

3 Gerenciamento de Mobilidade

A Internet não foi originalmente projetada para suportar a mobilidade de dispositivos. A infra-estrutura existente e o conjunto dos principais protocolos utilizados no modelo de camadas TCP/IP possuem limitações que dificultam sua utilização para cenários de mobilidade. Dentre algumas das principais limitações da arquitetura da Internet para lidar com estes cenários, podemos destacar, por exemplo, a utilização dos endereços IP tanto para a identificação quanto para o endereçamento de dispositivos (pois o roteamento de pacotes é baseado nos prefixos dos endereços IP). A utilização dos endereços IP para estes propósitos amarra um dispositivo identificado por um endereço IP à localização (i.e. ponto de conexão com a rede) que este mesmo endereço representa como caminho de roteamento (Eddy, 2004). Para que a mobilidade de dispositivos possa ser oferecida na Internet, e para que toda a arquitetura existente seja aproveitada, é necessária uma adaptação da mesma através da implementação de mecanismos, tanto na rede como nos dispositivos finais, que contornem suas principais limitações. Tais mecanismos são geralmente chamados de mecanismos de Gerenciamento de Mobilidade (Henderson e Works, 2003).

Neste capítulo, é apresentada uma investigação das principais soluções de Gerenciamento de Mobilidade existentes, com o objetivo de escolher qual a mais adequada para suportar a mobilidade de dispositivos na Internet.

A mobilidade pode ser classificada por diferentes tipos. Dependendo da solução de Gerenciamento de Mobilidade empregada, alguns tipos podem ser suportados e outros não. Alguns dos tipos de mobilidade mais comumente retratados na literatura, tais como apresentados por (Pandya, 1995) e (Wang e Abu-Rgheff, 2003), são:

- **Mobilidade de Terminais.** Refere-se à habilidade do usuário de utilizar seu terminal (i.e. dispositivo) para se mover entre redes heterogêneas (i.e. se mover entre diferentes sub-redes, possivelmente de diferentes domínios administrativos) enquanto continua a ter acesso ao mesmo conjunto de

serviços em que está subscrito e permanece alcançável para mensagens ou requisições de outros dispositivos.

- **Mobilidade Pessoal.** Significa que um usuário deve poder ser globalmente alcançável através de um único identificador pessoal e ser capaz de iniciar ou aceitar sessões a partir de qualquer dispositivo.
- **Mobilidade de Sessão.** Refere-se à habilidade do usuário de manter uma sessão ativa enquanto muda de dispositivo.

A implementação do Gerenciamento de Mobilidade pode ser oferecida em diferentes camadas na pilha de protocolos na arquitetura da Internet. A questão de em que camada a mobilidade deve ser apropriadamente oferecida é uma discussão ainda aberta na comunidade acadêmica (Eddy, 2004). Alguns dos critérios básicos para avaliar soluções de Gerenciamento de Mobilidade existentes são:

- **Identificador independente de localização.** “O identificador utilizado por um dispositivo deve ser independente da localização do mesmo e deve permanecer estático mesmo que o dispositivo mude de localização” (Henderson e Works, 2003).
- **Compatibilidade com roteamento IP.** “O Gerenciamento de Mobilidade deve apresentar um funcionamento compatível com o roteamento IP. Por exemplo, o dispositivo móvel deve adquirir um endereço IP topologicamente correto ao se mover.” (Henderson e Works, 2003)
- **Gerenciamento de localização.** Um dispositivo móvel deve ser capaz de ser localizado globalmente mesmo após uma mudança de localização (Eddy, 2004).
- **Transições transparentes.** Significa que a mobilidade não afeta conexões estabelecidas acima da camada de rede. Por exemplo, quando ocorre uma perda na conectividade devido à mobilidade, as conexões na camada de transporte são retomadas transparentemente para a aplicação (Eddy, 2004).
- **Liberdade de Infra-Estrutura.** Se uma solução de mobilidade requer que mudanças sejam feitas na infra-estrutura da Internet existente, isso significa que em todo o mundo, todas as sub-redes, roteadores e outras entidades da arquitetura de redes deverão também ser modificados. Dependendo das mudanças propostas, é fácil verificar a inviabilidade de sua implantação, considerando que a solução deve oferecer a mesma disponibilidade e

generalidade que a Internet atual oferece. Os usuários finais devem ser capazes de se movimentar para redes independentemente se elas são especificamente projetadas para mobilidade ou que seus administradores escolham configurá-las para suportar tal cenário. Portanto, é mais desejável que o Gerenciamento de Mobilidade seja implementado nas “bordas da Internet”, ao invés de implementá-la no núcleo da rede (Eddy, 2004).

Segundo diversos autores, tais como (Banerjee, Wu et al., 2003), (Wang e Abu-Rgheff, 2003) e (Eddy, 2004), juntamente com as principais características de cada camada na pilha de protocolos utilizados na Internet, podemos observar diversas soluções de Gerenciamento de Mobilidade existentes, propostas para serem implementadas em cada uma delas de acordo com os requisitos apresentados acima. As Seções a seguir apresentam algumas destas soluções.

3.1. Suporte Abaixo da Camada de Rede

Abaixo da camada de rede, não há a capacidade de utilização de endereços globalmente utilizáveis (roteamento entre redes) tais como o endereçamento IP. Um dispositivo conectado em uma sub-rede não é capaz de se comunicar com outro dispositivo localizado em outra sub-rede utilizando protocolos deste nível isoladamente. Para o roteamento entre diferentes sub-redes, é necessário no mínimo um protocolo na camada de rede. Por isso, nenhuma solução de mobilidade é implementada abaixo da camada de rede (soluções que oferecem mobilidade dentro de uma sub-rede apenas podem ser implementadas neste nível, mas não são o foco desta discussão). Entretanto, para qualquer solução de mobilidade em camadas superiores, é necessário um suporte da camada de Enlace para anexar o dispositivo móvel a diferentes sub-redes que possam estar disponíveis, e fazer isso também dinamicamente durante a transição entre localizações. Os dispositivos móveis de hoje em dia são equipados com uma variedade de tecnologias de interfaces de rede. Em um mesmo dispositivo, podem ser encontradas interfaces para Wi-Fi, Ethernet, etc. A camada de Enlace deve ser capaz de verificar quais meios estão disponíveis para utilização em uma determinada área e se anexar à topologia local permitindo que o usuário possa obter o máximo de conectividade enquanto se move. Além disso, quando a

camada de Enlace realiza uma transição entre redes distintas, a camada de rede necessita ser re-configurada para que as configurações IP sejam adequadas à nova rede a qual o dispositivo foi anexado. O protocolo DHCP (Dynamic Host Configuration Protocol) pode ser utilizado para adaptar as configurações IP para novas redes, fornecendo suporte a soluções de mobilidade implementadas em um nível mais alto na pilha de protocolos (Eddy, 2004).

3.2. Gerenciamento de Mobilidade na Camada de Rede

O Mobile IP (Perkins, 2002) é a solução de Gerenciamento de Mobilidade na camada de rede padronizada pelo IETF (Internet Engineering Task Force). Os endereços IP, apesar de serem dependentes de localização, são utilizados nesta solução como identificadores dos dispositivos. Isto é possível porque o endereço IP atribuído a cada dispositivo é o endereço IP da rede original a que ele pertence e não muda mesmo que o dispositivo migre para outra sub-rede. O Mobile IP é baseado no conceito de um Home Agent (HA) e um Foreign Agent (FA) para o roteamento de pacotes de um ponto de conexão para o próximo.

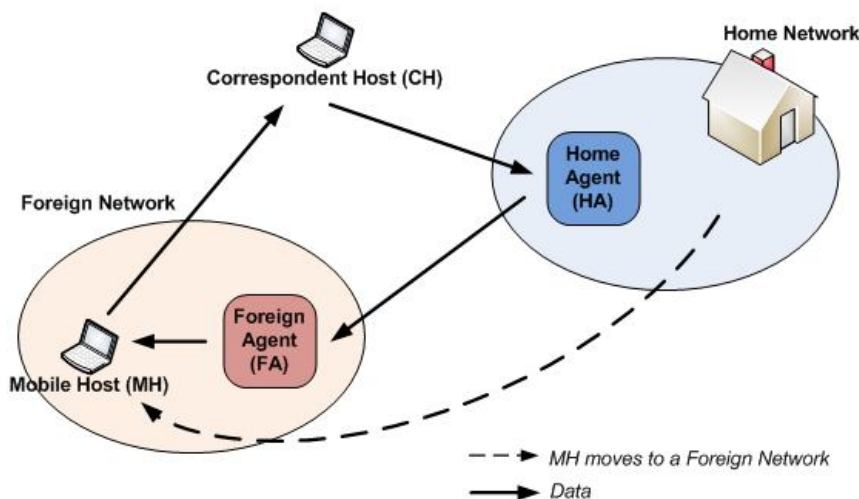


Figura 12 – Funcionamento da solução de Gerenciamento de Mobilidade na camada de rede Mobile IP

O Home Agent (HA) fica na rede a qual o endereço do nó móvel pertence. Atualizações de localização são enviadas pelo nó móvel (Mobile Host ou MH) para o HA quando seu ponto de conexão com a rede muda. O HA encaminha qualquer pacote destinado ao MH, que ele recebe na rede original (home network) do MH, através de um tunelamento IP, para o FA na localização física corrente do

MH. Não importa onde o MH está fisicamente conectado, ele envia pacotes para outros nós utilizando seu endereço IP da sua rede original no campo de remetente do pacote. Desta forma as conexões existentes são mantidas já que os vínculos de endereços não mudam (Eddy, 2004).

O gerenciamento de localização desta solução é implícito, já que o nó móvel retém seu endereço global e o HA mapeia sempre aquele endereço para a localização corrente do nó móvel. O Mobile IP oferece transparência total de mobilidade, o que nem sempre é desejado, visto que a ciência da mobilidade pode ser necessária para adaptações específicas em camadas superiores. Quanto à necessidade de mudanças na infra-estrutura, é necessário que os administradores das redes configurem os HAs e FAs, roteadores e firewalls para que essa solução possa ser implantada. Os usuários finais podem não ser atendidos em lugares em que os administradores das redes não estiverem dispostos ou sejam capazes de suportar o Mobile IP. Além disso, os usuários precisam de um endereço IP permanente e para isso precisam pedir a seu servidor de acesso à Internet que ofereça a ele esse serviço. Para o IPv4, a maioria dos dispositivos hoje em dia não possui um endereço IP fixo, ao invés disso adquirem um dinamicamente via DHCP sempre que se conectam à rede. Portanto, apesar de o Mobile IP ser um conceito amplamente difundido, sua aplicação tem indícios de ser inviável para atender aos crescentes requisitos das novas aplicações na Internet, devido várias limitações na prática para sua implantação (Eddy, 2004). Mesmo com as diversas melhorias e otimizações incluídas na versão para IPv6 do Mobile IP (e.g. não é mais necessária a implantação de Foreign Agents) esta solução ainda não apresenta aceitação para implantação em uma escala que ofereça um alto nível de disponibilidade.

3.3. Gerenciamento de Mobilidade Entre as Camadas de Rede e de Transporte

O Host Identity Protocol (HIP) (Moskowitz e Nikander, 2006) é um protocolo que separa o papel do endereço IP como identificador de um dispositivo do papel de localização do mesmo, introduzindo um novo namespace (host identity), baseado em criptografia de chaves públicas, em uma camada interposta entre as camadas de rede e de transporte (Bondareva e Baumann, 2006).

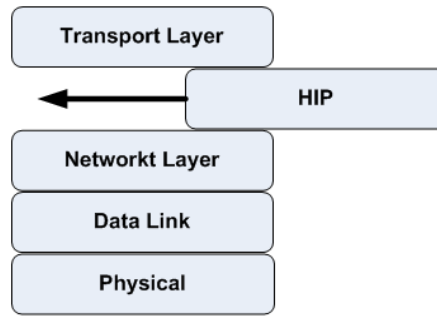


Figura 13 – Camada HIP entre as camadas de rede e de transporte (Bondareva e Baumann, 2006)

No HIP, para cada host é atribuído um nome globalmente único (i.e. estatisticamente único). Como este nome é baseado em criptografia de chaves públicas, esta identidade de host pode ser utilizada também para autenticar transações. A camada de protocolo HIP permite o desacoplamento de conexões de transporte dos endereços IP, e todos os pacotes carregam uma representação da identidade dos hosts. Na troca de pacotes, os nomes únicos dos hosts são utilizados para identificar nós e estabelecer e atualizar associações HIP entre eles. O HIP, para oferecer um gerenciamento de localização independente, possui uma entidade chamada Rendezvous Server, que basicamente é um servidor fixo de localização. Como diversos serviços de gerenciamento de localização, o Rendezvous Server é semelhante ao Home Agent na arquitetura Mobile IP (Bondareva e Baumann, 2006).

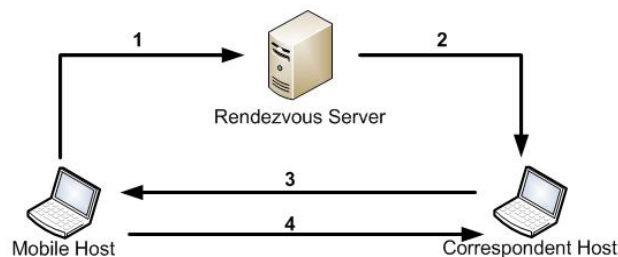


Figura 14 – Estabelecimento de uma associação HIP (Bondareva e Baumann, 2006)

Quando um nó inicia uma conexão com outro nó correspondente, a primeira mensagem de sinalização é roteada através do Rendezvous Server, que a encaminha para a localização corrente do nó correspondente (vide **Figura 14**). O resto das mensagens HIP são trocadas diretamente entre os nós. Quando o nó móvel muda sua localização, ele informa ao Rendezvous Server e o nó correspondente de seus novos localizadores (Bondareva e Baumann, 2006).

Esta solução é capaz de oferecer um Gerenciamento de Mobilidade entre múltiplos domínios de endereçamento IP e oferece uma mobilidade IP

transparente para a camada de transporte. Entretanto, apesar da necessidade de um Rendezvous Server não ser um requisito de grande impacto na infra-estrutura, a implantação do HIP requer mudanças radicais na camada de rede e nos pontos finais de comunicação, o que pode ser considerado como um grande limitador para a implantação e aceitação deste tipo de solução (Henderson e Works, 2003). Maiores detalhes sobre o protocolo HIP podem ser encontrados em (Moskowitz e Nikander, 2006).

3.4. Gerenciamento de Mobilidade na Camada de Transporte

Para implementar a mobilidade na camada de transporte, deve existir uma forma de re-vincular dinamicamente os endereços de uma conexão quando um dispositivo se move. O protocolo TCP (Transmission Control Protocol), que oferece transmissão confiável de dados pela Internet, vincula os dispositivos participantes de uma conexão com endereços IP únicos e estáticos, através de blocos de controle nomeados por tuplas compostas pelo endereço IP fonte, número de porta fonte, endereço IP destino e número de porta destino. Para que seja possível a re-vinculação dinâmica por parte da camada de transporte, várias modificações nessa camada devem ser realizadas. Além disso, a ciência da mobilidade pela camada de transporte é importante, pois diversos controles e estatísticas são computados no interior dessa camada e caso uma mobilidade ocorra esses controles devem ser reinicializados (Eddy, 2004).

Algumas extensões para o protocolo TCP foram propostas, tais como (Huitema, 1995) e (Funato, Yasuda et al., 1997). Entretanto, alterar um protocolo já consolidado, robusto e implantado numa escala global teria uma complexidade alta e talvez não viável, pois o funcionamento de todo código existente e a operação com todos os demais protocolos existentes não poderia ser prejudicada (Henderson e Works, 2003).

Um novo protocolo, chamado SCTP, tem se mostrado um forte candidato para implementar esta funcionalidade. Projetado inicialmente com o objetivo de oferecer confiabilidade e disponibilidade a sistemas de sinalização para telefonia de voz sobre IP (VoIP), o SCTP despertou um grande interesse por parte da comunidade acadêmica e da indústria, devido a uma série de novas características,

que podem resolver uma série de problemas existentes. Ultimamente, o IETF o adotou como um protocolo de propósito geral, a ser estabelecido ao lado do TCP e do UDP, como uma opção para utilização na pilha de protocolos da Internet (Ma, Yu et al., 2004).

O SCTP contempla muitas características do já estabelecido TCP, oferecendo, por exemplo, conexões confiáveis ponto a ponto orientadas à conexão em redes IP. Sua principal nova característica é o multihoming, que permite que múltiplas interfaces de rede, identificadas por diferentes endereços IP, sejam utilizadas em uma mesma associação de conexão entre dispositivos. As interfaces dos dispositivos participantes de uma associação SCTP são configuradas como primárias ou secundárias. Em uma conexão SCTP, para enviar pacotes para o outro dispositivo participante da associação, o protocolo utiliza o endereço destino primário configurado na associação. Caso este endereço se torne inalcançável, o protocolo automaticamente utiliza um endereço secundário. Um caminho entre dois dispositivos composto por endereços primários é chamado de caminho primário e um caminho entre dois dispositivos composto por endereços secundários é chamado de caminho secundário. Só pode existir um único caminho primário em uma associação SCTP entre dois dispositivos, mas vários caminhos secundários são permitidos (Ma, Yu et al., 2004).

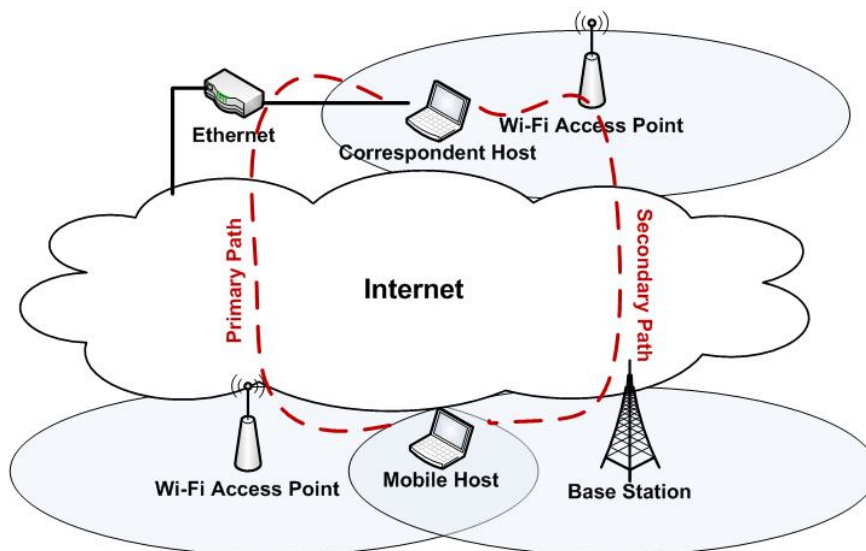


Figura 15 – Exemplo de associação SCTP com um caminho primário e um caminho secundário (Fu, Atiqzaman et al., 2004)

A capacidade de estabelecer caminhos alternativos, utilizados automaticamente pelo protocolo é uma característica muito interessante para dispositivos móveis. Por exemplo, um dispositivo móvel que possui diversas

interfaces de rede disponíveis pode se conectar com um caminho primário através de uma rede Wi-Fi, e ao mesmo tempo ter um caminho alternativo através de uma conexão 3G de sua operadora de celular. O dispositivo, enquanto tivesse dentro da cobertura da rede Wi-Fi, poderia se beneficiar da maior largura de banda, e menor custo de utilização. Quando ele saísse da área de cobertura desta rede, não perderia sua conectividade e as aplicações que estivessem utilizando conexões estabelecidas na camada de transporte não seriam afetadas, pois o endereço secundário de sua interface de rede 3G manteria sua conectividade (Ma, Yu et al., 2004).

No SCTP, os endereços IP pertencentes a uma associação devem ser inicialmente trocados entre os dispositivos antes do estabelecimento da mesma e não podem ser alterados enquanto a associação está ativa. Para que este protocolo possa beneficiar cenários de mobilidade, onde um dispositivo móvel não conhece de antemão os endereços IP de sub-redes que pode se anexar, uma extensão, chamada mSCTP (Mobile SCTP) (Riegel e Tuexen, 2003) foi proposta, que permite a adição, remoção e modificação dos endereços IP dinamicamente durante uma associação (Ma, Yu et al., 2004). Algumas soluções de Gerenciamento de Mobilidade baseadas no mSCTP foram propostas, tais como o TraSH (Transport Layer Seamless Handover) (Fu, Atiquzzaman et al., 2004) e Cellular SCTP (Aydin, Seok et al., 2003).

Soluções de Gerenciamento de Mobilidade na camada de transporte em geral não implementam todas as funcionalidades isoladamente. O gerenciamento de localização, por exemplo, deve ser oferecido por uma camada superior (e.g. a solução Cellular SCTP utiliza o serviço de localização do protocolo SIP para seu gerenciamento de localização). Implementar a mobilidade na camada de transporte oferece diversas vantagens, pois sua não necessita de alterações na infra-estrutura, nem de agentes redirecionadores (tais como no Mobile IP). Além disso, apresentam um menor potencial para problemas com mecanismos de segurança das redes (Eddy, 2004).

3.5. Gerenciamento de Mobilidade na Camada de Aplicação

A implementação da mobilidade na camada de aplicação apresenta grande parte das vantagens oferecidas por uma implementação na camada de transporte. A necessidade de mudanças na infra-estrutura é igualmente pequena, e a complexidade é ainda menor visto que não são necessárias mudanças em protocolos de transporte. Além disso, o gerenciamento de localização em geral é oferecido embutido em soluções nesta camada (Eddy, 2004).

Uma importante diferença entre as soluções de gerenciamento de mobilidade na camada de transporte das soluções em camadas acima desta (e.g. camada de aplicação), é que uma solução na camada de transporte é capaz de suportar transições transparentes mesmo para conexões (e.g. conexões TCP) em andamento, ou seja, aplicações orientadas a conexões de longa duração necessitam de uma solução como esta. Soluções puramente na camada de aplicação não possuem a capacidade de manter conexões de transporte ativas durante uma transição, pois camadas acima da camada de transporte não possuem acesso direto aos mecanismos de controle contidos dentro das conexões abertas na camada de transporte (Eddy, 2004). Entretanto, alguns dos tipos principais de aplicações/serviços/protocolos que são utilizados na Internet hoje em dia apresentam uma forma de comunicação peculiar, tal que mesmo que utilizem conexões TCP, podem se beneficiar da simplicidade de implementação de uma solução na camada de aplicação. Isto porque eles são baseados em transações pequenas recuperáveis, tais como aquelas presentes em aplicações como navegação na web, upload e download de correio eletrônico ou serviços publish/subscribe em que as subscrições e notificações são interações de curta duração. Nestes casos, as conexões de transporte são usualmente pequenas o suficiente para tornar o custo de uma tentativa de re-envio pequeno, no caso de uma interrupção (e.g. no caso de uma mobilidade) (Wedlund e Schulzrinne, 1999).

Considerando a menor complexidade para a implementação de uma solução de Gerenciamento de Mobilidade na camada de aplicação, assim como sua facilidade de implantação, podemos nos beneficiar da simplicidade de implementação de uma solução neste nível das camadas de protocolos. No

capítulo 3 é apresentada a implementação de uma solução de Gerenciamento de Mobilidade na camada de aplicação que utiliza o protocolo SIP (Session Initiation Protocol).

O protocolo SIP possui suporte inerente para a mobilidade pessoal, já que os usuários se registram com endereços independentes de localização através da utilização dos Registrars e Location Services e podem ser encontrados mesmo que estejam utilizando dispositivos diferentes, tais como um notebook, um PDA ou seu computador pessoal. Qualquer requisição é corretamente direcionada para o dispositivo que o usuário está utilizando (Popescu, 2003).

O trabalho de (Wedlund e Schulzrinne, 1999) propõe uma solução para que o SIP suporte também a mobilidade de terminais, permitindo que um dispositivo de um usuário mude de endereço durante uma sessão ativa. Primeiramente, é assumido que o usuário pertence a uma rede doméstica, na qual existe um Registrar doméstico que recebe registros a cada vez que o endereço do User Agent que ele está utilizando muda, de forma semelhante ao registro no Home Agent no Mobile IP (mas com o SIP a entidade sendo rastreada pelo gerenciamento de localização é o usuário e não o dispositivo). Quando um User Agent de um dispositivo correspondente envia um INVITE para um usuário, o Proxy Server possui informações correntes sobre a localização do usuário (endereço real do User Agent que o usuário está utilizando) e redireciona o INVITE para lá.

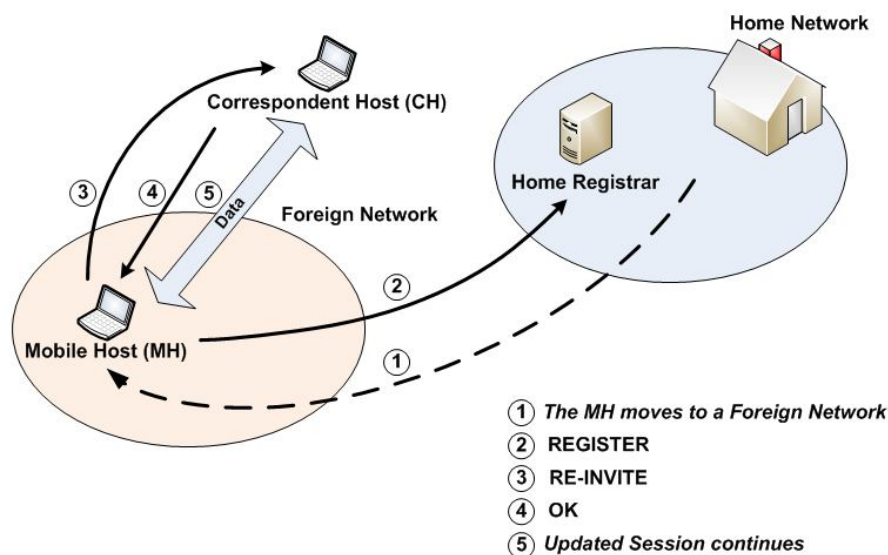


Figura 16 – Exemplo de mobilidade com o protocolo SIP (Wedlund e Schulzrinne, 1999)

Quando um dispositivo se move (i.e. detecta uma mudança em seu endereço IP, através de consultas periódicas ao Sistema Operacional ou de uma notificação

assíncrona recebida do mesmo) durante uma sessão ativa, ele deve enviar um novo convite INVITE (também chamado de REINVITE) diretamente para o UA do outro participante da sessão, indicando no campo Contact da requisição INVITE o novo endereço IP em que o dispositivo deseja receber novas mensagens SIP. Além disso, na parte SDP da sessão, ele deve atualizar os endereços referentes à troca de dados da sessão. O outro UA deve detectar que este novo INVITE não tem o objetivo de iniciar uma nova sessão, mas de atualizar a sessão em andamento. Assim, ele deverá utilizar os dados recebidos para atualizar a sessão e continuar a trocar dados a partir de então.