

## Organizando os dados em um Banco de Dados

Um Banco de Dados deve ser projetado para armazenar todos os dados relevantes para a empresa e fornecer acesso rápido e modificações fáceis. Na construção de um Banco de Dados, deve-se dar cuidadosa consideração a estas questões:

- Conteúdo: Que dados devem ser coletados e a que custo?
- Acesso: Que dados devem ser fornecidos a qual usuário quando apropriado?
- Estrutura Lógica: Como os dados devem ser arrumados de forma que façam sentido para um determinado usuário?
- Organização Física: Onde os dados devem estar fisicamente localizados?

Uma das ferramentas usadas pelos projetistas de Bancos de Dados para mostrar a relação lógica entre os dados é o modelo de dados. A estrutura das relações na maioria dos bancos de dados segue um dos três modelos lógicos: hierárquico, em rede e relacional. Os modelos hierárquicos e em rede ainda são usados atualmente, mas os modelos relacionais são os mais populares.

### Evolução de Modelos de Dados

- Redes (1964)
- Hierárquico
- ambos modelos fortemente influenciados por características físicas do banco de dados
- Relacional
  - proposto em 1970
  - disponível comercialmente em meados dos anos 80
- Modelos mais recentes
  - Semânticos
  - Funcionais
  - Relacionais aninhados
- Foco de pesquisa Atual
  - Objetos-relacionais, Orientados-a-Objetos
  - Objetos & Funções, Objetos & Lógica

## 1. Modelo Hierárquico ou de árvore

É aquele no qual os dados estão organizados de cima para baixo ou estrutura de árvore invertida. Por exemplo, os dados sobre um projeto para uma empresa podem seguir este tipo de modelo. Este método de ligação é semelhante à relação entre pais e filhos: a criança não existira sem os pais. É o que mais bem se adapta a situações nas quais as relações lógicas entre os dados podem ser representadas com a abordagem (um-para-muitos).

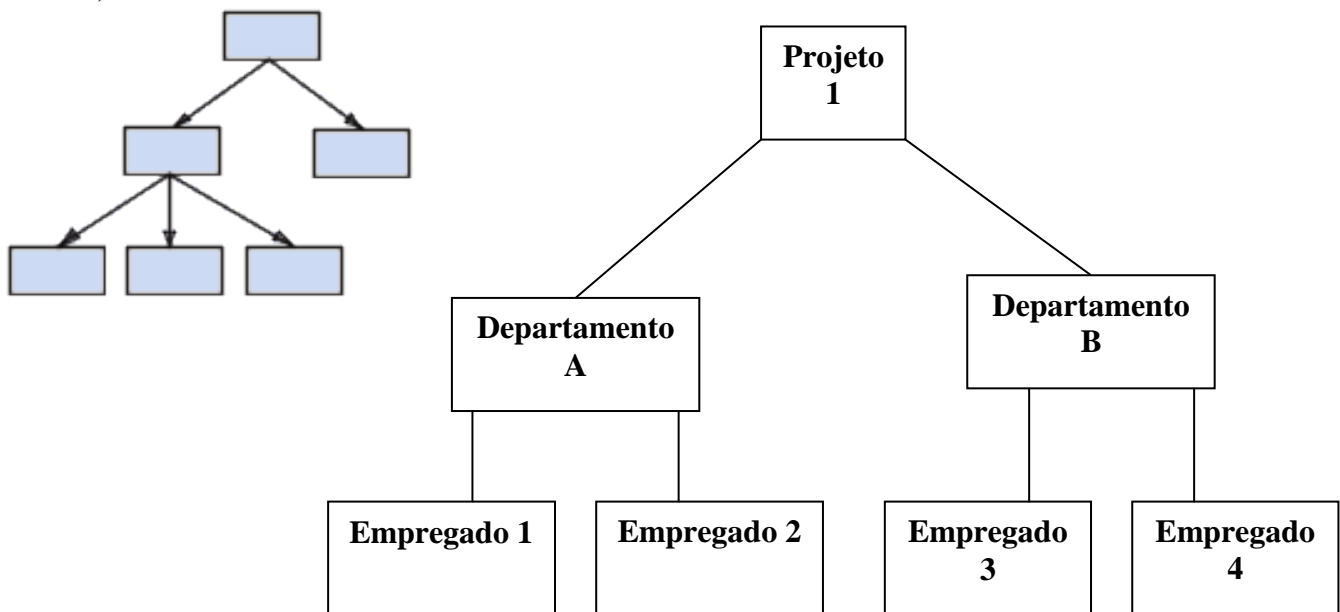


Figura 1 – Modelos de banco de dados Hierárquico

- Características
  - Os dados são organizados em vários tipos de registros
  - Relacionamentos explícitos entre registros: pai, filhos, irmãos
  - Registros devem ser vistos dentro de um contexto
- Restrições de integridade inerentes ao modelo
  - Ocorrências registros filhos só podem existir ligadas a algum pai –se um registro filho tem dois ou mais pais do mesmo tipo de registro, o registro filho deve ser duplicado para cada pai
  - Para um registro ter mais de um pai de tipos de registros diferentes, um deve ser o pai real e os outros devem ser pais virtuais.
- Vantagens do modelo hierárquico
  - Natural para organizações hierárquicas de dados (taxonomias, estruturas organizacionais, objetos compostos nos domínios da biologia e manufatura.)
  - Restrições de integridade garantidas pelo modelo adequadas a uma realidade hierárquica
  - Simplicidade de processamento
  - Agrupamento hierárquico natural (clustering)
- Sistemas
  - IBM's IMS (DL/1, IMS DB, IMS DC)
  - SYSTEM 2000

## 2. Modelo em Rede

Um modelo em rede é uma extensão do modelo hierárquico. Em vez de se terem apenas vários níveis de relações um-para-muitos, o modelo em rede é uma relação membro-proprietário, na qual um membro pode ter muitos proprietários. Nesse modelo, há frequentemente mais de um caminho pelo qual um determinado elemento de dado pode ser acessado.

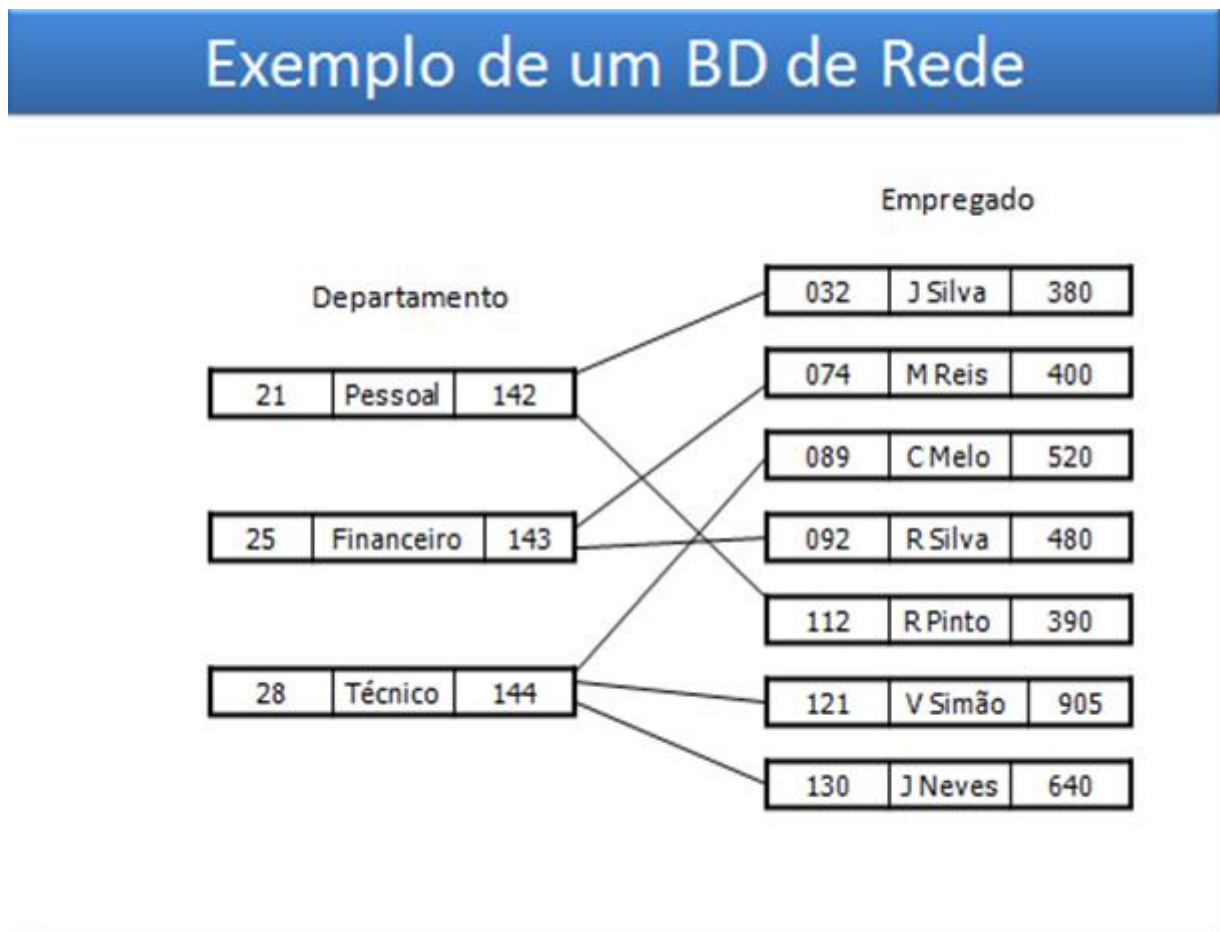
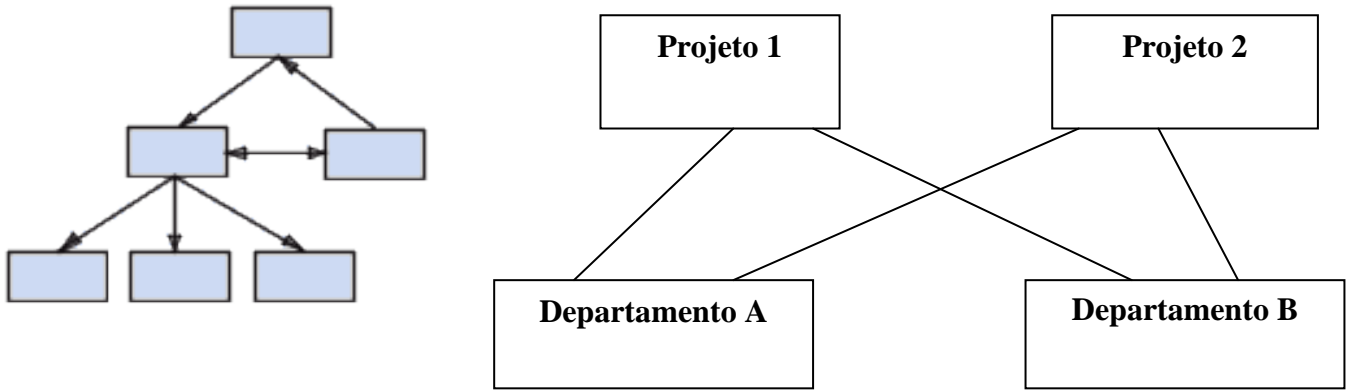


Figura 2 – Modelos de banco de dados em Rede

- Inicialmente apresentado no relatório do grupo de trabalho
  - CODASYL (Conference on Data Systems Language), 1971
  - Chamado modelo DBTG (CODASYL Data Base Task Group)
  - Revisado em 1978 e 1981
- ANSI (American National Standards Institute) apresentou uma proposta de padronização de uma linguagem de definição de redes (NDL) em 1984
- Características
  - Os dados são organizados em vários tipos de registros e tipos de SETs (estabelecem os relacionamentos)
  - Maior simetria nas consultas
  - Problema de estratégia de acesso aos registros “filhos”
- Restrições de integridade inerentes ao modelo
  - Permite especificar restrições de integridade relativas a
    - Inserção de registros em relacionamentos
      - Automática X manual
  - Permanência de registros no banco de dados, dependendo de seus relacionamentos
- Vantagens
  - Melhor apresentação de dados estruturados como grafo
  - Modela comportamento de relacionamentos, especificando
  - Restrições de inserção/remoção/modificação
  - Otimização através de “navegação inteligente”
- Sistemas
  - IDS (Honeywell), DMS 1100 (Univac), TOTAL (cincom), IDMS (Goodrich-Cullinet-CA), VAX-DBMS (digital)

### 3. Modelo Relacional

Os modelos relacionais se tornaram os mais populares. A finalidade global deste modelo é descrever o dado usando um formato tabular padrão (todos os elementos são localizados em tabelas bidimensionais). As tabelas organizam os dados em linhas e colunas, simplificando o acesso e a manipulação dos dados.

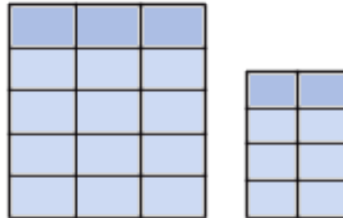


Figura 3 – Modelo de banco de dados Relacional

Uma vez colocados os dados no Banco de Dados relacional, pode-se fazer perguntas e manipular dados. As manipulações básicas de dados incluem a sua seleção, projeção e agrupamento.

Num. do projeto	Descrição	Num. do Departamento
155	Atlanta	257
498	Gênesis	632
226	Futurus	598

Num. do Departamento	Nome do Departamento	CPF do Gerente
257	Contabilidade	616369595-04
632	Vendas	595042158-75
598	Marketing	452158754-26

CPF do Gerente	Nome	Data Contratação	Num Departamento
258485745-22	Fred	15/06/1975	257
152425857-14	Jorge	15/04/1978	662
452158754-26	Beto	15/05/2001	598

Suponhamos que o presidente da empresa queira encontrar o nome do gerente do projeto Futurus e há quanto tempo ele está na empresa. Devido à ligação entre as tabelas, a resposta é precisa e rápida (BETO – 10 anos).

Vejam os exemplos abaixo: Será que podemos assegurar que estes dados são “bons”? Ou seja, não redundantes, flexíveis, simples.

Nome	Telefone	Sexo	Mensalidade	Data de Pagamento
Jorge	4221326	Masculino	R\$ 30,00	15/09
Marcelo	4211585	Masculino	R\$ 15,00	15/09
Marcelo	4214747	Masculino	R\$ 15,00	25/09

Este Banco de Dados controla o pagamento das mensalidades de uma academia. Perceba que ele está cheio de falhas ou anomalias. Precisamos normalizar as tabelas, que é um processo que envolve o desdobramento de um arquivo em dois.

Num Sócio	Nome	Telefone	Sexo
001	Jorge	4221326	Masculino
002	Marcelo	4211585	Masculino

Num Sócio	Mensalidade Paga	Data de Pagamento
001	R\$ 30,00	15/09
002	R\$ 15,00	15/09
002	R\$ 15,00	25/09

- Introdução por E.F. Codd (1970)
- Características
  - Base de dados é visualizada como um conjunto de tabelas, cada uma representando uma relação
  - Relacionamentos representados por valores de dados
  - Tabelas consistem de linhas → tuplas e colunas → atributos
  - Simetria nas consultas
  - Linguagens definidas:
    - álgebra e cálculo relacionais
    - “transformados” em SQL com ampla aceitação
    - SQL -linguagem padrão para bancos de dados relacionais

### Comparação entre os modelos

Cada um dos modelos que discutimos é usado por uma variedade de organizações, e cada um tem suas vantagens e desvantagens. Um Banco de Dados Hierárquico pode tomar menos tempo para ser manipulado do que os outros modelos, porque as relações dos dados são menos complexas, com cada filho tendo apenas um pai. Este modelo não é flexível em termos de como os dados são organizados e são difíceis de modificar.

O modelo em rede é mais flexível do que o hierárquico em termos de organização de dados. Entretanto, é mais difícil de desenvolver e usar, devido à complexidade das relações dos dados.

O modelo relacional é o mais usado, pois é fácil de controlar, mais intuitivo, porque organiza os dados em tabelas.