

Arquitetura Adaptável para Execução de Transações Móveis

Daniela Eloise Flôr

Aluna do curso de pós-graduação do INPE em Computação Aplicada (CAP)
danielaflor@lac.inpe.br

Maurício Ferreira
Gonçalves Vieira

Professor do curso de pós-graduação do INPE em Computação Aplicada (CAP)
mauricio@ccs.inpe.br

Nilson Sant'Anna

Professor do curso de pós-graduação do INPE em Computação Aplicada (CAP)
nilson@lac.inpe.br

Resumo

Atualmente é crescente o interesse em empregar dispositivos móveis no acesso aos sistemas de informação. Em conjunto com a tecnologia de comunicação sem fio, estes dispositivos provêm a base para o paradigma da Computação Móvel (CM). A partir de então, os cenários que vem sendo descerrados influenciaram várias áreas da ciência da computação, inclusive os Sistemas de Banco de Dados (SBDs). A arquitetura infraestruturada da CM, exige dos SBDs um novo comportamento, o que requer adaptações relacionadas principalmente ao gerenciamento de transações, já que os equipamentos móveis acrescentaram novas formas de processamento. Este trabalho abordará as particularidades do ambiente móvel, as plataformas de processamento de transações em tal cenário, e os modelos de transações já encontrados na literatura. Ao final é apresentada uma arquitetura adaptável formada por serviços que avaliam as características do ambiente móvel e indicam o modelo de transação mais adequado para a plataforma de processamento disponível.

Palavras-Chave: computação móvel, sistema de banco de dados, transações, propriedades ACID.

1. Introdução

A computação móvel é considerada por [1] como uma ampliação do conceito tradicional da computação distribuída. Essa extensão da classificação inicial se apóia na própria arquitetura de um sistema de computação móvel, que considera ter na rede fixa um sistema distribuído.

Na concepção da computação móvel alguns aspectos foram acrescentados ao conceito tradicional

da computação distribuída, sejam eles, a comunicação sem fio, mobilidade e portabilidade.

A analogia também se aplica aos Sistemas de Banco de Dados Distribuídos (SBDDs) e aos Sistemas de Banco de Dados Móveis (SBDMs).

Uma classificação feita por Özsu e Valduriez, estendida por [1] e visualizada na figura 1, reforça a extensão comentada, considerando o ambiente de computação dos sistemas gerenciadores de banco de dados sob três aspectos, autonomia, distribuição e heterogeneidade.

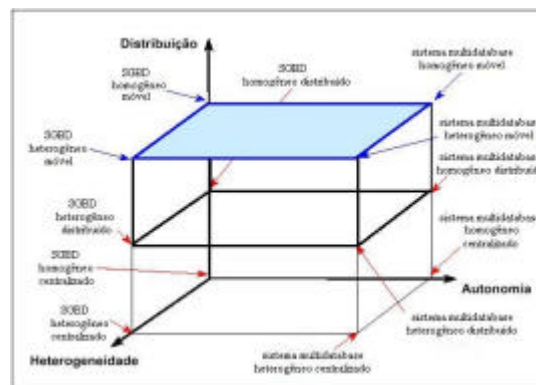


Figura 1. Classificação para os sistemas de banco de dados móveis

Diante da migração de aplicações e dados para os dispositivos móveis, faz-se necessário reavaliar questões sobre a gerência de dados nesse novo paradigma.

Entre essas questões estão a replicação e distribuição dos dados, o processamento de consultas, a recuperação e tolerância a falhas e o gerenciamento de transações. Sendo o último tópico o objetivo principal do trabalho.

Para compreender as dificuldades que os bancos de dados enfrentam para operar em um ambiente tão instável, faz-se necessário o entendimento das características desse ambiente e suas influências na execução de uma transação.

1.1. Arquitetura de um Sistema Móvel

A arquitetura de um sistema móvel é composta por computadores estacionários que conectados por cabos formam a rede fixa, esses equipamentos também são conhecidos como *hosts* fixos.

Os *hosts* fixos tem finalidade genérica, sendo que alguns deles, denominados estações de base ou estações de suporte móvel, são equipados com interface de comunicação sem fio e gerenciam os outros componentes da arquitetura, os computadores móveis.

As estações móveis são dispositivos portáteis que se movimentam livremente e acessam a rede fixa através das estações de base. A figura 2, extraída de [1] ilustra uma arquitetura genérica para um sistema de computação móvel.

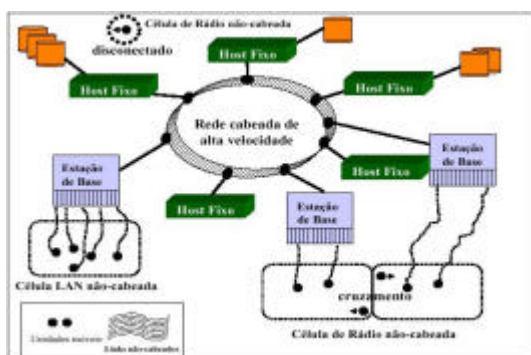


Figura 2. Arquitetura de um sistema de CM

Cada estação de base possui a capacidade de gerenciar as unidades móveis que estão dentro da sua área de abrangência, essa área é uma célula.

A computação móvel sofre importantes restrições devido às limitações dos dispositivos portáteis e da própria interface de comunicação sem fio. Essas restrições são tratadas a seguir.

1.2. Características e Restrições da Computação Móvel

Unidades móveis, como os assistentes pessoais digitais (PDA) e telefones celulares, possuem pouca memória, pouco poder de processamento e armazenamento, interface reduzida, são mais propensos a perdas, roubos ou quedas, e são alimentados por uma fonte de energia finita.

Diante dessa escassez de recursos e visando a economia dos mesmos, é possível que estes operem em quatro modos diferentes.

Quando uma estação móvel estiver fortemente conectada à rede fixa, diz-se que ela está completamente conectada. Parcialmente conectada ocorre quando a largura de banda é suficiente apenas para uma comunicação *downlink*, ou seja, da rede fixa para a estação móvel, e não ao contrário, que seria uma comunicação *uplink*.

Completamente desconectado ocorre quando a estação móvel não se encontra no domínio de nenhuma célula e não consegue estabelecer conexão com nenhuma estação de base.

O modo *doze*, ou soneca, é um estado de economia de recursos, a figura 3 ilustra os modos de operação e os protocolos responsáveis pelas possíveis mudanças de um modo para outro.

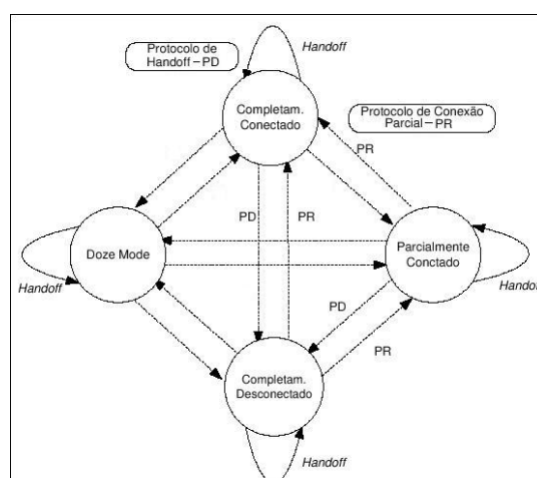


Figura 3. Modos de operação de uma estação móvel

A comunicação sem fio também acrescenta alguns problemas como a variabilidade da conexão, a segurança na transmissão dos dados, largura de banda reduzida e a latência da comunicação.

No cenário móvel, a própria mobilidade dos dispositivos portáteis acrescenta novas particularidades como o modo de endereçamento, a troca entre áreas de cobertura, a gerência de localização dos elementos móveis, consultas dependentes de localização, gerência de dados entre outras coisas.

De acordo com [4], as questões tratadas anteriormente sobre o tripé da computação móvel, ou seja, os dispositivos, o meio de comunicação e as implicações da mobilidade, afetam diretamente os SBD's.

Nesse contexto a forma de gerenciar dados no ambiente distribuído precisa ser reavaliada para verificar o que continua válido e o que precisa ser pesquisado para o ambiente da computação móvel.

1.3. Gerenciamento de dados em ambientes móveis

As aplicações que rodam em clientes móveis carregam as informações necessárias para o seu funcionamento autônomo, se for dotada dessa capacidade. A figura 4 exemplifica essa situação.

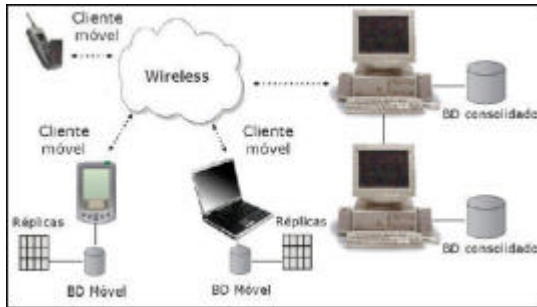


Figura 4. Cenário de um sistema da CM

A idéia é que as unidades móveis, muitas vezes heterogêneas, possam continuar trabalhando mesmo estando fora do alcance da rede sem fio, e, portanto, desconectada da rede fixa. Posteriormente as informações atualizadas no cliente móvel deverão ser validadas com as outras estações do sistema.

Várias questões envolvem o tratamento dos dados distribuídos entre os clientes móveis e fixos como caching, replicação, processamento de consultas, difusão de dados, processamento de transações entre outras.

1.3.1. Caching

O mecanismo de caching é utilizado para otimizar o tempo de resposta nas consultas da computação móvel e dar suporte quando a unidade móvel estiver desconectada.[4]

1.3.2. Replicação

A replicação consiste na manutenção de cópias inteiras ou parciais dos dados em várias estações da rede, inclusive as móveis. Essa técnica é muito valiosa para a computação móvel considerando a variação da conexão e da garantia de trabalho em modo desconectado.[4]

1.3.3. Difusão de dados

A difusão de dados, segundo [5], trata da troca de dados. Quando a rede fixa atende uma solicitação de dados por parte de um computador móvel a difusão é chamada de *pull-based*.

Quando a entrega dos dados é iniciada pela rede fixa a um conjunto de clientes o modelo utilizado é o *push-based*.

1.3.4. Processamento de consulta

O processamento de consultas em um sistema de computação móvel além do intuito tradicional, pode depender da localização do usuário ou do momento em que a mesma foi solicitada.

1.3.5. Processamento de transação

As transações aplicam um conjunto de operações aos dados de um SBD. Segundo [6], cada transação é uma unidade atômica, isolada que não viola nenhuma regra de consistência e tem o seu resultado persistido após o seu término.

No caso de um banco de dados em um ambiente de computação móvel, as características das unidades móveis e do próprio ambiente impedem que uma transação tenha as mesmas características ACID do modelo tradicional.

No ambiente da computação móvel uma transação poderá ser dividida em operações a serem executadas em estações de diferentes tipos – fixas ou móveis. Essa possibilidade inviabiliza a estrutura atômica de uma transação no ambiente móvel.

As unidades móveis dotadas de réplicas de dados podem usufruir da autonomia de trabalho mesmo durante as desconexões. Essa vantagem provoca inconsistência nos dados.

Além disso, uma transação iniciada por uma estação móvel em movimento pode iniciar a sua execução numa célula e terminá-la em outra.

Outra situação a ser considerada é a ocorrência de desconexões durante a transação, nesse caso, para evitar o retrabalho, as transações podem ser divididas e os resultados parciais poderiam ser considerados, mas essa proposta também exige uma revisão das propriedades ACID.

O tempo é outro fator que difere as transações em ambientes móveis das demais, a variabilidade da conexão e a latência da rede podem caracterizar uma transação como de “longa duração”.

Diante da realidade imposta pelo recente paradigma faz-se necessário uma reavaliação das questões relativas às transações.

2. Transações móveis

De acordo com [7], uma transação no cenário móvel, onde participam estações móveis e fixas é conhecida por transação móvel.

Com a participação das unidades móveis faz-se necessário considerar novas características e verificar as plataformas de execução disponíveis, os novos modelos de processamento ou possíveis adaptações nos modelos propostos para os ambientes centralizado e distribuído.

2.1. Características das transações móveis

As restrições do ambiente e dos componentes da computação móvel exigem que uma transação em tal cenário seja capaz de:

- Suportar desconexões (curtas ou longas) das estações móveis;
- Suportar a mobilidade de uma estação móvel ao longo da execução da transação;
- Controlar o deslocamento da transação quando a estação móvel se movimenta;
- Oferecer flexibilidade de atomicidade;
- Suportar transações com de longa duração;
- Fornecer autonomia local para permitir que as transações sejam processadas e guardadas na estação móvel quando há desconexões temporárias;
- Suportar comunicações e computações sem fios, além de concorrência;
- Restabelecer uma conexão em caso de desconexão ou modo soneca;
- Suportar a reflexão das atualizações realizadas na estação móvel nos dados da rede fixa para manter a consistência mútua de dados replicados.

2.2. Plataformas de execução de transações móveis

Considerando as estações móveis e estacionárias que formam a arquitetura de um sistema de computação móvel, além da heterogeneidade de tais dispositivos, foram identificadas em [7] cinco plataformas de execução de uma transação móvel:

- Execução total na rede fixa;
- Execução total na unidade móvel;
- Distribuição do processamento entre a rede fixa e a unidade móvel;
- Distribuição do processamento entre várias estações móveis;
- Distribuição do processamento entre computadores móveis e fixos;

Na primeira opção a transação móvel é solicitada pela unidade móvel, o processamento totalmente realizado na rede fixa e os resultados enviados de volta ao cliente.

Outra alternativa trata do processamento da transação móvel no próprio equipamento portátil. Em seguida alguma estratégia de reconciliação dos dados deve ser adotada para homogeneizar os dados.

O terceiro modelo é mais flexível, o dispositivo móvel processa os dados que armazena e para concluir a transação móvel solicita à rede fixa o processamento dos dados que não possui.

O cenário mais ambicioso e complexo está em distribuir o processamento das transações móveis entre as unidades móveis, uma unidade atua como servidor para outra desde que estejam na mesma área de cobertura.

A última opção considera um cenário totalmente distribuído estando o processamento dividido entre vários equipamentos portáteis e unidades fixas, considerando um ambiente de trabalho cooperativo, a conexão de múltiplos bancos de dados (*multidatabase*) e a participação das unidades móveis no controle da execução global.

Algumas propostas de transações para a computação móvel consideram algumas das plataformas citadas, a seção a seguir apresenta uma classificação desses modelos.

2.3. Modelos de transações para ambientes móveis

São muitas as características de uma transação móvel, conforme apresentado em 2.1, cada um dos modelos de transações móveis propostos enfocam algumas dessas características ignorando outras, apresentando soluções específicas para situações que se enquadram no modelo, deixando de lado as particularidades de outras aplicações.

Devido às restrições das propriedades ACID, alguns modelos propõem o relaxamento de tais propriedades como forma de viabilizar a transação em tal ambiente.

Os modelos pesquisados foram classificados de acordo com as características que focalizaram e quanto à plataforma de processamento.

2.3.1. Modelos que executam seu processamento na rede fixa

Os modelos apresentados nessa subseção indicam a requisição da transação a partir da estação móvel, ficando para a rede fixa a responsabilidade de executar e coordenar a execução.

As freqüentes desconexões e a localização inconstante das estações móveis deram origem a modelos que tratam uma transação móvel como um conjunto de subtransações, sendo que em alguns modelos o resultado parcial pode ser compartilhado com as demais subtransações.

Nesses modelos sempre há a indicação de um controle global das subtransações, que podem inclusive serem executadas em sites diferentes de acordo com o movimento da estação móvel.

Com essas características estão os seguintes modelos apresentados em [7], *Kangaroo*, *MDSTPM*, *MOFLEX* e *Pre-serialization*.

2.3.2. Modelos que executam seu processamento na unidade móvel

Outra possibilidade adotada por alguns modelos é permitir ao cliente móvel a execução local das transações e em seguida a incorporação das alterações na rede fixa, a característica predominante nesses modelos é o suporte à autonomia.

Para ser ter autonomia é necessário que nos dispositivos móveis estejam contidos a aplicação e os dados, provavelmente uma réplica total ou parcial do mesmo, dependendo da capacidade do dispositivo.

Nesses casos as transações são divididas em dois tipos, às que ocorrem no modo desconectado e as que ocorrem quando a conexão é restabelecida.

Os modelos que adotam essa estratégia possuem algum tipo de proposta para resolver possíveis conflitos na hora de refletir as atualizações locais na rede fixa.

De acordo com [7], os modelos que sugerem o processamento na unidade móvel são *HiCoMo*, *Pro-motion*, Baseado em semântica e *Prewrite*.

2.3.3. Modelos que podem distribuir o seu processamento entre a rede fixa e o dispositivo móvel

Esses modelos possuem maior flexibilidade e estão aptos a executar uma transação totalmente na unidade móvel caso não haja conexão com a rede fixa. Ou, caso contrário distribuir o processamento entre a rede fixa e o próprio dispositivo, conforme [7] os modelos propostos para essa plataforma de execução são o *Clustering* e *Two-tier replication*.

Pode-se verificar que nos últimos anos, muitos modelos para a execução de uma transação em um ambiente móvel foram propostos, [7] apresenta um estudo detalhado sobre esses modelos, e pode-se perceber que as opções não consideram todas as particularidades do ambiente móvel, como a mobilidade do usuário, a variação do meio de comunicação sem fio e principalmente a heterogeneidade do dispositivo móvel, além de contemplarem uma forma específica de processamento.

3. Arquitetura adaptável para execução de transações móveis

Uma transação móvel terá maiores chances de obter êxito se as questões envolvidas na sua execução forem devidamente avaliadas.

O propósito geral desta arquitetura é indicar a forma que uma transação móvel solicitada deve

ser executada, considerando todas as questões que influenciam os dispositivos, o meio e as chances que a mesma possui de ser finalizada com sucesso. A figura 5 ilustra a arquitetura.

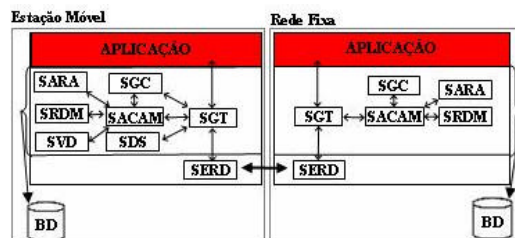


Figura 5. Arquitetura de adaptação para transações móveis

A aplicação irá interagir diretamente com a arquitetura, quer nas estações móveis ou na rede fixa. Após a chamada dos serviços responsáveis por verificar as características do meio, do dispositivo, de experiências anteriores, de verificação dos dados locais o serviço de avaliação (SACAM) irá indicar ao serviço de gerenciamento de transação (SGT) qual a melhor forma para executar a transação móvel que originou a chamada do serviço de adaptação de transações móveis.

3.1. Serviços que compõe a arquitetura

A arquitetura proposta é composta por serviços que possuem funções distintas, como:

3.1.1. Serviço de avaliação das características do ambiente móvel (SACAM)

- recebe informações sobre os recursos da estação móvel e do meio de comunicação;
- recebe estatísticas de casos parecidos com o que estiver sendo avaliado;
- compõe a tabela de valores dinâmicos para as variáveis do ambiente móvel, interpretando e indicando a forma de processamento mais adequada para a transação móvel;

3.1.2. Serviço de gerenciamento de comunicação (SGC)

- reconhece o meio de comunicação sem fio;
- reconhece a largura de banda;
- verifica taxa de transmissão;
- avalia o nível do sinal;
- reconhece e gerencia a conexão sem fio;

3.1.3. Serviço de reconhecimento do dispositivo móvel (SRDM)

- verifica a quantidade de bateria restante;
- identifica memória e espaço disponível;

3.1.4. Serviço de gerenciamento de transação (SGT)

- controla o tempo (configurado) de espera do retorno de uma transação ou subtransação;
- atua como cliente da aplicação;
- coordena a execução de transações ou subtransações;

3.1.5. Serviço de desconexão e modo soneca (SDS)

- altera o modo de operação da unidade móvel;

3.1.6. Serviço de envio e recepção de dados (SERD)

- responsável pela troca de dados entre estação móvel e rede fixa;

3.1.7. Serviço de verificação de dados (SVD)

- verifica se existem os dados no dispositivo móvel e a sua completude;

3.1.8. Serviço de armazenamento de resultados anteriores (SARA)

- armazena o resultado das indicações sobre a execução da transação móvel;
- disponibiliza estatística sobre casos parecidos;

3.2. Escopo e limitações da arquitetura

A área de gerência de dados em SBDM inclui vários assuntos, este trabalho enfocou as questões relacionadas ao gerenciamento de transação.

Não é de responsabilidade da arquitetura de adaptação de transações móveis decidir sobre o grau e o tipo de replicação dos dados, aplicar técnicas de validação ou invalidação de dados em cache, indicar a técnica para resolução de conflitos durante a integralização de transações locais e globais.

Outro assunto que vem sendo alvo de pesquisas é a gerência de localização, algumas vantagens e desvantagens sobre de quem é essa responsabilidade, da estação móvel ou da estação base, tem sido identificadas em recentes trabalhos. A arquitetura precisa da comunicação posterior ao movimento sem entrar em detalhes sobre quem está gerenciando a movimento da estação móvel.

O movimento entre células (*handoff*) adjacentes ou não, dificulta a entrega dos resultados de uma transação à estação móvel. O cálculo do melhor caminho a ser percorrido para a entrega do conteúdo após a identificação da nova estação base da unidade móvel é outro ponto não tratado pela arquitetura.

4. Conclusões e atividades futuras

Neste artigo foi possível verificar como a portabilidade, mobilidade e o meio sem fio influenciam na execução de uma transação no cenário móvel.

Adicionalmente, uma arquitetura para adaptação das transações móveis às variações do meio foi proposta, o intuito da mesma é a minimização do número de cancelamentos e a manutenção da confiabilidade das transações em sistemas de banco de dados no ambiente móvel.

As próximas etapas deste trabalho incluem:

- A implementação dos serviços da arquitetura;
- A realização de testes; e
- A avaliação da funcionalidade da proposta.

5. Referências Bibliográficas

- [1] DUNHAM, Margaret H. e HELAL, Abdelsalam. *Mobile Computing and Databases: Anything New?*. Department of Computer Sciences and Engineering, Southern Methodist University, Dallas, Texas, 1995.
- [2] CHRYSANTHIS, Panos. *Transaction Processing in Mobile Computing Environment*. Dans IEEE Workshop on Advances in Parallel and Distributed Systems (APADS), Princeton, USA, October 1993.
- [3] MATEUS, Geraldo Robson; LOUREIRO, Antonio A. Ferreira. *Introdução a Computação Móvel*. DCC/IM, COPPE/Sistemas, NCE/UFRJ, 11a . Escola de Computação, 1998.
- [4] BADRINATH, B.R.; ACHARYA, Arup e IMIELINSKI, Tomasz. *Impact of Mobility on Distributed Computations*. Rutgers University: New Brunswick, NJ, 1993.
- [5] BARBARÁ Daniel. "Mobile Computing and Databases - A Survey", IEEE Transactions on Knowledge and Data Engineering, Vol. 11, No. 1, Jan/Feb, 1999, pp. 108-117.
- [6] SILBERSCHATZ, Abraham.; KORTH, Henry.; SUDARSHAN S.; *Sistema de Banco de Dados*. São Paulo: Makron Books, 1999.
- [7] ALVARADO, Patrícia S.; *Transactions Adaptables pour les Environments Mobiles*. Université Joseph Fourier, 2004.