

Área do Conhecimentos de
Ciências Exatas e Engenharias

Introdução ao Sistema de Banco de Dados

Prof. Daniel L. Notari

Agosto - 2022

Dados

Dados

- ▶ São um conjunto de informações (quantitativas, qualitativas, categóricas ou indefinidas) podendo ser organizadas ou não.
- ▶ Podem ter
 - um tipo de dados associado.
 - um tamanho associado.

Dados brutos

- ▶ Designam os dados/valores recolhidos e estocados tal qual foram adquiridos, sem terem sofrido o menor tratamento.
- ▶ Apresentam-se como um conjunto de números, caracteres, imagens ou outros dispositivos de saídas para converter quantidades físicas em símbolos, num sentido muito extenso.

Dado: à luz da Ciência da Informação

- ▶ Uma sequencia de símbolos quantificados ou quantificáveis.
- ▶ Um texto é um dado.
- ▶ As letras são símbolos quantificados, já que o alfabeto por si só constitui uma base numérica.
- ▶ Tabela ASCII.

O dado, a informação e o conhecimento

- ▶ O dado na é a representação física de um evento no tempo e espaço que não agrega fundamento para quem o sente ou recebe, não podendo ser possível entender o que ele representa ou para que ele existe.
- ▶ A informação é o dado organizado, ou seja, um dado bruto que sofreu um tratamento.

O dado a informação e o conhecimento

Dado	Informação	Conhecimento
Simple observações sobre o estado do mundo	Dados dotados de relevância e propósito	Informação valiosa da mente humana Inclui reflexão, síntese, contexto
<ul style="list-style-type: none">• Facilmente estruturado• Facilmente obtido por máquinas• Frequentemente quantificado• Facilmente transferível	<ul style="list-style-type: none">• Requer unidade de análise• Exige consenso em relação ao significado• Exige necessariamente a mediação humana	<ul style="list-style-type: none">• De difícil estruturação• De difícil captura em máquinas• Frequentemente tácito• De difícil transferência

Requisitos Funcionais

- ▶ Descrevem o comportamento e a informação que a solução irá gerenciar.
- ▶ Descrevem capacidades que o sistema será capaz de executar em termos de comportamentos e operações - ações ou respostas específicas de aplicativos de tecnologia da informação.

Requisitos Não Funcionais

- ▶ Capturam condições que não se relacionam diretamente ao comportamento ou funcionalidade da solução, mas descrevem condições ambientais sob as quais a solução deve permanecer efetiva, ou qualidades que os sistemas precisam possuir.
- ▶ Também são conhecidos como requisitos de qualidade ou suplementares.
 - Podem incluir requisitos relacionados à capacidade, velocidade, segurança, disponibilidade, arquitetura da informação e apresentação da interface com o usuário.

Sistemas de Arquivos

Conceitos

- Um arquivo
 - é formado por uma coleção de registros lógicos, cada um deles representando um objeto ou entidade.
- Um registro lógico
 - uma sequencia de itens, sendo cada item chamado campo ou atributo.
- Cada campo
 - possui um nome, um tipo (numérico, lógico, literal) e um comprimento (número de palavras, bytes ou bits).

Operações Básicas de Manipulação de Arquivos

- Criação, Abertura e Fechamento
- Inserção de Registros
- Consulta de Registros (Acesso a Registros)
- Alteração de Registros
- Exclusão dos Registros
 - Física ou lógica
- Reorganização
- Leitura Exaustiva do Arquivo

Arquivo Sequencial

- Os registros são dispostos ordenadamente
- Obedecendo à sequência determinada por uma chave primária,
 - chamada chave de ordenação.
- A ordem lógica e física dos registros armazenados é a mesma.
- Já que os registros em arquivos sequenciais são armazenados em sucessão contínua,
 - acessar o registro n do arquivo (começando no início do arquivo) requer que os registros $n-1$ também sejam lidos.

Exemplo

NÚMERO	NOME	IDADE	FUNÇÃO
1000	ADEMAR	25	PROGRAMADOR I
1050	CLAUDIA	30	ANALISTA II
2010	IARA	27	PROGRAMADOR I
2046	MIGUEL	28	ANALISTA I

```
class Registro {  
    int numero;  
    String nome;  
    int idade;  
    String função;  
}
```

```
struct Registro {  
    int numero;  
    char nome[30];  
    int idade;  
    char função[30];  
}
```

Abertura de um arquivo

- Um arquivo sequencial
 - pode ser aberto tanto nos modos de leitura e escrita.
- Ao abrir-se um arquivo sequencial,
 - o ponteiro de leitura sempre aponta para o primeiro registro do arquivo.

Consulta a Registros

- O acesso aleatório a registros
 - se caracteriza pela identificação do registro através da especificação de um argumento de pesquisa,
 - tem como resultado registros que não estão armazenados de forma contínua dentro de um arquivo sequencial.
- Pode-se considerar dois tipos de acessos:
 - aqueles cujas chave de acesso coincide com a chave de ordenação (chave primária) e
 - aqueles onde a chave de acesso envolve qualquer atributo.

Inserção de Registros

- As inserções de registros
 - possuem um custo muito alto, pois estas consistem em:
 - a) Verificar-se qual a posição adequada para o novo registro (ordem de chave primária)
 - b) Deslocar-se todos os registros que possuírem chave maior do que aquele que está sendo incluído, para uma posição adiante
 - c) Incluir novo registro
- Este tipo de procedimento pode acabar tendo um custo proibitivo, dependendo do tamanho do arquivo. Pode-se usar um arquivo auxiliar.

Alteração de Registros

- Essa operação
 1. consistirá na identificação do registro a ser alterado
 2. através de uma consulta (leitura exaustiva ou pesquisa binária)
 3. modificação dos valores de alguns atributos
 4. regravação do registro.

Exclusão de Registros

- A exclusão física
 - assim como a inserção, pode possuir um custo proibitivo
 - Precisa-se deslocar todos os registros que são posteriores ao registro excluído.
- A exclusão lógica
 - incluir um campo ao registro do arquivo que conterá um indicador de exclusão
 - Para excluir-se um registro é suficiente alterar o valor deste campo e regravar-se no arquivo.

Reorganização do Arquivo

- A reorganização é necessária para:
 - Eliminar registros excluídos logicamente
 - Intercalação do arquivo auxiliar utilizado para as inclusões
 - Processamento do arquivo em lote das alterações (ou outra operação) que não puderam ser realizadas em tempo real.

Banco de Dados – Exemplos e Conceitos

Banco de Dados - Exemplo

Planilha Eletrônica

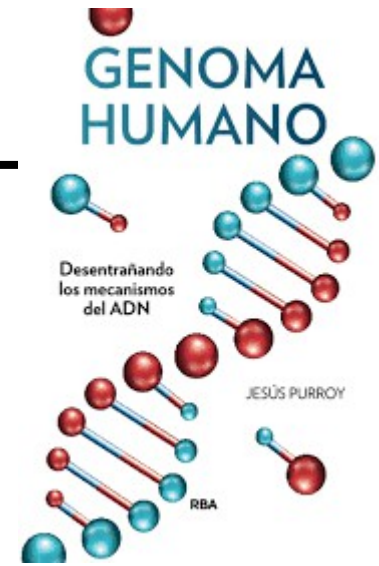
GC	D250	D500	D1000	P3	Sn	Nome	S250	S500	S1000
0.404301	1.550346	2.178267	0.962607	1.527230E-06	-0.0158915	Acine.fasta	1.441921	1.532032	0.732808
0.593780	1.862869	2.534370	1.236893	3.845661E-06	-0.0246687	Agro-tumefaciens.fasta	1.371953	1.387472	0.719663
0.592820	1.813819	2.426378	1.208882	3.520418E-06	-0.0231381	Agro-tume-linear.fasta	1.301741	1.349404	0.710608
0.497647	1.834696	2.427979	0.920974	1.168649E-06	-0.011112	Ana-marginale.fasta	1.430074	1.318752	0.797506
0.434774	2.190121	3.044996	1.196388	1.169521E-06	-0.0224146	Aqui-aeolicus.fasta	1.499050	1.504284	0.700958
0.651216	1.374920	1.530161	0.658361	6.219282E-06	-0.029948	Azoarcus.fasta	1.253902	1.277639	0.790516
0.353771	1.088610	1.350399	0.580717	1.950687E-06	-0.0163845	Bacillus-anthraxis.fasta	1.248574	1.276289	0.654815

Arquivo com Registros

1danielM1
2mauriciaF1
3luisaF1
4rafaelM1
5leandroM2
6lucianoM4
7danielM3
8lucianaF4
9danielaF3

Banco de Datos - Ejemplo

Banco de Datos Biológico, parte do genoma humano



>gi|50083297|ref|NC_005966.1| Acinetobacter sp. ADP1, complete genome
GATATTTTATCCACAGCATGTGTGAAATCTTATCCACAAGTGGTGTGAATTTAAATTTATTGTGCGAGCG
CTTCCTTTTTTAAGCAAAATCAAGTGAATCTTTCATGACAGATGTGGATAACTTGGTTAGAATGGCGACCC
CTTTCATCTGAGGGGGATGTTGTGGACTATTTTTGAATTTAAACGGACAGGGGATTACATGCTTTGGA
CAGACTGCTTAACCTCGCTTGCGACAAGAGCTATCTGATAACGTCTTTGCGATGTGGATTCGTCCTTTAGT
AGCTGAAGAGACGACAGATAGTCTACGCCTGTATGCGCCGAACCCTTACTGGACGCGTTATATTCAGGAA
CATCATCTTGAGTTGATCTCTATTCTGGTTGAACAACCTTTCTGAAGGGCGGATTCGCCAGGTCGAAATTT
TGGTTGACTCGCGTCCTGGTGCAATTTTATCCCCAGCAGAACAGCCAGCGACAACCACTGCTGCTCTGTC
AAGTACACCAGTTGTACCTCAGCGCGTTAAAAAAGAAGTTGTCGAACCAGCAGCTACGCAAAGTAATAAA
ATTTTAAATTCAAAAAACGTTTACTTAACCCACTGTTTACCTTCTCATTGTTTGTGCGAGGGCCGTTCTA
ACCAGATGGCTGCTGAGACTTGCCGAAAAGTTTTAACTCAACTTGGTGCCTCCCAGCACAATCCTTTGTT
TTTATATGGCCCGACAGGTTTAGGTAAACTCACTTGATGCAGGCTGTGGGTAAACGCATTACTTCAAGCC
AAACCAAATGCACGTGTGATGTATATGACGGCAGAAAGTTTTGTACAGGATTTTGTGAGCTCGTTACAAA
AAGGCAAGGTTGAAGAGTTTAAGAAAACTGTCGTTCACTTGATCTACTTTTGGTCGATGACATTCACTT
GCTGGCGGGTAAAGAAGCCAGTCTGGTCGAATTTTTCTATACCTTTAACGCCCTATTGGATGAATCTAAA
CAGATTATTTTAACTTCAGACCGTTATCCAAAAGAATTAACCGAGCTGGATCCGCGTCTGGTGTCACGTT

Banco de Dados - Exemplo

Banco de Dados em formato XML (eXtensible Markup Language)



```
<?xml version="1.0" encoding="UTF-8" ?>
<!DOCTYPE artigo (View Source for full doctype...)>
<artigo>
  <cabecalho>
    <titulo>Emprestimo FMI</titulo>
    <autor>Anonimo</autor>
  </cabecalho>
  <corpo>
    <introducao>O FMI ...</introducao>
    <paragrafo>Ontem, diversos ...</paragrafo>
    <paragrafo>O ministro da economia ...</paragrafo>
    <paragrafo>Além disso, ...</paragrafo>
    <conclusao>Enfim, o Brasil ...</conclusao>
  </corpo>
</artigo>
```


Banco de Dados - Exemplo

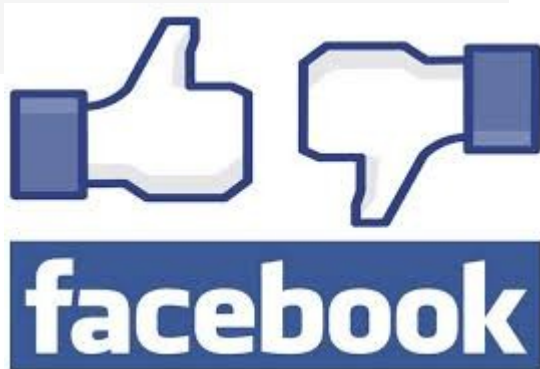
Banco de Dados em formato JSON (JavaScript Object Notation)



```
{
  "glossary": {
    "title": "example glossary",
    "GlossDiv": {
      "title": "S",
      "GlossList": {
        "GlossEntry": {
          "ID": "SGML",
          "SortAs": "SGML",
          "GlossTerm": "Standard Generalized Markup Language",
          "Acronym": "SGML",
          "Abbrev": "ISO 8879:1986",
          "GlossDef": {
            "para": "A meta-markup language, used to create markup languages such as
DocBook.",
            "GlossSeeAlso": ["GML", "XML"]
          },
          "GlossSee": "markup"
        }
      }
    }
  }
}
```

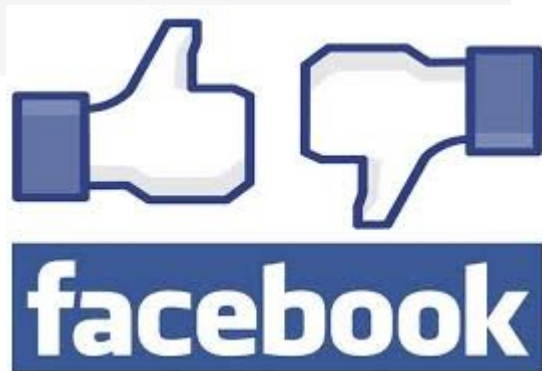
Banco de Dados - Exemplo

- Redes Sociais
- Big Data
- NoSQL



Banco de Dados - Exemplo

- Redes Sociais
- Big Data
- NoSQL



Banco de Dados - Resumo

- Banco de Dados
 - Planilha Eletrônica
 - Registro de arquivos
 - Banco de Dados Biológico
 - XML, JSON (dados semi-estruturados)
 - Big Data, NoSQL (dados não estruturados)
- Todos estes formatos de arquivos permitem armazenar diferentes tipos de informação.
- Todos são considerados bancos de dados?
- O QUE ELES TEM EM COMUM?

Banco de Dados - Introdução

- As informações precisam ser gerenciadas.
- São armazenados em arquivos.
- Realizam operações semelhantes.
- Gerência de Dados:
 - Criação e exclusão do arquivo
 - Inserção, alteração e exclusão dos dados no arquivo
 - Consulta sobre os dados
 - Indexação

Banco de Dados - Introdução

Um banco de dados é uma coleção de dados relacionados (Navathe, 2011)

RESTRIÇÕES ADVINDAS DO USO

Um banco de dados é uma coleção de dados logicamente coerentes e com algum significado inerente.

Um banco de dados é projetado, construído e populado com dados para um objetivo específico

Um banco de dados representa algum aspecto do mundo real, algumas vezes chamado minimundo

Pode ser mantido manualmente ou por máquina

Um Sistema de Gerência de Banco de Dados (SGBD) é uma coleção de programas que permite aos usuários criar e manter um banco de dados

- **Definição**: especificação dos tipos de dados
- **Construção**: armazenamento dos dados, controlado pelo SGBD
- **Manipulação**: consultas, atualização, geração de relatórios

Banco de Dados - Introdução

SGBD

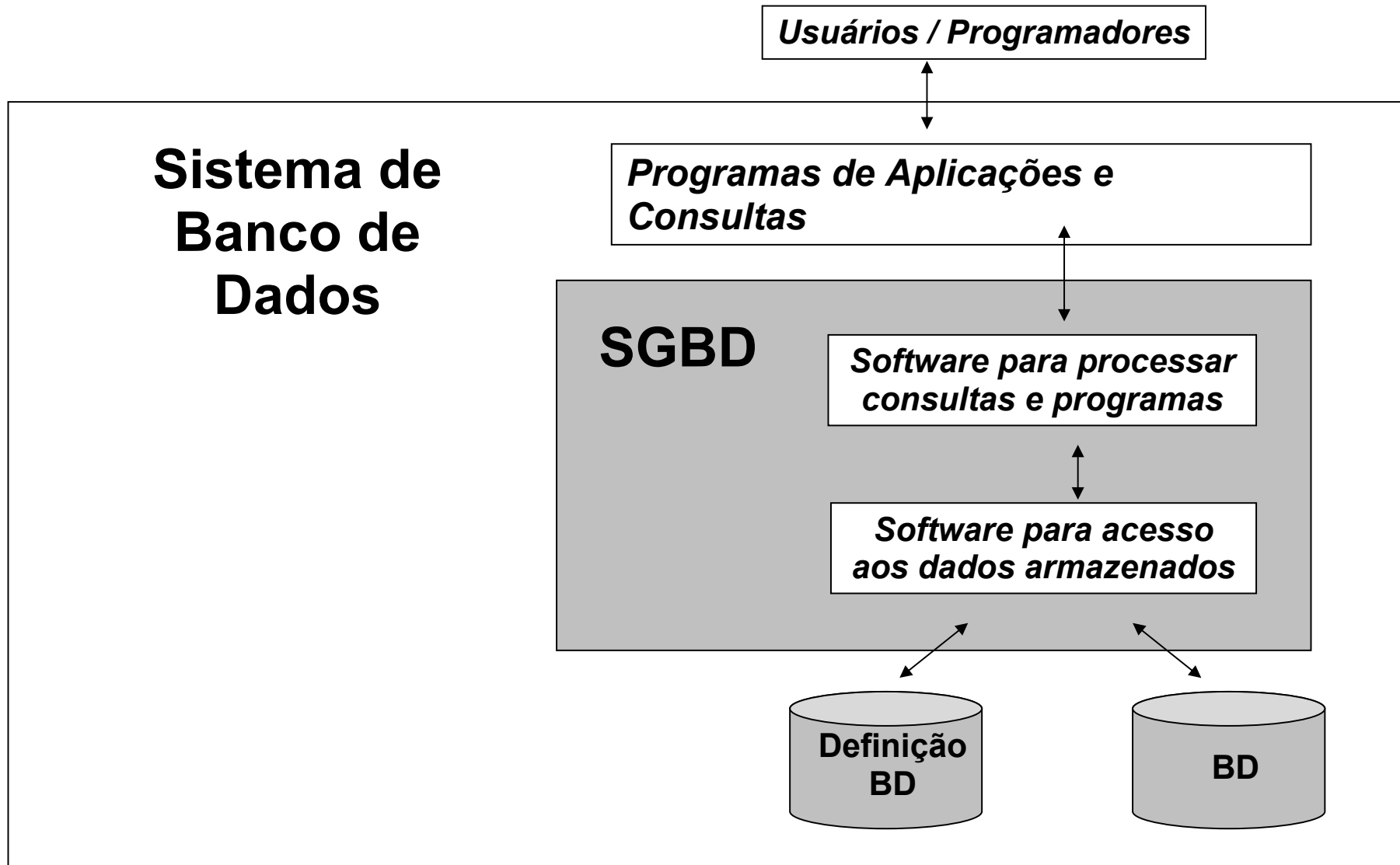
- Conjunto de dados associado a conjunto de programas para acesso aos dados

• Objetivo do SGBD: proporcionar um ambiente conveniente e eficiente para recuperação e armazenamento das informações do SGBD

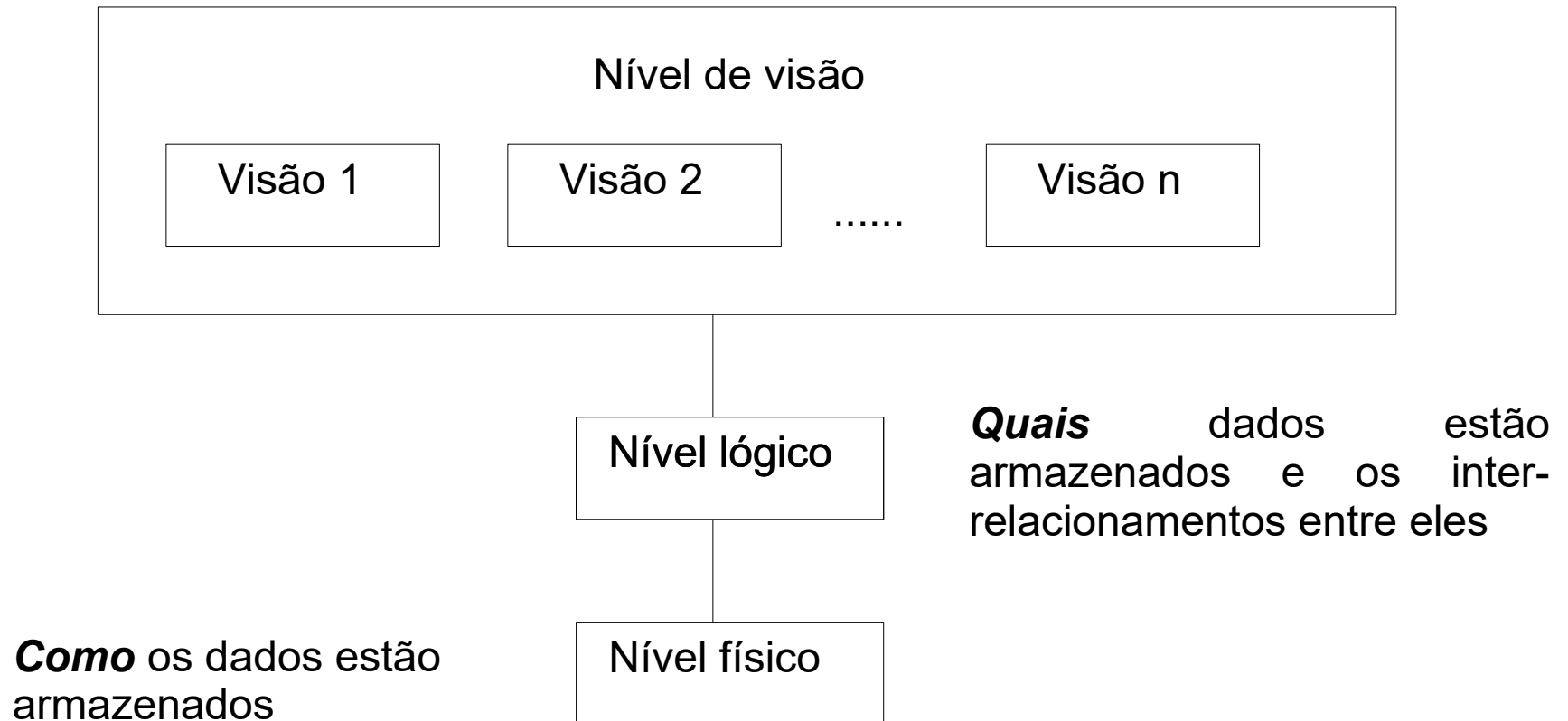


Banco de Dados - Introdução

SISTEMA DE BANCO DE DADOS



Banco de Dados - Introdução



Banco de Dados - Introdução

Usuários de um Banco de Dados

⇒ **Administrador do Banco de Dados:**

- responsável pelos dados, SGBD e programas relacionados
- autorização de acesso, coordenação e monitoração do uso

⇒ **Projetista do Banco de Dados (*Database Designer*):**

- identificação de requisitos dos usuários
- visão para cada usuário
- integração das visões
- escolha de estruturas apropriadas para armazenamento dos dados

⇒ **Usuário Final:**

- usuários ocasionais
- usuários parametrizados
- usuários sofisticados

⇒ **Analistas de Sistemas e Programadores de Aplicações**

Banco de Dados - Introdução

PREOCUPAÇÕES

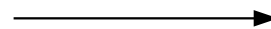
Segurança dos Dados

Separação LÓGICO x FÍSICO

- Esquema Interno
- Esquema Conceitual
- Esquema Externo

Linguagens de definição e manipulação de dados

Linguagens de Consulta mais poderosas, para “usuário final”



Serviço Autônomo do BD”

Banco de Dados - Introdução

PREOCUPAÇÕES

⇒ Independência Lógica dos Dados:

- Capacidade de mudar o esquema conceitual sem necessidade de modificar programas de aplicação e esquemas externos
- Apenas definição de visões e mapeamentos devem ser alterados
- Exemplo: Acrescentar um campo a um registro, acrescentar uma definição de tipo de registro

⇒ Independência Física de Dados:

- Capacidade de mudar o esquema interno sem necessidade de alterar o esquema conceitual (ou externos)
- Exemplo: reorganização física de arquivos, criação de estruturas de acesso adicionais

Banco de Dados - Introdução

FUNÇÕES DE UM SGBD

- Redundância Controlada
- Compartilhamento de Dados
- Restrições de Acesso
- Múltiplas Interfaces
- Restrições de Integridade
- Reestruturação em caso de falha

PREOCUPAÇÕES

The diagram shows a grey rectangular box labeled 'PREOCUPAÇÕES' on the right. Three arrows originate from its left side and point to the following items in the list: 'Restrições de Acesso', 'Restrições de Integridade', and 'Reestruturação em caso de falha'.

Arquiteturas de banco de dados

Arquiteturas de Sistemas de BD

- A arquitetura de sistemas de banco de dados é fortemente influenciada pelo sistema operacional sobre o qual o BD é executado e pelos aspectos da arquitetura de computadores (paralelismo, rede e distribuição):
 - Sistemas centralizados
 - Sistemas cliente-servidor
 - Sistemas paralelos
 - Sistemas distribuídos

Sistemas Centralizados

- Executados sobre um único sistema computacional
- Podem ser divididos conforme o modo como os computadores são usados:
 - Mono-usuário – computadores pessoais (uma única CPU, um ou mais discos rígidos, com um sistema operacional que pode dar suporte a apenas um usuário)
 - Multiusuário – número maior de discos rígidos e área de memória podendo ter diversas CPUs e um sistema operacional multiusuário.

Sistemas Cliente-Servidor

- Sistemas centralizados servindo como servidores:

- Clientes são computadores pessoais
- As funcionalidades de um SGBD podem ser superficialmente divididas em duas categorias : *front-end e back-end*:

Interface SQL
com o usuário

Interface para
o usuário

Gerador de
relatórios

Interface
gráfica

front-end

Interface SQL + API

Suporte a SQL

back-end

Sistemas Cliente-Servidor

- Podem ser caracterizados como:
 - Servidores de transações: Clientes enviam pedidos (SQL ou programas de aplicação) e o servidor executa a ação.
 - Servidores de dados: cliente solicita a leitura e/ou atualização de dados. Proporcionam meios para indexação de dados. Usados em redes locais de alta velocidade entre cliente e servidor. Cliente e servidor com poder de processamento compatível. Back-end reforçado no cliente.
- ***Nota: servidores de transação são mais utilizados.***

Sistemas Paralelos

- Uso de diversas CPU's e discos (compartilham memória e/ou disco)
- Sistemas que necessitam processar milhares de transações por segundo
- BD muito grande da ordem de exabytes
- Operações realizadas simultaneamente

Sistemas Paralelos

- Granularidade:
 - ⇒ Grossa = poucos e poderosos processadores
 - ⇒ Fina = usa milhares de pequenos processadores
- Reduzem o tempo de resposta das transações
(requerem SGBDs paralelos)
- Estratégia de paralelismo (1 consulta por processador,
uma consulta particionada em vários processadores)

Sistemas distribuídos

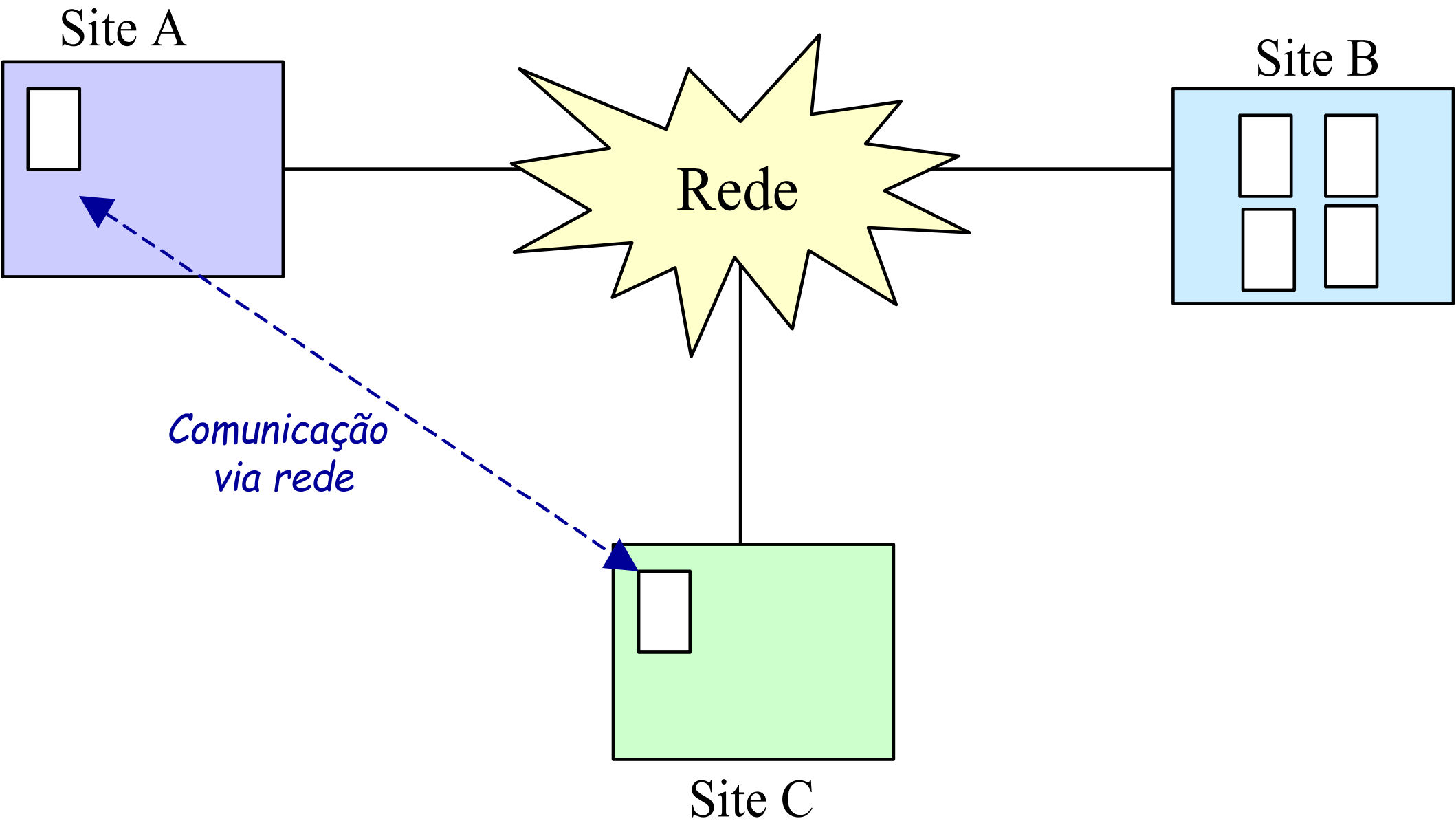
- BD armazenado em diversos computadores
- Comunicação por rede, ADSL, ...
- Computadores de sistemas distribuídos são conhecidos como *nós* ou *sites*
- Distribuição física e geográfica dos dados, administração separada e intercomunicação dos *sites*
- *Sites* podem acessar uns aos outros
- Esquema global comum (algumas partes em *sites*)
- Mesmo SGBD

Sistemas distribuídos

- Distinção de transações (conjunto de operações que tem que ser executadas atômicamente) locais e globais:
 - Locais = acessa um único *site* (no qual ela se inicia)
 - Globais = acessa diferentes *sites* (ou outro diferente daquele que se inicia).
- OBS: Sistemas de BD distribuídos heterogêneos (compostos por diferentes SGBDs nos sites). Dificuldade de gerenciamento de transações globais

Nota: sistemas de banco de dados distribuídos e cliente-servidor estão apoiados em redes de comunicação: Rede local – área geográfica restrita, processadores distribuídos; Rede de longa distância – grande área geográfica, processadores distribuídos.

Sistema Distribuído



Exercícios

Exercícios

- 1) O que é um banco de dados? Justifique. Cite exemplos.
- 2) Qual a diferença de um Banco de Dados para um Sistema de Arquivos?
- 3) Qual a diferença de um banco de dados para um sistema gerenciador de banco de dados?
- 4) A Casa de Vinhos XYZ está querendo armazenar os seus dados, para isto, defina os principais requisitos e defina o conjunto de dados principal envolvido. Agora, defina qual será a forma de armazenamento: sistema de arquivos, banco de dados ou um sistema gerenciador de banco de dados. Explique a sua escolha e o porque de não escolher as outras opções.

Exercícios

- 5) Cite um exemplo de software que pode ser utilizado para cada um: sistemas centralizados, sistemas cliente-servidor, sistemas paralelos, sistemas distribuídos.
- 6) Enumere as principais diferenças entre o processamento de dados com arquivos convencionais e o processamento de dados com um SGBD.
- 7) Cite softwares para serem usados com o modelo conceitual, modelo lógico e modelo físico