

Bancos de Dados

2. Conceitos e Arquitetura de Bancos de Dados

Arquitetura Moderna de SGBD

- SGBD antigos eram monolíticos e rígidos, voltados para funcionamento em ambientes centralizados (*mainframes* e etc.)
- Hoje os SGBD são projetados de maneira modular, para funcionar em redes de computadores de menor porte, dividindo tarefas
 - Módulos cliente: interface com o usuário, aplicações – rodam no computador do usuário
 - Módulos servidores: armazenamento de dados, pesquisas, recuperação – rodam em um ou mais computadores mais poderosos, centralizados ou distribuídos na rede
- Arquitetura cliente-servidor: vide EN cap. 25.

Modelos de Dados, Esquemas e Instâncias

- O mundo real é muito complexo, e não é possível reproduzi-lo com todos os seus detalhes em um sistema informatizado
- É necessário simplificar e generalizar a realidade:
abstração
- Técnica básica para a modelagem de dados
- Um *modelo de dados* é uma coleção de conceitos que podem ser usados para descrever a estrutura de um banco de dados
 - Necessariamente contém recursos para alcançar o grau desejado de abstração

Modelos de Dados, Esquemas e Instâncias

- *Estrutura de um banco de dados*: tipos de dados, relacionamentos e restrições que devem existir para que o banco de dados esteja correto e íntegro
- Pode incluir um conjunto de *operações básicas* para especificar recuperações e atualizações no BD
- Em modelos de dados mais modernos, costuma-se também especificar aspectos *dinâmicos* ou *comportamentais* das aplicações
 - Facilita a especificação de operações *definidas pelo usuário* e que são permitidas no BD
 - Necessários em BD orientados a objetos e em BD objeto-relacionais

Modelos de Dados, Instâncias e Esquemas

- Categorias de modelos de dados
 - Modelos de *alto nível* ou *conceituais*
 - Modelos de *dados representativos* ou de *implementação*
 - Modelos de *baixo nível* ou *físicos*

Modelos de Dados, Instâncias e Esquemas

- Modelos conceituais: conceitos utilizados
 - *Entidades*: objetos ou conceitos do mundo real
 - *Atributos*: características ou propriedades de uma entidade
 - *Relacionamentos*: interações entre entidades
- Modelos de implementação
 - Mais frequentemente usados nos SGBD comerciais
 - Amplo domínio do *modelo relacional*
 - Modelos legados: *de rede e hierárquico*
 - Nova família: *modelos de dados de objetos*
 - Mais próximos dos modelos conceituais, e em geral usados como tal em Engenharia de Software

Modelos de Dados, Instâncias e Esquemas

- Modelos físicos
 - Descrevem como os dados estarão armazenados no computador
 - Registros: formato, ordenação
 - Caminhos de acesso: estruturas de dados que tornam eficiente a busca de determinados registros no BD
 - Indexação

Modelos de Dados, Instâncias e Esquemas

- Esquema: *descrição* detalhada do banco de dados
 - Não é o mesmo que o banco de dados em si
 - Não é o *modelo de dados*, mas sim o *produto da modelagem*
 - Não se supõe que vá ser alterado com frequência, independente do nível de alterações que os dados possam ter
- Diagrama do esquema
 - Apresentação (gráfica) do esquema, usando as técnicas do modelo de dados
 - Exibe apenas alguns aspectos do esquema; outros podem requerer o uso de recursos adicionais para apresentação, como é o caso das *restrições*
- Componente do esquema: cada objeto presente nele

Figura 2.1 Diagrama esquemático para o banco de dados da Figura 1.2.

ALUNO

Nome	NumerodoAluno	Turma	Curso_Hab
------	---------------	-------	-----------

CURSO

NomedoCurso	NumerodoCurso	Creditos	Departamento
-------------	---------------	----------	--------------

PRE_REQUISITO

NumerodoCurso	NumerodoPre_requisito
---------------	-----------------------

DISCIPLINA

Identificador_Disciplina	NumerodoCurso	Semestre	Ano	Instrutor
--------------------------	---------------	----------	-----	-----------

RELATORIO_DE_NOTAS

NumerodoAluno	Identificador Disciplinas	Nota
---------------	---------------------------	------

Modelos de Dados, Instâncias e Esquemas

- Os dados contidos no BD em um determinado instante de tempo constituem o *estado* do BD naquele instante
- Outros nomes:
 - instantâneo ou retrato (*snapshot*) do BD
 - *Conjunto atual de ocorrências* ou *instâncias* do BD
- Num determinado estado do BD, cada componente do esquema terá seu próprio conjunto de instâncias atual
- Quando o esquema é definido e especificado para o SGBD, o estado do BD é o *estado vazio*
- Quando os dados iniciais são carregados, o estado passa a ser o *estado inicial*

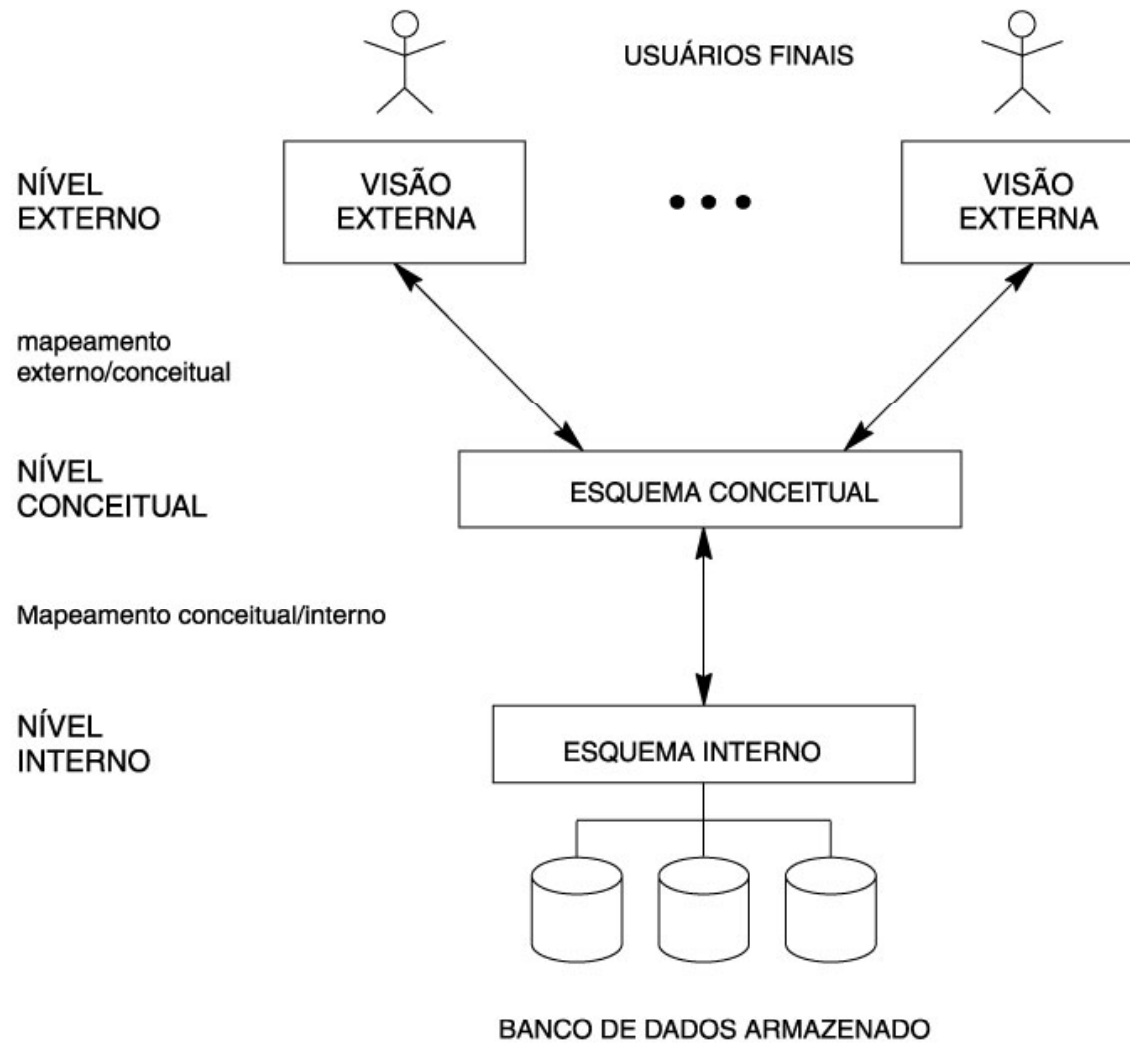
Modelos de Dados, Instâncias e Esquemas

- Toda inclusão, exclusão ou alteração de dados no BD altera seu estado
 - O SGBD é parcialmente responsável por garantir que cada estado do BD seja um *estado válido*, ou seja, um estado em que todas as restrições de integridade são cumpridas
- As descrições dos componentes e restrições do esquema (*metadados*) são mantidas no catálogo do SGBD
- Mudanças na estrutura do BD são chamadas de *evolução do esquema*
 - SGBD possuem recursos para aplicar evoluções no esquema mesmo com o BD em operação

Arquitetura do SGBD

- Características necessárias
 - Isolamento entre programas e dados
 - Suporte a múltiplas visões
 - Uso do catálogo para armazenar o esquema do BD
- Essas características são melhor analisadas no contexto da *arquitetura de três esquemas* ou *arquitetura ANSI/SPARC*

Figura 2.2 A arquitetura de três-esquemas



Arquitetura do SGBD

- Nível interno -> esquema interno
 - Descreve a estrutura de armazenamento do BD
 - Usa um modelo de dados físico e descreve detalhes completos sobre a forma de armazenamento dos dados e os caminhos de acesso a eles

Arquitetura do SGBD

- Nível conceitual -> esquema conceitual
 - Descreve a estrutura do BD para uma comunidade de usuários
 - Esconde detalhes de implementação e se concentra na descrição e entidades, tipos de dados, relacionamentos, operações definidas pelo usuário e restrições
 - Usa um modelo de dados conceitual ou um modelo de implementação

Arquitetura do SGBD

- Nível externo ou de visão
 - Inclui diversos *esquemas externos* ou *visões do usuário*
 - Cada esquema externo descreve a parte do BD na qual um determinado grupo de usuários está interessado e esconde o restante do BD
 - Usa um modelo de dados conceitual ou um modelo de implementação

Arquitetura do SGBD

- Todos os três esquemas são apenas descrições dos dados
 - Os dados encontram-se apenas no nível físico
- Cada grupo de usuários se refere apenas a seu próprio esquema externo
 - O SGBD deve transformar uma requisição que se refira a um esquema interno em uma solicitação baseada no esquema interno, de modo que possa ser processada e resolvida junto ao BD armazenado
 - Se a solicitação envolver uma recuperação de dados, o resultado deve ser reformatado para corresponder à visão do usuário
- Os processos de transformação de solicitações e resultados são chamados de *mapeamentos* (*mappings*)

Independência de Dados

- A arquitetura de três esquemas define independência de dados como sendo a capacidade de alterar o esquema em um determinado nível sem ter que alterar o esquema no próximo nível mais elevado
 - Independência lógica de dados: capacidade de alterar o esquema conceitual sem ter que alterar os esquemas externos
 - Ex.: acréscimo ou redução de um atributo
 - Independência física de dados: capacidade de alterar o esquema interno (físico) sem ter que alterar esquemas conceituais (ou externos)
 - Ex.: reorganização de índices e tabelas visando aumento de desempenho

Linguagens

- Linguagem de Definição de Dados (DDL – Data Definition Language)
 - Especificação dos esquemas conceitual e interno em SGBD que não possuem uma separação estrita entre níveis
- Linguagem de Definição de Armazenamento (SDL – Storage Definition Language)
 - Especificação do esquema interno em SGBD que mantêm separação entre os níveis conceitual e interno
- Linguagem de Especificação de Visões (VDL – View Definition Language)
 - Especificação de visões do usuário e seu mapeamento para o esquema conceitual
- Linguagem de Manipulação de Dados (DML – Data Manipulation Language)
 - Execução de operações de recuperação, inclusão, alteração e exclusão de dados

Linguagens

- Na prática, o SGBD provê apenas uma linguagem integrada com todas essas finalidades
 - A definição do armazenamento é geralmente mantida em separado, para que seja possível realizar ajustes de desempenho no sistema (*tuning*)
 - Ex. SQL = DDL + VDL + DML
 - SDL era parte de versões primitivas da SQL; hoje não é mais

Linguagens

- Tipos de DML
 - Alto nível (não-procedural)
 - Especificar operações complexas de BD
 - Incluída em programas em outras linguagens, porém em destaque para processamento pelo SGBD
 - Recupera muitos registros a cada operação: *DML orientada a conjunto*
 - Indica *que* dados recuperar, mas não *como* fazê-lo
 - *Linguagem de consulta*
 - Baixo nível (procedural)
 - Embutida em uma linguagem de programação de uso geral
 - Recupera e processa um registro de cada vez: *DML de um registro a cada vez*
 - *Sublinguagem* de uma linguagem *hospedeira*

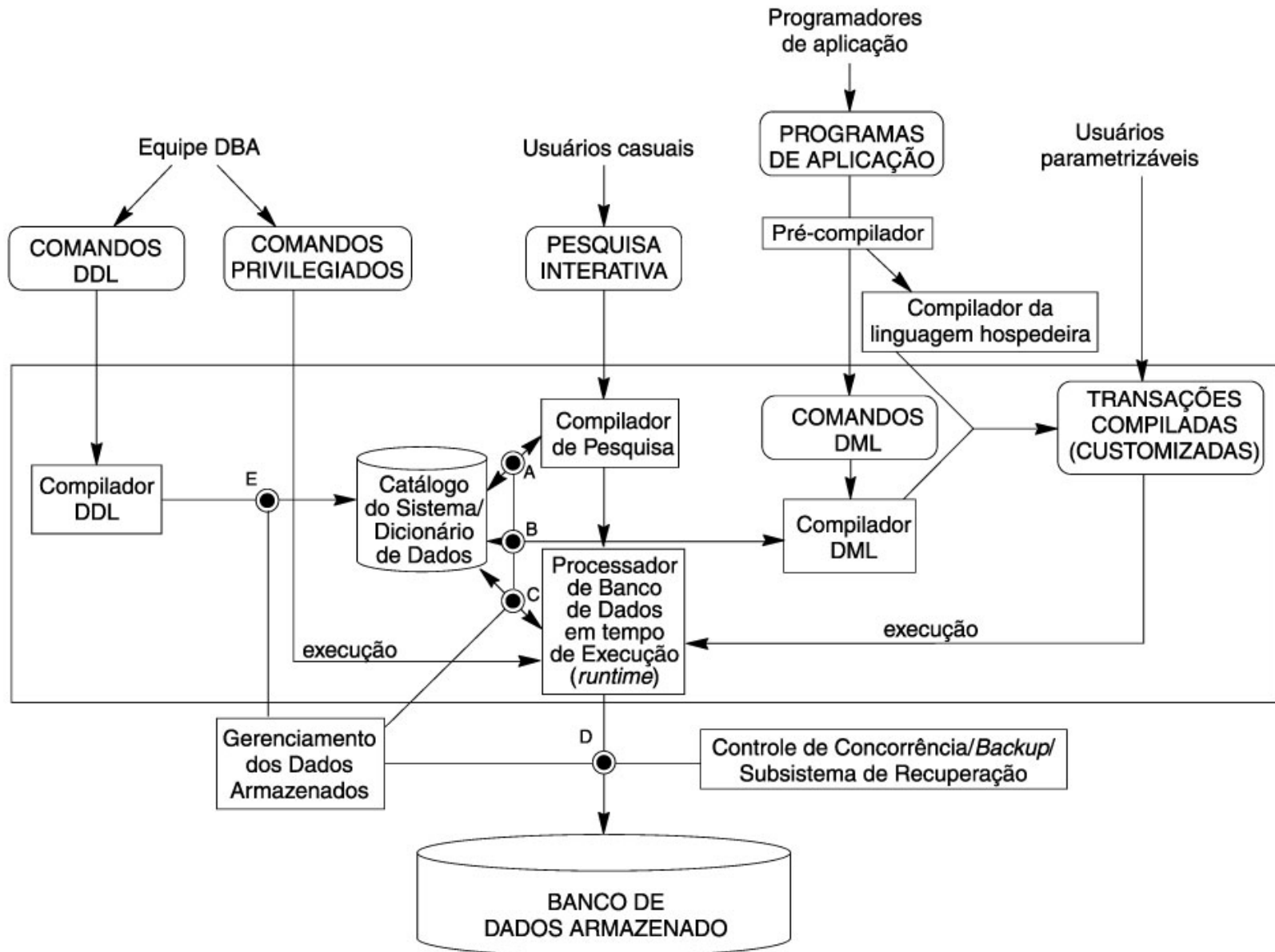
Interfaces

- Baseadas em menus para *browsing*
- Baseadas em formulários
- Gráficas (GUI)
- Linguagem natural
- Voltadas para usuários leigos
- Voltadas para DBA

O Ambiente do SBD

- Módulos ou componentes do SGBD
 - Gerenciador de dados armazenados (*stored data manager*)
 - Acesso aos dados armazenados em disco, através do S. O.
 - Gerenciamento de buffers em memória
 - Compilador da DDL
 - Processa definições do esquema e constrói o catálogo
 - Processador de *run-time* do BD
 - Processa comandos de recuperação ou atualização de dados
 - Compilador de consultas
 - Otimiza consultas recebidas interativamente
 - Pré-compilador
 - Extrai comandos de DML de um programa hospedeiro
 - Compilador de DML
 - Produz código-objeto para acesso ao BD

Figura 2.3 Módulos componentes de um SGBD e suas interações.



Oracle Architecture

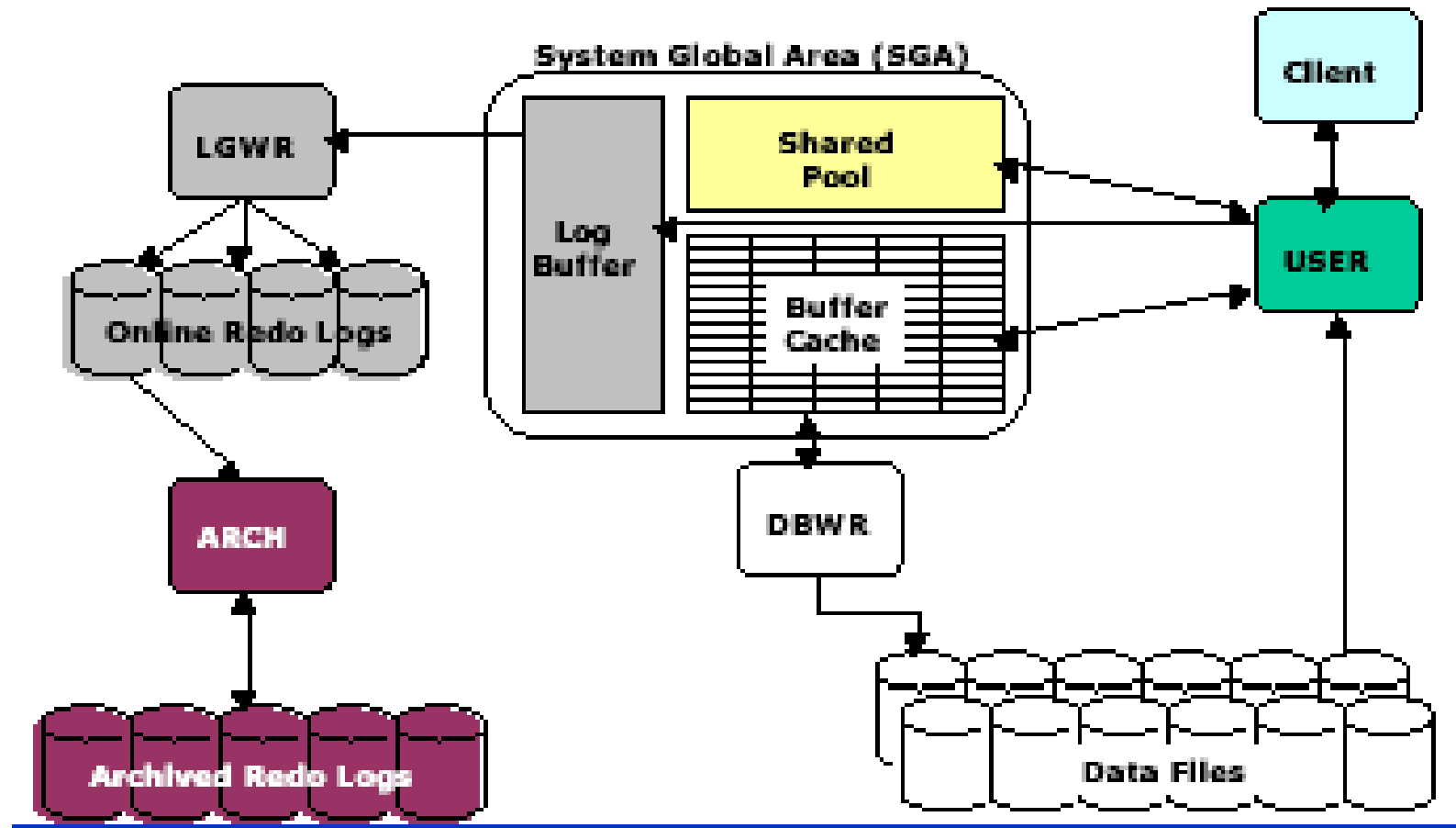


Figura 2.4 A arquitetura física centralizada.

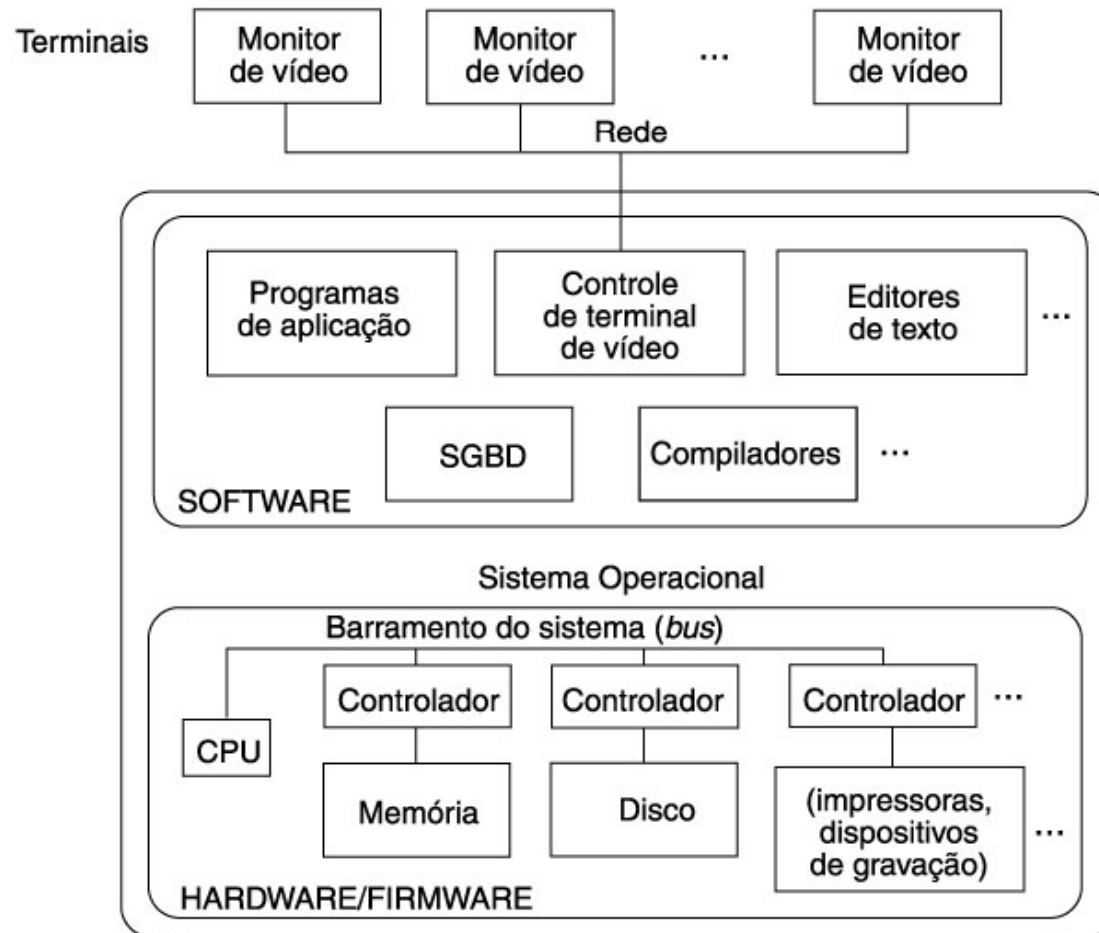


Figura 2.5 A arquitetura lógica de duas camadas cliente/servidor.

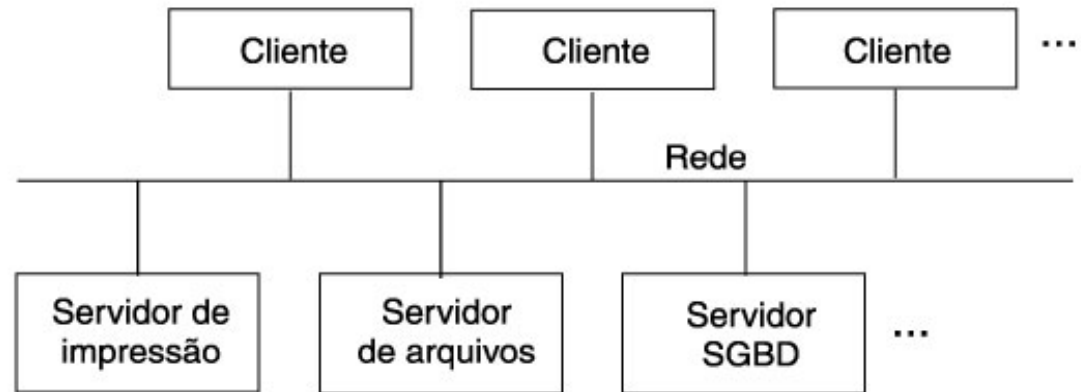


Figura 2.6 Arquitetura física cliente/servidor de duas camadas.

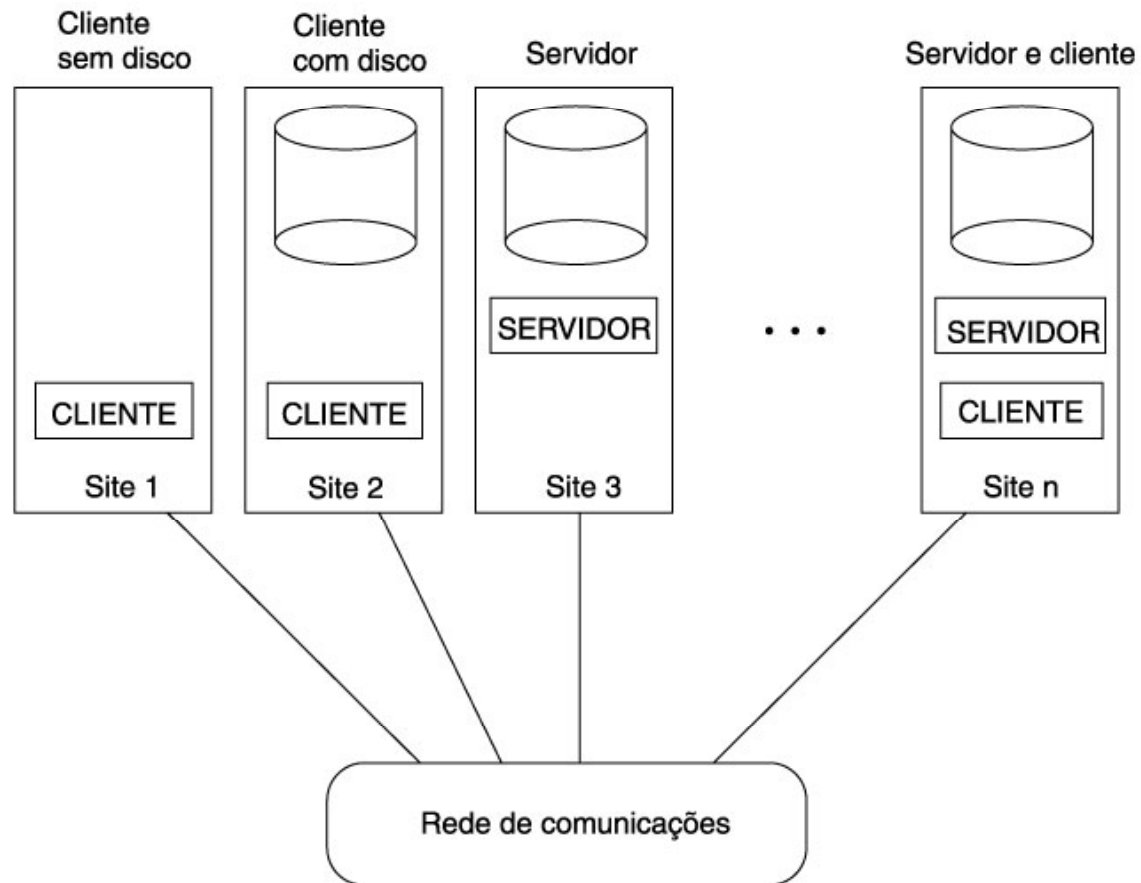
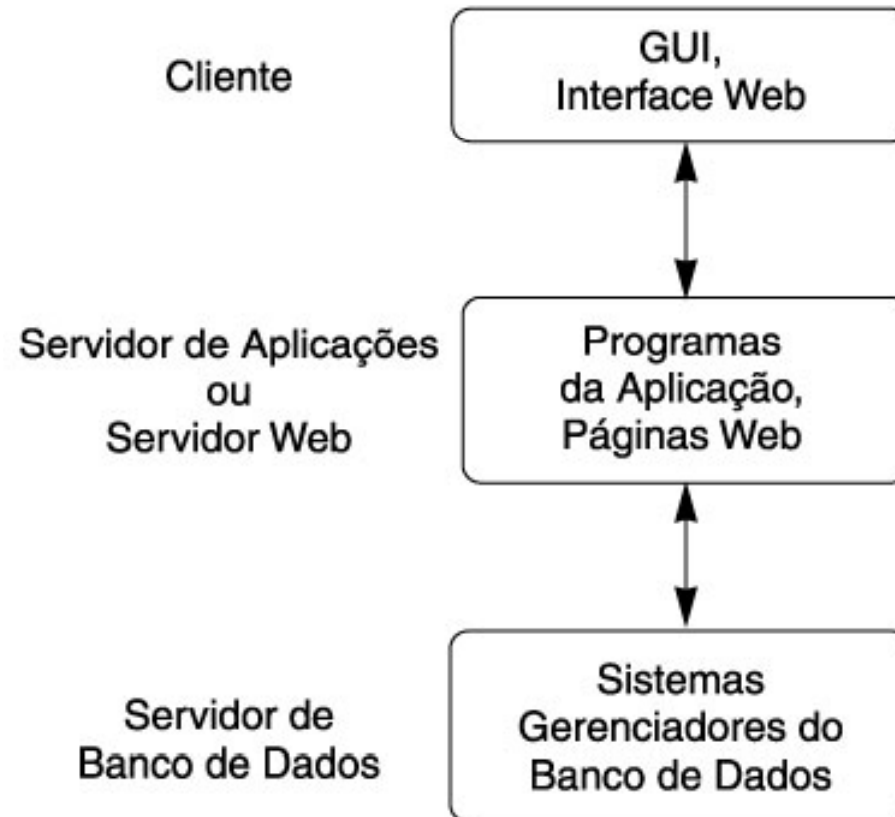


Figura 2.7 Arquitetura lógica cliente/servidor de três camadas.



O Ambiente do SBD

- Utilitários do SBD
 - Carga de dados (loading)
 - Ferramentas de conversão de formato
 - Cópia de segurança (backup)
 - Reorganização de arquivos
 - Monitoramento do desempenho
 - Ordenar arquivos
 - Compactar dados
 - Monitorar acesso dos usuários
 - Gerenciamento de perfis de usuários
 - Etc...

O Ambiente do SBD

- Ferramentas de desenvolvimento
 - CASE
 - Sistema de gerenciamento de dicionários de dados
 - Ambientes de desenvolvimento de aplicações e prototipação rápida
- Ferramentas de comunicação de dados e acesso remoto

Classificação dos SBD

- Quanto ao modelo de dados
 - Relacionais
 - Orientados a objetos
 - Hierárquicos
 - Em rede
 - Objeto-relacionais

Classificação dos SBD

- Quanto ao número de usuários
 - Monousuário
 - Multiusuário
- Quanto ao número de locais de acesso a dados
 - Centralizado
 - Distribuído
 - SGBDD homogêneos
 - SGBDD heterogêneos -> SGBD *federado* ou *sistema multi-banco de dados*

Classificação dos SBD

- Quanto ao custo
 - Custos baixos, operação em micros (ex. Access)
 - Custos elevados, operação em servidores (ex. Oracle)
 - Software livre (ex. MySQL)
- Quanto à finalidade
 - Finalidade genérica: uso de SGBD genérico
 - Finalidade especial: uso de SGBD desenvolvido para a aplicação
 - Ex.: reservas em companhias aéreas; catálogos telefônicos
 - OLTP: On-line Transaction Processing – grande número de transações concorrentes sem impor atrasos excessivos

Referências

- Capítulo 2 de Elmasri & Navathe