

Gerenciamento de Transações Móveis: Modelos para Computação Móvel

Giani Carla Ito
*Intituto Nacional de
Pesquisas espaciais – INPE*
E-mail: giito@lac.inpe.br

Maurício Ferreira
*Intituto Nacional de
Pesquisas espaciais – INPE*
E-mail: Mauricio@lac.inpe.br

Nilson Sant’Ana
*Intituto Nacional de
Pesquisas espaciais – INPE*
E-mail: nilson@lac.inpe.br

Resumo

O avanço da tecnologia de comunicação sem fio e a crescente utilização de computadores portáteis dá início a um novo paradigma denominado computação móvel, onde o usuário pode acessar dados e recursos do sistema de diferentes pontos de acesso, independente de sua localização ser fixa ou móvel. Comparada à computação tradicional, já incorporada naturalmente na vida das pessoas, a computação móvel promete uma revolução mais abrangente, devido à sua maior flexibilidade de localização podendo ocasionar grandes mudanças no comportamento social e corporativo.

Com base nestas colocações o artigo proposto tem como objeto de estudo bancos de dados móveis com ênfase em gerenciamento de transações móveis, reunindo conceitos sobre computação móvel, uma arquitetura de bancos de dados móveis.

Palavras Chaves : Bancos de dados móveis, computação móvel, gerenciamento de dados.

1. Introdução

A utilização de bancos de dados móveis propicia facilidades de locomoção física que possibilitam aos usuários acesso aos dados com poucas ou nenhuma restrição

Neste contexto, o termo computação onipresente (*Ubiquitous Computing*) refere-se a ambientes com objetos operados por computador e conectados em redes sem fio, ou seja, mobilidade, comunicação e poder de processamento integrado em vários objetos com finalidades diferentes.

A computação móvel consiste em um sistema composto por dois tipos distintos de entidades, denominadas computadores móveis e computadores fixos [1], relacionados a seguir.

2.1. Computadores Móveis

As unidades ou *hosts* móveis são computadores portáteis, interligados em rede através de ligações sem fio, que permitem a mobilidade estabelecendo uma conexão virtual

de tempo e lugar. Usuários têm a possibilidade de participar de congressos, convenções e reuniões de negócios e realizar visitas a clientes e fornecedores, tendo acesso a seus arquivos pessoais armazenados em uma estação de trabalho distante, participando de teleconferências e efetuando, normalmente, suas tarefas computacionais diárias, mesmo quando distantes de sua residência ou local de trabalho.

Outro exemplo da utilização de bancos de dados móveis é o de pesquisadores que estão viajando regularmente pelo mundo e precisam entrar em contato com laboratórios para atualização de dados. Através dos recursos móveis, o pesquisador não precisa enviar a todo o momento dados coletados para a rede fixa. Poderá fazer isso somente quando for necessário.

2. Arquitetura da Computação Móvel

A computação móvel utiliza computadores portáteis e comunicação sem fio para permitir aos usuários trabalharem fora de ambientes fixos. O ambiente móvel baseia-se na capacidade que os usuários têm, munidos de um dispositivo móvel (*laptops, handhelds, etc*), se comunicarem com a parte fixa da rede e possivelmente com outros dispositivos móveis, independentemente da sua localização. A comunicação sem fio pode ser realizada através de sistema

de qualquer localização dentro de uma área geográfica, como mostra a figura 1.

Computadores portáteis são dependentes de bateria, possuindo ainda algumas restrições físicas, como poder de armazenamento e processamento mais baixos se comparados aos computadores pessoais.

2.2. Computadores Fixos

Os computadores ou estações fixos fazem parte de uma rede fixa cuja localização e conectividade não mudam. São utilizados para fornecer vários tipos de informações para os usuários. As mensagens destinadas às unidades móveis são enviadas em primeiro lugar para a estação fixa, que

mantém a localização da unidade móvel e envia as mensagens através das estações base.

As estações-base também conhecidas como *mobile support station* (estações de suporte móveis) realizam a comunicação entre a rede fixa e computadores móveis. São identificadas como servidores e estão localizadas próximas ou em uma torre de antena, que transmite ou recebe sinais eletromagnéticos dos dispositivos numa área específica. A altura, o design e o tamanho da antena são fatores que determinam o alcance do sinal que será transmitido. As estações-base são conectadas via linhas fixas para os nós da rede sem fio.

Para gerenciar a mobilidade das unidades móveis o domínio geográfico é dividido em pequenos subdomínios, denominados células cobertas por interface sem fio. Cada célula é gerenciada por uma estação base, com transmissores e receptores para responder ao processamento de informações necessárias dos clientes localizados na célula. Caso os computadores móveis estejam próximos, a comunicação poderá ser realizada sem a intervenção de uma estação de suporte móvel. As células podem variar de tamanho, cobrindo desde um prédio até grandes áreas geográficas.

software de comunicação, o software responsável pelo gerenciamento do sistema e a aplicação de bancos de dados móvel.

Os usuários interagem com a aplicação do banco de dados e acessam informações armazenadas através de um SGBD-D. Os softwares para um ambiente distribuído devem tratar problemas de comunicação, desconexão, topologia dinâmica, entre outros fatores.

O banco de dados servidor normalmente é um computador de médio ou grande porte que é coordenado por um sistema de gerenciamento. O *host* que armazena o bancos de dados contém dados dinâmicos que podem ser usados pelos usuários móveis. A posição do *host* na organização depende do tipo de informações que está sendo processado.

O banco de dados móvel troca informações com o computador que contém o banco de dados, de modo a manter a atualização, realizar consultas, etc. A comunicação entre ambos acontece com interrupções e em intervalos irregulares, por curtos períodos de tempo, sendo que, algumas vezes, não estão conectados em uma mesma rede.

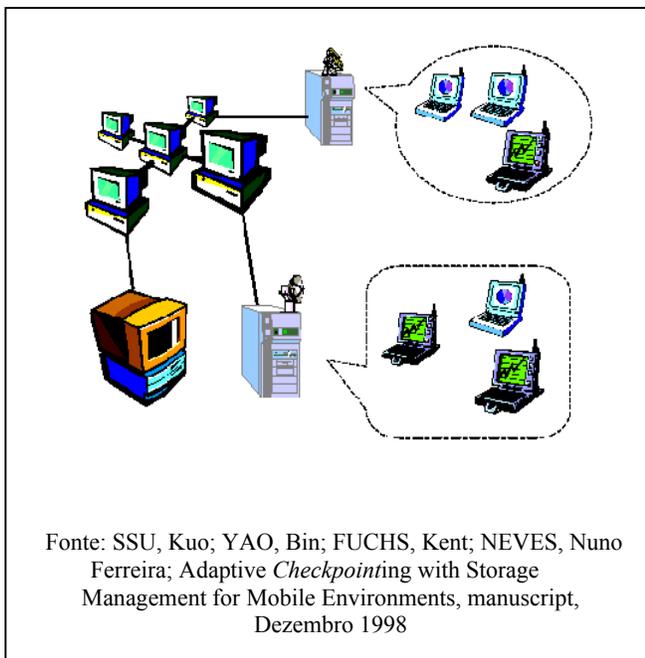


Figura 1- Arquitetura Computação Móvel

3. Arquitetura de Bancos de dados Móveis

Existem vários tipos de arquiteturas possíveis para bancos de dados móveis. Uma arquitetura básica [3], consiste de plataformas de aplicação, um *host* servidor e a comunicação realizada entre eles, como mostra figura 2.

Neste caso, a plataforma de aplicação de bancos de dados móveis é um computador portátil que contém parte do banco de dados, onde são armazenados informações, algumas camadas de software, o sistema operacional, o

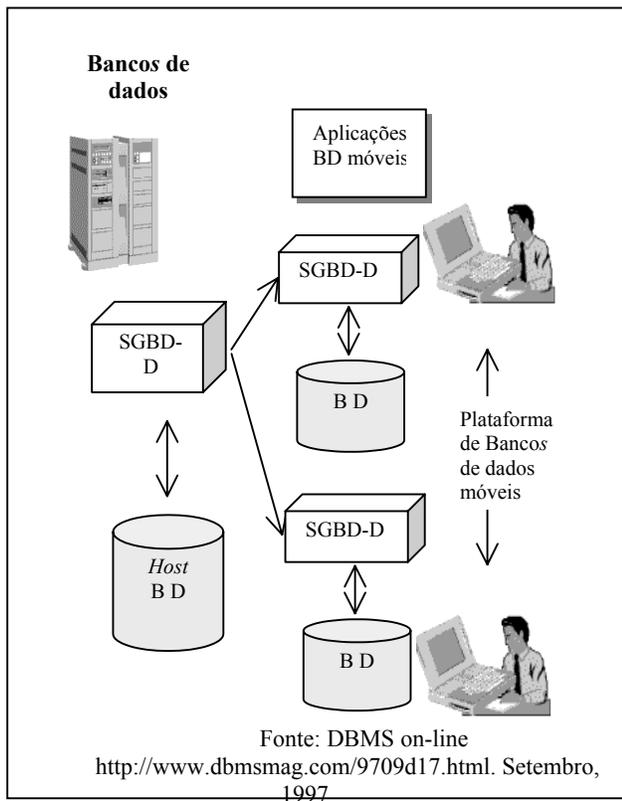


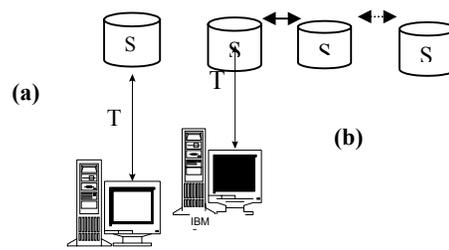
Figura 2- Arquitetura de Bancos de Dados Móveis

3.1. Gerenciamento de Transações Móveis

Para garantir a consistência de dados compartilhados propensos à falhas e desconexões, usuários de unidades móveis e fixas atualizam e recuperam dados através de transações. Em sistemas de bancos de dados tradicionais, usuários interagem com o banco de dados por transações atômicas, consistentes, isoladas e duráveis. Em aplicações móveis estas propriedades são consideradas restritivas por serem limitadas, não suportando as desconexões e operações parciais (commit e abort).

Em um ambiente centralizado, a transação é solicitada, executada e o resultado é mostrado no mesmo site, como mostra a figura 3 (a).

No caso de sistemas distribuídos, as transações são executadas concorrentemente em múltiplos processadores e conjuntos de dados. A execução é coordenada completamente pelo sistema incluindo controle de concorrência, replicação, entre outras operações. É executada concorrentemente por múltiplos processadores e conjuntos de dados sendo dividida em fragmentos e distribuída em diversos sites para a realização da sua execução, como demonstra a figura 3 (b).



Em um ambiente móvel, os modelos de transações móveis devem suportar as limitações da computação móvel. A transação não é totalmente gerenciada pelo sistema; o deslocamento de uma unidade é que controla a execução da transação [6]. Desta forma, ela é realizada sequencialmente através de várias estações base e em possíveis conjuntos de dados, dependendo do movimento da unidade móvel, figura 3 (c).

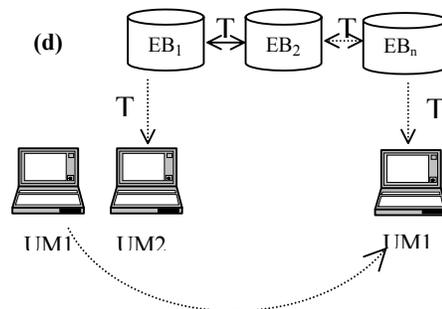
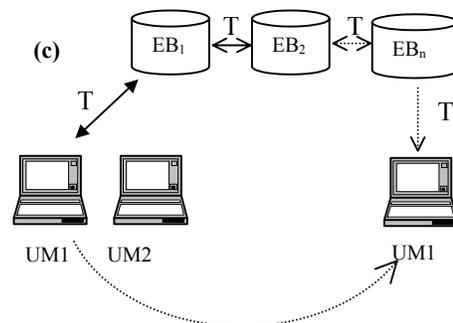


Figura 3 – Execução de Transações

Devido à mobilidade, as transações podem ser executadas de diversas localizações e conseqüentemente de diferentes servidores. Diferentemente de transações realizadas em uma rede fixa, as transações móveis podem acessar dados que estão em constantes mudanças de lugar, sendo assim dependente da localização deles.

A mobilidade resulta em transações que acessam sistemas heterogêneos de informações. A utilização de redes

sem fio acarreta transações de longa duração, devido à sobrecarga de dados na rede, sendo mais propensas a erros, pelo fato das frequentes desconexões e a duração das baterias. Podem ser baseadas em sessões para evitar transações demasiadamente demoradas. Devido ao longo tempo da duração das atividades, vários refinamentos de transações são necessários. A idéia básica é tratar cada transação como um conjunto de transações parciais.

Uma unidade móvel pode ter autonomia local para processar transações apesar das frequentes desconexões. Na reconexão, os efeitos das transações móveis encerradas durante a desconexão deverão ser incorporadas ao banco de dados, garantindo desta forma, o sucesso da transação [7].

3.2. Modelo de Transação Móvel

A figura 25 mostra o modelo genérico da arquitetura de banco de dados móveis. Neste modelo o servidor (suporta operações básicas como commit, abort e alocação do meio) e o banco de dados estão anexados em cada computador fixo.

Cada estação base tem um coordenador que recebe as operações das transações solicitadas pelas unidades móveis e monitora sua execução no banco de dados servidor dentro da rede fixa. As transações são entregues pela unidade móvel ao coordenador em sua estação base, o qual as envia para o servidor de banco de dados distribuídos na rede fixa para a execução (ELMAGARMID & RUSINKIEWICZ, 1995).

Uma unidade móvel pode realizar dois tipos de transações:

A transação completa pode ser submetida em uma única mensagem de solicitação. A unidade móvel entrega o controle da execução para seu coordenador e aguarda o retorno dos resultados para a execução da transação;

As transações podem ser submetidas em múltiplas mensagens de solicitação. A unidade móvel interativamente submete as operações da transação para o seu coordenador e a operação subsequente só pode ser submetida após os resultados terem retornado do coordenador.

Enquanto a primeira abordagem envolve um único coordenador para todas as operações de uma transação, a segunda pode envolver múltiplos coordenadores devido à mobilidade da unidade móvel.

3.3. Tipos de Mobilidade de uma Transação

Existem quatro tipos de mobilidade da transação definidos por (DUNHAM & KUMAR, 1998):

A transação é originada em uma unidade móvel, onde é executada completamente, não existindo qualquer tipo de movimento nesta operação.

A transação inteira completa sua execução localmente, para minimizar custos de comunicação e sobrecarga na rede. Qualquer dado solicitado que não está na unidade móvel é movido de outros sites (unidade móvel, estação base ou qualquer outro nó da rede fixa), sendo que, durante

sua execução, a unidade móvel pode se mover para uma célula estrangeira. Este tipo de execução é chamada de local atômica.

Uma transação pode saltar por um número de células estrangeiras durante seu tempo de execução. Este processo é caracterizado como migração de transação. Uma unidade móvel pode entrar em uma nova célula enquanto está em operação. Neste caso, pode ser necessário migrar parte do processamento que foi executado em uma unidade fixa para outra unidade fixa. Uma motivação para a migração é a melhora de desempenho. Além disso, movendo o processamento finalizado para uma unidade móvel, custos de comunicação podem ser minimizados. A migração de uma transação pode ser feita através de dupla relocação de dados, balanceando o carregamento entre estações base, de modo que, uma transação executa parcialmente em uma estação base e posteriormente, transfere os dados processados para a nova estação base.

Uma transação móvel é definida como um conjunto de fragmentos, que podem “visitar” apenas um subconjunto pré-definido de unidades móveis. Este tipo de restrição de movimento é dependente da distribuição do banco de dados entre as unidades móveis.

3.4. Graus de Mobilidade de uma Transação

De acordo com (DUNHAM & KUMAR, 1998), existem seis tipos ou graus de mobilidade. Estas categorias podem ser aplicadas em uma arquitetura ou transação. O grau mais alto suportado por uma arquitetura é o grau permitido por qualquer T executada nesta arquitetura.

A classificação é dividida em dois eixos (figura 28). O eixo horizontal representa a mobilidade geográfica e o eixo vertical representa a mobilidade dos dados.

O grau 0 (estático) assume que os processadores não se movem. O grau 1 (distribuído) assume que as unidades se movimentam. O grau 2 (*Ad-Hoc*) assume que qualquer nó pode se mover e comunicar diretamente. A seguir são descritos os seis graus de mobilidade:

- (0,0) – Ambiente centralizado com um único processador. Não há movimento geográfico.
- (0,1) – Ambiente distribuído tradicional. O movimento dos processadores e dos dados é permitido mas não há movimento geográfico.
- (1,0) – Esta categoria permite o deslocamento da unidade móvel, mas não há movimento dos dados nem dos processadores.
- (1,1) – Tradicional ambiente da computação móvel. As unidades móveis se movem mas não podem se comunicar diretamente uma com a outra.
- (2,0) – Neste caso qualquer nó pode se comunicar, mas não pode haver movimento dos dados ou processadores.

- (2,1) – Tipo de mobilidade genérica, onde se movimentam dados e processadores.

3.5.Requisitos de uma Transação Móvel

Para descrever o processamento de transações em uma unidade móvel, é viável a união das propriedades ACID e os conceitos de visibilidade, consistência, permanência e recuperação (WALBORN & CHRYSANTHIS, 1998).

A visibilidade refere-se a habilidade da transação para ver os efeitos causados nos dados por outra transação. Para reduzir os custos de recuperação, os efeitos da transação não estão visíveis até que a transação termine e as trocas sejam feitas permanentemente no banco de dados. Permitindo as novas transações a visão das operações não terminadas, o resultado pode gerar dependências de dados e encerramentos não desejados. Desde que, as atualizações de uma unidade móvel desconectada não sejam feitas no servidor, as transações subseqüentes usando os mesmos dados normalmente não podem continuar até que a conexão ocorra e a transação termine.

Um sistema de gerenciamento de banco de dados que suporte o processamento de transações em ambiente móvel desconectado deve incluir, segundo (WALBORN & CHRYSANTHIS, 1998):

- Uma forma de comunicação com o banco de dados servidor;
- Um subsistema de processamento da transação para executar na unidade móvel;
- Um método para reintegrar transações processadas enquanto desconectadas com o banco de dados servidor;
- *Logging* suficiente, *checkpoint* e sistema de recuperação para aliviar falhas do sistema;
- Um método para gerenciar a replicação e a consistência de dados.

3.6.Modelos de Transações Móveis

A seguir serão abordados alguns modelos de transações móveis. São eles (DUNHAM & KUMAR, 1998):

- Modelo Kangaroo;
- Modelo de Fundamento Semântico;
- Modelo de Reporting e Co-transações;

Modelo MDSTPM;

Transação Proxy.

Modelo Kangaroo: O modelo de transação Kangaroo é o único em redes móveis que inclui movimento em sua estratégia de execução, capturando o movimento natural de uma unidade móvel, suportando transações longas.

O modelo é construído assumindo um multi-banco de dados onde o gerenciamento da transação é desempenhado pela estação base. Quando a unidade móvel se movimenta, o gerenciamento da transação também se movimenta. Cada transação é dividida em subtransações, as quais serão executadas independentemente do sistema de banco de dados da rede fixa. Quando a solicitação de uma transação é feita por uma unidade móvel, a estação base associada cria uma transação móvel.

Modelo de Fundamento Semântico: Este modelo assume o processamento de uma transação móvel como um problema de conexão e concorrência, podendo ter longos atrasos no processamento de dados, bem como, desconexões. Esta abordagem utiliza a organização dos objetos para dividir grandes e complexos objetos em pequenos fragmentos. O servidor do banco de dados disponibiliza os fragmentos a pedido de uma unidade móvel. Para completar a transação, a unidade móvel retorna os fragmentos para o servidor.

Modelo de Reporting e Co-transações: Neste modelo, a origem das transações é representada em termos de reporting e co-transações, as quais podem ser executadas tanto na estação base como na unidade móvel. Uma transação reporting pode compartilhar seu resultado parcial com a transação de origem a qualquer hora e pode realizar a operação commit independentemente.

Uma co-transação é uma classe especial da transação anterior, na qual pode ser forçada a esperar por outra transação. Depois de entregar seu resultado, pode continuar a executar outras operações.

Modelo MDSTPM: Neste modelo, é pressuposta a existência de um MDSTPM (Multidatabase Transaction Processing Manager) em cada unidade móvel. Cada componente do sistema de banco de dados é responsável pelo gerenciamento das transações locais. Para facilitar a execução de transações globais, uma camada de software adicional pode ser implementada, permitindo o gerenciamento e coordenação de transações através dos múltiplos componentes do sistema de gerenciamento (YEO & ZASLAVSKY,1994).

Transação Proxy: Uma transação proxy é utilizada para gerenciar a recuperação de dados. Pode ser considerada como uma subtransação da transação original. Alternativamente, podem ser executadas no modo off-line, quando o tráfego da rede ou a carga na estação base é baixa. Neste caso, transações proxy correspondem a backups periódicos do processamento desempenhado em uma unidade móvel .

CONSIDERAÇÕES FINAIS

Por se tratar de um assunto relativamente novo, o ambiente móvel representa um desafio para mundo tecnológico. No campo dos desafios técnicos, um dos grandes problemas é a capacidade de armazenamento de energia, ainda pequena nos dias atuais. Outro fator é a confiabilidade da utilização da computação sem fio em operações que requerem elevado grau de segurança, pois, as informações são transmitidas pelo ar, através de ondas eletromagnéticas, fazendo com que a confiabilidade seja questionada.

A mobilidade também ocasiona uma série de problemas, tais como, altas taxas de erros, baixa velocidade de comunicação, desconexões freqüentes, baixa largura de banda, heterogeneidade de tecnologias sem fio, ausência de integração entre as redes móveis, etc.

Embora a capacidade dos equipamentos móveis esteja crescendo consideravelmente, a memória disponível e o poder de processamento são recursos muito valiosos na computação móvel. Desta forma, um sistema de bancos de dados relacional deve ser o mais compacto possível, para que possa ser utilizado em dispositivos móveis.

No campo das aplicações, muito se tem discutido a respeito das unidades móveis e mecanismos que permitam a elas se adaptarem às alterações do ambiente, através de softwares adequados, protocolos de acesso homogêneo, acesso uniforme à informação sobre localização, endereço universal para unidades móveis, alocação automática de largura de banda, entre outros fatores.

4. Referências

[1] DUNHAM, Margaret H; Helal A. **Mobile Computing and Databases: Anything New?**. SIGMOD Record, v. 24, n. 4, dezembro, 1995

[2] PITOURA, Evaggelia; SAMARAS, George. **Data Management for Mobile Computing**, Kluwer Academic Publishers, 1998.

[3] RENNHACKKAMP, Martin. **Mobile databases, Tetherless Computing Liberates End**

Users but Complicates the Enterprise. DBMS online, setembro 1997.

[4] OZU, M. Tamer; VALDURIEZ, Patrick. **Principles of distributed databases systems**, Upper Saddle River, ed. Prentice –Hall, New Jersey, 1999.

[5] KORTH, Henry F.; SILBERSCHATZ, Abraham. **Database Systems Concepts**. New York: McGraw-Hill, 1999.

[6] DUNHAM, Margaret H; KUMAR, V. **Defining Location Data Dependency, Transaction Mobility and Commitment**. Relatório Técnico 98-CSE-1, fevereiro, 1998

[7] WALBORN, D. Gary; CHRYSANTHIS, K. Panos. **Supporting Semantics-Based Transaction Processing in Mobile Database Applications**. Proceedings of the 14th Symposium on Reliable Distributed Systems (SRDS-95), Setembro, 1995.

[8] NASSU, Eugenio Akihiro; FINGER, Marcelo. **O Significado de “Aqui” em Sistemas Transacionais Móveis**, I Workshop SIDAM (Sistemas de Informação Distribuída de Agentes Móveis), outubro, 2000, p. 55-63.

[9] SHORES, Redwood. **Reviewr’s Guide Oracle8i™ Lite**. Oracle Corporation, CA 94065, 2000.

Disponível em : <http://www.Oracle.com>

[10] SYBASE, **Synchronization Technologies for Mobile and Embedded computing**, 2000.
Disponível em:
<http://www.Sybase.com/detail/1,3693,1009526,00.html>

[11] NOETHER; Angela . **Extending enterprise data to mobile devices**. Developer Tollbox Team, 2000.

Disponível em : <http://www.developer.ibm.com>

Observação: Esta versão será substituída cfe conversa com prof. Reinaldo