

Controle de Dados e Acesso

IN1128/IF694 – Bancos de Dados Distribuídos e Móveis
 Ana Carolina Salgado – acs@cin.ufpe.br
 Bernadette Farias Lôscio – bfl@cin.ufpe.br



Cin.ufpe.br

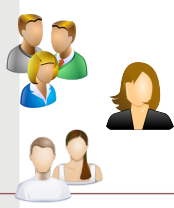
Roteiro

- O que é controle de dados e acesso?
- Gerenciamento de Visões
 - O que são visões?
 - Atualização de visões
- Segurança de dados
 - Proteção de dados
 - Controle de autorização
- Controle de integridade semântica



Controle de dados e acesso

- Garantia de que usuários autorizados realizarão operações corretas sobre o BD



Controle de dados e acesso

- Manutenção da consistência através de regras pré-definidas



Controle de dados e acesso

- Manutenção da consistência através de regras pré-definidas
- Envolve os conceitos:
 - Gerenciamento de visões: meio de limitar o acesso aos dados do banco
 - Controle de segurança: engloba a proteção dos dados e o controle de acesso
 - Controle de integridade semântica: diz respeito a como manter a consistência do banco de dados



Gerenciamento de Visões

- O que são visões?
 - Resultado de uma consulta sobre uma relação
 - Exemplo:

Tabela "QUADROS"

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	Persistence of Memory	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000



Gerenciamento de Visões

- Criando uma visão**
 CREATE VIEW QUADROS_LEONARDO (ID, NOME) AS
 SELECT Q.ID, Q.NOME FROM QUADROS Q WHERE
 ID_AUTOR= "01"

Tabela "QUADROS"

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	Persistence of Memory	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000

Gerenciamento de Visões

- Restrição de acesso usando visões**
 Exemplo:
 Suponha que um guia da Pinacoteca só possa visualizar informações de outros guias
 CREATE VIEW MESMO_CARGO AS SELECT * FROM
 FUNC F1, FUNC F2 WHERE F1.CARGO = F2.CARGO
 AND F1.ID_FUNC = USER

Tabela "FUNC"

ID_FUNC	NOME	CARGO
01	Maria Antônia	Gerente
02	Renato Moura	Recepcionista
03	Geraldo Motta	Guia
04	Julio Cabral	Guia
05	Carina Andrade	Recep
06	André Farias	Guia

Visão "MESMO_CARGO"

ID_FUNC	NOME	CARGO
03	Geraldo Motta	Guia
04	Julio Cabral	Guia
05	Carina Andrade	Recep
06	André Farias	Guia

Gerenciamento de Visões

- Atualização de visões**
 - Atualizáveis
 - Não-atualizáveis
- Exemplo**
 - Seja a visão criada com o ID e nome dos quadros pertencentes a Leonardo da Vinci

ID	NOME
01	Mona Lisa
05	A Última Ceia

Queremos inserir a tupla
<08, "A Virgem do Rochedo">

Gerenciamento de Visões

- Tabela QUADROS é atualizada adicionando o elemento <08, "A virgem dos Rochedos">**

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000
08	01	A Virgem do Rochedo	NULL

Gerenciamento de Visões

- Agora consideremos as seguintes relações**

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000
08	01	A Virgem dos	

Queremos:
Nome do quadro e o País_Origem do pintor no qual o ID é igual a 01

ID_AUTOR	NOME	PAIS_ORIGEM
01	Leonardo da Vinci	Itália
02	Edvard Munch	Noruega
03	Salvador Dalí	Espanha
04	Vincent Van Gogh	Holanda
05	Michelangelo	Itália

Tabela "QUADROS"

Tabela "AUTOR"

Gerenciamento de Visões

- O resultado da criação da visão será**

NOME	PAIS_ORIGEM
Mona Lisa	Itália
A Última Ceia	Itália
A Criação de Adão	Itália

- Suponha que agora queremos excluir a tupla <"A Última Ceia", "Itália">**

Gerenciamento de Visões

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000
08	01	A Virgem de Lencinhos	1.000.000

ID_AUTOR	NOME	PAIS_ORIGEM
01	Leonardo da Vinci	Itália
02	Edvard Munch	Noruega
03	Salvador Dalí	Espanha
04	Vincent Van Gogh	Holanda
05	Michelangelo	Itália

Tabela "QUADROS"

Tabela "AUTOR"

Gerenciamento de Visões

- **Visões em BDs distribuídos**
 - Nome e consulta de recuperação armazenados em catálogo
 - Catálogos centralizado, semi-replicado ou totalmente replicado

Gerenciamento de Visões

- **Visões em BDs distribuídos**

ID	NOME
01	Mona Lisa
05	A Última Ceia

```
SELECT NOME FROM
QUADROS_LEONARDO
WHERE ID = 01
```

Visão "QUADROS_LEONARDO"

- Internamente, é necessário que esta consulta seja modificada para que o processador de consultas possa achar seu resultado. Mas, modificada como?

Gerenciamento de Visões

- **Modificando a consulta:**

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000

ID	NOME
01	Mona Lisa
05	A Última Ceia

Consulta sobre a visão
SELECT ID FROM QUADROS_LEONARDO WHERE NOME = "Mona Lisa"

Consulta Modificada
SELECT Q.ID FROM QUADROS Q WHERE Q.NOME = "Mona Lisa"
AND Q.ID_AUTOR = "01"

Gerenciamento de Visões

- **A avaliação de visão é uma operação custosa em BDs distribuídos**
- **Atualização de visões para cada nova consulta**
- **Uso de Visões Materializadas**
 - A consulta é executada e os dados são armazenados localmente
 - Manutenção de visões materializadas
 - Quando e Como?
 - Quando: imediata (como parte da transação que atualiza os dados base da visão) ou adiada
 - Como: incremental ou rematerialização

Segurança de Dados

- **Protege dados contra acesso não autorizado**
- **Dois aspectos:**
 - Proteção de dados
 - Controle de autorização
- **Proteção de dados**
 - Evita que usuários não autorizados reconheçam o conteúdo dos dados
 - Utiliza-se da criptografia

Segurança de Dados

- **Controle de autorização**
 - Garantir que usuários autorizados realizem apenas operações sobre as quais ele tem permissão
 - Verifica a tripla <Usuário, Operação, Objeto> ;

Segurança de Dados

- **Controle de autorização**
 - Verifica a tripla <Usuário, Operação, Objeto> ;

Segurança de Dados

- **Controle de autorização**
 - As regras de autorização são armazenadas no catálogo
 - Uso de matriz de autorização para representar as regras de autorização

	QUADROS	FUNC	VALOR
Maria Antonia	UPDATE	INSERT	UPDATE
Renato Moura	NENHUM	UPDATE	NENHUM
Geraldo Motta	SELECT	SELECT WHERE CARGO = "Guia"	NENHUM

Segurança de Dados

- **Três formas de armazenar a matriz de autorização**
 - Por linha
 - Por coluna
 - Por elemento

	QUADROS	FUNC	VALOR
Maria Antonia	UPDATE	INSERT	UPDATE
Renato Moura	NENHUM	UPDATE	NENHUM
Geraldo Motta	SELECT	SELECT WHERE CARGO = "Guia"	NENHUM

Segurança de Dados

- **Três formas de armazenar a matriz de autorização**
 - Por linha

	QUADROS	FUNC	VALOR
Maria Antonia	UPDATE	INSERT	UPDATE
Renato Moura	NENHUM	UPDATE	NENHUM
Geraldo Motta	SELECT	SELECT WHERE CARGO = "Guia"	NENHUM

Vantagens? Desvantagens?

Segurança de Dados

- **Três formas de armazenar a matriz de autorização**
 - Por coluna

	QUADROS	FUNC	VALOR
Maria Antonia	UPDATE	INSERT	UPDATE
Renato Moura	NENHUM	UPDATE	NENHUM
Geraldo Motta	SELECT	SELECT WHERE CARGO = "Guia"	NENHUM

Vantagens? Desvantagens?

Segurança de Dados

- **Três formas de armazenar a matriz de autorização**
 - Por elemento

	QUADROS	FUNC	VALOR
Maria Antonia	UPDATE	INSERT	UPDATE
Renato Moura	NENHUM	UPDATE	NENHUM
Geraldo Motta	SELECT	SELECT WHERE CARGO = "Guia"	NENHUM



Segurança de Dados

- **Controle de autorização distribuído**
 - Autenticação remota de usuários
 - Gerenciamento de regras de autorização distribuídas
 - Tratamento de visões
 - Controle de grupo de usuários



Segurança de Dados

- **Autenticação remota dos usuários**
 - Autenticação por usuários (a inserção de novos usuários pode ser custosa)
 - Autenticação por site (se a informação do usuário não está replicada)
 - Autenticação em um site global ou em cada site (assume que os usuários sempre fazem acessos de um mesmo site)
- **Gerenciamento de regras de autorização**
 - Replicação total das informações nos sites
 - Replicação somente nos sites onde possuem objetos referenciados



Segurança de Dados

- **Tratamento de visões**
 - Visão como objeto composto
 - O acesso a uma visão implica o acesso aos seus objetos componentes
 - Replicação total (mais comum)
 - Armazenamento local, ou seja, a visão e seus componentes armazenados em sites distintos (autonomia de dados)
 - Necessidade de verificação de acesso em cada site
- **Controle de grupo de usuários**
 - Replicação total
 - Replicação parcial das autorizações
 - Replicação parcial dos grupos



Controle de Integridade Semântica

- **Garantia da consistência do banco de dados rejeitando transações que levam o banco de dados a um estado inconsistente**

- O banco de dados deve satisfazer um conjunto de restrições definidas.



Controle de Integridade Semântica

- **Restrições de integridade**
 - Restrições estruturais
 - Relacionadas a propriedades semânticas básicas como restrição de chave exclusiva
 - Restrições comportamentais
 - Relacionadas com o comportamento do aplicativo
- **Manutenção pode ser feita de forma procedural (embutida nas aplicações) ou declarativa (assertivas definidas no próprio SGBD)**
 - O conj. de assertivas de integridade semântica definem a consistência do banco de dados



Controle de Integridade Semântica

- **Como expressar estas restrições ao banco de dados?**
 - Uso de assertivas
- **No sistema de controle de integridade existem dois componentes principais**
 - Uma linguagem para manipular as assertivas
 - Um mecanismo de imposição cuja função é assegurar o cumprimento das regras



Controle de Integridade Semântica

- **Restrições de integridade**
 - Pré-definidas
 - Baseadas em keywords (not null, foreign key...)
 - Pré-condições
 - Expressam condições que devem ser satisfeitas por todas as tuplas de uma relação dado um determinado tipo de atualização
 - CHECK ON PROJETO WHEN DELETE (ORÇAMENTO = 0)
 - Gerais
 - Podem envolver mais de uma relação e todas as variáveis são quantificadas
 - CHECK ON e1.EMP, E2:EMP
(e1.NOME = e2.NOME IF e1.NUM = e2.NUM)



Controle de Integridade Semântica

- **Exemplos de assertivas que operam sobre o nosso sistema da Pinacoteca**
 - NOME NOT NULL IN QUADROS

Tabela "QUADROS"

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000



Controle de Integridade Semântica

- **Exemplos de assertivas que operam sobre o nosso sistema da Pinacoteca**
 - CHECK ON QUADROS (NOVO.VALOR > ANT.VALOR AND NOVO.ID = ANT.ID)

Tabela "QUADROS"

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000



Controle de Integridade Semântica

- **Imposição de integridade**
 - Rejeita atualizações que levam o banco de dados a estados inconsistentes
 - Dois métodos de rejeição
 - Detecção de inconsistências
 - Prevenção de inconsistências



Controle de Integridade Semântica

- **Detecção de inconsistências**
 - Solução baseada em posttests
 - Testes se aplicam no novo estado do banco de dados
 - Solução ineficiente se uma grande quantidade de trabalho tiver que ser desfeita no caso de uma falha



Controle de Integridade Semântica

■ Prevenção de inconsistências

- Solução baseada em pré-testes
- Mais eficientes, pois não é necessário desfazer o trabalho que levará o banco de dados a um estado incorreto.



Controle de Integridade Semântica

■ Modificação de consulta

- Método elegante para efetuarmos uma imposição de restrição
- Modifica a consulta com o acréscimo da restrição

ID	ID_AUTOR	NOME	VALOR
01	01	Mona Lisa	1.000.000
02	02	O Grito	300.000
03	03	A Persistência da Memória	450.000
04	04	Autoretrato	700.000
05	01	A Última Ceia	990.000
06	05	A Criação de Adão	500.000

Atualização

```
UPDATE QUADROS SET VALOR = VALOR * 1.1
WHERE NOME = "Mona Lisa"
```

Atualização Modificada com a restrição

```
UPDATE QUADROS SET VALOR = VALOR * 1.1
WHERE NOME = "Mona Lisa"
AND NOVO.VALOR > ANT.VALOR
```

Tabela "QUADROS"



Controle de Integridade Semântica

■ Assertivas compiladas

- Reduz a quantidade do banco de dados que deve ser verificada
- Definida pela tripla <Relação, Atualização, Assertiva >



Verifica apenas as assertivas relacionadas a essa atualização no conjunto de dados a ser inseridos



Controle de Integridade Semântica

■ Controle distribuído da integridade semântica

- Dois principais problemas
 - Definição e armazenamento de assertivas
 - Imposição das assertivas

- O objetivo é minimizar os custos com a comunicação, definindo então os melhores lugares para armazenamento de cada assertiva
- As estratégias a serem utilizadas para armazenamento e imposição depende do tipo da assertiva



Controle de Integridade Semântica

■ Assertivas individuais

- Envolvem uma relação única e uma variável única;
- Armazenamento
 - A definição da assertiva é enviada a todos os sites que contêm fragmentos da relação envolvida



Controle de Integridade Semântica

■ Assertivas individuais

- A restrição deve ser compatível com os dados de cada site
- A compatibilidade é checada em dois níveis:
 - Predicados
 - Dados
 - A compatibilidade de predicado é feita comparando o predicado da restrição com o predicado do fragmento
 - Se os predicados forem compatíveis então é feito o teste para verificar se a instância do fragmento não viola a restrição
 - Caso existe alguma incompatibilidade então a restrição é globalmente rejeitada



Controle de Integridade Semântica

■ Assertivas individuais - exemplo

- P1: $0 \leq \text{sal} < 1000$
- P2: $1000 \leq \text{sal} < 5000$
- P3: $5000 \leq \text{sal}$

- C: $\text{sal} \leq 4000$

■ A restrição C é compatível com P1 e P2!



Controle de Integridade Semântica

■ Imposição da restrição (garantia de que a restrição será obedecida)

- Se a atualização for uma inserção, todas as tuplas a serem inseridas são explicitadas pelo usuário e todas as assertivas individuais deverão ser avaliadas no site em questão
- Se a atualização for de exclusão ou modificação, as assertivas serão enviadas aos locais que armazenam a relação (ou fragmentos da relação)



Controle de Integridade Semântica

■ Assertivas orientadas a conjuntos

- Envolvem várias relações e variáveis
- Consideram predicados de junção
- Pré-testes são realizados em relações individuais
- A restrição precisa ser enviada a todos os sites que possuem fragmentos associados a variáveis usadas na restrição
- É feita apenas a checagem de compatibilidade de dados (não adianta checar a compatibilidade dos predicados: o do fragmento envolve uma relação e o da restrição é de junção)
 - É feita a junção de cada fragmentos de R com cada um dos fragmentos de S envolvidos no predicado de junção, onde R e S são relações que fazem parte da restrição



Controle de Integridade Semântica

■ Assertivas orientadas a conjuntos

- É feita apenas a checagem de compatibilidade de dados (não adianta checar a compatibilidade dos predicados: o do fragmento envolve uma relação e o da restrição é de junção)
 - É feita a junção de cada fragmentos de R com cada um dos fragmentos de S envolvidos no predicado de junção, onde R e S são relações que fazem parte da restrição



Controle de Integridade Semântica

■ Assertivas orientadas a conjuntos

- Imposição (checagem do cumprimento da restrição)
 - Qualificação da atualização transformando em relações temporárias
 - Envio das relações temporárias para os sites onde as junções devem ser realizadas
 - Junção dos resultados no site mestre de consulta
 - Resposta do sistema depende da implementação



Controle de Integridade Semântica - resumo

- A especificação das restrições de integridade distribuídas é uma etapa importante do projeto de um BDD
 - Os fragmentos devem ser definidos de forma cuidadosa
- Os custos de comunicação e processamento podem ser proibitivos no controle de restrições de integridade distribuídas
 - As restrições podem envolver tabelas (ou fragmentos) distribuídos ao longo da rede
- O controle pode ser feito a partir da definição de pré-testes

