
 universidade
de aveiro **deti** departamento de
electrónica, telecomunicações
e informática

Introdução aos Sistemas de Base de Dados

Base de Dados - 2022/23
Carlos Costa

(Adapted from several DB courses and Books - see bibliography)

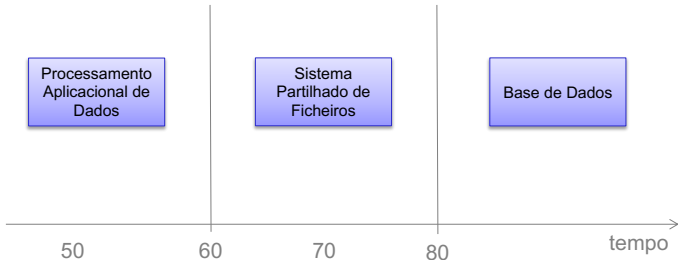
1

 deti

Base de Dados - Conceito

- **Base de Dados (BD)**: uma coleção organizada de dados que estão relacionados e que podem ser partilhados por múltiplas aplicações.

Evolução



2

2

deti

Processamento Isolado de Dados

- **Dados isolados** - cada aplicação gere os seus próprios dados.
- Os mesmos dados podem estar replicados.
- Diferentes organizações e formatos de dados.
- Problemas de “sincronismo” -> incoerências.

Relacionados

3

3

deti

Sistema de Gestão de Ficheiros

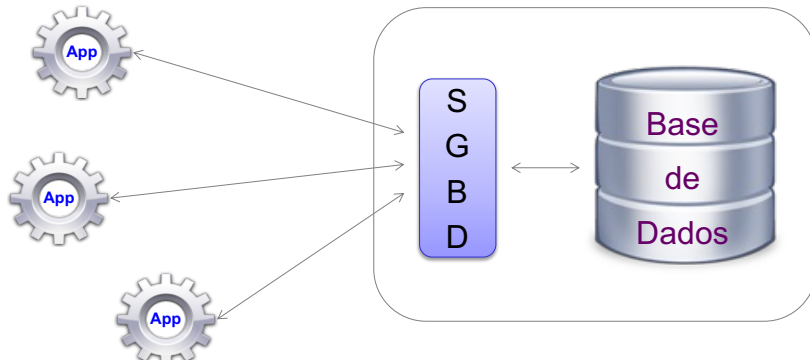
- Dados organizados e armazenados em ficheiros partilhados por várias aplicações.
- Cada aplicação acede diretamente aos ficheiros.
- Cada aplicação usa uma interface proprietária.
- Problemas:
 - Acesso concorrente aos dados
 - Integridade
 - Segurança

4

4

deti

Sistema de Gestão de Base de Dados (SGBD)



Database Management System (DBMS): “is a general-purpose software system that facilitates the processes of *defining*, *constructing*, *manipulating*, and *sharing* databases among various users and applications.”

5

5

deti

Sistema de Gestão de Base de Dados (SGBD)

- Definição (*Defining*)
 - Especificação do tipo de dados, estruturas de dados e restrições
 - database catalog or dictionary
 - Tipo
 - Estrutura
 - Restrição
- Construção (*Constructing*)
 - Processo de armazenamento de dados
- Manipulação (*Manipulating*)
 - Envolve operações como a pesquisa e obtenção de dados
- Partilha (*Sharing*)
 - Acesso simultâneo aos dados por parte de vários utilizadores e programas

6

6

SGBD - Características Gerais



- Entidade única que opera com a BD
 - O acesso à BD é sempre mediado pelo SGDB
- Existe uma interface de acesso que esconde os detalhes de armazenamento físico dos dados
- Elevada abstração ao nível aplicativo
- Os dados estão integrados (nível lógico) numa mesma unidade de armazenamento
- Suporta uma ou mais BD
- Keyword - Data Independence

7

7

SGBD - Vantagens



- Independência entre programas e dados
- Integridade dos dados
 - Controlo de alteração de dados de acordo com as regras de integridade definidas
- Consistência dos dados
 - Nos processos de transações e mesmo em falhas de software/hardware
- Eficiência no acesso aos dados
 - Especialmente em cenários de manipulação de grandes quantidades de dados, por um ou mais utilizadores
- Isolamento utilizadores
 - Cada utilizador tem a “sensação” de ser o único

8

8

SGBD - Vantagens (cont.)



- Melhor gestão do acesso concorrential
- Serviços de Segurança
 - Controlo de Acessos / Permissões
 - Codificação de Dados
- Mecanismos de backup e recuperação de dados
- Administração de dados
 - Disponibilidade de ferramentas desenvolvidas pelo fabricante e/ou terceiras entidades
- Linguagem de desenho e manipulação de dados

Nota: Muitas das vantagens anteriores são também requisitos funcionais de um SGBD.

9

9

SGBD - Potenciais Desvantagens

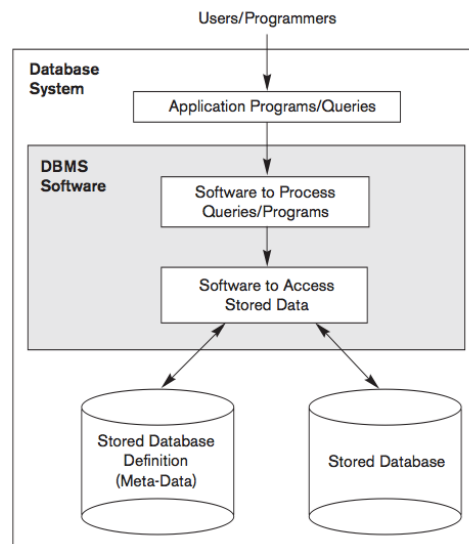


- Maiores custos e complexidade na instalação e manutenção
 - Especial em soluções empresariais
- Não respondem aos requisitos de alguns cenários aplicacionais
 - Por exemplo, pesquisa de texto
 - Motivou o aparecimento de novos modelos (NoSQL, IndexEngine, etc)
- Centralização dos dados mais suscetível a problemas de tolerância a falhas (software e hardware) e de escalabilidade

10

10

SGBD - Vista Simplificada



11

11

SGBD - Utilizadores

- **Utilizadores Finais**
 - aqueles que usam o sistema com determinada finalidade com recurso a ferramentas disponibilizadas pelo fabricante do sistema ou aplicações de terceiras entidades.
- **Programadores de Aplicações**
 - Desenvolvem aplicações que permitem que os utilizadores interajam com a base de dados. Podem utilizar várias linguagem de programação.
- **Administradores da Base de dados**
 - Tratam dos processos de gestão e manutenção da base de dados.

12

12



SGBD - Metadados

- Metadados (dados sobre dados)
- O SGBD armazena uma descrição da própria estrutura da base de dados, restrições de integridade e condições de acesso.
 - Descritores de objetos da base de dados (tabelas, utilizadores, regras, vistas, indexes, etc)
 - Informação sobre dados em uso e por quem (locks).
 - *Schemas e mappings*

13

13



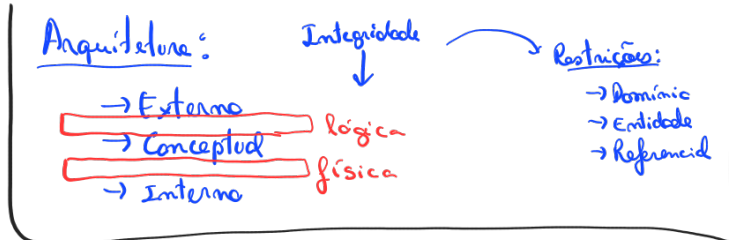
Interfaces (Aplicações)

- Web-based
- Form-based (desktop)
- GUI (Graphical User Interface)
 - Manipulação visual de esquemas de BD com recurso a diagramas. Possibilidade de construção e execução de queries.
- Natural Query Language
- DBMS Command Line
 - Criar contas de utilizadores, parametrizar o sistema, definir permissões e privilégios, definir/alterar estruturas de dados, definir tipos de dados, etc.
 - Utilizando uma linguagem própria - SQL

14

14

Vamos aprender mais à frente →

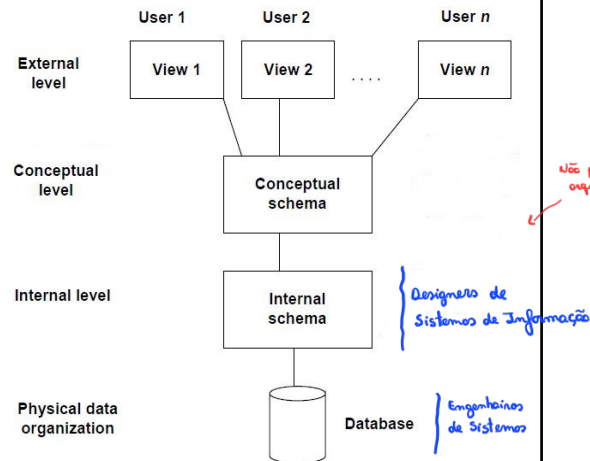


2/16/23

SGBD - Arquitetura ANSI/SPARC¹

Three-level architecture:

- **External level**
database users
- **Conceptual level**
database designers and administrators
- **Internal level**
systems designers



1. ANSI/X3/SPARC Study Group on Data Base Management Systems: (1975), Interim Report. FDT, ACM SIGMOD bulletin, Volume 7, No. 2

15

15

ANSI/SPARC - Nível Interno

- Lida com a implementação física da BD
 - Estrutura dos registos em disco - files, pages, blocks
 - Indexes e ordenação dos registos
- Domínio: Programadores de sistemas de BD
- Exemplo de Esquema

```

RECORD FUNCIONARIO
  LENGTH=44
  HEADER: BYTE (5)
  OFFSET=0
  NOME: BYTE (25)
  OFFSET=5
  SALARIO: FULLWORD
  OFFSET=30
  DEPARTAMENTO: BYTE (10)
  OFFSET=34
  
```

16

16



ANSI/SPARC - Nível Conceptual

- Esquema Conceptual - descreve a estrutura da base de dados para os utilizadores
 - Descreve entidades, tipo de dados, relações, operações, restrições, etc
 - Utiliza (tipicamente) um modelo de dados para descrição do esquema conceptual
- Oculta detalhes de implementação física(abstração)
- Domínio: Administrador BD e prog. de aplicações
- Exemplo de esquema

```
CREATE TABLE FUNCIONARIO
  (Nome VARCHAR(25) ,
   Salario REAL, Dept_Nome VARCHAR(10))
```

17

17



ANSI/SPARC - Nível Externo

- Oferece vistas da base de dados adaptadas a casa utilizador
 - Apresentação dos dados pode ser trabalhada, parte dos dados pode ser ocultada, etc.
- Domínio: Utilizadores finais e prog. de aplicações
- Exemplo de Esquema

```
FolhaPagamentos:
  char *Nome
  double Salario

Funcionarios:
  char *Nome
  char *Departamento
```

18

18

ANSI/SPARC - Independência dos dados

- A alteração do esquema (*schema*) de um nível não tem impacto no esquema do nível acima.

=> Dois níveis de independência

■ Nível Físico

- Alterações do nível físico não devem ter impacto no esquema conceptual.
- Por exemplo, podemos alterar a forma como armazenamos os dados no sistema de ficheiros por razões de desempenho.

■ Nível Lógico

- Alterações no esquema conceptual (modelo de dados) não devem repercutir-se nos esquemas externos ou aplicações já desenvolvidas.

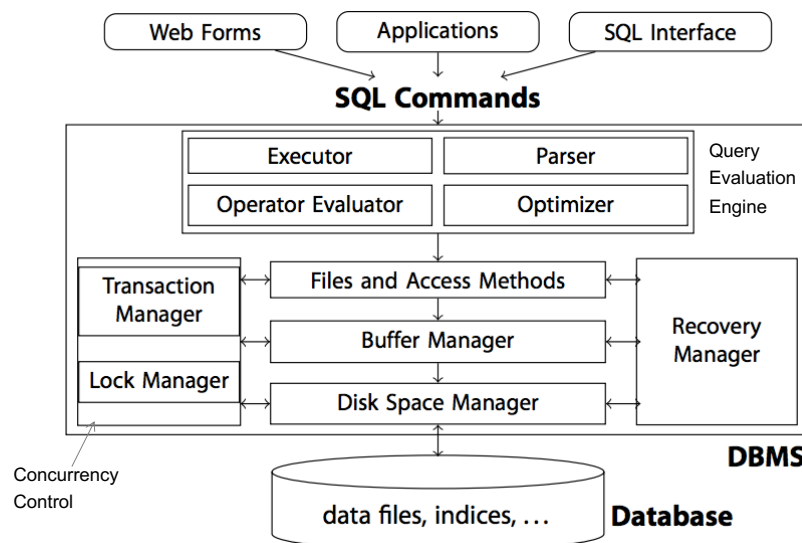
A responsabilidade da implementação é nossa!

Podemos mudar a lógica mas não as saídas...

19

19

SGBD - Arquitetura Típica



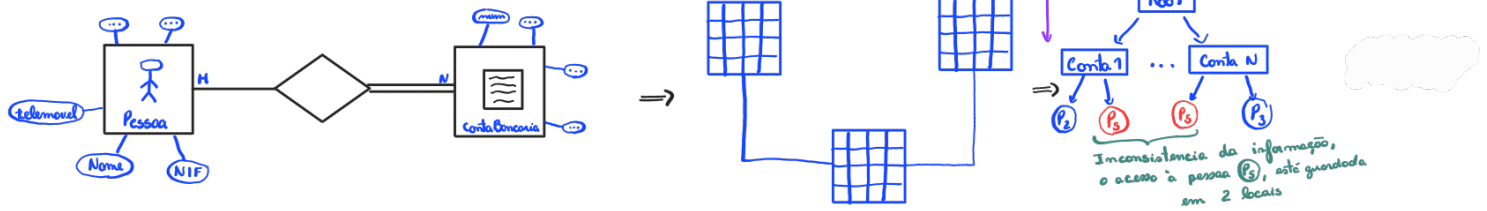
20

R. Ramakrishnan and J. Gehrke. Database Management Systems. McGraw-Hill.

20

I/O → limita escalon na vertical um equipamento

• Desenho conceptual:



Modelo de Base de Dados

- Modelo de BD - coleção de conceitos para descrição lógica de dados (Modelo Lógico)
- **Esquema (Schema)**: a descrição de um conjunto particular de dados com recurso a um determinado modelo
- Um bom modelo de dados é fundamental para garantir a independência dos dados
- O **Modelo Relacional** é um dos mais utilizados nos dias de hoje.
 - Bancos, Hospitais, Finanças, Seguradoras, etc

21

21

Modelos de Base de Dados

- 1ª Geração (Pré-relacional)
 - Hierárquico
 - Rede
- 2ª Geração
 - **Relacional**
- 3ª Geração (Pós-relacional)
 - Object-relational
 - Object-oriented
 - Key-value store
 - Document-oriented
 - Column-oriented
 - Graph database



22

22

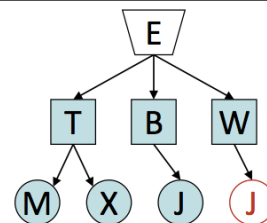
Modelos NoSQL

Phases	Targets	
Conceptual Design	Conceptual Data Model	ERD(Entity Relationship Diagram) UML(Unified Modeling Language) ORM(Object Role Modeling) FCO-IM(Fully Communication Oriented Information Modeling)
Logical Design	NoSQL Data Model	Key-Value , Document, Column Family , Graph
Physical Design	NoSQL Database	Key-Value Riak, Redis, Memcached ,Berkeley DB ,Hamster DB, Amazon Dynamo DB ,Project Voldemort
		Document MongoDB, Couch DB, Terrastore, Orient DB, Raven DB
		Column Family Cassandra, HBase, Hypertable, Amazon Simple DB
		Graph Neo4J, Infinite Graph, Orient DB, Flock DB

23

23

Modelo Hierárquico

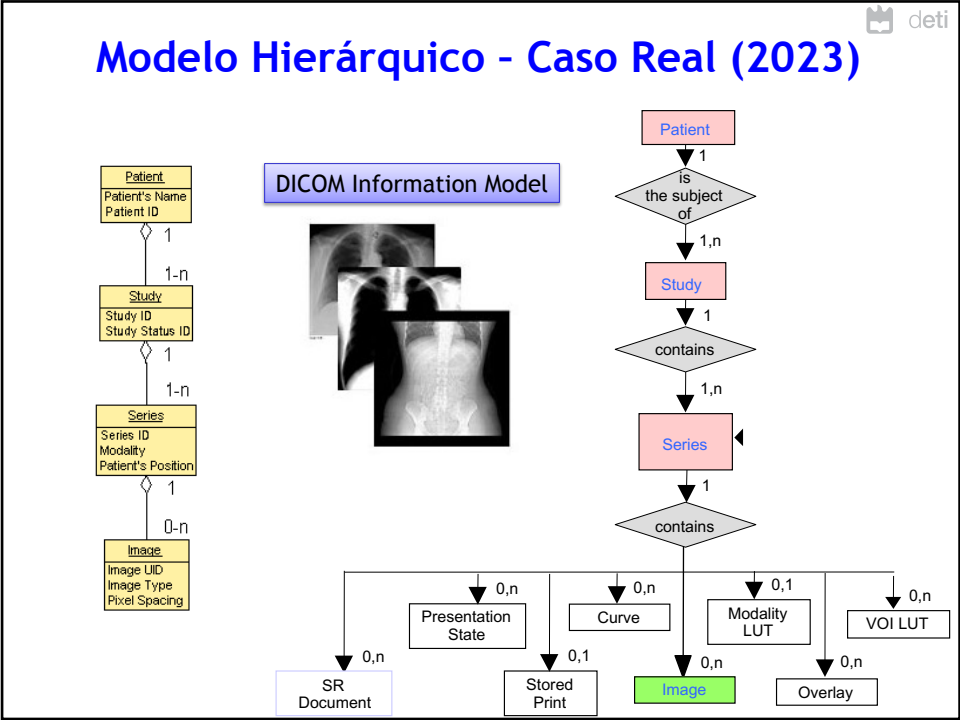


- Dados estão armazenados numa **estrutura** hierárquica (**árvore**).
- Os nós da árvore designa-se como **registos** que estão ligados por **ponteiros** (**links**).
- Um registo é composto por um conjunto de atributos.
- Um link é uma associação entre dois registos do tipo **pai-filho**.
- Um registo pai encontra-se associado a N registos filhos (**1:N**).

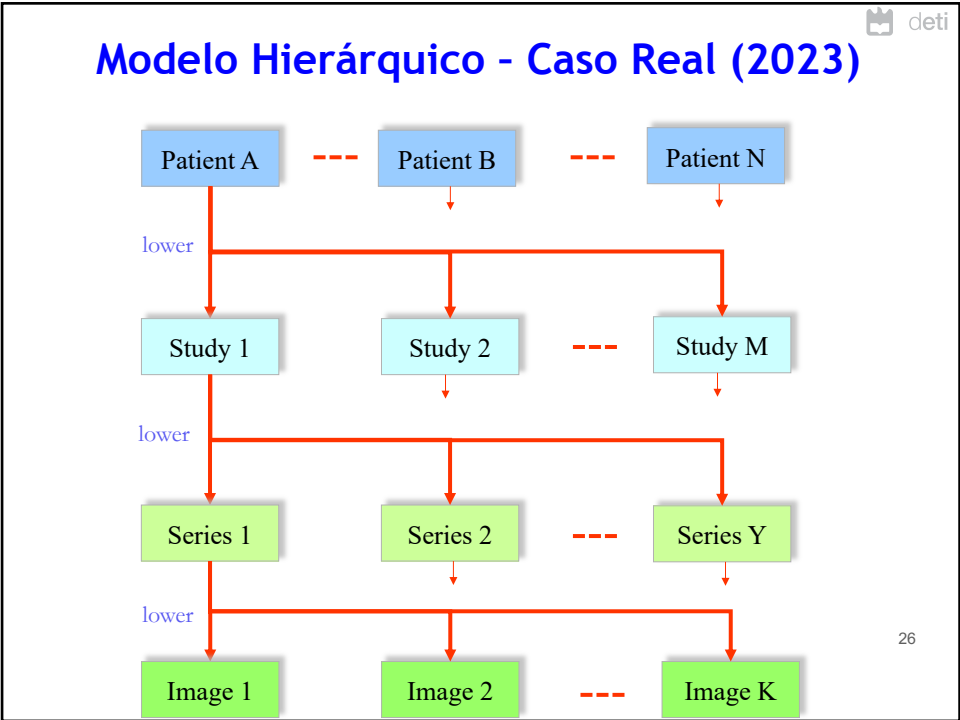
24

24

Não é bom para relações
N:M



25



26

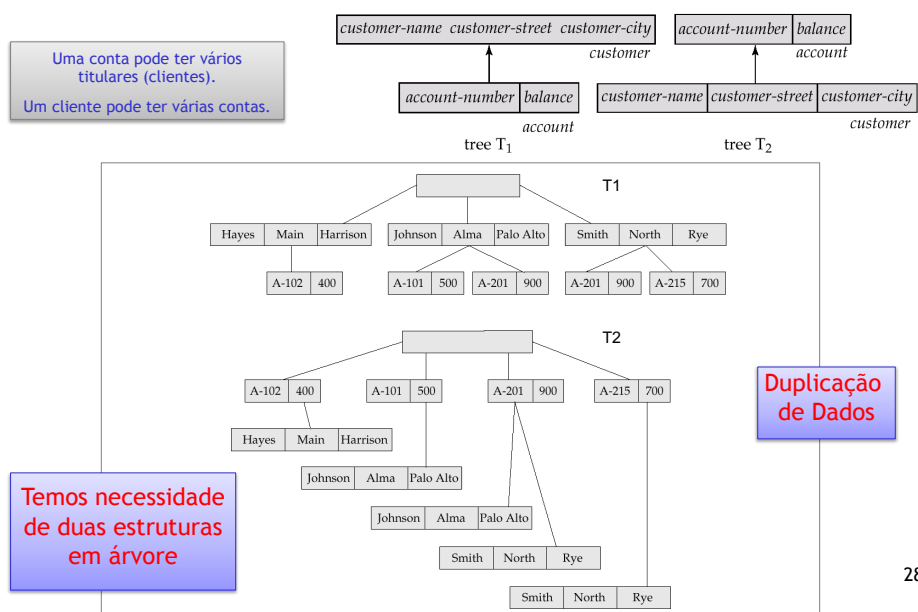
Modelo Hierárquico - (Des)vantagens

- Adaptado a cenários de acesso sequencial aos dados.
 - Qualquer acesso aos dados passa sempre pelo segmento raiz.
 - A maior parte das necessidades atuais requer acesso aleatório!
- Redundância de informação
 - Desperdício de espaço e inconsistências de dados
- Restrições de integridade, exemplo:
 - A eliminação de um segmento pai, implica a remoção de todos os segmentos filhos associados.
- Não permite estabelecer associações N:M

27

27

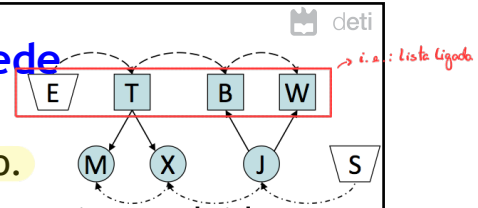
Modelo Hierárquico - Relação N:M



28

28

Modelo de Rede

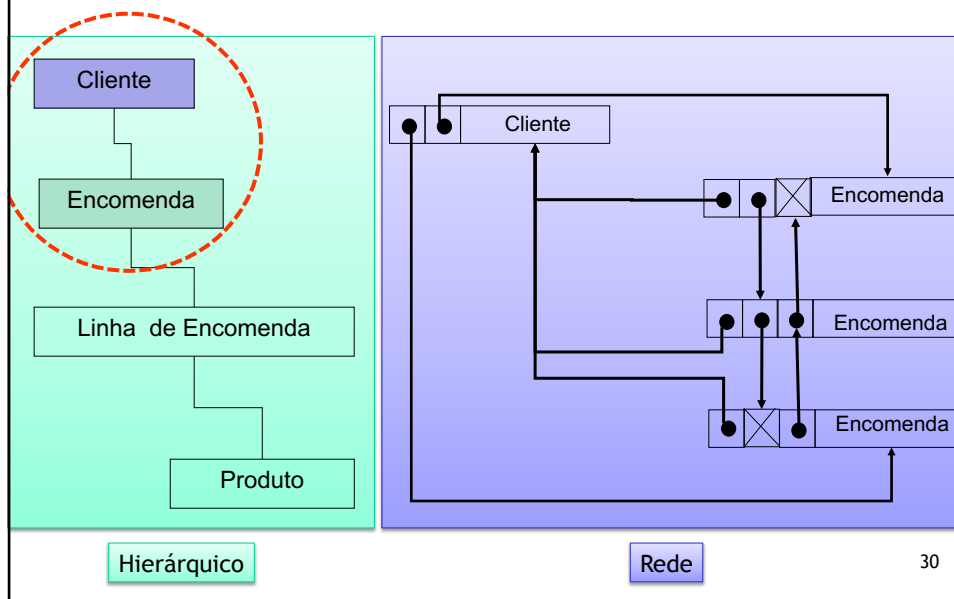


- Extensão do modelo hierárquico.
- Permite que um mesmo registro esteja envolvido em várias associações -> visão de rede.
- Melhorias na capacidade de navegação na estrutura de dados.
- Relações representadas através de grafos.
- Um conjunto (set) suporta associação entre registros do mesmo tipo
 - Tipicamente implementados com listas ligadas circulares
 - Relacionamento 1:N entre dois tipos de registro.

29


29

Modelo Hierárquico -> Rede



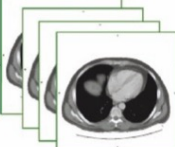
30

30




Modelo Hierárquico versus Relacional

Hierarchical		Relational
1. Find all patients with name "John Smith"	Patient Level key: Patient ID, (0010,0020)	1. Find all images with Patient Name=John Smith and Study Date=last month and Modality=CT
2. For each found patient, find all studies from the last month	Study Level key: Study Instance UID, (0020,000D)	
3. For each found study, select those with CT series	Series Level key: Series Instance UID, (0020,000E)	
4. For each found series, get the images	Image Level key: SOP Instance UID, (0008,0018)	



31

31



SGBD

ALGUMAS CURIOSIDADES...

32

32

deti

DB-Engines Ranking - Engine

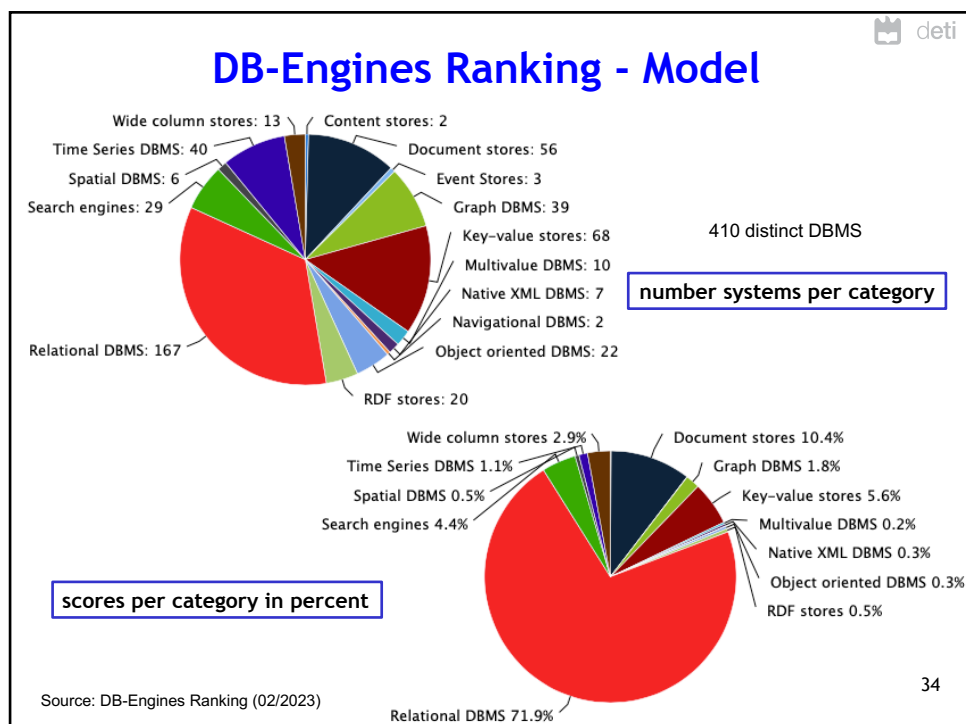
410 systems in ranking, February 2023

Rank			DBMS	Database Model	Score		
Feb 2023	Jan 2023	Feb 2022			Feb 2023	Jan 2023	Feb 2022
1.	1.	1.	Oracle	Relational, Multi-model	1247.52	+2.35	-9.31
2.	2.	2.	MySQL	Relational, Multi-model	1195.45	-16.51	-19.23
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	929.09	+9.70	-19.96
4.	4.	4.	PostgreSQL	Relational, Multi-model	616.50	+1.65	+7.12
5.	5.	5.	MongoDB	Document, Multi-model	452.77	-2.42	-35.88
6.	6.	6.	Redis	Key-value, Multi-model	173.83	-3.72	-1.96
7.	7.	7.	IBM Db2	Relational, Multi-model	142.97	-0.60	-19.91
8.	8.	8.	Elasticsearch	Search engine, Multi-model	138.60	-2.56	-23.70
9.	10.	10.	SQLite	Relational	132.67	+1.17	+4.30
10.	9.	9.	Microsoft Access	Relational	131.03	-2.33	-0.23
11.	12.	11.	Cassandra	Wide column	116.22	-0.09	-7.76

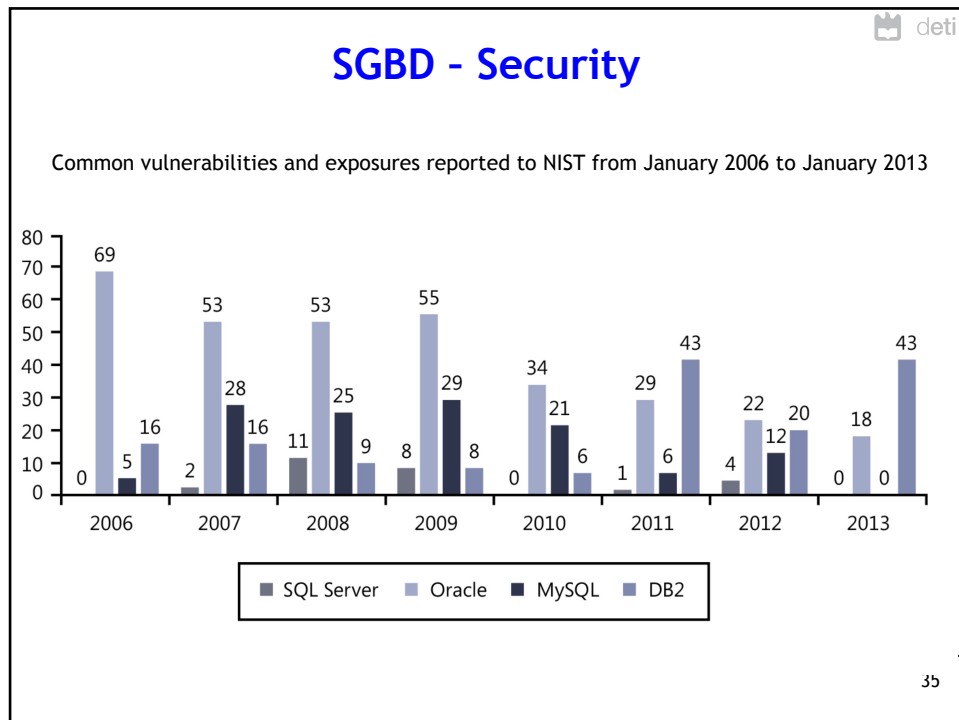
Ranks database management systems according to their popularity.
Source: DB-Engines Ranking (02/2023)

33

33



34



35

Resumo

- Introdução aos Sistemas de Base de Dados
- Sistemas Gestores de Base de Dados
- Modelos de Base de Dados

36

36