitectura preterida em prol das mais recentes
- Servidores de ficheiros (file-servers)
 - Processamento distribuído pela rede
 - As aplicações e os SGBD residem em cada posto
 - O servidor de ficheiros actua simplesmente como um dispositivo de armazenamento partilhado.
 - Tráfego significativo na rede
 - Cópias completas do SGBD em cada posto
 - Processos de gestão (recuperação, integridade, ...) mais complexos

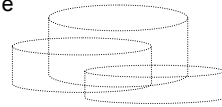


2003, O.Belo, *Sistemas de Bases de Dados, Introdução.*

28

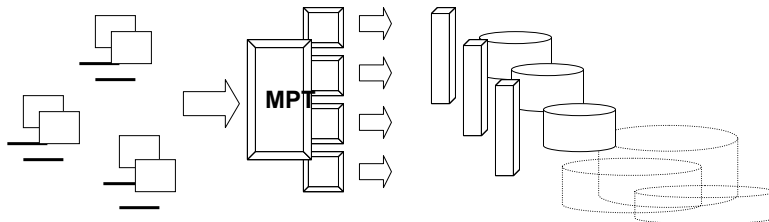
Arquitecturas Multiutilizador para SGBD

- Clientes-Servidor
 - Dois componentes: o cliente e o servidor.
 - Acesso alargado a dados e serviços
 - Aumento de desempenho e da consistência
 - Custos de hardware mais reduzidos
 - Custos de comunicação menores
 - Cliente – “Front-end”
 - Gere o interface
 - Suporta a lógica das aplicações
 - Geram os pedidos para o servidor
 - Servidor – “back-end”
 - Aceitam e processam os pedidos de base de dados
 - Verificação as autorizações e as regras de integridade
 - Execução de interrogações e actualizações
 - Faz a manutenção do catálogo dos sistema
 - Garante o acesso concorrente e recuperação.



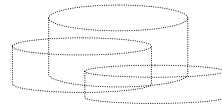
Monitores de Processamento de Transacções

- Direcção de transacções (“routing”).
- Gestão de transacções distribuídas.
- Balanceamento de pedidos (“load balancing”).
- Canalização de pedidos (“funneling”).
- Aumento da fiabilidade do sistema.



Catálogos de Sistema

- O catálogo de um SBD é constituído pela informação necessária à descrição dos dados contidos numa BD. Esta informação também é designada por metadados ou dados acerca dos dados.
- A informação contida no catálogo é constituída por:
 - As designações dos utilizadores com autorização para acederem ao SGBD e aos dados contidos na BD.
 - As designações dos diversos itens de dados contidos na BD.
 - ...



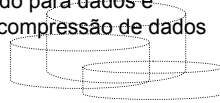
A arquitectura ANSI-SPARC de 3 níveis

- Este modelo de arquitectura foi introduzido em 1975 pela American National Standards Institute (ANSI) e Standards Planning and Requirements Committee (SPARC) que reconheceram a necessidade de um modelo a três níveis juntamente com um catálogo do sistema.
- O modelo ANSI-SPARC não se tornou um standard tal como era desejado, mas fornece as bases para o entendimento de algumas das funcionalidade de um SGBD.
- Este modelo é importante porque nos possibilita a categorização de três níveis de abstracção para um SGBD, o que significa três níveis distintos para a descrição de elementos de dados.

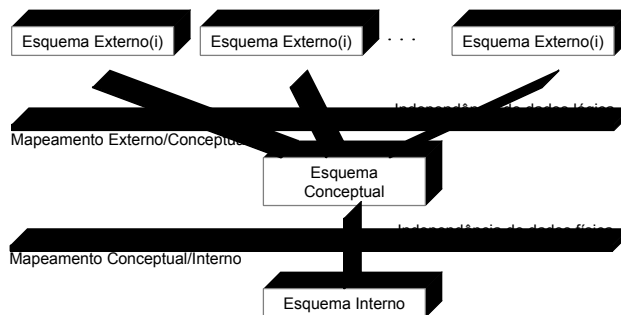


Os 3 níveis de abstracção

- Esquema Externo
 - Este nível representa todas as vistas dos utilizadores sobre a base de dados, descrevendo todas as partes da BD que são importantes para os utilizadores.
- Esquema Conceptual
 - Apresenta o esquema global da base de dados à sua comunidade de utilizadores; é uma descrição relacionada com os dados que estão armazenados na base de dados e os relacionamentos entre si.
 - Representa todas as entidades assim como os seus atributos e relacionamentos, as restrições sobre os dados, informação semântica acerca dos dados e informação e critérios relacionados com a segurança e a integridade dos dados.
- Esquema Interno
 - Está relacionado com o espaço de armazenamento alocado para dados e índices, descrições de registos, localização dos registos, compressão de dados ou técnicas de encriptação de dados.

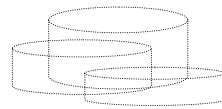


A arquitectura dos 3 níveis



Independência de dados

- Independência lógica de dados.
- Independência física de dados.



Bibliografia

- Connolly, T., Begg, C., Database Systems - A Practical Approach to Design, Implementation, and Management, III Edição, Addison-Wesley, 2002.
- Teorey, T., Database Modeling and Design: The Fundamental Principles, II Edição, Morgan Kaufmann, 1994.
- Date C., An Introduction to Database Systems , Volume I, VI Edição, Addison-Wesley Systems Programming Series, 1996.
- Hansen, G., Hansen, J., Database Management and Design, II Edition, Prentice-Hall, Inc, 1996.
- Ramakrishnan, R., Database Management Systems, McGraw-Hill International Editions, 1998.

