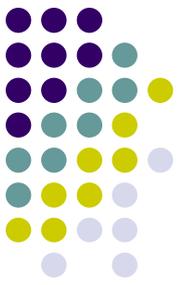
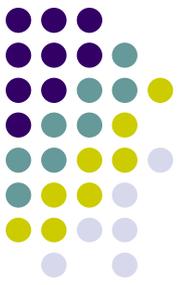


Processamento de Consultas Distribuídas



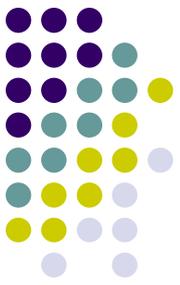
- Preocupação em minimizar o tempo total para a apresentação do resultado alcançado.
- Crescimento dos problemas: além do acesso a disco, existe também a preocupação com a localização e transmissão de dados na rede.
- Para recuperar as todas as tuplas da relação CONTA, o processamento pode não ser trivial - CONTA pode estar fragmentada, replicada ou ambas.

Processamento de Consultas Distribuídas



- Pode haver a necessidade de diversas junções e uniões para a reconstrução da relação original (CONTA).
- Lembrando que:
O processamento de consultas em SGBDs tem como aspecto mais importante a escolha da forma de junção a ser realizada.

Processamento de Consultas Distribuídas



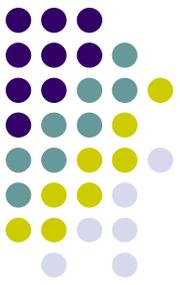
- Consideremos a junção

CONTA X DEPOSITANTE X AGENCIA

no seguinte contexto:

- nenhuma das três relações foi replicada ou fragmentada
- CONTA esta no Sítio 1
- DEPOSITANTE no Sítio 2
- AGÊNCIA no Sítio 3
- origem da consulta é no Sítio 1, onde deverá ser produzido o resultado da mesma.

Processamento de Consultas Distribuídas



Algumas das estratégias para realização da consulta são:

- enviar cópia das três relações para o Sítio 1, onde será aplicada a melhor estratégia de consulta local
- enviar cópia de CONTA para o Sítio 2, e lá obter $\text{temp1} = (\text{CONTA} \times \text{DEPOSITANTE})$; em seguida enviar temp1 do Sítio 2 para o Sítio 3 obtendo lá o $\text{temp2} = (\text{temp1} \times \text{AGÊNCIA})$; por fim enviar o resultado (temp2) para o Sítio 1
- aplicar estratégia similar à anterior, alterando as regras entre os Sítios 1, 2 e 3

Processamento de Consultas Distribuídas



Primeira estratégia: necessidade de re-processamento dos índices da relação transportada, configurando um **processamento extra**.

Segunda estratégia: ocorre a transmissão de uma relação potencialmente maior.

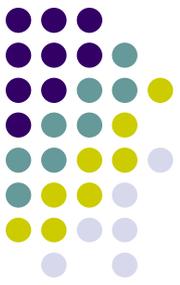
Necessariamente, não existe a melhor estratégia, mas deve-se considerar:

1 - volume de dados a ser transportado

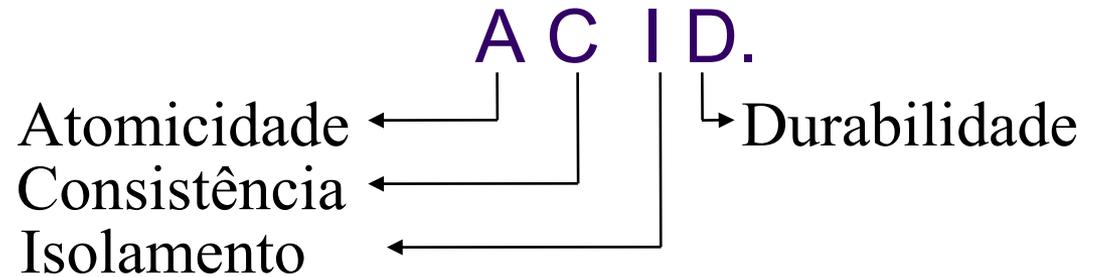
2 - custo de transmissão de um Sítio para outro

3 - velocidade relativa de processamento em cada Sítio

Processamento de Transações Distribuídas



Transações distribuídas devem preservar as propriedades:



Podem ocorrer dois tipos de transações:

- Locais - mantém acesso e atualizam somente a base de dados local
- Globais – mantém acesso e atualizam diversas bases de dados locais

Como garantir as propriedades ACID em um SBDD?

Processamento de Transações Distribuídas



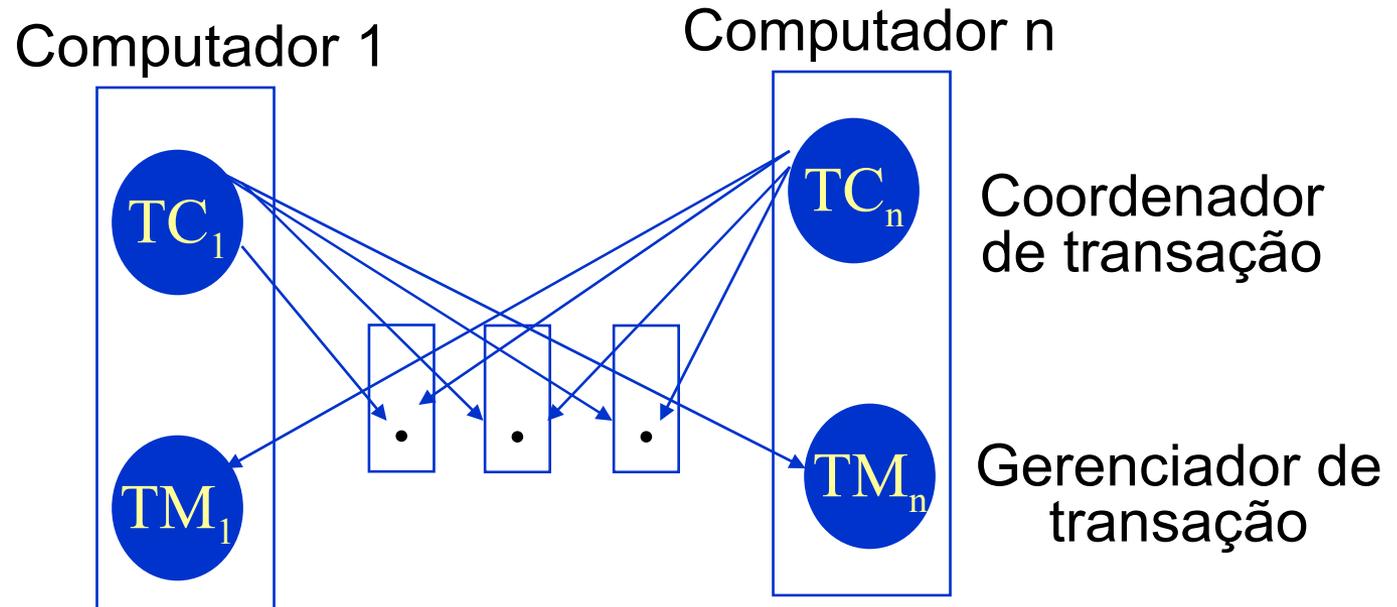
Exemplo de Modelo para Gerência Distribuída:

- Cada Sítio possui o seu próprio gerenciador de transações local, cuja função é garantir as propriedades ACID das transações executadas localmente.
- Os diversos gerenciadores cooperam para executar as transações globais
- Pode-se definir um modelo abstrato para o sistema de transações. Nesta definição cada Sítio possui dois subsistemas:
 - Gerenciador de transações
 - Coordenador de transações

Processamento de Transações Distribuídas



- O gerenciador administra a execução daquelas transações que mantêm acesso aos dados locais. Pode ser uma transação local ou parte de uma transação global.
- O coordenador administra a execução de várias transações (locais ou globais) iniciadas naquele Sítio.



Processamento de Transações Distribuídas



Tarefas do Gerenciador de Transações

- Manutenção de um *log* para propósito de recuperação
- Participação em um esquema de controle de concorrência adequado para coordenação da execução de transações concorrentes em um mesmo Sítio

Tarefas do Coordenador de Transações

- Iniciar a execução da transação
- Quebrar uma transação em um número determinado de subtransações e distribuí-las pelos Sítios apropriados para a execução
- Coordenar o término das transações, que podem resultar em efetivações em todos os Sítios ou em interrupção em todos os Sítios

Processamento de Transações Distribuídas



- SBDD: Sujeito aos mesmos tipos de falhas que ocorrem em um sistema centralizado

- Falhas adicionais podem ocorrer, como:

- falha em um Sítio,
- falha na comunicação entre eles,
- perda de mensagens
- Etc...

Cada um desses problemas deve ser considerados no projeto de recuperação de um SBDD.

Para garantir robustez, é necessário:

- detectar qualquer uma dessas falhas
- reconfigurar-se de modo a manter o processamento
- recuperar-se enquanto a falha é corrigida.

Processamento de Transações Distribuídas



Preocupações relevantes na manipulação de dados neste contexto são:

- Suporte a Protocolos de efetivação (garantir a atomicidade)
 - duas fases (obstrução) – 2PC
 - três fases – 3PC
- Controle de concorrência

Algumas alterações no modo de funcionamento e de controle do sistema centralizado podem permitir a funcionalidade deles em um sistema distribuído.