

Introdução a SGBD



Centro de Informática
UFPE

Valéria Times

TÓPICOS

- ◆ Introdução à Implementação de SGBD
- ◆ Armazenamento de Dados
- ◆ Representação de Elementos de Dados
- ◆ Estruturas de Índice
- ◆ Aspectos Operacionais de SGBD
- ◆ Gerenciamento de Transações
- ◆ Controle de Concorrência
- ◆ Recuperação após Falhas
- ◆ Segurança
- ◆ Processamento e Otimização de Consulta
- ◆ Sintonia em BD

3/5/2012 © CIn/UFPE 2

Introdução à SGBD

- ◆ Revisão de Terminologia
 - **Dados:** qualquer informação que valha a pena preservar, mais provavelmente em formato eletrônico.
 - **Banco de Dados (BD):** coleção de dados organizados para facilitar o acesso e a modificação, preservada durante um longo período.
 - **SGBD:** Coleção de dados interrelacionados e um conjunto de programas para acessá-los.
 - **Consulta:** Operação que extrai dados especificados do BD.

3/5/2012 © CIn/UFPE 3

Introdução à SGBD

- ◆ Revisão de Terminologia (cont.)
 - **Relação:** organização de dados em uma tabela bidimensional, na qual:
 - **linhas(tuplas):** representam entidades básicas ou fatos de algum tipo.
 - **colunas(atributos):** representam propriedades destas entidades.
 - **Esquema:** descrição da estrutura dos dados em um BD (metadados ou intenção).

3/5/2012 © CIn/UFPE 4

Introdução à SGBD

- ◆ SGBD são caracterizados por:
 - Admitir o acesso **eficiente** a grandes quantidades de dados.
 - Permitir que os dados **persistam** por muito tempo com **segurança**.
 - Suportar linguagens de consultas poderosas.
 - Gerenciar transações **duráveis** que podem ser executadas de forma **concorrente** e de um modo que parece ser **atômico** e **independente** de outras transações.

3/5/2012 © CIn/UFPE 5

Introdução à SGBD

- ◆ Comparação com Sistemas de Arquivos:
 - São inadequados porque não suportam:
 - recuperação eficiente e complexa de dados,
 - alterações eficientes em pequenos fragmentos de dados,
 - bufferização controlada de dados úteis na memória principal,
 - execução atômica e independente de transações.

3/5/2012 © CIn/UFPE 6

Introdução à SGBD

◆ Principais Recursos de um SGBD

- **Armazenamento Persistente:** registro e acesso a grandes quantidades de dados de forma eficiente.
- **Interface de Programação:** flexibilidade para manipular dados armazenados.
- **Recuperação após Falhas:** capacidade de se recuperar de muitos tipos de falhas ou erros.
- **Segurança:** possibilita a criação de várias contas de usuário e de diferentes tipos de acesso aos dados.

3/5/2012 © CIn/UFPE 7

Introdução à SGBD

◆ Principais Recursos de um SGBD (Cont.)

- **Controle de Concorrência:** admite o acesso simultâneo a dados por muitos processos distintos (transações).
- **Gerenciamento de Transações:** para evitar algumas das consequências indesejáveis da concorrência:
 - **Isolamento:** aparência de que as transações são executadas uma de cada vez.
 - **Atomicidade:** requisito de que as transações não sejam executadas de forma alguma ou sejam completamente.

3/5/2012 © CIn/UFPE 8

Introdução à SGBD

◆ Principais Recursos de um SGBD (Cont.)

- **Otimização de Consultas:** determina uma estratégia para execução da consulta de forma otimizada.

3/5/2012 © CIn/UFPE 9

Introdução à SGBD

◆ Por que Controle de Concorrência?

- Ausência de Concorrência degrada o desempenho do sistema.
- Quando vários usuários acessam (e possivelmente modificam) um BD simultaneamente, o SGBD deve **ordenar** as transações de usuário cuidadosamente para evitar **conflitos**.

3/5/2012 © CIn/UFPE 10

Introdução à SGBD

◆ Que Conflitos?

- Considere um BD contendo informações sobre reservas aéreas e vários vôos.
 - Um usuário pode identificar uma poltrona disponível em um certo vôo, enquanto outro usuário pode simultaneamente estar reservando a mesma poltrona.
- Considere uma aplicação bancária.
 - Enquanto um programa de aplicação estiver calculando o saldo de uma dada conta, outra aplicação pode estar transferindo um valor desta mesma conta para outra.

3/5/2012 © CIn/UFPE 11

Introdução à SGBD

◆ Por que Recuperação após Falhas?

- SGBD deve proteger os usuários dos **efeitos** das falhas, garantindo que todos os dados são consistentes mesmo após estes erros e a reinicialização do sistema.

◆ Que efeitos?

- Se uma operação de BD solicitada por um usuário é realizada e confirmada pelo SGBD, então o resultado desta operação não deve ser perdido se o sistema falha.

3/5/2012 © CIn/UFPE 12

Introdução à SGBD

- ◆ **Que efeitos? (Cont.)**
 - Por outro lado, se o SGBD não tiver confirmado ainda a realização da operação, mas já estiver realizando algumas mudanças aos dados no momento da falha, então estas modificações parciais devem ser desfeitas.

3/5/2012 © CIn/UFPE 13

Introdução à SGBD

- ◆ **Principais componentes de um SGBD**
 - **Processamento de Consultas**
 - **Gerenciamento de Transações**
 - **Gerenciamento de Armazenamento**

↓

Diferentes **Tipos de Usuários**: submetendo diversos comandos

3/5/2012 © CIn/UFPE 14

Tipos de Usuários

- ◆ **Fontes de Comandos para o SGBD**
 - **Usuários Convencionais e Programas Aplicativos**: solicitam ou modificam dados.
 - **Administradores de BD**: fazem a manutenção da estrutura/esquema do BD via DDL.
 - Necessitam de autoridade especial.
 - São analisados por um processador de DDL e repassados ao mecanismo de execução.

3/5/2012 © CIn/UFPE 15

Processamento de Consultas

- ◆ **Procedimentos do SGBD para:**
 - **Resolução da Consulta**
 - É analisada gramaticalmente e aperfeiçoada por um compilador de consultas.
 - Plano de consulta resultante é repassado ao mecanismo de execução.
 - Em seguida, pequenos itens de dados são solicitados ao gerenciador de disco.
 - Estas solicitações de dados são convertidas em páginas e repassadas ao gerenciador de buffers.
 - Dados (página ou bloco de disco) são transferidas do disco para a memória.

3/5/2012 © CIn/UFPE 16

Processamento de Consultas

- ◆ **Processador de Consultas**: Possui os seguintes componentes:
 - **Compilador de Consultas**:
 - Converte a consulta para uma forma interna chamada de **Plano de Consulta**.
 - Plano de Consulta: seqüência de operações a serem executadas sobre os dados.
 - **Mecanismo de Execução**:
 - É responsável por executar cada uma das etapas do plano de consulta escolhido.

3/5/2012 © CIn/UFPE 17

Processamento de Consultas

- ◆ **Processador de Consultas (Cont.):**
 - **Compilador de Consultas**: Consiste em:
 - **Analisador de Consultas**: Constrói uma estrutura em árvore a partir da forma textual da consulta.
 - **Pré-processador de Consultas**: Valida semanticamente a consulta. (e.g. vendo se todas as relações citadas na consulta existem realmente).
 - **Otimizador de Consultas**: Transforma o plano de consulta inicial na melhor seqüência possível de operações sobre os dados.

3/5/2012 © CIn/UFPE 18

Processamento de Consultas

- ◆ **Processador de Consultas (Cont.):**
 - **Compilador de Consultas:**
 - **Otimizador de Consultas:** Usa metadados e estatísticas sobre os dados para decidir qual seqüência de operações será provavelmente mais rápida.
 - Por exemplo, a existência de um índice pode tornar um plano bem mais eficiente do que outro.

3/5/2012 © Cin/UFPE 19

Gerenciamento de Transações

- ◆ **Processamento de Transações:**
 - Uma ou mais operações de BD podem ser agrupadas em **transações**.
 - **Definição:**
 - **Transações** descrevem unidades lógicas de processamento de BD.
 - Consistem em unidades de trabalho que devem ser executadas:
 - **atomicamente**
 - em **isolamento** aparente de outras transações
 - de modo que toda ação concluída nunca deve ser perdida.

3/5/2012 © Cin/UFPE 20

Gerenciamento de Transações

- ◆ **Processamento de Transações (Cont.):**
 - Comandos de transação são repassados ao gerenciador de transações pelos aplicativos, informando:
 - início e fim das transações
 - preferências dos aplicativos
 - Durante sua execução, uma transação pode acessar dados compartilhados por outras transações.
 - Em caso de operações de atualização, o dado não deve ser compartilhado (**bloqueio**)
 - Para evitar inconsistência

3/5/2012 © Cin/UFPE 21

Gerenciamento de Transações

- ◆ **Processamento de Transações (Cont.):**
 - Se uma transação não puder ser concluída satisfatoriamente, ela deve ser cancelada (**aborted**)
 - Voltar os dados ao estado original
 - Se a transação termina corretamente, ela é efetivada (**committed**)
 - Disponibilizar seu resultado para transações subsequentes.
 - Transação como unidade de consistência:
 - Uma operação que, dado um BD consistente, depois de completada deixa o BD de novo em estado consistente.

3/5/2012 © Cin/UFPE 22

Gerenciamento de Transações

- ◆ **Processador de Transações:**
 - É dividido em:
 - **Gerenciador de Controle de Concorrência:** assegura a atomicidade e o isolamento de transações.
 - **Gerenciador de registro de log e recuperação:** responsável pela durabilidade de transações.
 - Executa as seguintes tarefas:
 - Registro de **log**
 - Controle de Concorrência
 - Resolução de Impasse

3/5/2012 © Cin/UFPE 23

Gerenciamento de Transações

- ◆ **Registro de Log:**
 - Toda mudança no BD é registrada separadamente em disco para garantir durabilidade.
 - **Gerenciador de Log do SGBD:**
 - Inicialmente atualiza o **registro de log** em buffers em seguida negocia com o **gerenciador de buffers** o momento da transferência deles para o disco.
 - Assegura que caso ocorra uma falha, o **gerenciador de recuperação** possa examinar o log de mudanças e estabelecer o BD a algum estado consistente.

3/5/2012 © Cin/UFPE 24

Gerenciamento de Transações

- ◆ **Controle de Concorrência:** Gerenciador de controle de concorrência do SGBD:
 - Assegura que as ações individuais de várias transações sejam executadas em uma dada **ordem** de modo que o efeito final seja igual ao que haveria se as transações fossem de fato executadas em sua totalidade, uma de cada vez.
 - Mantém **bloqueios** sobre certos itens de dados, impedindo que duas transações tenham acesso ao mesmo item de dados por meios que interajam de forma incorreta.

3/5/2012 © CIn/UFPE 25

Gerenciamento de Transações

- ◆ **Controle de Concorrência (Cont.):**
 - **Bloqueios** são armazenados em uma **tabela de bloqueios** em memória principal.
 - **Gerenciador de controle de concorrência** afeta a execução das operações de BD, proibindo que o **mecanismo de execução** tenha acesso a partes bloqueadas do BD.

3/5/2012 © CIn/UFPE 26

Gerenciamento de Transações

- ◆ **Propriedades das Transações:** Se são implementadas de forma apropriada então elas satisfazem as propriedades **ACID**, onde:
 - **A (Atomicidade):** execução de transações na base do tudo ou nada.
 - **C (Consistência):** transações devem preservar a consistência do BD.
 - **I (Isolamento):** fato de que cada transação deve aparentar ser executada como se nenhuma outra transação estivesse em execução ao mesmo tempo.

3/5/2012 © CIn/UFPE 27

Gerenciamento de Transações

- ◆ **Propriedades das Transações (Cont.):**
 - **D (Durabilidade):** condição de que o efeito de uma transação sobre o BD nunca deve ser perdido, mesmo depois da transação ter sido concluída.

3/5/2012 © CIn/UFPE 28

Gerenciamento de Transações

- ◆ **Resolução de Impasse:**
 - Competindo por recursos, as transações podem chegar a uma situação em que nenhuma delas pode prosseguir porque cada uma necessita de algo que a outra tem.
 - Gerenciador de transação é responsável por intervir e cancelar (*abortar*) uma ou mais transações, a fim de permitir que as outras prossigam.

3/5/2012 © CIn/UFPE 29

Gerenciamento de Armazenamento

- **Transferir Dados para Memória Principal:**
 - Todos os componentes do SGBD que necessitam de informação do disco devem interagir com os buffers e o gerenciador de buffers, seja diretamente ou através do mecanismo de execução.
 - **Gerenciador de Buffers:** responsável por particionar a memória principal disponível em **buffers**.
 - **Buffers:** regiões com dimensões de página para as quais os blocos de disco podem ser transferidos.

3/5/2012 © CIn/UFPE 30

Gerenciamento de Armazenamento

- **Transferir Dados para Memória Principal (Cont.):**
 - Tipos de informação que podem ser necessários:
 - **Dados:** o conteúdo do próprio BD.
 - **Metadados:** esquema do BD que descreve sua estrutura e restrições.
 - **Estatísticas:** informações obtidas e armazenadas pelo SGBD sobre propriedades dos dados (e.g. tamanho das relações).
 - **Índices:** estruturas de dados que admitem o acesso eficiente aos dados.

3/5/2012 © CIn/UFPE 31

