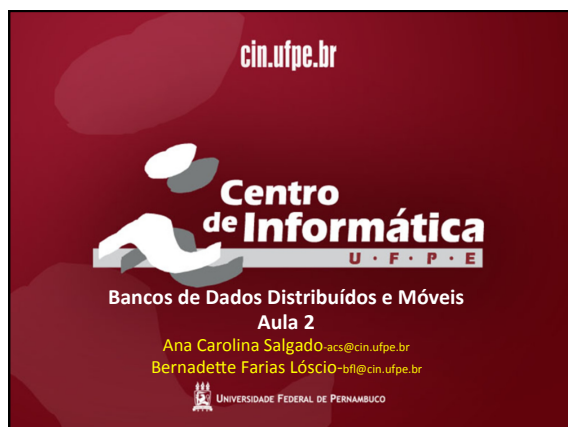




Revisão

- Um dos importantes objetivos da tecnologia de bancos de dados distribuídos é...
 - alcançar integração sem ter centralização
- Por que distribuir?
 - organização inerentemente distribuída
 - Dividir para conquistar



Revisão

- Definições
 - Banco de dados distribuído
 - Múltiplos bancos de dados interrelacionados e distribuídos ao longo de uma rede de computadores
 - SGBD Distribuído
 - Torna possível o gerenciamento dos bancos de dados distribuídos e torna a distribuição transparente para o usuário
 - SBDD
 - BDD + SGBDD



Revisão

- Promessas de um SBDD
 - Gerenciamento de dados distribuídos e replicados
 - Facilidade de expansão
 - Melhorias de desempenho
 - Acesso confiável a partir de transações distribuídas



Revisão

- Gerenciamento de dados distribuídos e replicados
 - Existem diferentes tipos de transparência...
 - Fragmentação
 - Replicação
 - Rede
 - Independência de dados
- Confiabilidade por meio de transações distribuídas
 - Não existe apenas um único ponto de falha!
- Melhoria de desempenho
 - Os dados ficam mais próximos do seu “uso” (localização de dados)
 - Redução do volume de dados e dos acessos remotos



Agenda

- Arquitetura – visão geral
- Modelos de arquitetura para SGBD distribuído
- Arquiteturas de SGBD distribuídos

Objetivo da aula: apresentar os fundamentos para a definição das arquiteturas de SGBDDs, bem como as principais arquiteturas (cliente/servidor, peer-to-peer e bancos de dados múltiplos)



Arquitetura – visão geral

- A arquitetura de um sistema define sua estrutura
 - Definição dos componentes, a função de cada componente, o inter-relacionamento e interações entre esses componentes
- A especificação da arquitetura consiste em
 - Identificar os vários módulos com suas interfaces e inter-relacionamentos de dados e fluxos de controles



Arquitetura de SGBDD

- Definir uma arquitetura consiste em estruturar o SGBDD, de modo a prover as funcionalidades
 - Gerenciamento transparente de dados distribuídos
 - Confiabilidade em transações distribuídas
 - Desempenho otimizado
 - Expansão de sistemas mais fácil



Padronização de SGBD

- Modelo de referência
 - Uma estrutura conceitual cujo propósito é dividir o trabalho de padronização em partes gerenciáveis e mostrar em um nível geral como estas partes se relacionam umas com as outras
- Abordagens para definição de um sistema
 - Baseada em componentes
 - Componentes de um sistema são definidos junto com o inter-relacionamento entre os componentes



Padronização de SGBD

- Abordagens (Cont.)
 - Baseada em funções
 - Classes de usuários são identificados junto com a funcionalidade que o sistema deverá prover para cada classe
 - Baseada em dados
 - Identificar os diferentes tipos de descrição dos dados e especificar as unidades funcionais que obterão e/ou usarão os dados de acordo com essas visões



Padronização de SGBD

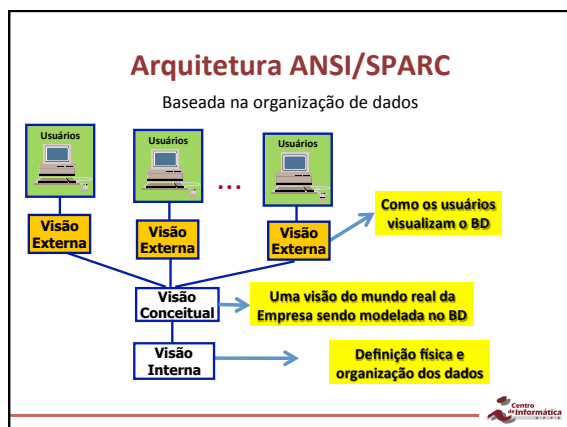
- Análise comparativa
 - Baseada em componentes
 - Vantagem: possui uma visão completa de todos os inter-relacionamentos do sistema, facilitando o projeto
 - Desvantagem: não oferece informação sobre os objetivos e dados a serem utilizados
 - Baseada em funções
 - Vantagem: possui clareza na especificação dos objetivos do sistema
 - Desvantagem: oferece pouca informação sobre como os objetivos serão atingidos ou o nível de complexidade do sistema



Padronização de SGBD

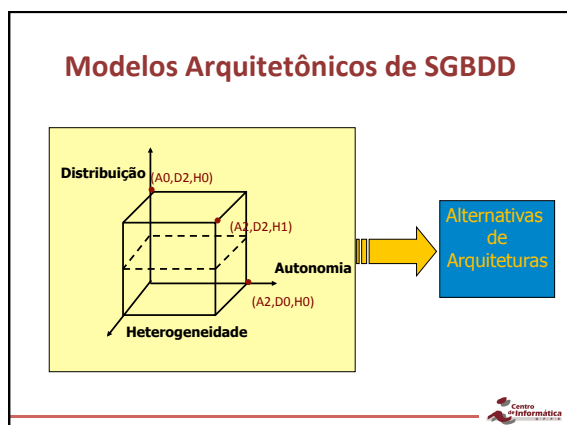
- Análise (Cont.)
 - Baseada em dados
 - Vantagem: dá importância central ao dado. É especificada uma estrutura arquitetônica que define as unidades funcionais que perceberão ou usarão os dados.
 - Desvantagem: não dá para criar um modelo de arquitetura completo sem os módulos funcionais





Modelos arquitetônicos para SGBDD

- A arquitetura define
 - os modos pelos quais vários bancos de dados podem interoperar
 - como um SGBDD pode ser estruturado
- A arquitetura dos sistemas pode ser classificada de acordo com as seguintes características:
 - Autonomia
 - Distribuição
 - Heterogeneidade



Modelos Arquitetônicos de SGBDD

- Autonomia
 - Refere-se à distribuição de controle, indicando o grau no qual SGBDs individuais podem operar de forma independente



Modelos Arquitetônicos de SGBDD

- Requisitos de um sistema que preserve autonomia
 - As operações locais dos SGBDs individuais não são afetadas pela sua participação no sistema distribuído
 - O processamento de consultas local não deve ser afetado pelo processamento de consultas global
 - A consistência do sistema não deve ser afetada pela sua entrada ou saída no sistema distribuído



Modelos Arquitetônicos de SGBDD

- Dimensões da Autonomia
 - **Autonomia de projeto** - Os SGBDs são livres para usar os modelos de dados e as técnicas de gerenciamento de transações que preferirem
 - **Autonomia de comunicação** - habilidade de um SGBD componente decidir como se comunicar com outros SGBDs
 - **Autonomia de execução** - habilidade de um SGBD componente executar operações locais da maneira que ele queira fazer



Modelos Arquitetônicos de SGBDD

- Classificação de acordo com a autonomia
 - Fortemente integrados
 - Os dados estão logicamente centralizados em um BD
 - Um dos SGBDs fica no controle do processamento das requisições dos usuários



Modelos Arquitetônicos de SGBDD

- Classificação de acordo com a autonomia
 - Semi-autônomos
 - SGBDs que operam independentemente, mas decidiram participar de uma federação para compartilhar seus dados locais
 - Cada SGBD determina que partes de seus próprios BDs serão acessíveis por usuários de outros SGBDs.
 - precisam ser modificados para habilitar a troca de informações com os demais



Modelos Arquitetônicos de SGBDD

- Classificação de acordo com a autonomia
 - Isolamento total
 - os sistemas individuais são SGBDs isolados (stand-alone) e não se comunicam com outros SGBDs
 - Não existe um controle global sobre a execução de SGBDs individuais.



Modelos Arquitetônicos de SGBDD

- Distribuição física dos dados
 - Lida com a distribuição física de dados sobre vários sites
 - **Cliente/servidor** - Concentra a tarefa de gerenciamento de dados em servidores, enquanto os clientes se concentram em fornecer o ambiente do aplicativo, incluindo a interface com o usuário
 - **Não hierárquica (peer-to-peer ou distribuição total)** - Cada máquina tem toda a funcionalidade de SGBD e pode se comunicar com outras máquinas para executar consultas e transações

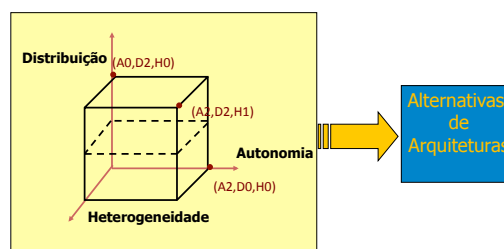


Modelos Arquitetônicos de SGBDD

- Heterogeneidade
 - Vários níveis (hardware, subsistemas de comunicação, sistemas operacionais)
 - SGBD
 - Modelo de dados, linguagem de consulta, algoritmos (protocolos) de gerenciamento de transações



Modelos Arquitetônicos de SGBDD



Alternativas de Arquiteturas

- Identificadas por dimensões
 - Dimensões: A (**A**utonomia), D (**D**istribuição) e H (**H**eterogeneidade)
 - Alternativas: números 0, 1 ou 2
 - A: 0 (firme integração), 1 (sistemas semi-autônomos) e 2 (isolamento total)
 - D: 0 (sem distribuição), 1 (sistemas cliente/servidores) e 2 (distribuição ponto-a-ponto)
 - H: 0 (sistemas homogêneos) e 1 (sistemas heterogêneos)



Alternativas de Arquiteturas

- SGBD Distribuído Cliente/Servidor (**A0, D1, H0**)
- SGBD Distribuído Peer-to-Peer (**A0, D2, H0**)
- Sistema de múltiplos bancos de dados heterogêneos (**A2, D2, H1**)

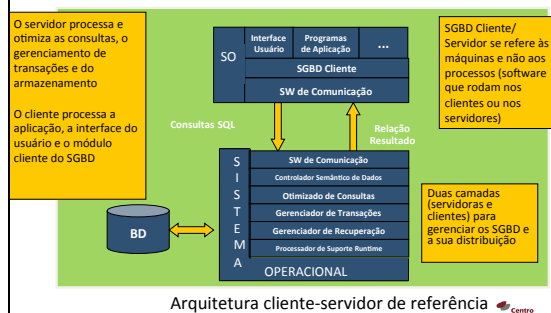


Alternativas de Arquiteturas

- Arquitetura cliente/servidor
 - Estamos interessados no software que deverá rodar na máquina cliente e qual software deverá rodar no servidor
 - Como distribuir
 - Processamento e otimização de consultas
 - Gerenciamento de transações
 - Gerenciamento de cache
 - ...



Arquiteturas de SGBD Distribuído



Arquiteturas de SGBD Distribuído

- Tipos de arquitetura cliente/servidor
 - Mais simples - um servidor que é acessado por múltiplos clientes (múltiplos clientes-único servidor)
 - Mais sofisticada - múltiplos servidores no sistema (múltiplos clientes-múltiplos servidores) com o gerenciamento das conexões para o servidor em cada cliente ou cada cliente conhece apenas "seu" servidor que se comunica com outros servidores



Arquiteturas de SGBD Distribuído

- Arquitetura cliente/servidor
 - A visão lógica dos dados é a mesma que os sistemas ponto-a-ponto, ou seja, tem a aparência de um BD único
 - O que distingue uma arquitetura da outra é a forma como a transparência é alcançada!

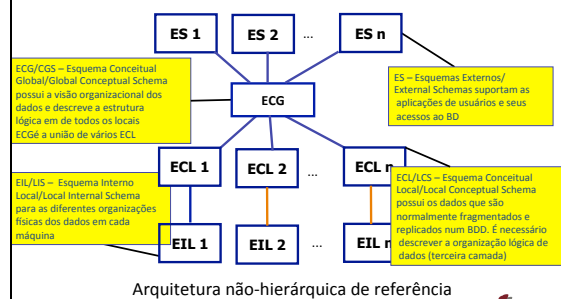


Arquiteturas de SGBD Distribuído

- Arquitetura peer-to-peer (não-hierárquica)
 - Implica na mesma funcionalidade em cada site da rede, ou seja, não existe distinção entre funções do cliente e do servidor
 - Visão **tradicional** do conceito peer-to-peer x Visão **moderna** do conceito peer-to-peer



Arquiteturas de SGBD Distribuído



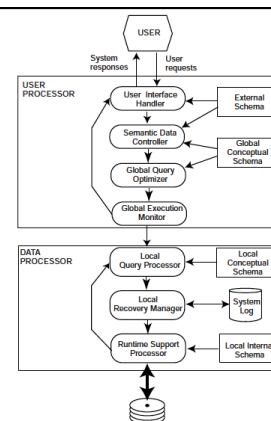
Arquiteturas de SGBD Distribuído

- Arquitetura não-hierárquica (Cont.)
 - Fornece níveis de transparência
 - A independência dos dados
 - A transparência de localização e replicação pela definição do esquemas conceituais locais e global, e o mapeamento entre eles
 - A transparência da rede, com suporte oferecido pela definição do esquema conceitual global (o SGBDD traduz as consultas globais em grupos de consultas locais, que são executadas em diferentes locais)



Arquiteturas de SGBDD

- Componentes de um SGBDD



Arquiteturas de SGBD Distribuído

- *User interface handler* (Gerenciador da interface do usuário)
 - Interpretação dos comandos do usuário à medida que eles chegam
 - Formatação dos dados do resultado conforme eles são enviados para o usuário



Arquiteturas de SGBD Distribuído

- *Semantic Data Controller* (Controlador de dados semânticos)
 - Utiliza as restrições de integridade e as autorizações definidas como parte do esquema conceitual global, a fim de verificar se a consulta do usuário pode ser processada



Arquiteturas de SGBD Distribuído

- *Global Query Optimizer* (Otimizador e decompositor de consultas globais)
 - Determina uma estratégia de execução para minimizar uma função de custo
 - Converte as consultas globais em consultas locais, usando os esquemas conceituais globais, assim como o diretório global



Arquiteturas de SGBD Distribuído

- *Global Execution Monitor* (Monitor de execução global)
 - Coordena a execução distribuída da solicitação do usuário
 - Durante a execução de consultas de forma distribuída, os monitores de execução em vários sites podem se comunicar uns com os outros



Arquiteturas de SGBD Distribuído

- *Local Query Processor* (Processador (Otimizador) de consulta local)
 - É responsável pelo melhor caminho para acessar qualquer item de dados
 - Atua como um seletor de caminho de acesso
- *Local Recovery Manager* (Gerenciador de recuperação local)
 - É responsável pela garantia de que o banco de dados local permanecerá consistente, mesmo quando ocorrerem falhas



Arquiteturas de SGBD Distribuído

- *Runtime Support Processor* (Processador de suporte runtime)
 - Acessa fisicamente o banco de dados, de acordo com os comandos físicos no escalonamento gerado pelo otimizador de consultas
 - É a interface para o sistema operacional e contém o gerenciador de buffers(ou cache) do BD, que é responsável por manter o cache em memória principal e gerenciar o acesso aos dados



Arquiteturas de SGBD Distribuído

- Componentes de um SGBDD (arquitetura peer-to-peer)
 - A divisão entre processador do usuário e processador de dados é meramente organizacional
 - Cada máquina deverá ter os dois componentes!

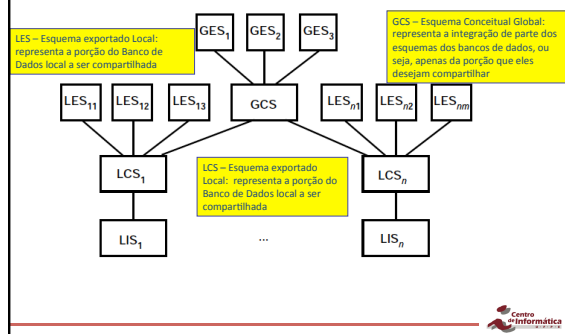


Arquiteturas de SGBD Distribuído

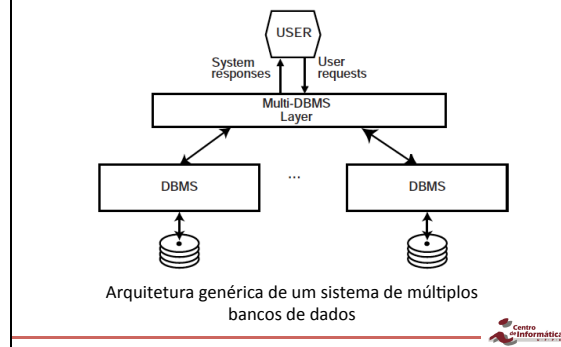
- Sistema de múltiplos bancos de dados heterogêneos (A2, D2, H1)
 - SGBDs individuais são completamente autônomos e não tem conhecimento de outros SGBDs
 - Também conhecidos como sistemas de integração de dados
 - Sistemas de integração podem envolver outros tipos de fontes de dados



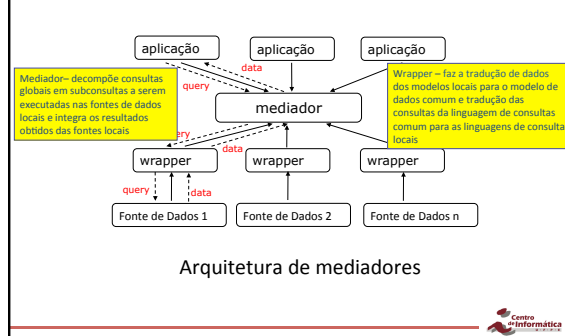
Arquiteturas de SGBD Distribuído



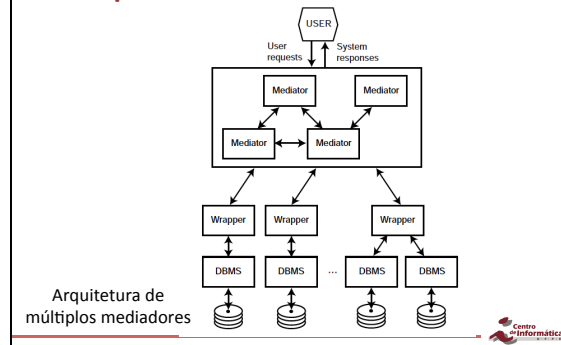
Arquiteturas de SGBD Distribuído



Arquiteturas de SGBD Distribuído



Arquiteturas de SGBD Distribuído



Conclusões

- Definir a arquitetura de um sistema é importante para especificar a sua estrutura
- Existem arquitetura de referências (geralmente são apenas ideais)
- Arquiteturas de SGBDDs são definidas com base em três fatores:
 - Autonomia, Distribuição e Heterogeneidade
- Possíveis arquiteturas
 - SGBD Distribuído Cliente/Servidor
 - SGBD Distribuído Peer-to-Peer Sistema de múltiplos bancos de dados heterogêneos